

GENERAL MATHEMATICS

UNITS 3 & 4

SECOND EDITION

CAMBRIDGE SENIOR MATHEMATICS
FOR QUEENSLAND

KAY LIPSON | MICHAEL EVANS | ROSE HUMBERSTONE
PETER JONES | KYLE STAGGARD

INCLUDES INTERACTIVE
TEXTBOOK POWERED BY
CAMBRIDGE HOTMATHS





Shaftesbury Road, Cambridge CB2 8EA, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press & Assessment is a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org

© Peter Jones, Michael Evans, Kay Lipson and Kyle Staggard 2019

© Kay Lipson, Michael Evans, Rose Humberstone, Peter Jones and Kyle Staggard 2025

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press & Assessment.

First published 2019

Second Edition 2025

20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

Cover designed by Sardine Design

Text designed by diacriTech

Typeset by diacriTech

Printed in Singapore by C & C Offset Printing Co. Ltd.

A catalogue record for this book is available from the National Library of Australia at www.nla.gov.au

ISBN 978-1-009-57800-4

Additional resources for this publication at www.cambridge.edu.au/GO

Reproduction and communication for educational purposes

The Australian *Copyright Act 1968* (the Act) allows a maximum of one chapter or 10% of the pages of this publication, whichever is the greater, to be reproduced and/or communicated by any educational institution for its educational purposes provided that the educational institution (or the body that administers it) has given a remuneration notice to Copyright Agency Limited (CAL) under the Act.

For details of the CAL licence for educational institutions contact:

Copyright Agency Limited

Level 12, 66 Goulburn Street

Sydney NSW 2000

Telephone: (02) 9394 7600

Facsimile: (02) 9394 7601

Email: memberservices@copyright.com.au

Reproduction and communication for other purposes

Except as permitted under the Act (for example a fair dealing for the purposes of study, research, criticism or review) no part of this publication may be reproduced, stored in a retrieval system, communicated or transmitted in any form or by any means without prior written permission. All inquiries should be made to the publisher at the address above.

Cambridge University Press & Assessment has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate. Information regarding prices, travel timetables and other factual information given in this work is correct at the time of first printing but Cambridge University Press & Assessment does not guarantee the accuracy of such information thereafter.

Cambridge University Press & Assessment acknowledges the Aboriginal and Torres Strait Islander peoples of this nation. We acknowledge the traditional custodians of the lands on which our company is located and where we conduct our business. We pay our respects to ancestors and Elders, past and present. Cambridge University Press & Assessment is committed to honouring Aboriginal and Torres Strait Islander peoples' unique cultural and spiritual relationships to the land, waters and seas and their rich contribution to society.

Contents

About the lead author and consultants	viii
Introduction and overview	ix
Acknowledgements	xii
1 Bivariate data analysis 1	1
1A Bivariate data – classifying the variables	2
1B Investigating associations between two categorical variables	8
1C Displaying bivariate data from two numerical variables – the scatterplot	20
1D Interpreting a scatterplot	27
1E A measure of strength for a linear relationship – the correlation coefficient	37
1F The coefficient of determination	46
Review of Chapter 1	52
Key ideas and chapter summary	52
Skills checklist	52
Multiple-choice questions	54
Short-response questions	57
2 Bivariate data analysis 2	61
2A Fitting a least squares line to numerical data	62
2B Using the least squares line to model a linear relationship . .	72
2C Association and causation	92
2D Solving practical problems by identifying, analysing and describing associations	97
Review of Chapter 2	104
Key ideas and chapter summary	104
Skills checklist	105
Multiple-choice questions	106
Short-response questions	110

3	Time series analysis	115
3A	Time series data	116
3B	Smoothing a time series using moving means	129
3C	Smoothing a time series plot using moving medians	137
3D	Seasonal indices	143
3E	Fitting a trend line and forecasting	157
	Review of Chapter 3	163
	Key ideas and chapter summary	163
	Skills checklist	164
	Multiple-choice questions	166
	Short-response questions	170
4	Arithmetic and geometric sequences	174
4A	Using recursion to generate an arithmetic sequence	175
4B	Defining an arithmetic sequence by recursion	179
4C	A general rule for finding the n th term of an arithmetic sequence	188
4D	Application of arithmetic sequences	194
4E	Geometric sequences	203
4F	A general rule for finding the n th term of a geometric sequence	210
4G	Applications of geometric sequence	215
	Review of Chapter 4	225
	Key ideas and chapter summary	225
	Skills checklist	227
	Multiple-choice questions	229
	Short-response questions	232
5	Earth geometry and time zones	235
5A	Angle measurement and arc length	236
5B	Latitude and longitude	241
5C	Time zones and time differences	260
	Review of Chapter 5	266
	Key ideas and chapter summary	266
	Skills checklist	266

	Multiple-choice questions	267
	Short-response questions	268
6	Revision of Unit 3 Chapters 1–5	270
6A	Topic 1: Bivariate data analysis 1	271
6B	Topic 2: Bivariate data analysis 2	279
6C	Topic 3: Time series analysis	288
6D	Topic 4: Growth and decay in sequences	295
6E	Topic 5: Earth geometry and time zones	298
6F	List of Unit 3 assessment and examination practice online items	302
7	Loans, investments and annuities 1	303
7A	Using a recurrence relation to model compound interest loans and investments	304
7B	Using a compound interest formula to model a compound interest loan or investment	317
7C	Effective annual rate of interest	322
7D	Practical problems involving compound interest loans and investments	328
7E	Using a recurrence relation to model the present value of an ordinary annuity	334
7F	Investigating reducing-balance loans	342
7G	Using the present value annuity formula to model the present value of an ordinary annuity	348
7H	Solving practical problems involving the present value annuity formula	352
	Review of Chapter 7	358
	Key ideas and chapter summary	358
	Skills checklist	360
	Multiple-choice questions	362
	Short-response questions	365
8	Loans, investments and annuities 2	370
8A	A recursive model for annuities	371
8B	Investigating annuities	382
8C	Using the future value annuity formula	389

8D	Solving practical problems involving the future value annuity formula	395
8E	Perpetuities	402
8F	Solving practical problems involving perpetuities	406
	Review of Chapter 8	411
	Key ideas and chapter summary	411
	Skills checklist	412
	Multiple-choice questions	413
	Short-response questions	417
9	Graphs and networks	422
9A	Graphs and associated terminology	423
9B	The adjacency matrix	432
9C	Directed graphs and their adjacency matrices	437
9D	Planar graphs and Euler's formula	443
9E	Exploring a graph	450
9F	Eulerian graphs and applications	455
9G	Hamiltonian graphs and applications	461
9H	Weighted graphs, networks and shortest path problems	467
	Review of Chapter 9	472
	Key ideas and chapter summary	472
	Skills checklist	474
	Multiple-choice questions	476
	Short-response questions	481
10	Networks and decision mathematics 1	486
10A	Trees and connector problems	487
10B	Project planning – precedence tables and activity networks	497
10C	Scheduling problems	507
10D	Applications of critical path analysis	527
10E	Altering the duration of an activity	531
	Review of Chapter 10	541
	Key ideas and chapter summary	541
	Skills checklist	543
	Multiple-choice questions	544
	Short-response questions	548

11	Networks and decision mathematics 2	555
11A	Flow networks	556
11B	Bipartite graphs and assignment problems	567
11C	The Hungarian algorithm	573
	Review of Chapter 11	584
	Key ideas and chapter summary	584
	Skills checklist	585
	Multiple-choice questions	585
	Short-response questions	587
12	Revision of Unit 4 Chapters 7–11	591
12A	Topic 1: Loans, investments and annuities 1	592
12B	Topic 2: Loans, investments and annuities 2	596
12C	Topic 3: Graphs and networks	600
12D	Topic 4: Networks and decision mathematics 1	608
12E	Topic 5: Networks and decision mathematics 2	616
13	Revision of Units 3 & 4 Chapters 1–11	623
13A	Paper 1 revision questions	624
13B	Paper 2 revision questions	644
13C	List of Unit 4 and Units 3 & 4 assessment and examination practice online items	656
A	Appendix A: The problem-solving and modelling task	657
A1	About the problem-solving and modelling task	657
A2	A content guide for a PSMT report	658
	Glossary	672
	Answers	679

Online appendices accessed through the Interactive or PDF Textbook

Included in the Interactive and PDF Textbook only

Appendix B Online guides to using technology

Appendix C Guide to PSMTs

About the lead author and consultants

About the lead author

Kay Lipson has a PhD in Statistics Education, and Diploma of Education. Her experience includes over 30 years teaching mathematics and statistics at both secondary and tertiary level, as well as extensive periods as a VCE examiner. Kay is an experienced author of textbooks for the Years 11 and 12 Mathematics courses, specialising in the areas of Probability and Statistics.

About the consultants

As at the time this edition was published:

Joel Speranza is currently working with the Toowoomba Catholic Schools Office, delivering General Mathematics hybridly to students from eight different regional schools. He also creates supporting material for the senior mathematics curriculum through the website Maths Video Australia.

Nicole Ross is a Teacher of Mathematics, Somerville House

Danielle Galpin is a Teacher of Mathematics, Villanova College

Fiona Steele is Head of Mathematics & Digital Technology, St Michael's College

Melanie Chin is a senior mathematics teacher at Brisbane Grammar School, and Secretary of the Queensland Association of Mathematics Teachers

Introduction and overview

Cambridge Senior Mathematics for Queensland General Mathematics Units 3 & 4 provides complete and aligned coverage of the QCAA syllabus to be implemented in Year 12 from 2025. Its four components – the print book, downloadable PDF textbook, online Interactive Textbook (ITB) and Online Teaching Resource (OTS) – contain a huge range of resources, available to schools in a single package.

New features in the second edition include:

- **Learning intentions** complemented by a Skills Checklist at the end of each chapter that allows you to check your understanding and tick off your achievements
- **Section summaries** at the end of each section further learning intentions by highlighting the key points and skills needed to progress
- **The problem-solving and modelling task (Appendix A):** This is a new appendix at the end of the book, written by consultant Joel Speranza, providing advice on how to complete problem-solving and modelling tasks (PSMTs). This is supported by video resources accessed through QR codes and in the Interactive Textbook.

The Second Edition also features significantly revised and updated material from the first edition, including:

Detailed worked examples which go step by step through a problem together with precise explanations. For every worked example also has a video to encourage independent learning and is linked to exercises.

Degree of difficulty classification of questions: in the exercises, questions are classified as simple familiar **SF**, complex familiar **CF**, or complex unfamiliar **CU** questions and are indicated by a strip along the margin. The revision chapters described below also contain model questions for each of these categories.

Two revision chapters each covering one unit. Revision chapters are divided into simple familiar multiple-choice and short-response questions that have been designed with each degree of difficulty in mind to help prepare students for assessment.

Assessment: Examination practice questions and various assessment tasks are provided at the end of each exercise, in the revision chapters and the Online Teaching Suite.

Interactive Textbook (ITB)

The Interactive Textbook (ITB) is an online HTML version of the print textbook powered by the HOTmaths platform, included with the print book or available as a separate purchase.

Updated and revised for the new syllabus, the Interactive Textbook includes:

- **Video demonstrations** of all worked examples
- **Quick quizzes** containing auto-marked multiple-choice questions that have been thoroughly updated and revised, enabling students to check their understanding.
- A **success criteria** checklist at the end of each chapter with linked questions and examples available for download
- Comprehensive **worked solutions** for all questions are provided in the Interactive Textbook as an option that teacher can choose to enable for their students.
- **Downloadable skillsheets** can be used for homework or in class to focus on a single skill or small set of related skills.
- **Definitions** pop up for key terms in the text, and are also provided in a dictionary.

The Online Teaching Suite (OTS)

The Online Teaching Suite is automatically enabled with a teacher account and is integrated with the teacher's copy of the Interactive Textbook. All the teacher resources are in one place for easy access. The features include:

- A teacher's view of a student's working and self-assessment which enables them to modify the student's self-assessed marks, and respond where students flag that they had difficulty.
- The **task manager** allowing to direct students on a custom activity sequence based on their scores in measurable activities
- Quickly create customised tests from a bank of multiple-choice questions using the **test generator**. Tests are auto-marked in the Interactive Textbook or can be printed and used for homework or assessment practice.
- An expanded and revised suite of **chapter tests** and **assignments**
- Editable **curriculum grids** and **teaching programs**.
- A brand-new **Exam Generator**, allowing the creation of customised printable and online trial exams (see the following page for more).

More about the Exam Generator

A new Exam Generator, available at no extra charge within the Online Teaching Suite, will include a comprehensive bank of QCAA exam questions, augmented by exam-style questions written by experts, to allow teachers to create custom trial exams.

Custom exams can model end-of-year exams, or target specific topics or types of questions that students may be having difficulty with.

Features include:

- Filtering by question-type, topic, chapter and degree of difficulty
- Searchable by key words
- Answers provided to teachers
- Worked solutions for all questions
- QCAA marking scheme
- Multiple-choice exams can be auto-marked if completed online, with filterable reports
- All custom exams can be printed and completed under exam-like conditions or used as revision.

Acknowledgements

Thanks to **Joel Speranza** for the creation of the PSMT appendix and supplementary supporting videos.

The author and publisher wish to thank the following sources for permission to reproduce material:

Images: © Getty Images / Andriy Onufriyenko, Chapter 1 Opener / oxygen, Chapter 2 Opener / SENEZ, Chapter 3 Opener / kirstypargeter, Chapter 4 Opener / by Ruhey, Chapter 5 Opener / Anna Efetova, Chapter 6 Opener / Sylverarts, Chapter 7 Opener / gremlin, Chapter 8 Opener / PM Images, Chapter 9 Opener / filo, Chapter 10 Opener / SEGU JEON, Chapter 11 Opener / Photo by Alex Tihonov, Chapter 12 Opener / Olena Malik, Chapter 13 Opener.

Every effort has been made to trace and acknowledge copyright. The publisher apologises for any accidental infringement and welcomes information that would redress this situation.

General Mathematics Syllabus 2025, © State of Queensland (QCAA) 2019, licensed under CC BY 4.0, <https://www.qcaa.qld.edu.au/copyright>.

Chapter 1

Bivariate data analysis 1

Chapter questions

UNIT 3 BIVARIATE DATA AND TIME SERIES ANALYSIS, SEQUENCES AND EARTH GEOMETRY

Topic 1: Bivariate data analysis 1

- ▶ How do we define bivariate data?
- ▶ How do we construct two-way tables?
- ▶ How do we interpret and identify patterns in two-way tables using percentages?
- ▶ How do we construct a scatterplot?
- ▶ How do we describe an association between two numerical variables in terms of direction, form and strength?
- ▶ How do we calculate and interpret the correlation coefficient?

Questions such as ‘Is the new treatment for a cold more effective than the old treatment’, or ‘Do younger people spend more time using social media than older people’ are concerned with understanding the association between two variables. In this chapter we begin our study of **bivariate data**, data which is recorded on two variables for the same subject.

1A Bivariate data – classifying the variables

Learning intentions

- ▶ To introduce bivariate data.
- ▶ To be able to classify data as categorical or numerical.
- ▶ To be able to identify explanatory and response variables.

To determine how to answer questions involving two variables requires the variables to be clearly defined.

Categorical and numerical variables

You will recall from General Mathematics in Year 11 we defined two classifications of variables, **categorical** and **numerical** variables:

- Categorical variables generate data values that are names or labels, such as *sex* (male, female) or *coffee size* (small, medium, large).
- Numerical variables generate data values that are numbers, usually resulting from counting or measuring, such as *number of brothers* (0, 1, 2, ...) or *hand span* (cm).



Example 1 Identifying variables as categorical or numerical

Identify each of the following variables as either categorical or numerical:

- a** weight (kg)
- b** favourite colour
- c** support for same sex marriage (yes, no)
- d** number of pine trees per acre of forest
- e** attitude to lowering the driving age (strongly agree, agree, no opinion, disagree, strongly disagree)

Solution

- a** Weight - Numerical
- b** Favourite colour - Categorical
- c** Support for same sex marriage - Categorical
- d** Number of pine trees per acre of forest - Numerical
- e** Attitude to lowering the driving age - Categorical

Explanation

The data values arise from measuring.
The data values are labels such as red or blue.
The data values are labels.
The data values arise from counting.
The data values are labels.

The first step in investigating the association between two variables is to classify each variable as either categorical or numerical. What can we say about the types of variables involved in the following question?

‘Is the new treatment for headache more effective than the old treatment?’

To investigate this question requires firstly a definition of ‘effective’. Suppose that the effectiveness of the treatment is to be measured by the time it takes for the headache to be relieved, measured in minutes. Then the two variables in this question are *type of treatment*, a categorical variable taking the values ‘new’ and ‘old’, and *time taken for the headache to be relieved*, a numerical variable. Thus, investigation of a question like this can be classified as **investigating the association between a categorical variable and a numerical variable**.

‘Are city voters more likely to vote for the Greens party than country voters?’

This question involves two variables, *place of residence*, which is a categorical variable taking the values ‘city’ and ‘country’, and *vote for the Greens*, which also is a categorical variable taking the values ‘yes’ and ‘no’. Investigation of a question like this can be classified as **investigating the association between two categorical variables**.

‘Are younger people more knowledgeable about environmental issues than older people?’

If the *age* of the respondent is measured in years, then *age* is a numerical variable. If *knowledge of environmental issues* is measured with a sequence of questions and the respondent is given a score out of 100, then *knowledge of environmental issues* is also a numerical variable. Investigation of a question like this can be classified as **investigating the association between two numerical variables**.



Example 2 Identifying associations as categorical or numerical

For each of the following questions, determine if they involve investigating associations between:

- one numerical variable and one categorical variable or
 - two categorical variables or
 - two numerical variables.
- a** Are younger people (age measured in years) more likely to believe in astrology (measured as ‘yes’ or ‘no’) than older people?
 - b** Do students who spend more hours studying each week get higher test scores?
 - c** Are people who have a driver’s licence (measured as ‘yes’ or ‘no’) more likely to be in favour of lowering the driving age (measured as ‘yes’ or ‘no’)?

Solution

- a** One numerical variable (*age*) and one categorical variable (*belief in astrology*)
- b** Two numerical variables (*hours studied per week* and *test score*)
- c** Two categorical variables (*have a driver's licence* and *support for lowering the driving age*)

Identifying response and explanatory variables

The second step in investigating the association between two variables is to determine which of the two variables is the explanatory variable and which is the response variable. We use the explanatory variable to explain the associated changes in the response variable. For example:

- 'Is the new treatment for headache more effective than the old treatment?'
 - *Type of treatment* is the **explanatory variable** as it may explain any changes in *time taken for the headache to be relieved*.
 - *Time taken for the headache to be relieved* is the **response variable** as changes could occur in response to the *type of treatment* used.
- 'Are city voters more likely to vote for the Greens party than county voters?'
 - *Place of residence* is the explanatory variable as knowing a person's place of residence might be useful in explaining *vote for the Greens*.
 - *Vote for the Greens* is the response variable as differences could occur in response to the *place of residence* of the voter.
- 'Are younger people more knowledgeable about environmental issues than older people?'
 - *Age* is the explanatory variable as it may explain any changes in *knowledge of environmental issues*.
 - *Knowledge of environmental issues* is the response variable as changes could occur in response to the *age* of the person.



Example 3 Identifying response and explanatory variables

We wish to investigate the question, 'Does the time it takes a student to travel to school depend on their mode of transport?' The variables here are *time* and *mode of transport*. Which is the response variable (RV) and which is the explanatory variable (EV)?

Solution

EV: *mode of transport*
RV: *time*

Explanation

In asking the question in this way, we are suggesting that a student's *mode of transport* might explain the differences we observe in the time it takes students to travel to school.



Example 4 Identifying response and explanatory variables

Can we predict people's height (in cm) from their wrist circumference (cm)? The variables in this investigation are *height* and *wrist measurement*. Which is the response variable (RV) and which is the explanatory variable (EV)?

Solution

EV: *wrist measurement*

RV: *height*

Explanation

Since we wish to predict height from wrist circumference, we are using *wrist measurement* as the predictor or explanatory variable. *Height* is then the response variable.

It is important to note that, in Example 4, we could have asked the question the other way around; that is, 'Can we predict people's wrist measurement from their height?' In that case *height* would be the explanatory variable, and *wrist measurement* would be the response variable. The way we ask our statistical question is an important factor when there is no obvious explanatory variable.

Response and explanatory variables

When investigating the association between two variables, the **explanatory variable (EV)** is the variable we expect to **explain** or **predict** the value of the **response variable (RV)**.

Note: The explanatory variable is sometimes called the independent variable (IV) and the response variable the dependent variable (DV).

Section Summary

- ▶ **Bivariate data** is obtained when information about two variables is recorded for each subject.
- ▶ Each variable can be classified as **categorical** or **numerical**.
- ▶ Three different situations are possible which analysing bivariate data:
 - ▷ both variables are categorical, or
 - ▷ one variable is categorical and one is numerical, or
 - ▷ both variables are numerical.
- ▶ When investigating the association between two variables, it is helpful to identify which is the **explanatory variable** and which is the **response variable**.

Exercise 1A

Identifying variables as categorical or numerical

Example 1

- 1** Identify each of the following variables as either categorical or numerical.
 - a** income (low, medium, high)
 - b** favourite TV show
 - c** time taken to drive to work in minutes
 - d** emotional intelligence, as measured on a standardised psychological test on a scale from 1–100
 - e** self-assessed state of health
(1 = excellent, 2 = good, 3 = satisfactory, 4 = poor, 5 = very poor)
 - f** temperature ($^{\circ}\text{C}$)
 - g** weekly salary (\$)
 - h** weekly salary (1 = below average, 2 = average, 3 = above average)
 - i** weekly salary (less than \$500, \$500–\$999, \$1000–\$1999, more than \$2000)

Identifying associations as categorical or numerical

Example 2

- 2** For each of the following questions, determine if they involve investigating associations between:
 - one numerical and one categorical variable or
 - two categorical variables or
 - two numerical variables.
 - a** Are males and females equally likely to be in favour of same sex marriage?
 - b** Do Year 11 students watch more hours of television each week than Year 12 students?
 - c** Do countries with higher household incomes (\$) tend to have lower infant mortality rates (deaths/1000 births)?
 - d** Is there a relationship between attitude to gun control and country of birth?

Identifying response and explanatory variables

Example 3

For each of the following situations identify the explanatory variable (EV) and the response variable (RV). In each situation the variable names are *italicised*.

Example 4

- 3**
 - a** We wish to investigate whether a fish's *toxicity* can be predicted from its *colour*.
 - b** The relationship between *weight loss* and *type of diet* is to be investigated.
 - c** We wish to investigate the relationship between a used car's *age* and its *price*.
 - d** It is suggested that the *cost* of heating in a house depends on the type of *fuel* used.
 - e** The relationship between the *house price* and its *location* is to be investigated.

- 4 The following pairs of variables are related. Which is likely to be the explanatory variable? The variable names are italicised.
- a *exercise level* and *age*
 - b *years of education* and *salary level*
 - c *comfort level* and *temperature*
 - d *time of year* and *incidence of hay fever*
 - e *age group* and *musical taste*
 - f *AFL team supported* and *state of residence*
- 5 For each of the following pairs of variables, determine:
- which are numerical and which are categorical, and
 - which is the explanatory variable and which is the response variable.
- a *sex* and *attitude to lowering the legal drinking age*
 - b *hours of study per week* and *hours spent per week using social media*
 - c *gestation time* and *birth weight of babies*
 - d *sex* and *hours spent per week using social media for Year 12 students*
 - e *voting preference* (Liberal, Labor, Greens, other) and *support for tax cuts*

Paper 1-style multiple-choice questions

- 6 Respondents to a survey question ‘How concerned are you about climate change’ were asked to select from the following responses:
- 1 = not at all, 2 = a little, 3 = moderately, 4 = extremely
- The data which was collected in response to this question is:
- A numerical B metric C categorical D continuous
- 7 The variables *weight* (light, medium, heavy) and *height* (less than 160 cm, 160 – 175 cm, over 175 cm) are:
- A both categorical variables
 - B a categorical and a numerical variable respectively
 - C a numerical and a categorical variable respectively
 - D both numerical variables
- 8 Researchers believe that reaction time might be lower in colder temperatures. They devise an experiment where *reaction time* in seconds is measured at three different *temperature* levels (1 = less than 8°C, 2 = from 8°C to 18°C, 3 = more than 18°C). The response variable, and its classification are:
- A *reaction time*, categorical
 - B *temperature*, categorical
 - C *reaction time*, numerical
 - D *temperature*, numerical

1B Investigating associations between two categorical variables

Learning intentions

- ▶ To summarise data from two categorical variables using **two-way frequency tables**.
- ▶ To appropriately percentage a two-way frequency table.
- ▶ To use a percentaged two-way frequency table to identify patterns that suggest the presence of an association.

To begin our analysis of data arising from two categorical variables we will introduce a table used to summarise bivariate data.

Constructing a two-way frequency table

It has been suggested that city and country people have differing attitudes to gun control; that is, that support for gun control depends on where a person lives. How might we investigate this relationship? Suppose we ask a sample of three people about their *attitude to gun control*, and we also record their *residence*. The resulting data for the three people might look like this:

Subject no.	Residence	Attitude to gun control
1	City	For
2	Country	For
3	City	Against

The first thing to note is that these two variables, *attitude to gun control* (for or against) and *residence* (city or country), are both categorical variables. Categorical data are usually presented in the form of a frequency table.

Suppose we continue until we have interviewed a sample of 100 people, and we find that there are 58 who live in the country and 42 who live in the city. We can present this result in a frequency table as shown to the right.

Residence	Frequency
Country	58
City	42
Total	100

From this table, we can see that there were more country than city people in our sample.

Suppose also when we record the attitude to gun control, we might have 62 ‘for’ and 38 ‘against’ gun control. Again, we could present these results in a frequency table as shown to the right.

Attitude to gun control	Frequency
For	62
Against	38
Total	100

From this table, we can see that more people in the sample were for gun control than against gun control. However, we cannot tell from the information contained in the tables whether *attitude to gun control* depends on the *residence* of the person. To do this we need to construct a **two-way frequency table**, which gives both the *attitude to gun control* and the *residence* for each person in the sample.

We begin by counting the number of people in the sample who are:

- from the country and for gun control
- from the city and for gun control
- from the country and against gun control
- from the city and against gun control.

Suppose again from our sample of 100 people we find the following frequencies:

- 32 country people are for gun control
- 30 city people are for gun control
- 26 country people are against gun control
- 12 city people are against gun control.

Explanatory and response variables in two-way frequency tables

Before we set up the two-way frequency table, we need to decide which of the two variables is the explanatory variable and which is the response variable. Since we think that a person's attitude to gun control might depend on their place of residence, but not the other way around, then:

- *residence* is the explanatory variable (EV)
- *attitude to gun control* is the response variable (RV).

In two-way frequency tables, it is conventional to let the categories of the *response variable* label the *rows* of the table and the categories of the *explanatory variable* label the *columns* of the table. Following this convention, we can create the following two-way frequency table.

	<i>Residence</i>	
<i>Attitude to gun control</i>	Country	City
For	32	30
Against	26	12

To complete the table, it is usual to calculate the row and column sums.

	<i>Residence</i>		
<i>Attitude to gun control</i>	Country	City	Total
For	32	30	62
Against	26	12	38
Total	58	42	100

Row sum
Row sum

Column sum Column sum

The shaded regions in the table are called the **cells** of the table. It is the numbers in these cells that we look at when investigating the relationship between the two variables.



Example 5 Constructing a two-way frequency table

The following data were obtained when gender and intention to go to university (*university*) were recorded for 10 Year 9 students. Create a two-way frequency table from these data.

Student	Gender	University	Student	Gender	University
1	Female	Yes	6	Male	Yes
2	Male	Yes	7	Female	Yes
3	Female	No	8	Male	No
4	Female	Yes	9	Female	No
5	Male	No	10	Female	Yes

Solution

	Gender	
University	Male	Female
Yes		
No		

	Gender	
University	Male	Female
Yes		
No		

	Gender	
University	Male	Female
Yes		
No		

Explanation

It is possible that a student's intention to go to university may depend on their gender, but not the other way around. Thus, *gender* is the explanatory variable and *intends to go to university* is the response variable. Create the table showing the values of *gender* labelling the columns, and *university* labelling the rows.

Consider Student 1, who is female and indicated 'yes' to go to university. Place a mark in the corresponding cell of the table.

Go through the data set one person at a time, placing a mark in the appropriate cell for each person.

	<i>Gender</i>		
<i>University</i>	Male	Female	Total
Yes	2	4	6
No	2	2	4
Total	4	6	10

Finally, tally the marks in each cell, and calculate the row and columns sums. Make sure the total adds to the number of students in the sample.

Consider again the two-way frequency table created to investigate the association between place of residence and attitude to gun control. This table tells us that more country people are in favour of gun control than city people. But is this just due to the fact that there were more country people in the sample? To help us answer this question we need to express the frequencies in each cell as **percentage frequencies**.

Percentaged two-way frequency tables

When the two-way frequency table has been constructed so that the values of the explanatory variable label the rows, we calculate **column percentages** to help us investigate the association. This will give us separately the percentages of country and city people for and against gun control, which can then be compared.

Column percentages are determined by dividing each of the cell frequencies by the relevant column sums. Thus, the percentage of:

- country people who are for gun control is: $\frac{32}{58} \times 100 = 55.2\%$
- country people who are against gun control is: $\frac{26}{58} \times 100 = 44.8\%$
- city people who are for gun control is: $\frac{30}{42} \times 100 = 71.4\%$
- city people who are against gun control is: $\frac{12}{42} \times 100 = 28.6\%$

Note: Unless small percentages are involved, it is usual to round percentages to one decimal place in tables.

	<i>Residence</i>	
<i>Attitude to gun control</i>	Country	City
For	55.2%	71.4%
Against	44.8%	28.6%
Total	100.0%	100.0%

Using percentages to identify relationships between variables

Calculating the values in the table as percentages enables us to compare the attitudes of city and country people on an equal footing. From the table, we see that 55.2% of country people in the sample were for gun control compared to 71.4% of the city people. This means that

the city people in the sample were more supportive of gun control than the country people. This reverses what the frequencies showed.

The fact that the percentage of ‘country people for gun control’ differs from the percentage of ‘city people for gun control’ indicates that a person’s attitude to gun control depends on their residence. Thus, we can say that the variables *attitude to gun control* and *residence* are **associated**.

If the variables *attitude to gun control* and *residence* were not associated, we would expect approximately equal percentages of country people and city people to be ‘for’ gun control. Finding a single row in the two-way frequency distribution in which percentages are clearly different is sufficient to identify a relationship between the variables.

We could have also arrived at this conclusion by focusing our attention on the percentages ‘against’ gun control. We might report our findings as follows.

Report: In this sample of 100 people, a higher percentage of city people (71.4%) supported gun control than country people (55.2%). This indicates that a person’s attitude to gun control is associated with their place of residence.

Note: Finding a single row in the two-way frequency distribution in which percentages are clearly different is sufficient to identify a relationship between the variables.

We will now consider a two-way percentage frequency table that shows no evidence of a relationship. Consider the following table that summarises responses to the question ‘Should mobile phones be banned in cinemas?’ These responses were obtained from 100 students in Year 10 and Year 12 – we are interested in investigating whether there is an association between these variables.

	<i>Year level</i>	
<i>Should mobile phones be banned in cinemas?</i>	Year 10	Year 12
Yes	87.9%	86.8%
No	12.1%	13.2%
Total	100.0%	100.0%

When we look across the first row of the table, we see that the percentages in favour are very similar. In this case, we might report our findings as follows.

Report: In this sample of 100 Year 10 and Year 12 students, we see that the percentage of Year 10 and Year 12 students in support of banning mobile phones in cinemas is similar: 87.9% to 86.8%. This indicates that a person’s support for banning mobile phones in cinemas is not associated with their year level.


Example 6 Identifying and describing an association from a two-way percentage frequency table

Are males and females in Year 9 equally likely to indicate an intention to go to university? Data from interviews with 200 Year 9 students are summarised in the following table. Write a brief report addressing this question and quoting appropriate percentages.

	Gender		
University	Male	Female	Total
Yes	50	54	104
No	55	41	96
Total	105	95	200

Solution

	Gender	
University	Male	Female
Yes	47.6%	56.8%
No	52.4%	43.2%
Total	100.0%	100.0%

We can see from the top row that a higher percentage of females than males (56.8% compared with 47.6%) intend to go to university.

Report:

In this sample of 200 Year 9 students, a higher proportion of females than males (56.8% compared with 47.6%) were intending to go to university. There is an association between gender and intention to go to university.

Explanation

Determine the column percentages as follows:

$$\% \text{ of males} = \frac{50}{105} \times 100 = 47.6\%$$

Complete the table as shown.

Select an appropriate row to compare the male and female percentages.

Construct a report.

Two-way frequency tables for categorical variables with more than two categories

The two-way percentage frequency table on the following page displays the smoking status for 500 adults (smoker, past smoker, never smoked) by highest level of education (Year 9 or less, Year 10 or 11, Year 12, University).

<i>Smoking status</i>	<i>Education level</i>			
	Year 9 or less	Year 10 or 11	Year 12	University
Smoker	34.0	31.7	26.5	18.4
Past smoker	36.0	33.8	30.9	28.0
Never smoked	30.0	34.5	42.6	53.6
Total	100.0	100.0	100.0	100.0

The following report has been prepared comparing the percentages in the ‘Smoker’ row.

Report: In this sample of 200 adults the percentage of smokers steadily decreases with education level, from 34.0% for Year 9 or below, decreasing to 31.7% for those with Year 10 or 11, 26.5% for those with Year 12, and lowest at 18.4% for those with university level education. This indicates that smoking is associated with level of education.



Example 7 Identifying and describing associations from a percentaged two-way table (3×3)

A survey was conducted with 1000 males under 50 years old. As part of this survey, they were asked to rate their interest in sport as ‘high’, ‘medium’, and ‘low’. Their age group was also recorded. The results are displayed in the table.

<i>Interest in sport</i>	<i>Age group (%)</i>			
	Under 18 years	19–25 years	26–35 years	36–50 years
High	56.5	50.2	40.7	35.0
Medium	30.1	34.4	36.8	44.7
Low	13.4	15.4	22.5	20.3
Total	100.0	100.0	100.0	100.0

- Which is the explanatory variable, *interest in sport* or *age group*?
- Write a brief report describing the association between *interest in sport* and *age group*.

Solution

- Age group* is the EV.
- Report:** In this sample of 100 males there is an association between the level of interest in sport and age. The percentage of males with a high level of interest in sport is seen to decrease steadily across the age categories from 56.5% for under 18 years, 50.2% for 19–25 years, 40.7% for 26–35 years to, at its lowest, 35% for 36–50 years.

Explanation

Age is a possible explanation for the level of interest in sport, but interest in sport cannot explain age.

If we look across all rows, we can see that the percentages are different for each age group. Select one row to compare and discuss – here we have chosen ‘high’.

Section Summary

- ▶ When both variables are categorical, bivariate data can be summarised in a **two-way frequency table**.
- ▶ If values of the explanatory variable label the columns, and the value of the response variable labels the rows, then **column percentages** should be calculated.
- ▶ When column percentages are calculated, the association is identified by **comparing the percentages across a row**.



Exercise 1B

Constructing a two-way frequency table

Example 5

- 1 The following data were obtained when a sample of 20 Year 12 students were asked if they intended to go to university (*University*). *Gender* of the student was also recorded.

Student	Gender	University	Student	Gender	University
1	F	Yes	11	F	Yes
2	M	Yes	13	M	Yes
3	F	No	13	F	No
4	F	Yes	14	F	Yes
5	M	No	15	M	No
6	M	Yes	16	M	Yes
7	F	Yes	17	F	Yes
8	M	No	18	M	No
9	F	No	19	F	No
10	F	Yes	20	F	Yes

- a Identify which variable is the explanatory and which is the response variable.
- b Create a two-way frequency table from the data, with the values of the explanatory variable labelling the columns.

- 2 The data in the following table were obtained when a sample of 30 adults were asked if they supported reducing university fees (*reduce fees*). They were also classified by their *age group* 17–18 years, 19–25 years, or 26 years or more.

Age group	Reduce fees	Age group	Reduce fees	Age group	Reduce fees
17–18	Yes	26 or more	Yes	26 or more	No
19–25	Yes	17–18	Yes	19–25	Yes
26 or more	No	19–25	Yes	17–18	No
17–18	Yes	17–18	Yes	26 or more	Yes
19–25	Yes	17–18	Yes	17–18	No
26 or more	Yes	26 or more	No	26 or more	Yes
17–18	Yes	19–25	Yes	19–25	Yes
19–25	No	26 or more	Yes	17–18	Yes
26 or more	No	17–18	No	19–25	No
19–25	No	17–18	Yes	26 or more	Yes

- Identify which variable is the explanatory and which is the response variable.
- Create a two-way frequency table from these data, with the values of the explanatory variable labelling the columns.
- Calculate the column percentages for the table.

Using two-way tables to describe associations between two categorical variables

Example 6

- 3 A survey was conducted with 242 university students. For this survey, data were collected on the students' *enrolment status* (full-time, part-time) and whether or not each *drinks alcohol* ('Yes' or 'No'). Their responses are summarised in the table below.

Drinks alcohol	Enrolment status (%)	
	Full-time	Part-time
Yes	80.5	81.8
No	19.5	18.2
Total	100.0	100.0

- Which variable is the explanatory variable?
- What percentage of part-time students drink alcohol?
- Use the information in the table to complete the following report.

Report: In this sample of students we see that the percentage of part-time students who drink alcohol is similar to the percentage of full-time students are who drink alcohol . This indicates that whether or not a student drinks alcohol is associated with their enrolment status.

- 4 Are those people who are satisfied with their job more likely to be satisfied with their life? Data collected from a survey of 110 adults are summarised in the two-way frequency table below.

	Satisfaction with job		
Satisfaction with life	Dissatisfied	Satisfied	Total
Dissatisfied	36	14	50
Satisfied	12	48	60
Total	48	62	110

- a Identify the explanatory variable and response variables.
- b Determine appropriate percentages.
- c Use the information in the table to complete the following report. Report: In this sample of 110 adults there is an association between job satisfaction and life satisfaction. % of people who were satisfied with their job were satisfied with their lives, much higher than the % of people who were dissatisfied with their job who were satisfied with their lives.
- 5 It has been suggested that females might be more satisfied with their lives overall than males. Data were collected from a sample of 360 adults and are summarised in the two-way frequency table below.

	Gender		
Satisfied with life?	Female	Male	Total
Yes	153	155	308
No	24	28	52
Total	177	183	360

- a Identify the explanatory and response variables.
- b Determine appropriate percentages.
- c Use the information in the table to write a brief report describing the association between gender and satisfaction with life.
- 6 The table below was constructed from data collected to see if *handedness* (left, right) was associated with *gender* (male, female).

	Gender %	
Handedness	Male	Female
Left	22	16
Right	222	147

- a Determine appropriate percentages.
- b Write a brief report describing the association between *handedness* and *gender*.

- 7** A survey was conducted with 59 male and 51 female university students to determine whether they exercised ‘regularly’, ‘sometimes’ or ‘rarely’.

	<i>Gender %</i>	
<i>Exercised</i>	Male	Female
Rarely	28.8	39.2
Sometimes	52.5	54.9
Regularly	18.6	5.9
Total	99.9	100.0

- a** Identify the explanatory variable.
- b** What percentage of females exercised sometimes?
- c** Write a brief report describing the association between how regularly these students exercised and their gender.
- 8** It was suggested that students in Dr Evan’s mathematics class would achieve better grades than students in Dr Smith’s mathematics class. Write a brief report on the association between teacher and grade, based on the data in the following table.

	<i>Class</i>		
<i>Exam grade</i>	Dr Evans	Dr Smith	Total
Fail	2	3	5
Pass	11	20	31
Credit or above	5	9	14
Total	18	32	50

- 9** Researchers predicted that using a special pillow would be more effective in curing snoring than treatment with drugs. Discuss the association between outcome of treatment and type of treatment shown in the following table.

	<i>Type of treatment</i>		
<i>Outcome of treatment</i>	Drug	Pillow	Total
Complete cure	4	10	14
Partial cure	11	12	23
No improvement	26	10	36
Total	41	32	73

Example 7

- 10** As part of the General Social Survey conducted in the US, respondents were asked to say whether they found life *exciting*, *pretty routine* or *dull*. Their marital status was also recorded as married, widowed, divorced, separated or never married. The results are organised into a table as shown.

Attitude to life	Marital status (%)				
	Married	Widowed	Divorced	Separated	Never
Exciting	47.6	33.8	46.7	45.9	52.3
Pretty routine	48.7	54.3	47.6	44.6	44.4
Dull	3.7	11.9	6.7	9.5	3.2
Total	100.1	100.0	100.0	100.0	99.0

- What percentage of widowed people found life ‘dull’?
- What percentage of people who were never married found life ‘exciting’?
- Write a brief report describing the association between a person’s attitude to life and their marital status.

Paper 1-style multiple-choice questions

Use the following information to answer Questions 11–13.

The data in the following table was collected to investigate the association between tertiary qualifications and happiness.

	Tertiary qualification		Total
	Yes	No	
Happy with life			
Yes	116	138	254
No	12	34	46
Total	128	172	300

- The percentage of participants in the study who do not have a tertiary education is closest to:
A 57.3% **B** 80.2% **C** 54.3% **D** 11.3%
- Of those people in the study who did not have a tertiary education, the percentage who are happy with their lives is closest to:
A 57.3% **B** 80.2% **C** 54.3% **D** 11.3%

- 13** The data in the table supports the contention that there is an association between *tertiary qualifications* and *happiness* because:
- A** 84.7% of people are happy.
 - B** more people without a tertiary qualification are happy than people with a tertiary qualification.
 - C** 90.6% of people with a tertiary qualification are happy, compared to 80.2% of those without a tertiary qualification.
 - D** 54.3% of happy people do not have a tertiary qualification.

1C Displaying bivariate data from two numerical variables – the scatterplot

Learning intentions

- ▶ To introduce the scatterplot for displaying data from two numerical variables.
- ▶ To construct a scatterplot using a spreadsheet.

In this and the following sections of this chapter, we will look at techniques for investigating and understanding the relationship between **two numerical variables**.

The first step in investigating the association between two numerical variables is to construct a scatterplot. We will illustrate the process by constructing a scatterplot to display average *hours worked* (the RV) against university *participation rate* (the EV) in nine countries. The data are shown below.

<i>Participation rate (%)</i>	26	20	36	1	25	9	30	3	55
<i>Hours worked</i>	35	43	38	50	40	50	40	53	35

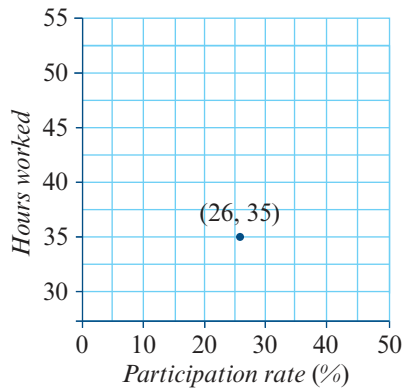
The scatterplot

- A scatterplot is a plot which enables us to display bivariate data when **both of the variables are numerical**.
- In a scatterplot, each point represents a single case.
- When constructing a scatterplot, it is conventional to use the **vertical** or **y-axis** for the response variable (RV) and the **horizontal** or **x-axis** for the explanatory variable (EV).

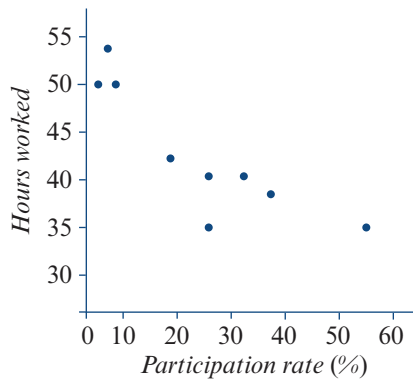
Constructing a scatterplot

In a **scatterplot**, each point represents a single case; in this instance, a country. The horizontal or x -coordinate of the point represents the university participation rate (the EV). The vertical or y -coordinate represents the average hours worked (the RV).

The scatterplot below shows the point for a country for which the university participation rate is 26% and average hours worked is 35.



The points for each of the remaining countries are then plotted, as shown below.



Note: Following the convention to label the vertical axis with the RV and the horizontal axis with the EV will become very important when we begin fitting lines to scatterplots in the next chapter, so it is a good habit to get into from the start.

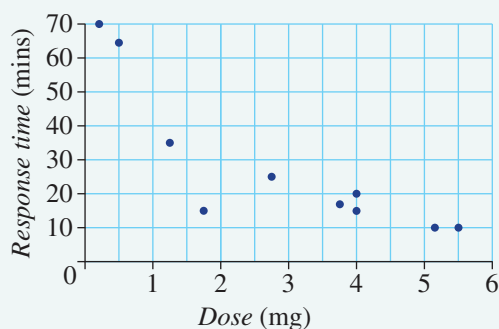
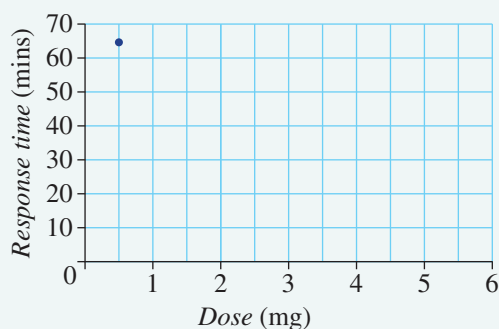

Example 8 Constructing a scatterplot using graph paper

The following table gives the time in minutes it takes for a headache to respond to medication together with the dose of the medication received by a group of 10 patients. Construct a scatterplot of these data.

Patient	1	2	3	4	5	6	7	8	9	10
Dose (mg)	0.5	1.2	4.0	5.5	2.6	3.7	5.1	1.7	0.2	4.0
Response time (mins)	65	35	15	10	22	16	10	18	70	20

Solution

Dose is the EV – this will label the horizontal axis. Response time is the RV – this will label the vertical axis.


Explanation

Which variable will be on which axes?

It is likely that the response time may be explained by the drug dosage.

Determine the scales for each axis.

Dose ranges from 0.3 mg to 5.3 mg.

A horizontal scale from 0 to 6 with intervals of 1 mg would be suitable.

Response time ranges from 10 mins to 70 mins. A vertical scale from 0 mins to 70 mins with intervals of 10 mins would be suitable.

Set up the axes, and then plot the data from the first patient (0.5, 65).

Complete the graph, adding all ten data points.



Example 9 Constructing a scatterplot using Excel

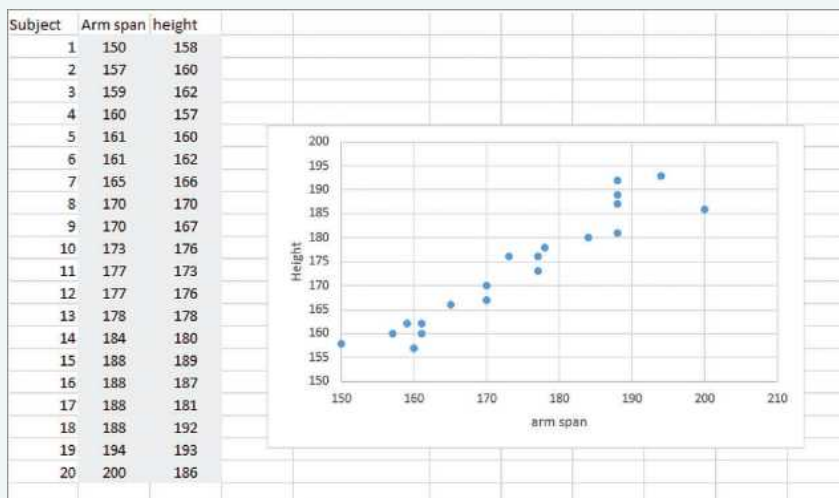
Use the data in the following table to construct a scatterplot of the height (RV) of 20 adults against arm span, which is the distance between a person's fingertips on each hand when the arms are outspread (EV).

Subject	Arm span	Height
1	150	158
2	157	160
3	159	162
4	160	157
5	161	160
6	161	162
7	165	166
8	170	170
9	170	167
10	173	176

Subject	Arm span	Height
11	177	173
12	177	176
13	178	178
14	184	180
15	188	189
16	188	187
17	188	181
18	188	192
19	194	193
20	200	186

Solution

Enter the data into two columns B and C as shown below. Make sure that you save the data for use in later examples. Select both columns (including heading) and on the **Insert** tab, in the **Charts** group, click **Scatter**. Double click on each scale separately and edit to cover the range of the data. Axis labels can be added using the Add Chart Elements option when editing the scale.



Section Summary

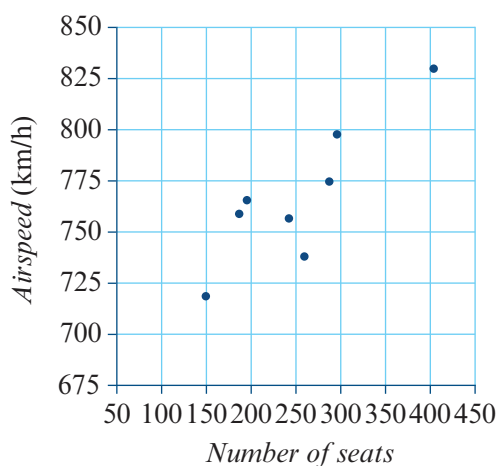
- ▶ When both variables are numerical, bivariate data can be displayed in a scatterplot.
- ▶ The explanatory variable (EV) should label the horizontal axis, and the response variable (RV) should label the vertical axis.

Exercise 1C

Save any scatterplots constructed in this section for use in later exercises.

The elements of a scatterplot

- 1 The scatterplot below has been constructed to investigate the association between the airspeed (in km/h) of commercial aircraft and the number of passenger seats.



Use the scatterplot to answer the following questions.

- a Which is the explanatory variable?
- b What type of variable is *airspeed*?
- c How many aircraft were investigated?
- d What was the airspeed of the aircraft that could seat around 300 passengers?

Constructing a scatterplot using graph paper

Example 8

- 2 The table below shows the maximum and minimum temperatures in Toowoomba during one six-day period.

Day	1	2	3	4	5	6
Minimum temperature($^{\circ}\text{C}$)	17.7	19.8	23.3	22.4	22.0	25.6
Maximum temperature ($^{\circ}\text{C}$)	29.4	34.0	34.5	35.0	36.9	36.4

Plot the maximum temperature (on the vertical axis) against the minimum temperature (on the horizontal axis) for each of the six days.

- 3 The price and age of several secondhand dirt bikes is listed in the table.

<i>Age (years)</i>	<i>Price (\$)</i>	<i>Age (years)</i>	<i>Price (\$)</i>
7	4800	11	1650
10	5700	4	6900
7	3900	3	9600
9	1950	8	6500
8	4275	4	8400
9	3300	1	11400
9	3900	7	6600

Identify the RV and the EV and construct a scatterplot of these data.

- 4 The table below shows the number of people in a movie theatre at 5-minute intervals after the advertisements started.

<i>Number in theatre</i>	87	102	118	123	135	137
<i>Time (minutes)</i>	0	5	10	15	20	25

Identify the RV and the EV and construct a scatterplot of these data.

Constructing a scatterplot using Excel

Example 9

- 5 The proprietor of a hairdressing salon recorded the amount spent advertising in the local paper, and the volume of business undertaken for each month for a year, with the following results.

<i>Month</i>	<i>Advertising (\$)</i>	<i>Volume of business (\$)</i>
1	3500	28350
2	4500	30210
3	4000	28140
4	5000	27330
5	2500	15660
6	1500	9300
7	3500	24180
8	3000	21090
9	5500	34500
10	6000	38610
11	5500	31680
12	4500	29550

- a** Determine which is the explanatory variable and which is the response variable.
- b** Construct a scatterplot of the volume of business conducted against the amount spent on advertising.
- 6** The table below shows the number of runs scored and the number of balls faced by batsmen in a one-day international cricket match. Identify the RV and the EV and construct a scatterplot of these data.

<i>Balls faced</i>	29	16	19	62	13	40	16	9	28	26	6
<i>Runs scored</i>	27	8	21	47	3	15	13	2	15	10	2

- 7** The table below shows the changing diameter of a metal ball as it is heated. Identify the RV and the EV and construct a scatterplot of these data.

<i>Temperature ($^{\circ}\text{C}$)</i>	0	10	50	75	100	150
<i>Diameter (cm)</i>	2.00	2.02	2.11	2.14	2.21	2.28

Paper 1-style multiple-choice questions

- 8** For which one of the following pairs of variables would it be appropriate to construct a scatterplot?
- A** *eye colour* (blue, green, brown, other) and *hair colour* (black, brown, blonde, other)
- B** *test score* and *gender* (male, female)
- C** *political party preference* (Labor, Liberal, Other) and *age* in years
- D** *age* in years and *blood pressure* in mmHg
- 9** A scatterplot is constructed to investigate the association between two variables, a person's weekly salary and the amount of money they spend each week on entertainment. Which statement is correct?
- A** The horizontal axis should show the amount of money they spend each week on entertainment as the response variable.
- B** The horizontal axis should show a person's weekly salary as the explanatory variable.
- C** The vertical axis should show a person's weekly salary as the explanatory variable.
- D** The vertical axis should show the amount of money they spend each week on entertainment as the explanatory variable.

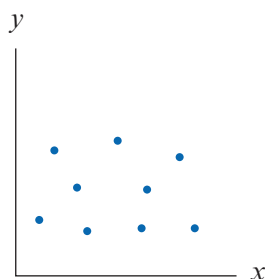
1D Interpreting a scatterplot

Learning intentions

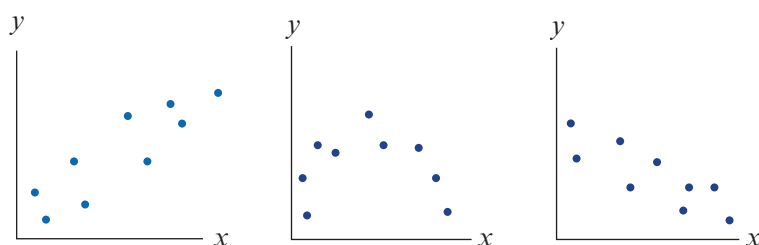
- ▶ To use a scatterplot to identify an association between two variables.
- ▶ To use a scatterplot to classify an association according to:
 - ▷ direction, which may be positive or negative
 - ▷ form, which may be linear or non-linear
 - ▷ strength, which may be weak, moderate or strong.

What features do we look for in a scatterplot that will help us to identify and describe any relationships present? First, we look to see if there is a **clear pattern** in the scatterplot.

In the example below, there is **no clear pattern** in the points. The points are **randomly scattered** across the plot, so we conclude that there is **no relationship**.



For the three examples below, there is a clear (but different) pattern in each set of points, so we conclude that there is an association in each case.



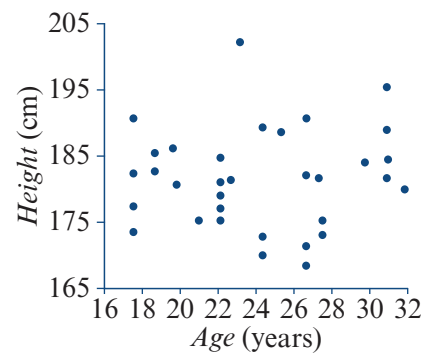
After finding a clear pattern, we need to be able to describe these associations clearly, as they are obviously different. To do this, there are several things we look for in the pattern of points:

- direction and **outliers** (if any)
- form
- strength.

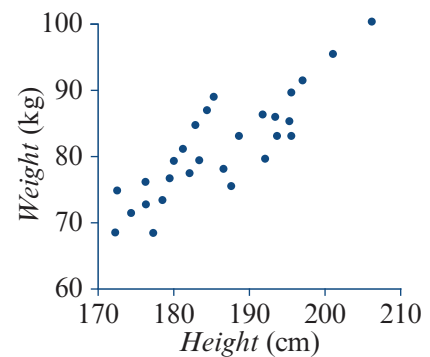
We will consider each of these attributes of the scatterplot in turn.

Direction of an association and outliers

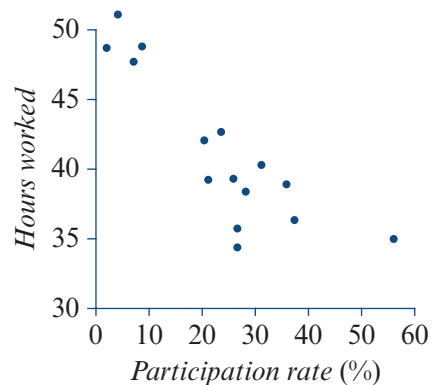
The scatterplot of height against age for a group of footballers (shown opposite) is just a *random* scatter of points. This suggests that there is *no association* between the variables *height* and *age* for this group of footballers. However, there is a possible outlier; the footballer who is 201 cm tall who seems to be much taller than the others.



In contrast, there is a clear pattern in the scatterplot of *weight* against *height* for these footballers (shown opposite). The two variables are associated. If the points in the scatterplot trend upwards as we go from left to right, we say there is a **positive association** between the variables. In this example, the positive association means that taller players tend to be heavier. In this scatterplot, there are no outliers.



Likewise, the scatterplot of working hours against university participation rates for 15 countries shows a clear pattern. The two variables are associated. If the points in the scatterplot trend downwards as we go from left to right, we say there is a **negative association** between the variables. In this example, the negative association means that countries with higher university participation rates tend to work fewer hours. In this scatterplot, there are no outliers.



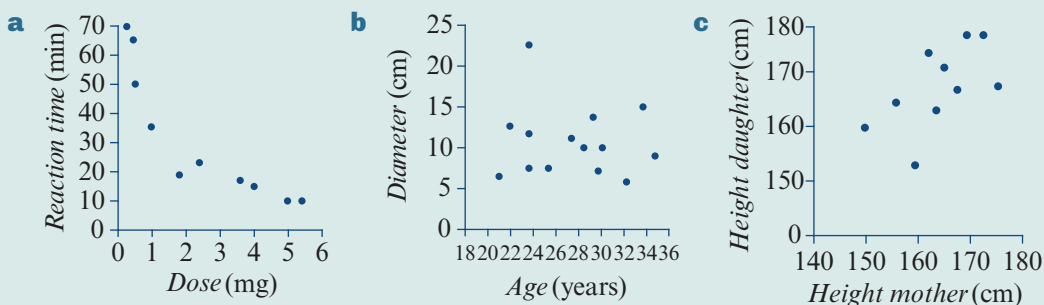
In general terms, we can classify the direction of an association as follows.

Direction of an association

- Two variables have a **positive association** when the value of the response variable tends to increase as the value of the explanatory variable increases.
- Two variables have a **negative association** when the value of the response variable tends to decrease as the value of the explanatory variable increases.
- Two variables have **no association** when there is no consistent change in the value of the response variable when the values of the explanatory variable increase.


Example 10 Classifying the direction of an association

Classify the direction of each of the following scatterplots.


Solution

- a** The direction of the association is negative.
- b** There is no association between diameter and age.
- c** The direction of the association is positive.

Explanation

There is a clear pattern in the scatterplot. The points in the scatterplot trend *downwards* from left to right.

There is no pattern in the scatterplot of *diameter* against *age*.

There is a clear pattern in the scatterplot. The points in the scatterplot trend *upwards* from left to right.

Once we have identified the direction of an association we can interpret this specifically in terms of the variables under investigation. So for, example:

- If there is a positive association between *mother's height* and *daughter's height* then we can say that taller mothers tend to have taller daughters.
- If there is a negative association between *reaction time* and *dose* then we can say reaction time tends to decrease as the drug dose increases.
- If there is no association between *age* and *diameter* then we are saying that the diameter of an individual's wrist does not seem to relate to their age.


Example 11 Interpreting the direction of an association

Write a sentence interpreting each of the following associations:

- a** There is a positive association between *study time* and *score on the exam*.
- b** There is a negative association between *study time* and *time spent watching TV*.

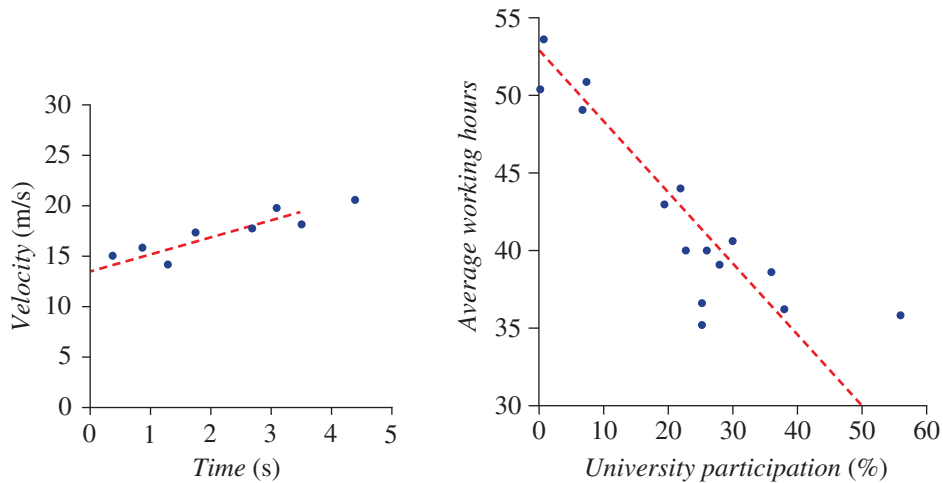
Solution

- a** Those people who spend more time studying tend to score higher marks on the exam.
- b** Those people who spend more time studying tend to spend less time watching TV.

Form of an association

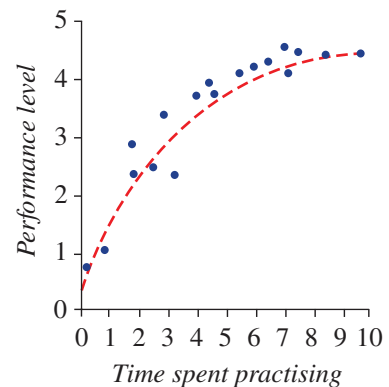
The next feature that interests us in an association is its general form. Do the points in a scatterplot tend to follow a linear pattern or a curved pattern? If the scatterplot has a linear form then we say that the association between the variables is **linear**.

For example, both of the scatterplots below can be described as having a linear form; that is, the scatter in the points can be thought of as random fluctuations around a straight line. We can say that the associations between the variables involved are linear. (The dotted lines have been added to the graphs to make it easier to see the linear form.)



By contrast, consider the scatterplot opposite, plotting performance level against time spent practising a task. There is an association between performance level and time spent practising, but it is clearly non-linear.

This scatterplot shows that while level of performance on a task increases with practice, there comes a time when the performance level will no longer improve substantially with extra practice.



While non-linear relationships exist (and we must always check for their presence by examining the scatterplot), many of the relationships we meet in practice are linear. For this reason we will restrict ourselves to the analysis of scatterplots with linear forms for now.

In general terms, we can describe the **form of an association** as follows.

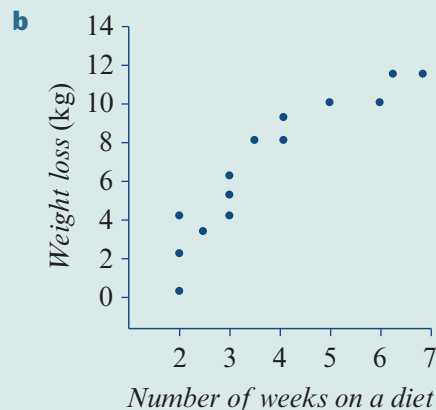
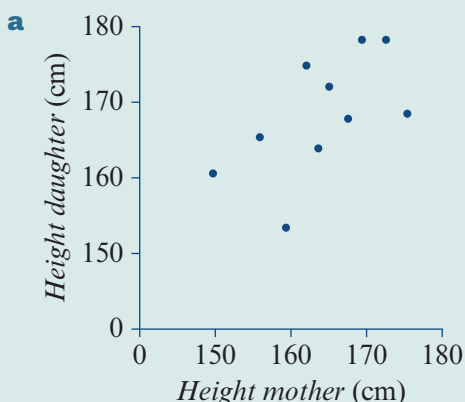
Form of an association

A scatterplot is said to have a **linear form** when the points tend to follow a straight line. A scatterplot is said to have a **non-linear form** when the points tend to follow a curved line.



Example 12 Classifying the form of an association

Classify the form of the association in each of the following scatterplots.



Solution

- a** The association is linear.
- b** The association is non-linear.

Explanation

There is a clear pattern. The points in the scatterplot can be imagined to be scattered around a straight line.

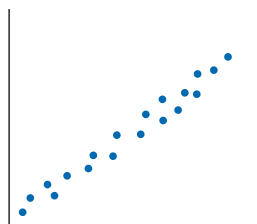
There is as a clear pattern. The points in the scatterplot can be imagined to be scattered around a curved line rather than a straight line.

Strength of an association

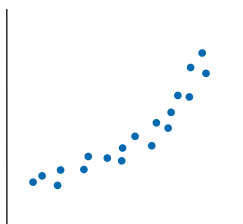
The **strength of an association** is the measure of how much scatter there is in the scatterplot.

Strong association

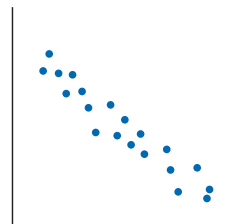
When there is a **strong association** between the variables, a pattern is clearly seen. There is only a small amount of scatter in the plot.



Strong positive association



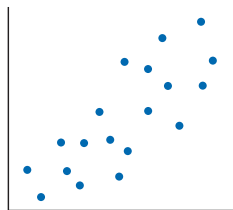
Strong positive association



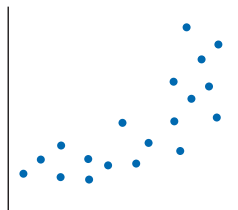
Strong negative association

Moderate association

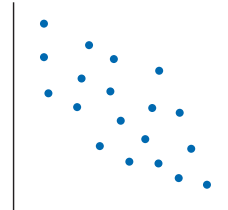
As the amount of scatter in the plot increases, the pattern becomes less clear. This indicates that the association is less strong. In the examples below, we might say that there is a **moderate association** between the variables.



Moderate positive association



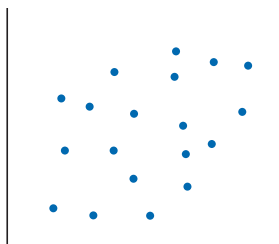
Moderate positive association



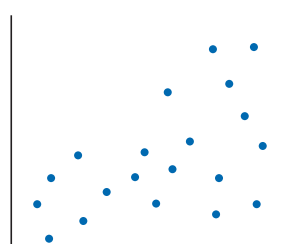
Moderate negative association

Weak association

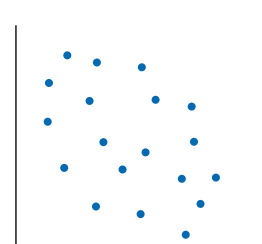
As the amount of scatter increases further, the pattern becomes even less clear. This indicates that any association between the variables is weak. The scatterplots below are examples of **weak association** between the variables.



Weak positive association



Weak positive association



Weak negative association

No association

Finally, when all we have is scatter and no pattern can be seen, we say that there is **no association** between the variables.

In general terms, we can describe the **strength of an association** as follows.

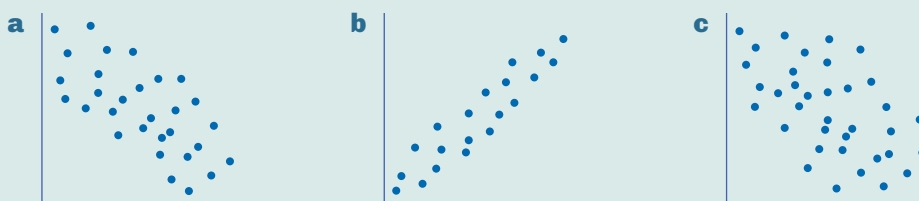
Strength of an association

An association is classified as:

- **Strong** if the points on the scatterplot tend to be tightly clustered about a trend line.
- **Moderate** if the points on the scatterplot tend to be broadly clustered about a trend line.
- **Weak** if the points on the scatterplot tend to be loosely clustered about a trend line.
- When no pattern can be seen we say that there is **no association**.


Example 13 Assessing the strength of an association

Assess the strength of the relationship in each of the following scatterplots as no association, weak association, moderate association or strong association.


Solution

- a** moderate
- b** strong
- c** weak

Explanation

Compare each of these scatterplots to the previous examples to classify the strength of the association.

Section Summary

From a scatterplot we can describe key features of a bivariate association.

- ▶ **Direction:** The two variables in the scatterplot have:
 - ▷ a positive association when the value of the response variable tends to increase as the value of the explanatory variable increases
 - ▷ a negative association when the value of the response variable tends to decrease as the value of the explanatory variable increases
 - ▷ no association when there is no consistent change in the value of the response variable when the values of the explanatory variable increase.
- ▶ **Form:** An association is classified as:
 - ▷ linear when the points tend to follow a straight line
 - ▷ non-linear when the points tend to follow a curved line.
- ▶ **Strength:** An association is classified as:
 - ▷ strong if the points on the scatterplot tend to be tightly clustered about a trend line
 - ▷ moderate if the points on the scatterplot tend to be moderately clustered about a trend line
 - ▷ weak if the points on the scatterplot tend to be loosely cluster about a trend line.

Exercise 1D

Assessing the direction of an association from the variables

Example 10

- 1 For each of the following pairs of variables, indicate whether you expect an association to exist between the variables. If associated, say whether you would expect the variables to be positively or negatively associated.

- a *intelligence* and *height*
- b *level of education* and *salary level*
- c *salary* and *tax paid*
- d *frustration* and *aggression*
- e *population density* and *distance from the city centre*
- f *time using social media* and *time spent studying*

Example 11

- 2 Write a sentence interpreting each of the following associations:

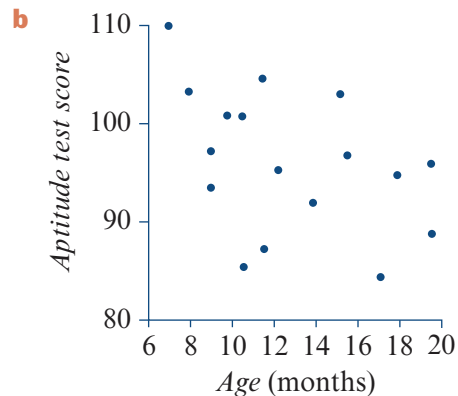
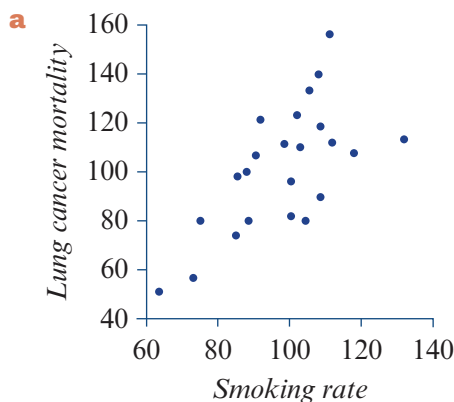
- a There is a positive association between *fitness level* and *amount* of daily exercise.
- b There is a negative association between *time* taken to run a marathon and *speed* of the runner.

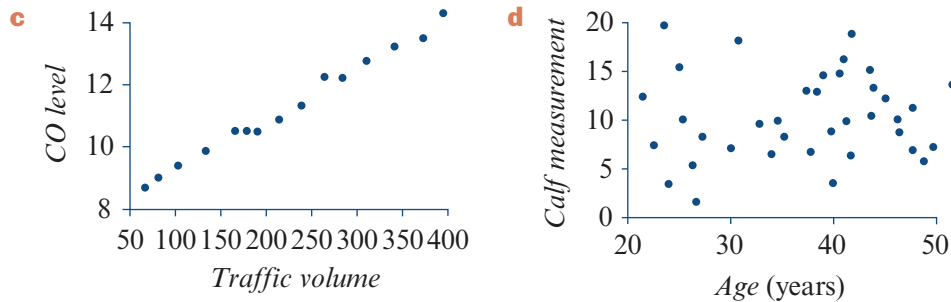
Using a scatterplot to classify the direction, form and strength of an association

Example 10–13

- 3 For each of the following scatterplots, state whether the variables appear to be related. If the variables appear to be related:

- state whether the association is positive or negative
- classify the association as linear or non-linear
- classify the strength of the association as weak, moderate, strong or no association.

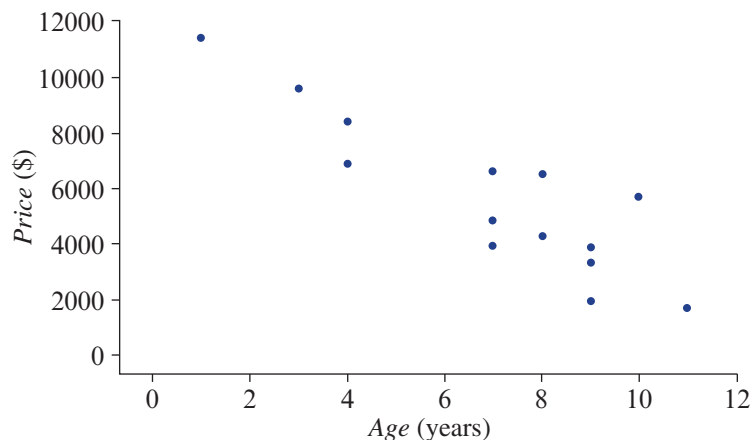




- 4** Use the scatterplot of maximum and minimum temperatures in Toowoomba during one six-day period constructed in Exercise 1C, Question **2** to complete the following:
- a** State whether the association is positive or negative.
 - b** Is the association linear or non-linear?
 - c** Classify the strength of the association as weak, moderate, strong or no association.
- 5** Use the scatterplot of the amount a hairdressing salon spent on advertising in the local paper, and the volume of business undertaken for each month for a year constructed in Exercise 1C, Question **5** to complete the following.
- a** State whether the association is positive or negative.
 - b** Is the association linear or non-linear?
 - c** Classify the strength of the association as weak, moderate, strong or no association.

Using a scatterplot identify and describe of an association

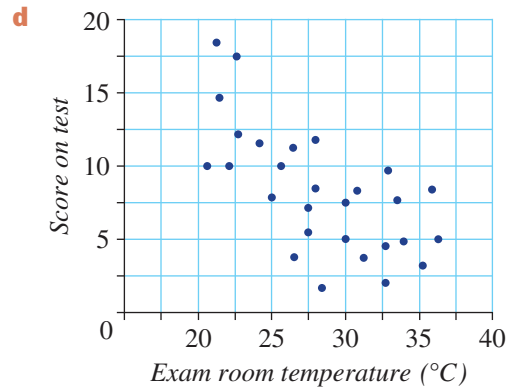
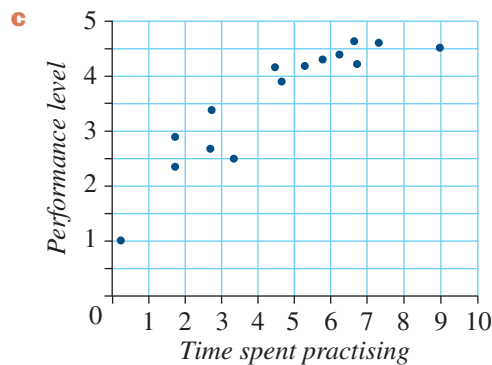
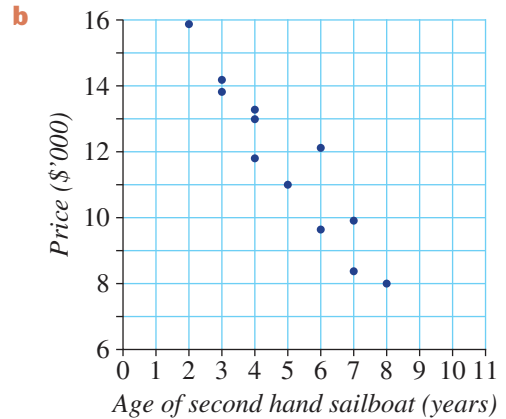
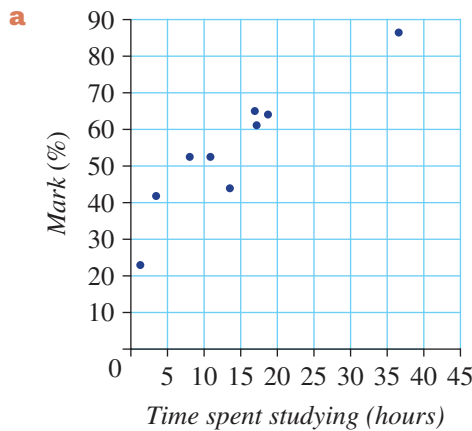
- 6** The following scatterplot shows the price and age of a sample of secondhand dirt bikes.



Complete the following description of the association between the price and age of secondhand dirt bikes by choosing the correct alternative:

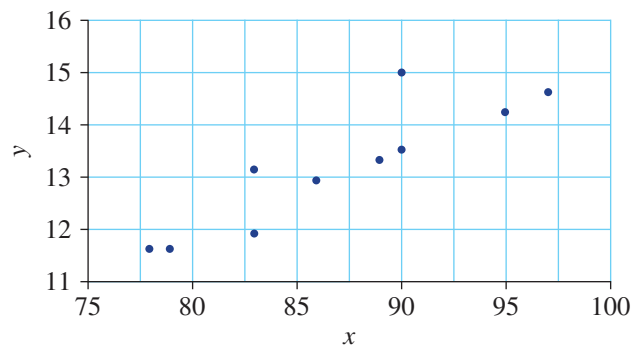
There is a weak, moderate, strong, positive, negative, linear, non-linear association between the age of a dirt bike and its price. Dirt bikes which are older tend to be higher, lower in price.

- 7 The variables in each of the following scatterplots are associated. In each case describe the association in terms of the variables in the scatterplot.



Paper 1-style multiple-choice questions

- 8 The following scatterplot shows a linear association between two numerical variables.



Choose the best description for the direction and strength of the association.

- A** strong positive **B** strong negative
C weak positive **D** weak negative

1E A measure of strength for a linear relationship – the correlation coefficient

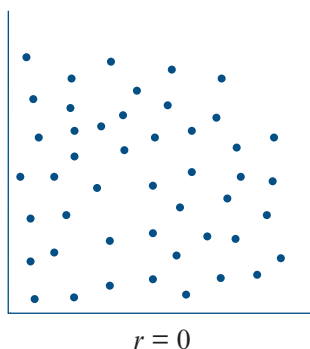
Learning intentions

- ▶ To introduce Pearson's correlation coefficient r as a measure of the strength of a linear association between two variables.
- ▶ To use technology to determine the value of Pearson's correlation coefficient r .
- ▶ To classify the strength of a linear association as weak, moderate or strong based on the value of Pearson's correlation coefficient r .

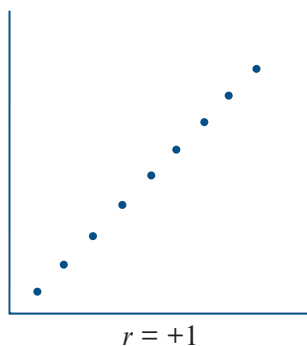
The strength of a linear association is an indication of how closely the points in the scatterplot fit a straight line. If the points in the scatterplot lie exactly on a straight line, we say that there is a perfect linear association. If there is no fit at all, we say there is no association. In general, we have an imperfect fit, as seen in all of the scatterplots to date.

To measure the **strength of a linear relationship**, a statistician called Karl Pearson developed a **correlation coefficient**, r , which has the following properties.

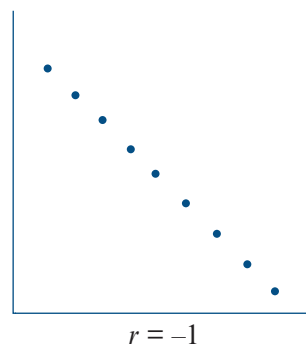
- If there is *no linear* association, $r = 0$.



- If there is a *perfect positive* linear association, $r = +1$.



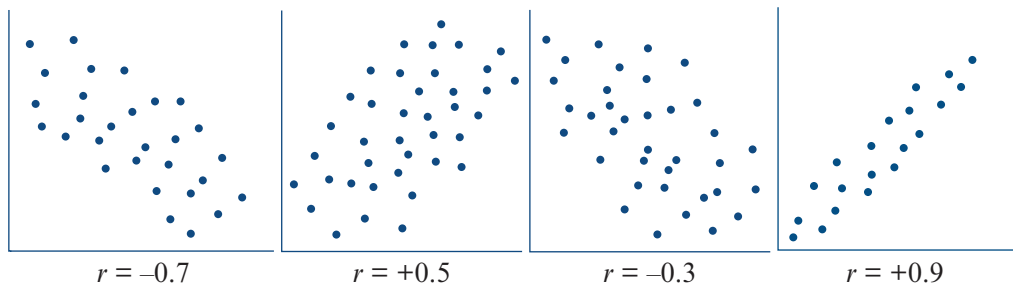
- If there is a *perfect negative* linear association, $r = -1$.



Pearson's correlation coefficient:

- measures the strength of a linear relationship, with larger values indicating stronger relationships
- has a value between -1 and $+1$
- is positive if the direction of the linear relationship is positive
- is negative if the direction of the linear relationship is negative.

If there is a less than perfect linear association, then the correlation coefficient, r , has a value between -1 and $+1$, or $-1 < r < +1$. The following scatterplots show approximate values of r for linear associations of varying strengths.



Pearson's correlation coefficient

The Pearson's correlation coefficient, r :

- measures the **strength of a linear association**, with larger values indicating stronger relationships
- has a value between -1 and $+1$
- is positive if the direction of the linear association is positive
- is negative if the direction of the linear association is negative
- is close to zero if there is no association.

Calculating the correlation coefficient

Pearson's correlation coefficient, r , gives a numerical measure of the degree to which the points in the scatterplot tend to cluster around a straight line.

Formally, if we call the two variables we are working with x and y , and we have n observations, then r is given by:

$$r = \frac{1}{n-1} \left(\frac{\sum (x - \bar{x})(y - \bar{y})}{s_x s_y} \right)$$

In this formula, $\bar{x}$ and s_x are the mean and standard deviation of the x -values, and $\bar{y}$ and s_y are the mean and standard deviation of the y -values.

In practice, you would always use your calculator to determine the value of the correlation coefficient. However, to understand what is involved when you use your calculator, it is best that you know how to calculate the correlation coefficient from the formula first.



Example 14 Calculating the correlation coefficient from first principles

Use the formula to calculate the correlation coefficient, r , for the following data.

x	1	3	5	4	7
y	2	5	7	2	9

$$\bar{x} = 4, s_x = 2.236$$

$$\bar{y} = 5, s_y = 3.082$$

Give the answer correct to two decimal places.

Solution

$$\bar{x} = 4 \quad s_x = 2.236$$

$$\bar{y} = 5 \quad s_y = 3.082 \quad n = 5$$

x	$(x - \bar{x})$	y	$(y - \bar{y})$	$(x - \bar{x}) \times (y - \bar{y})$
1	-3	2	-3	9
3	-1	5	0	0
5	1	7	2	2
4	0	2	-3	0
7	3	9	4	12
Sum	0		0	23

$$\therefore \sum (x - \bar{x})(y - \bar{y}) = 23$$

$$r = \frac{23}{4 \times 2.236 \times 3.082} = 0.83 \text{ (to two decimal places)}$$

Explanation

Write down the values of the means, standard deviations and n .

Set up a table like that shown opposite to calculate $\sum (x - \bar{x})(y - \bar{y})$.

Determine the value of the correlation coefficient r .

Calculator activity 1E Using a calculator to find the correlation coefficient

Use a calculator to find the correlation coefficient for the following set of bivariate data.

x	1	3	5	4	7
y	2	5	7	2	9

Casio fx82

Change the mode to statistics with two random variables.

Press **(Mode)** > **(2)** > **(2)** and you should get a table like this:

	x	y
1		
2		
3		

Insert the first x data value by pressing **(1)** then **(=)**. Continue inserting the rest of the x data values similarly until they are all entered.

Use the arrow keys (on the blue circular button near the screen) to navigate to the y column and insert the y data values similarly, ensuring that each observation aligns correctly.

Leave table by pressing **(AC)**

Press **(Shift)** > **(1)** [STAT] > **(5)** [Reg] > **(3)** [r] > **(1)** to get the following result:

Continued on next page

r
0.8342976876

TI-30XB

Press **[data]** to move to a data entry table.

Insert the first x data value by pressing **[1]** then **[v]** and continue inserting the rest of the x data values. Move to the second column by pressing **[>]** and then put in the y data values, ensuring that they match up with their respective x values. When finished, press

[C] to exit the table.

Press **2nd > [data] [stat] > [2] > [Enter] [Enter] [Enter]**

Use the arrow keys on the top right circular button to navigate to the correlation coefficient r .

Press **[Enter] [Enter]** and you should get:

r
0.8342976876

Sharp

Enter statistics mode for (x, y) observations by pressing **[Mode] [1] [one]**

Insert the first (x, y) observation by pressing **[1] > [STO] [(x, y)] > [2] > [M+] [DATA]** and continue following the same process for the other observations.

Press **[ALPHA] > [÷] [r] > [=]** and you should get:

r
0.8342976876

Excel can be used to find the correlation coefficient between two variables, as shown in the following example.



Example 15 Calculating the correlation coefficient using Excel

Use the data from Example 9 to calculate the correlation coefficient between *height* and *arm span* for a group of 20 adults.

Solution

- Enter the data into two columns B and C as shown.
- In an empty cell enter the formula ‘= CORREL(B2:B21,C2:C21)’.
- Press Enter and the value of the correlation will be shown in that cell.

E10					
	A	B	C	D	E
1	Subject	Arm span	height		
2	1	150	158		
3	2	157	160		
4	3	159	162		
5	4	160	157		
6	5	161	160		
7	6	161	162		
8	7	165	166		
9	8	170	170		
10	9	170	167		0.950667531
11	10	173	176		
12	11	177	173		
13	12	177	176		
14	13	178	178		
15	14	184	180		
16	15	188	189		
17	16	188	187		
18	17	188	181		
19	18	188	192		
20	19	194	193		
21	20	200	186		

The correlation coefficient is 0.95 (to two decimal places).

Guidelines for classifying the strength of a linear relationship using the correlation coefficient

The correlation coefficient, r , can be used to classify the strength of a linear association as follows.

strong positive linear association	$0.75 \leq r \leq 1$
moderate positive linear association	$0.50 \leq r < 0.75$
weak positive linear association	$0.25 \leq r < 0.50$
no linear association	$-0.25 < r < 0.25$
weak negative linear association	$-0.50 < r \leq -0.25$
moderate negative linear association	$-0.75 < r \leq -0.50$
strong negative linear association	$-1 \leq r \leq -0.75$



Example 16 Classifying the strength of a linear relationship using the correlation coefficient

Classify the strength of each of the following values of the correlation coefficient according to the previous table:

a $r = -0.08$

b $r = 0.80$

c $r = 0.56$

d $r = -0.3$

Solution

- a** No linear association
- b** Strong positive linear association
- c** Moderate positive linear association
- d** Weak negative linear association

Explanation

$r = -0.08$ lies in the interval $-0.25 < r < 0.25$

$r = 0.80$ lies in the interval $0.75 \leq r \leq 1$

$r = 0.56$ lies in the interval $0.50 \leq r < 0.75$

$r = -0.3$ lies in the interval $-0.50 < r \leq -0.25$

Section Summary

- ▶ **Pearson's correlation coefficient r** is a measure of the strength of a linear association when the variables are **numeric** and the association is **linear**.
- ▶ Assumptions when using Pearson's correlation coefficient, r are:
 - ▷ the data from both variables are **numerical**
 - ▷ the association is **linear**.
- ▶ Pearson's correlation coefficient, r :
 - ▷ has a value between -1 and $+1$, with larger values indicating stronger associations
 - ▷ is close to zero if there is **no** association
 - ▷ is **positive** if the direction of the linear association is positive
 - ▷ is **negative** if the direction of the linear association is negative.
- ▶ The strength of the correlation coefficient is classified according to the following table:

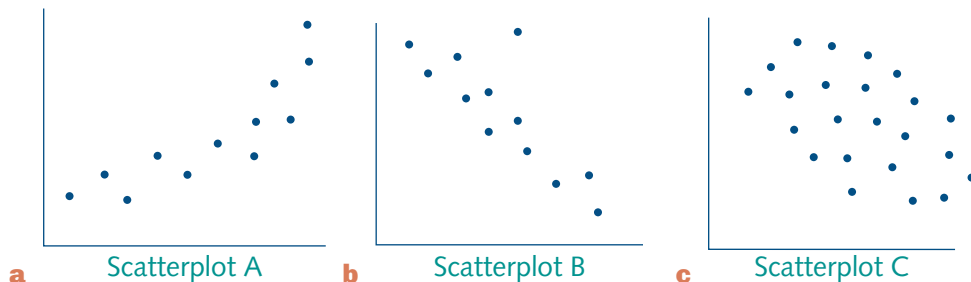
$0.75 \leq r \leq 1$	strong positive association
$0.50 \leq r < 0.75$	moderate positive association
$0.25 \leq r < 0.50$	weak positive association
$-0.25 < r < 0.25$	no association
$-0.50 < r \leq -0.25$	weak negative association
$-0.75 < r \leq -0.50$	moderate negative association
$-1 \leq r \leq -0.75$	strong negative association



Exercise 1E

Assumptions

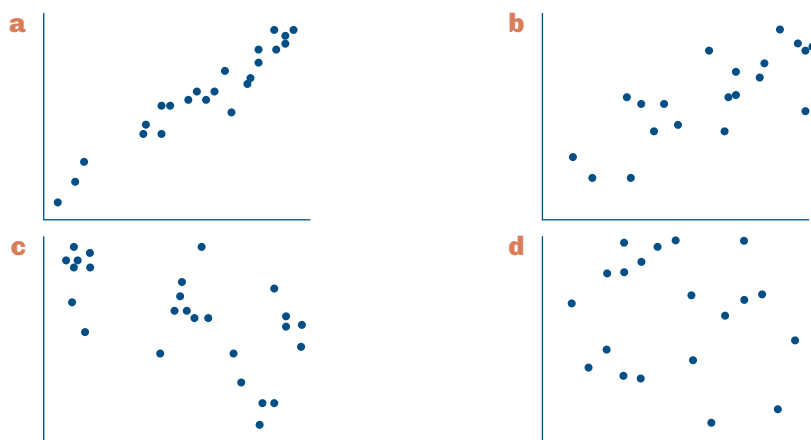
- 1 The scatterplots of three sets of related variables are shown.



- For each scatterplot, describe the association in terms of strength, direction, form and outliers (if any).
- For which of these scatterplots would it be inappropriate to use the correlation coefficient, r , to give a measure of the strength of the association between the variables? Give reasons.

Estimating r

- 2 Estimate the value of the correlation coefficient, r , in each of the following plots, using the plots on pages 37–38 as a guide.



Calculating r using the formula

Example 14

- 3 Use the formula to calculate the correlation coefficient, r , correct to two decimal places.

x	2	3	6	3	6
y	1	6	5	4	9

$$\bar{x} = 4, s_x = 1.871$$

$$\bar{y} = 5, s_y = 2.915$$

- 4 Use the formula to calculate the correlation coefficient, r , correct to three decimal places.

x	1	2	4	6	7
y	0	2	3	7	10

$$s_x = 2.5495, s_y = 4.0373$$

Calculating r using technology

Example 15

- 5 The table below shows the *maximum* and *minimum* temperatures during a heat-wave. The *maximum* and *minimum* temperature each day are linearly associated. Use appropriate technology (calculator or spreadsheet) to show that $r = 0.818$, correct to three decimal places.

Day	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday
Maximum ($^{\circ}\text{C}$)	29.4	34.0	34.5	35.0	36.9	36.4
Minimum ($^{\circ}\text{C}$)	17.7	19.8	23.3	22.4	22.0	22.0

- 6 This table shows the number of runs scored and balls faced by batsmen in a cricket match. *Runs scored* and *balls faced* are linearly associated. Use appropriate technology to show that $r = 0.8782$, correct to four decimal places.

Batsman	1	2	3	4	5	6	7	8	9	10	11
<i>Runs scored</i>	27	8	21	47	3	15	13	2	15	10	2
<i>Balls faced</i>	29	16	19	62	13	40	16	9	28	26	6

- 7 This table shows the hours worked and university participation rate (%) in six countries. *Hours worked* and *university participation rate* are linearly associated. Use appropriate technology to show that $r = -0.6727$, correct to four decimal places.

Country	Australia	Britain	Canada	France	Sweden	US
<i>Hours worked</i>	35.0	43.0	38.2	39.8	35.6	34.8
<i>Participation rate (%)</i>	26	20	36	25	37	55

- 8 The table below shows the *weight* (in kg) and blood *glucose* level (in mg/100 mL) of eight adults.

<i>Weight</i>	82.1	70.1	76.6	82.1	83.9	73.2	66.0	77.5
<i>Glucose</i>	101	89	98	100	108	104	94	89

Use appropriate technology to determine the value of Pearson's correlation coefficient for this data set. Give your answer rounded to three decimal places.

- 9 The table below shows the scores a group of nine students obtained on two class tests, *Test 1* and *Test 2*, as part of their school-based assessment.

<i>Test 1</i>	33	45	27	42	50	38	17	35	29
<i>Test 2</i>	43	46	36	34	48	34	29	41	28

Use appropriate technology to determine the value of Pearson's correlation coefficient for this data set. Write your answer rounded to three decimal places.

- 10 The table below shows the carbohydrate content (*carbs*) and the fat content (*fat*) in 100 g of nine breakfast cereals.

<i>Carbs (g)</i>	88.7	67.0	77.5	61.7	86.8	32.4	72.4	77.1	86.5
<i>Fat (g)</i>	0.3	1.3	2.8	7.6	1.2	5.7	9.4	10.0	0.7

Use appropriate technology to determine the value of Pearson's correlation coefficient for this data set. Give your answer rounded to three decimal places.

Classifying the strength of the association based on the value of r

Example 16

- 11 Use the guidelines on page 41 to classify the strength of a linear association for which Pearson's correlation coefficient is calculated to be:

- a** $r = 0.205$ **b** $r = -0.303$ **c** $r = -0.851$ **d** $r = 0.333$
e $r = 0.952$ **f** $r = -0.740$ **g** $r = 0.659$ **h** $r = -0.240$
i $r = -0.484$ **j** $r = 0.292$ **k** $r = 1$ **l** $r = -1$

Paper 1-style multiple-choice questions

- 12 The table shows the time spent watching the news on television each week and the age (in years) of the viewer.

<i>Hours TV news</i>	5	2	3	3	6	1	15	4	28
<i>Age (years)</i>	20	22	23	32	32	36	41	47	48

The value of the correlation coefficient for this data is closest to:

- A** 0.3 **B** 0.4 **C** 0.5 **D** 0.6
- 13 A correlation coefficient of -0.83 between two variables indicates that there is:
- A** no relationship between the variables
B a weak negative relationship between the variables
C a moderate negative relationship between the variables
D a strong negative relationship between the variables

1F The coefficient of determination

Learning intentions

- ▶ To calculate the value of the coefficient of determination.
- ▶ To use the coefficient of determination to assess the strength of the association in terms of the explained variation.

Introduction

If two variables are associated, it is possible to estimate the value of one variable from that of the other. For example, people's weights and heights are associated. Thus, given a person's height, we can roughly predict their weight. The degree to which we can make such predictions depends on the value of r . If there is a perfect linear association ($r = 1$) between two variables, we can make an exact prediction.

For example, when you buy cheese by the gram there is an exact association between the weight of the cheese and the amount you pay ($r = 1$). At the other end of the scale, there is no association between an adult's height and their IQ ($r \approx 0$). So knowing an adult's height will not enable you to predict their IQ any better than guessing.

The coefficient of determination

The degree to which one variable can be predicted from another linearly related variable is given by a statistic called the **coefficient of determination**.

The coefficient of determination is denoted as R^2 and calculated by squaring the correlation coefficient:

$$\text{coefficient of determination, } R^2 = r^2$$

R^2 is usually expressed as a percentage.



Example 17 Calculating the coefficient of determination

If the correlation between *weight* and *height* is $r = 0.8$, then calculate the coefficient of determination.

Solution

$$R^2 = \text{coefficient of determination} = 0.8^2 = 0.64 \text{ or } 0.64 \times 100 = 64\%$$

We now know how to calculate the coefficient of determination, but what does it tell us?

Interpreting the coefficient of determination

The coefficient of determination tells us the percentage of **variation in the response variable** that is **explained** by the **variation in the explanatory variable**.



Example 18 Interpreting the coefficient of determination

In the previous example we found the coefficient of determination between height and weight to be 0.64 (or 64%). Interpret this value in terms of the variables *weight* and *height*.

Solution

The coefficient of determination tells us that 64% of the variation in people's *weight* is explained by the variation in their *height*.

What do we mean by 'explained'?

If we take a group of people, their weights and heights will vary. One explanation is that taller people tend to be heavier and shorter people tend to be lighter. The coefficient of determination tells us that 64% of the variation in people's weights can be explained by the variation in their heights. The rest of the variation (36%) in their weights will be explained by other factors, such as diet, lifestyle, build. We could say that 36% of the variation in weight is NOT explained by the variation in height.



Example 19 Calculating and interpreting the coefficient of determination

The level of carbon monoxide (CO) in the air measured at the roadside, and the traffic volume at the same location are linearly related, with $r = +0.985$. Determine the value of the coefficient of determination, write it in percentage terms and interpret. In this relationship, *traffic volume* is the explanatory variable.

Solution

The coefficient of determination is:

$$r^2 = (0.985)^2 = 0.9702$$

Written as a percentage: $0.9702 \times 100 = 97.0\%$ rounded to one decimal place.

Therefore, 97.0% of the variation in carbon monoxide levels in the air can be explained by the variation in traffic volume.

Clearly, traffic volume is a very good predictor of carbon monoxide levels in the air. Thus, knowing the traffic volume enables us to predict carbon monoxide levels with a high degree of accuracy. This is not the case with the next example.


Example 20 Calculating and interpreting the coefficient of determination

Scores on tests of verbal and mathematical ability are linearly related with correlation coefficient $r = 0.275$. Determine the value of the coefficient of determination, write it in percentage terms, and interpret. In this relationship, *verbal ability* is the explanatory variable.

Solution

The coefficient of determination is:

$$r^2 = (0.275)^2 = 0.0756 = 7.6\% \text{ as a percentage rounded to one decimal place.}$$

Therefore, only 7.6% of the variation observed in scores on the mathematical ability test can be explained by the variation in scores obtained on the verbal ability test.

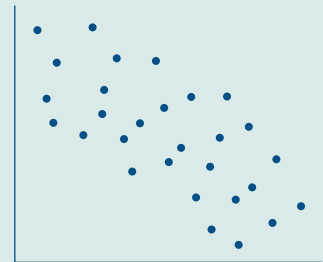
Clearly, the score on the verbal ability test is not a good predictor of the score on the mathematical ability test; 92.4% of the variation in mathematical ability is explained by other factors.

Given the value of the coefficient of determination, we can reverse the calculation and find the value of the correlation coefficient. However, since the square root of a number can be positive or negative, we need more information to be able to do this correctly, such as a scatterplot.


Example 21 Calculating the correlation coefficient from the coefficient of determination

For the relationship described by this scatterplot, the coefficient of determination = 0.5210.

Determine the value of the correlation coefficient, r .


Solution

$$R^2 = 0.5210$$

$$r = \pm\sqrt{0.5210} = \pm 0.7218$$

$$\therefore r = -0.7218$$

Explanation

Since we know the value of the coefficient of determination (R^2), we need to find the square root of this value to find r .

There are two solutions, one positive and the other negative. Use the scatterplot to decide which applies.

The scatterplot indicates a negative association.

Section Summary

- ▶ The **coefficient of determination** R^2 is equal to the square of the correlation coefficient r (r^2). It is usually expressed as a percentage.
- ▶ The coefficient of determination tells us the percentage of variation in the response variable which is explained by the variation in the explanatory variable.

Exercise 1F

Calculating the coefficient of determination from r

Example 17

- 1 For each of the following values of r , calculate the value of the coefficient of determination and convert to a percentage (correct to one decimal place).
- | | | |
|-----------------------|----------------------|-----------------------|
| a $r = 0.675$ | b $r = 0.345$ | c $r = -0.567$ |
| d $r = -0.673$ | e $r = 0.124$ | f $r = 0.019$ |

Calculating and interpreting the coefficient of determination

Example 18

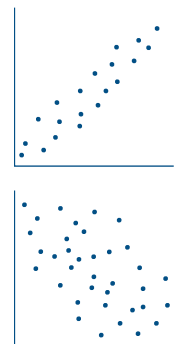
Example 19

- 2 For each of the following, determine the value of the coefficient of determination, write it in percentage terms, and interpret in terms of the variables in the question.
- a** Scores on hearing tests and age (EV) are linearly related, with $r = -0.611$.
 - b** Mortality rates and smoking rates (EV) are linearly related, with $r = 0.716$.
 - c** Life expectancy and birth rates (EV) are linearly related, with $r = -0.807$.
 - d** Daily maximum (RV) and minimum temperatures are linearly related, with $r = 0.818$.
 - e** Runs scored (RV) and balls faced by a batsman are linearly related, with $r = 0.8782$.

Calculating r from the coefficient of determination given a scatterplot

Example 21

- 3 **a** For the relationship described by the scatterplot shown, the coefficient of determination, $R^2 = 0.8215$. Determine the value of the correlation coefficient, r (correct to three decimal places).
- b** For the relationship described by the scatterplot shown, the coefficient of determination $R^2 = 0.1243$. Determine the value of the correlation coefficient, r (correct to three decimal places).



Comparing coefficients of determination

- 4** A study was conducted where a group of 100 adults were asked to record their weekly income, their weekly expenditure on food, and their weekly expenditure on leisure.

The researcher determined the following correlation coefficients:

Income and expenditure on food: $r = 0.6$

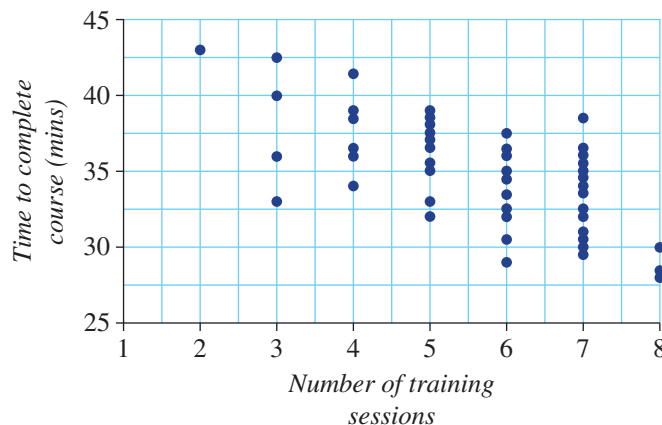
Income and expenditure on leisure: $r = 0.5$

- Calculate and interpret the coefficient of determination relating income and expenditure on food.
- Calculate and interpret the coefficient of determination relating income and expenditure on leisure.
- Write a sentence comparing the importance of income in understanding expenditure on food and expenditure on leisure.

Paper 1-style multiple-choice questions

Use the following information to answer Questions 5–7.

The association between the *number of training sessions* attended by participants before undertaking an obstacle course, and the *time* in minutes it took them to complete the course, is described by the scatterplot shown. The coefficient of determination is 0.3969.



- 5** The value of the correlation coefficient, r (rounded to two decimal places) is closest to:
- 0.16
 - 0.63
 - 0.40
 - 0.63

- 6** The percentage of variation in *time* explained by the variation in the *number of practice sessions* is closest to:
- A** 39.7% **B** 63.0%
C 15.8% **D** 37.0%
- 7** The percentage of variation in *time* NOT explained by the variation in the *number of practice sessions* is closest to:
- A** 39.7% **B** 63.0%
C 37.0% **D** 60.3%
- 8** Suppose that in a certain industry the correlation between *years spent studying* and *income* for employees is 0.73, and the correlation between *age* and *income* is 0.45. Given this information, which one of the following statements is true?
- A** Older employees tend to have spent more years studying.
B The correlation between *age* and *years spent studying* is 0.32.
C *Years spent studying* explains a higher percentage of the variation in *income* than *age*.
D *Age* explains a higher percentage of the variation in *income* than *years spent studying*.
- 9** Which of following statements could be true?
- A** The correlation coefficient between *height* (in centimetres) and *weight* (1 = light, 2 = medium, 3 = heavy) was determined to be 0.68.
B The correlation coefficient between *height* (in centimetres) and *head circumference* (in centimetres) was determined to be 1.45.
C The correlation coefficient between *age* (in years) and *salary* (in \$000's) was determined to be 0.68.
D The correlation coefficient between *height* (in centimetres) and *head circumference* (in centimetres) was found to be 0.49, and the coefficient of determination was determined to be 70%.

Key ideas and chapter summary

Bivariate data

Bivariate data are data in which each observation involves recording information about two variables for the same person or thing. An example would be the heights and weights of the children in a preschool.

Categorical and numerical variables

Categorical variables generate data values which are labels; **numerical variables** generate values which are numbers.

Explanatory and response variables

The **explanatory variable** (EV) may explain the associated changes in the **response variable** (RV).

Two-way frequency table

A **two-way frequency table** summarises **bivariate data** obtained from two categorical variables. In a two-way frequency table, the columns are labelled with the values of the EV, and the rows are labelled with the values of the RV.

Percentaged two-way frequency table

Associations between two categorical variables are identified by comparing percentages in a **percentaged two-way frequency table**. If column percentages have been calculated, then percentages are compared across a row to identify the association.

Scatterplots

A **scatterplot** is used to display bivariate data from two numerical variables. In a scatterplot, the EV is plotted on the horizontal axis, and the RV is plotted on the vertical axis.

Feature of interest in a scatterplot

From a scatterplot the **direction**, **form** and **strength** of an association can be determined.

Correlation coefficient r

The **correlation coefficient**, r , gives a measure of the **strength of a linear relationship** between two numerical variables.

Coefficient of determination R^2

The **coefficient of determination**, R^2 , gives the percentage of variation in the RV that can be explained by the variation in the EV.

Skills checklist



Download this checklist from the Interactive Textbook, then print it and fill it out to check your skills.



1A

1 I can identify categorical and numerical variables in bivariate data.



See Example 1 and Exercise 1A Question 1

1A	2 I can identify associations as categorical or numerical.	☐
	See Example 2 and Exercise 1A Question 2	
1A	3 I can identify explanatory and response variables.	☐
	See Example 3, Example 4 and Exercise 1A Question 3	
1B	4 I can construct a two-way frequency table.	☐
	See Example 5 and Exercise 1B Question 1	
1B	5 I can percentage a two-way frequency table.	☐
	See Example 6 and Exercise 1B Question 2	
1B	6 I can describe an association from a percentaged two-way frequency table.	☐
	See Example 6, Example 7 and Exercise 1B Questions 3 and 10	
1C	7 I can construct a scatterplot using graph paper.	☐
	See Example 8 and Exercise 1C Question 1	
1C	8 I can construct a scatterplot using technology.	☐
	See Example 9, Exercise 1C Question 4	
1D	9 I can classify an association from a scatterplot according to direction, form and strength.	☐
	See Example 10, Example 12, Example 13 and Exercise 1D Question 3	
1E	10 I can calculate the correlation coefficient, r , from first principles.	☐
	See Example 14 and Exercise 1E Question 3	
1E	11 I can calculate the correlation coefficient, r , using technology.	☐
	See Calculator Activity 1E, Example 15 and Exercise 1E Question 5	
1E	12 Classifying the strength of a linear relationship using the correlation coefficient.	☐
	See Example 16 and Exercise 1E Question 8	
1F	13 I can calculate the coefficient of determination R^2 .	☐
	See Example 17 and Exercise 1F Question 1	
1F	14 I can interpret the coefficient of determination R^2 .	☐
	See Example 18, Example 19, Example 20 and Exercise 1F Question 2	

Multiple-choice questions

Use the information in the following frequency table to answer Questions 1–4.

<i>Plays sport</i>	<i>Gender</i>	
	Male	Female
Yes	68	79
No	34	
Total	102	175

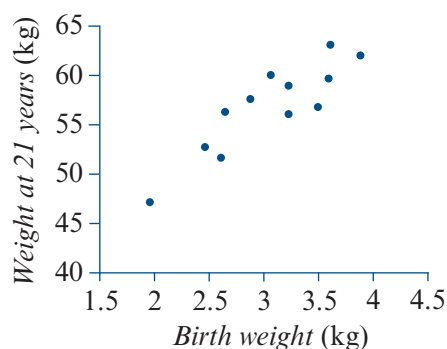
- The variables *plays sport* and *gender* are:
 - both categorical variables
 - a categorical and a numerical variable, respectively
 - a numerical and a categorical variable, respectively
 - both numerical variables
- The number of females who do not play sport is:
 - 21
 - 45
 - 79
 - 96
- The percentage of males who do not play sport is:
 - 19.4%
 - 33.3%
 - 34.0%
 - 66.7%
- The variables *plays sport* and *gender* appear to be associated because:
 - more females play sport than males
 - fewer males play sport than females
 - a higher percentage of females play sport compared to males
 - a higher percentage of males play sport compared to females

Use the information in the following frequency table to answer Questions 5 and 6.

The results of a survey conducted with 578 secondary students are summarised in the following table.

<i>How exciting is your life?</i>	<i>How important is it to obey rules?</i>			Total
	Very important	Important	Not important	
Very exciting	50	84	44	178
Routine	103	169	56	328
Dull	25	42	5	72
Total	178	295	105	578

- 8** The scatterplot shows the weights of 12 women at birth and at the age of 21. The association is best described as:
- A** weak, positive and linear
B weak, negative and linear
C strong, positive and non-linear
D strong, positive and linear



- General Mathematics Units 3 & 4 ISBN 978-1-009-57800-4 © Lipson et al. 2025 Cambridge University Press
Photocopying is restricted under law and this material must not be transferred to another party.

- 10** For which one of the following pairs of variables would it be appropriate to construct a scatterplot?
- A** *eye colour* (blue, green, brown, other) and *hair colour* (black, brown, blonde, other)
 - B** *test score* and *gender* (male, female)
 - C** *political party preference* (Labor, Liberal, Other) and *age* in years
 - D** *weight* in kg and *blood pressure* in mmHg

- 11** The variables *response time* to a drug and *drug dosage* are linearly associated, with $r = -0.9$. From this information, we can conclude that:
- A** response times are -0.9 times the drug dosage
 - B** response times tend to decrease with decreased drug dosage
 - C** response times tend to decrease with increased drug dosage
 - D** response times are 81% of the drug dosage

- 12** The birth weight and weight at age 21 of eight women are given in the table below.

<i>Birth weight (kg)</i>	1.9	2.4	2.6	2.7	2.9	3.2	3.4	3.6
<i>Weight at 21 (kg)</i>	47.6	53.1	52.2	56.2	57.6	59.9	55.3	56.7

The value of the correlation coefficient is closest to:

- A** 0.536 **B** 0.6182 **C** 0.7863 **D** 0.8232
- 13** For a dataset with 12 points the value of $\Sigma \left(\frac{x - \bar{x}}{s_x} \right) \left(\frac{y - \bar{y}}{s_y} \right)$ is equal to -3.2 . The value of the correlation coefficient is closest to:
- A** -0.320 **B** -0.291 **C** 0.320 **D** 0.291
- 14** The correlation between heart weight and body weight in a group of mice is $r = 0.765$. Using body weight as the explanatory variable, we can conclude that:
- A** 58.5% of the variation in heart weight is explained by the variation in body weight
 - B** 76.5% of the variation in heart weight is explained by the variation in body weight
 - C** heart weight is 58.5% of body weight
 - D** heart weight is 76.5% of body weight
- 15** Which of following statements could be true?
- A** The correlation coefficient between *weight* in kg and *blood pressure* (1 = low, 2 = average, 3 = high) is 0.68.
 - B** The correlation coefficient between *height* (in cm) and *foot length* (in cm) is 1.07.
 - C** The correlation coefficient between *blood pressure* (in cm) and *weight* (in kg) is -0.3 , and the coefficient of determination is $R^2 = -0.09$.
 - D** The coefficient of determination between *age* (in years) and *salary* (in \$000's) is 46%.

Short-response questions

- 1 One thousand drivers who had an accident during the past year were classified according to their age and the number of accidents they had.

<i>Number of accidents</i>	Age < 30	Age $\geq$ 30
At most one accident	130	170
More than one accident	470	230
Total	600	400

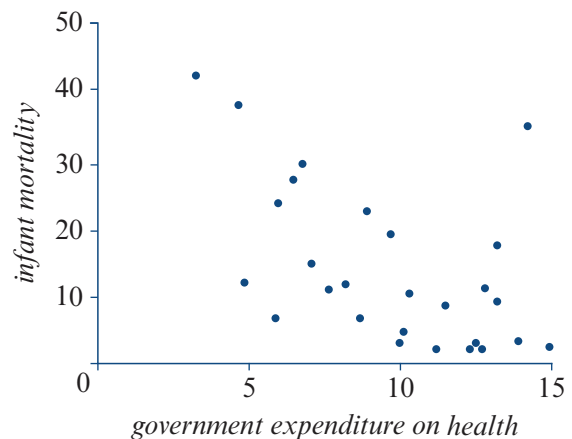
- What are the variables shown in the table? Are they categorical or numerical?
 - Determine the response and explanatory variables.
 - How many drivers under the age of 30 had more than one accident?
 - Convert the table values to percentages by calculating the column percentages.
 - Use these percentages to comment on the statement: 'Of drivers who had an accident in the past year, younger drivers (age < 30) are more likely than older drivers (age $\geq$ 30) to have had more than one accident.'
- 2 A retailer recorded the number of ice-creams sold and the day's maximum temperature over eight consecutive Saturdays one summer.

<i>Temperature ($^{\circ}\text{C}$)</i>	22	25	36	34	21	28	41	31
<i>Number of ice-creams sold</i>	145	155	200	198	150	179	230	180

- Construct a scatterplot of these data.
- From the scatterplot describe the association between temperature and the number of ice-creams sold in terms of direction, form and strength.
- Determine the value of the correlation coefficient.
- Classify the strength of this relationship based on the value of r .

- 3 The scatterplot on the right shows the Government expenditure on health against the infant mortality for a group of countries.

- Describe the association between government expenditure on health and the infant mortality in terms of direction, form and strength.
- Give reasons why it is appropriate to calculate the correlation coefficient for these data.



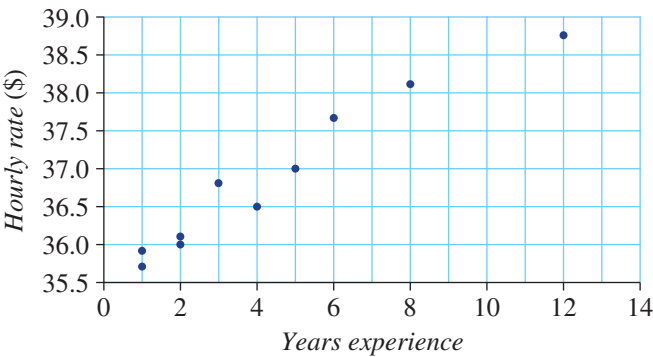
- 4 Are males and females in Year 12 equally likely to have a part-time job? Data from interviews with 300 Year 12 students are summarised in the following table.

<i>Part-time job</i>	Male	Female	Total
Yes	94	123	208
No	46	37	92
Total	140	160	300

- a Identify which is the explanatory variable and which is the response variable.
- b Construct an appropriately percentaged two-way frequency table.
- c Use the information in the table to write a brief report describing the association between gender and having a part-time job.
- 5 Suppose that the correlation between age and scores on a hearing test are linearly related, and that $r = -0.77$.
- a Determine the value of the coefficient of determination, R^2 .
- b Interpret R^2 in terms of the variables *age* and *score* on the hearing test.
- 6 The data below are the hourly pay rates (in dollars per hour) of 10 production-line workers along with their years of experience on initial appointment.

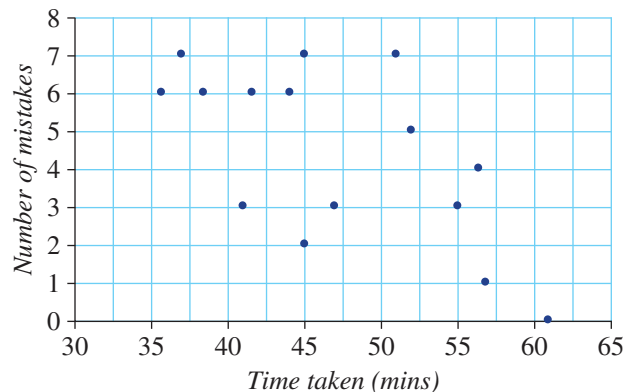
<i>Rate (\$h)</i>	35.90	35.70	36.10	36.00	36.79	36.45	37.00	37.65	38.10	38.75
<i>Experience (years)</i>	1	1	2	2	3	4	5	6	8	12

- a The following scatterplot of the data has been constructed. Explain why the variable *rate* is plotted on the vertical axis and the variable *experience* on the horizontal axis.

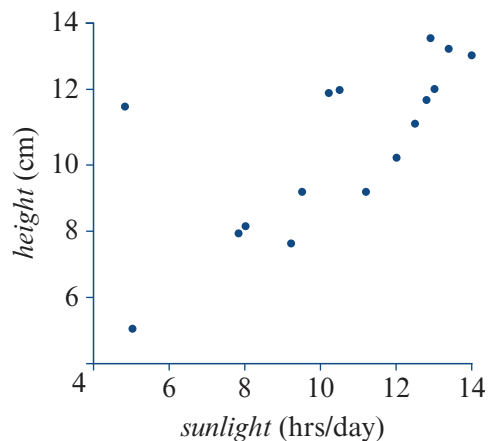


- b Comment on direction, outliers, form and strength of any association shown in the scatterplot.
- c Determine the value of the correlation coefficient correct (r) to three decimal places.
- d Determine the value of the coefficient of determination (R^2) and interpret.

- 7** The manager of a large manufacturing plant was worried about the quality of the products coming from the plant. She noticed that some workers made far more mistakes than others in assembling the gizmos and was concerned that they were trying to work too quickly. The manager collected information from a random sample of 15 workers, recording the time taken to assemble a gizmo and the number of mistakes made.



- a** Use the scatterplot to describe the relationship between the time taken and the number of mistakes made.
- b** Explain why it is appropriate to calculate the correlation coefficient, r , for these data.
- c** The value of correlation coefficient, r , for the data is -0.6345 . Determine the value of the coefficient of determination (R^2) and interpret for these variables.
- 8** Miller conducted a study to investigate the height of his seedlings and the average number of hours of daily sunshine each plant received over a 14-day period. He planted his seedlings in 10 different locations in his garden, and his results are shown in the following scatterplot.



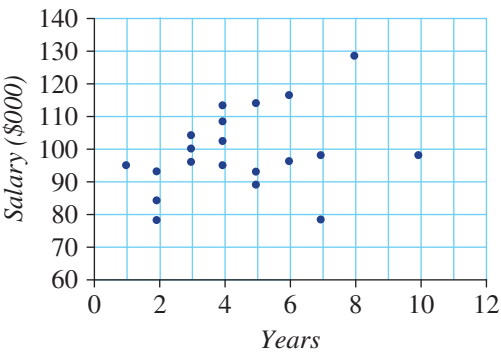
- a** Describe the relationship in the scatterplot.
- b** Miller checks his data and realises he has made a mistake in recording the data.

The actual height of the plant which received an average of 4.8 hrs/day was not the 11.5 cm he had used in his analysis. He decides to remove this data value. Explain how the value of the correlation coefficient calculated with the outlier excluded would compare to that calculated with the outlier included.

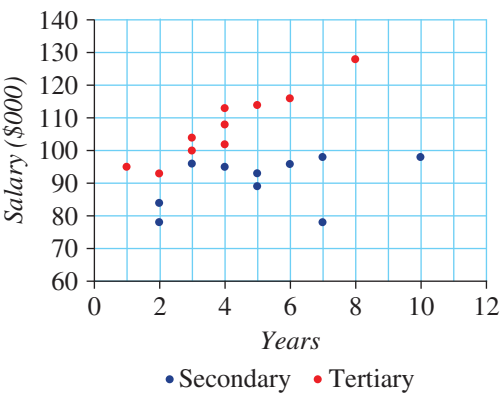
- 9 The following table gives the educational level (*education*), the number of years the person has worked for the company (*years*) and their current salary to the nearest thousand dollars (*salary*) for a group of current employees of a particular company.

<i>Education</i>	<i>Year</i>	<i>Salary</i>	<i>Education</i>	<i>Years</i>	<i>Salary</i>
Secondary	2	78	Tertiary	2	93
Secondary	3	96	Tertiary	3	104
Secondary	2	84	Tertiary	4	113
Secondary	4	95	Tertiary	5	114
Secondary	7	98	Tertiary	4	108
Secondary	6	96	Tertiary	4	102
Secondary	7	78	Tertiary	1	95
Secondary	10	98	Tertiary	8	128
Secondary	5	89	Tertiary	3	100
Secondary	5	93	Tertiary	6	116

- a A scatterplot of the data for all 20 employees, with the variable *salary* as the response variable and the variable *years* as the explanatory variable is shown. Describe the association between *salary* and *years* for all employees from the scatterplot.



- b In this scatterplot of the variable *salary* against *years*, different colour dots are used for the educational level of the employees. Describe the association between *salary* and *years* separately for those with secondary and those with tertiary levels of education from the scatterplot.



- c Discuss the merits of using *years* in the prediction of *salary* for each of these groups of employees of the company, quoting relevant statistics.

Chapter 2

Bivariate data analysis 2

Chapter questions

UNIT 3 BIVARIATE DATA AND TIME SERIES ANALYSIS, SEQUENCES AND EARTH GEOMETRY

Topic 2: Bivariate data analysis 2

- ▶ How do we fit a least squares line to data using technology?
- ▶ How do we fit a least squares line to data using sample statistics?
- ▶ How do we construct a residual plot and use it to assess whether a linear model is appropriate for the data?
- ▶ How do we use the equation of the least squares line to make predictions?
- ▶ How do we differentiate between correlation and causation?

Once we identify a linear association between two numerical variables, we can fit a linear model to the data and determine its equation. This equation gives us a better understanding of the nature of the relationship between the two variables, and we can also use the linear model to make predictions based on this understanding of the relationship.

2A Fitting a least squares line to numerical data

Learning intentions

- ▶ To define linear regression.
- ▶ To define a residual.
- ▶ To introduce the least squares line.
- ▶ To determine the equation of the least squares line using summary statistics.
- ▶ To determine the equation of the least squares line using technology.

The process of modelling an association with a straight line is known as **linear regression** and the resulting line is often called the **regression line**.

The equation of a line relating two variables x and y is of the form

$$y = mx + c$$

where m and c are constants. When the equation is written in this form:

- c represents the coordinate of the point where the line crosses the y -axis (the y -intercept)
- m represents the slope (gradient) of the line.

In order to summarise any particular (x, y) data set, numerical values for c and m are needed that will ensure the line passes close to the data.

The best approach to fitting a straight line to data is to use the **least squares method**. This method assumes that the variables are linearly related, and works best when there are no clear outliers in the data.

Some terminology

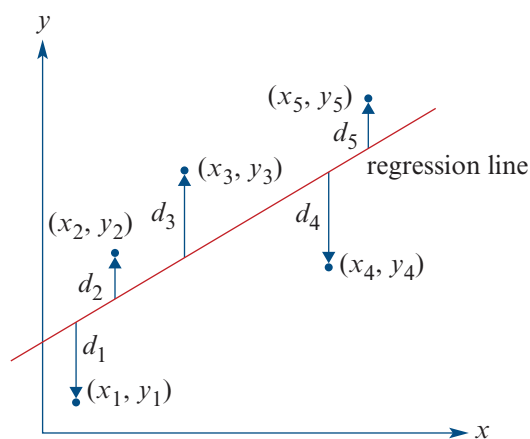
To explain the least squares method, we need to define several terms.

The scatterplot shows five data points, (x_1, y_1) , (x_2, y_2) , (x_3, y_3) , (x_4, y_4) and (x_5, y_5) .

A regression line (not necessarily the least squares line) has also been drawn on the scatterplot.

The vertical distances d_1 , d_2 , d_3 , d_4 and d_5 of each of the data points from the regression line are also shown.

These vertical distances, d , are known as **residuals**.



The least squares line

The least squares regression line is the line where the sum of the squares of the residuals is as small as possible; that is, it minimises:

$$\text{the sum of the squares of the residuals} = d_1^2 + d_2^2 + d_3^2 + d_4^2 + d_5^2$$

Why do we minimise the sum of the *squares* of the residuals and not the sum of the residuals? This is because the sum of the residuals for the least squares regression line is always zero. Some residuals are positive and some negative, and in the end they add to zero. Squaring the residuals solves this problem.

The least squares line

The **least squares line** is the line that minimises the sum of the squares of the **residuals**.

The assumptions for fitting a least squares regression line to data are the same as for using the correlation coefficient, r . These are that:

- the data are numerical
- the association is linear
- there are no clear outliers.

Determining the equation of the least squares regression line

To determine exactly the equation of the least squares regression line we need to determine the values of the intercept (c) and the slope (m) that define the line. The mathematics is beyond the scope of this course, but calculus can be used to give us rules for these values:

The equation of the least squares regression line

The equation of the least squares regression line is given by $y = mx + c$, where:

$$\text{the slope } (m) \text{ is given by } m = \frac{rs_y}{s_x}$$

and

$$\text{the intercept } (c) \text{ is then given by } c = \bar{y} - m\bar{x}$$

Here:

- r is the **correlation coefficient**
- s_x and s_y are the **standard deviations** of x and y
- $\bar{x}$ and $\bar{y}$ are the **mean** values of x and y .

Warning!

If you do not correctly decide which is the explanatory variable (the x -variable) and which is the response variable (the y -variable) before you start calculating the equation of the least squares regression line, you may get the wrong answer.


Example 1 Determining the equation of the least squares regression line using the formula

The height (x) and weight (y) of 11 people have been recorded, and the values of the following statistics determined:

$$\bar{x} = 173.3 \text{ cm} \quad s_x = 7.444 \text{ cm} \quad \bar{y} = 65.45 \text{ cm} \quad s_y = 7.594 \text{ cm} \quad \text{and} \quad r = 0.8502$$

Use the formula to determine the equation of the least squares regression line that enables weight to be predicted from height. Calculate the value of the slope and intercept correct to two decimal places.

Solution

EV: height (x)

RV: weight (y)

$$\bar{x} = 173.3 \quad s_x = 7.444$$

$$\bar{y} = 65.45 \quad s_y = 7.594$$

$$r = 0.8502$$

$$\begin{aligned} \text{Slope: } m &= \frac{rs_y}{s_x} = \frac{0.8502 \times 7.594}{7.444} \\ &= 0.867 \end{aligned}$$

$$\begin{aligned} \text{Intercept: } c &= \bar{y} - m\bar{x} \\ &= 65.45 - 0.8673 \times 173.3 \\ &= -84.853 \end{aligned}$$

$$y = 0.87x - 84.85$$

$$\text{weight} = 0.87 \times \text{height} - 84.85$$

Explanation

Identify and write down the explanatory variable (EV) and the response variable (RV). Label as x and y , respectively.

Write down the given information.

Calculate the slope.

Calculate the intercept.

Round the values of m and c to two decimal places and substitute into the equation.

Write the least squares equation using the variable names.

Mostly we will use technology to calculate the intercept and slope directly from the data. Note that the Casio and Sharp calculators use the symbol a for the intercept (c), and the symbol b for the slope (m), while the TI uses the symbol b for the intercept (c), and the symbol a for the slope (m). Make sure you are familiar with the symbols used on your calculator.

Calculator activity 2A Determining the equation of the least squares line using a scientific calculator

The data in the following table gives the height (y) of 20 adults and their arm span, which is the distance between a person's fingertips on each hand when the arms are outspread (x). Use the data to determine the equation of the least square line n .

x	150	157	159	160	161	161	165	170	170	173
y	158	160	162	157	160	162	166	170	167	176

x	177	177	178	184	188	188	188	188	194	200
y	173	176	178	180	189	187	181	192	193	186

Casio fx82

Change the mode to statistics with two random variables.

Press **[Mode]** > **[2]** > **[2]** and you should get a table like this:

	x	y
1		
2		
3		

Insert the first data by pressing 150 then **[=]**. Continue inserting the data similarly for the x variables.

Use the arrow keys (on the blue circular button near the screen) to navigate to the y column and insert the y data points similarly, ensuring that each observation aligns correctly.

Leave table by pressing **[AC]**

Press **[Shift]** > **[1]** [STAT] > **[5]** [Reg] > **[1]** [a] > **[=]** to get the intercept.

a
32.97194808

Press **[Shift]** > [STAT] > **[5]** [Reg] > **[2]** [b] > **[=]** to get the slope.

b
0.806640206

Rounding to three decimal places and substituting the values of the intercept and slope into the equation gives:

$$y = 0.807x + 32.972.$$

In terms of the variables in the question:

$$\text{height} = 0.807 \times \text{armspan} + 32.972$$

TI-30XB

Press **[data]** to move to a table.

Insert the first data by pressing **[1]** then **[▼]** and continue inserting the rest of the x data.

Move to the second column by pressing **[▶]** and then put in the y data, ensuring that they match up with their respective x data. When finished, press clear to exit the table.

Press **2nd > [data] [stat] > [2] > [enter] [enter] [enter]** and then use the arrow keys on the top right circular button to navigate to the intercept b , then press **[enter] [enter]**. Repeat the steps again but this time navigate to the slope a , then press **[enter] [enter]**. You should get:

b	32.97194808
a	0.806640206

Note that the TI-30XB labels the intercept and slope *differently*.

Rounding to three decimal places and substituting the values of the intercept and slope into the equation gives:

$$y = 0.807x + 32.972.$$

In terms of the variables in the question:

$$height = 0.807 \times armspan + 32.972$$

Sharp

Enter statistics mode for (x, y) observations by pressing **[Mode] [1] [1]**.

Insert the first (x, y) observation by pressing **150 > [STO] [(x, y)] > 158 > [M+] [DATA]** and continue following the same process for the other observations.

Press **[ALPHA] > [a] > [=]** to get the intercept a :

a=	32.97194808

Press **[ALPHA] > [b] > [=]** to get the slope b :

b=	0.806640206

Rounding to three decimal places and substituting the values of the intercept and slope into the equation gives:

$$y = 0.807x + 32.972.$$

In terms of the variables in the question:

$$height = 0.807 \times armspan + 32.972.$$

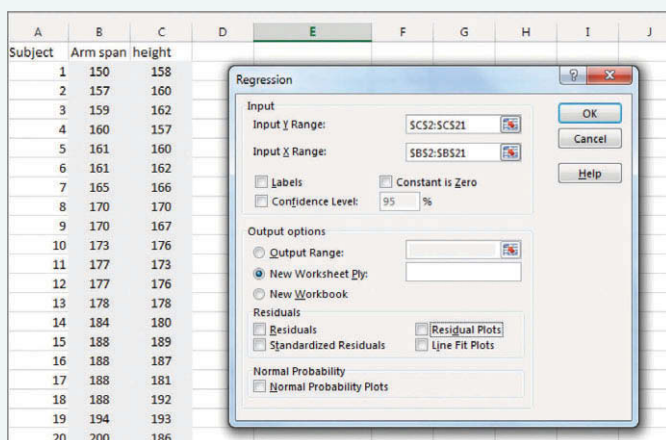


Example 2 Determining the equation of the least squares line using Excel

Use Excel to fit a least squares line to the data relating height (y) to arm span (x) for the group of 20 adults. Give your answers correct to three decimal places.

Solution

Select the Data Analysis command button on the Data tab. When Excel displays the **Data Analysis** dialog box, select the **Regression** tool from the Analysis Tools list and then click OK. Place the cursor on the **Input X Range** dialogue box, and then select the values of arm span. Next, place the cursor on the **Input Y Range** dialogue box, and then select the values of height.



Select OK to generate the following output.

SUMMARY OUTPUT						
Regression Statistics						
Multiple R	0.950668					
R Square	0.903769					
Adjusted R Square	0.898423					
Standard Error	3.788415					
Observations	20					
ANOVA						
	df	SS	MS	F	Significance F	
Regression	1	2426.212	2426.212	169.0494	1.37E-10	
Residual	18	258.3376	14.35209			
Total	19	2684.55				
Coefficients						
	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%
Intercept	32.97195	10.85292	0.007074	10.17082	55.77308	10.17082
X Variable	0.80664	0.06204	13.0019	1.37E-10	0.676299	0.936982

The intercept and slope are given in the bottom table of the output. Thus:

- intercept $c = 32.97195 = 32.972$ to three decimal places
- slope $m = 0.80664 = 0.807$ to three decimal places.

Note also that the value of the coefficient of determination R^2 is also given in the output; here $R^2 = 0.903769 = 90.4\%$.

The equation of the least squares regression line is: $height = 0.807 \times armspan + 32.972$

Section Summary

- ▶ The vertical distance between a point on a bivariate plot and a line fitted to the data is called a **residual**.
- ▶ The method of **least squares** is a method for fitting a straight line to a scatterplot, based on minimising the sum of the squared residuals.
- ▶ The equation of the least squares line is given by $y = mx + c$, where:

$$\text{the slope } (m) \text{ is given by: } m = \frac{rs_y}{s_x}$$

and

$$\text{the intercept } (c) \text{ is then given by: } c = \bar{y} - m\bar{x}$$

Here:

- ▷ r is the correlation coefficient
- ▷ s_x and s_y are the standard deviations of x and y
- ▷ $\bar{x}$ and $\bar{y}$ are the mean values of x and y .
- ▶ A calculator or computer can be used to determine the intercept and slope of the least squares line from the bivariate data.



Exercise 2A

Using a formula to determine the equation of a least squares regression line

Example 1

- 1 We wish to determine the equation of the least squares regression line that enables the *pollution level* beside a freeway to be predicted from *traffic volume*.

- a Which is the response variable (RV) and which is the explanatory variable (EV)?
- b Use the formula to determine the equation of the least squares regression line that enables the pollution level (y) to be predicted from the traffic volume (x), where:

$$r = 0.940 \quad \bar{x} = 11.4 \quad s_x = 1.87$$

$$\bar{y} = 231 \quad s_y = 97.9$$

Write the equation in terms of *pollution level* and *traffic volume* with the intercept and slope rounded to the nearest whole number.

- 2 We wish to determine the equation of the least squares regression line that enables *life expectancy* in a country to be predicted from *birth rate*.

- a Which is the response variable (RV) and which is the explanatory variable (EV)?
- b Use the formula to determine the equation of the least squares regression line that enables life expectancy (y) to be predicted from birth rate (x), where:

$$r = -0.810 \quad \bar{x} = 34.8 \quad s_x = 5.41 \quad \bar{y} = 55.1 \quad s_y = 9.99$$

Write the equation in terms of *life expectancy* and *birth rate* with the y-intercept and slope written correct to one decimal place.

- 3** We wish to determine the equation of the least squares regression line that enables the *distance* travelled by a car (in 1000s of km) to be predicted from its *age* (in years).

- a** Which is the response variable (RV) and which is the explanatory variable (EV)?
b Use the formula to determine the equation of the least squares regression line that enables distance travelled (y) by a car to be predicted from its age (x), where:

$$r = 0.947 \quad \bar{x} = 5.63 \quad s_x = 3.64 \quad \bar{y} = 78.0 \quad s_y = 42.6$$

Write the equation in terms of *distance travelled* and *age* with the y -intercept and slope written correct to one decimal place.

Using technology to determine the equation of a least squares line

Example 2

- 4** The table shows the number of sit-ups and push-ups performed by six students.

<i>Sit-ups</i> (x)	52	15	22	42	34	37
<i>Push-ups</i> (y)	37	26	23	51	31	45

Let the number of *sit-ups* be the explanatory (x) variable. Use appropriate technology to show that the equation of the least squares regression line is:

$$\text{push-ups} = 0.57 \times \text{sit-ups} + 16.45$$

- 5** The table shows average hours worked per week and university participation rate (%) in six countries.

<i>Hours</i>	35.0	43.0	38.2	39.8	35.6	34.8
<i>Rate</i>	26	20	36	25	37	55

Use appropriate technology to show that the equation of the least squares regression line that enables *hours* worked to be predicted from participation *rate* is:

$$\text{hours} = -0.17 \times \text{rate} + 43.50$$

- 6** The table shows the number of *runs* scored and *balls* faced by batsmen in a cricket match.

<i>Runs</i> (y)	27	8	21	47	3	15	13	2	15	10	2
<i>Balls</i> (x)	29	16	19	62	13	40	16	9	28	26	6

- a** Use appropriate technology to show that the equation of the least squares regression line enabling *runs* scored to be predicted from *balls* faced is:

$$y = 0.728x - 2.649$$

- b** Rewrite the regression equation in terms of the variables involved.

- 7 The table below shows the number of TVs and cars owned (per 1000 people) in six countries.

Number of TVs (y)	378	404	471	354	381	624
Number of cars (x)	417	286	435	370	357	550

We wish to predict the *number of TVs* from the *number of cars*. The following output obtained:

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.82272395							
R Square	0.6768747							
Adjusted R Square	0.59609338							
Standard Error	63.9990812							
Observations	6							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	34319.8037	34319.8037	8.37909891	0.04435459			
Residual	4	16383.5296	4095.8824					
Total	5	50703.3333						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	61.2008106	131.863267	0.46412327	0.66669383	-304.91031	427.311933	-304.91031	427.311933
X Variable 1	0.9295218	0.32111524	2.89466732	0.04435459	0.03796296	1.82108064	0.03796296	1.82108064

- a Which is the response variable?
- b Write down the regression equation in terms of the variables involved, giving the values of the intercept and slope correct to two decimal places.
- 8 The equation of a least squares line $y = mx + c$ is calculated for a set of bivariate data. Use the following information to determine the value of the correlation coefficient, r , rounded to three decimal places.

	x	y
mean	12.51	10.66
standard deviation	4.796	5.162
least squares equation	$y = 0.485x + 16.72$	

- 9 Use mathematical reasoning to answer the following questions:
- a A least squares regression line is calculated and the slope is found to be negative. What does this tell us about the sign of the correlation coefficient?
- b The correlation coefficient is zero. What does this tell us about the slope of the least squares regression line?
- c The correlation coefficient is zero. What does this tell us about the intercept of the least squares regression line?

Paper 1-style multiple-choice questions

- 10** A least squares line of the form $y = mx + c$ is fitted to a scatterplot. Which of the following statements is always true:
- A** The line will divide the data points so that there are as many points above the line as below the line.
 - B** The sum of the vertical distances from the line to each data point will be a minimum.
 - C** x is the explanatory variable and y is the response variable.
 - D** y is the explanatory variable and x is the response variable.
- 11** The statistical analysis of the set of bivariate data involving variables x and y resulted in the information displayed in the table below:

	x	y
mean	32.5	44.6
standard deviation	3.42	6.84
least squares equation	$y = 1.45x - 2.56$	

Using this information the value of the correlation coefficient r for this set of bivariate data is closest to

- A** 0.73
 - B** 0.34
 - C** 0.50
 - D** 0.53
- 12** A retailer recorded the number of cups of soup sold and the day's minimum temperature over eight consecutive Saturdays one summer.

Temperature ($^{\circ}\text{C}$)	10	7	-4	-2	11	4	-9	1
Number of cups of soup sold	60	79	100	95	75	90	130	90

The equation of the least squares regression line fitted to the data is closest to:

- A** $\text{number of cups of soup} = -0.32 \times \text{temperature} + 30.98$
- B** $\text{number of cups of soup} = -2.73 \times \text{temperature} + 96.02$
- C** $\text{number of cups of soup} = 96.0 \times \text{temperature} - 2.73$
- D** $\text{temperature} = 30.98 \times \text{number of cups of soup} - 0.32$

2B Using the least squares line to model a linear relationship

Learning intentions

- ▶ To interpret the intercept and slope of the least squares line.
- ▶ To use the equation of the least squares line to make predictions.
- ▶ To use a residual plot to investigate the linearity assumption.
- ▶ To use the coefficient of determination to tell us how well a statistical model predicts an outcome.
- ▶ To report a regression analysis.

Interpreting the slope and intercept of a fitted line

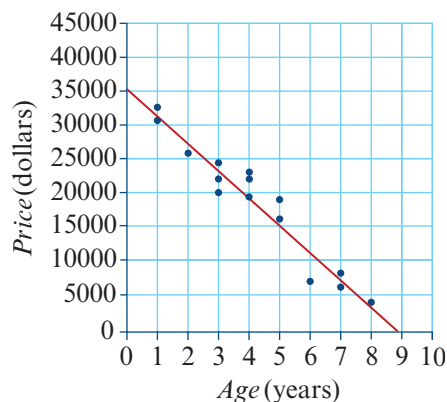
Suppose that we wish to investigate the nature of the association between the price of a secondhand car and its age. The ultimate aim is to determine a mathematical model that will enable the price of a secondhand car to be predicted from its age.

The age (in years) and price (in dollars) of a selection of secondhand cars of the same brand and model have been collected and are recorded in a table (shown).

Age (years)	Price (dollars)	Age (years)	Price (dollars)	Age (years)	Price (dollars)
1	32 500	3	22 000	5	18 400
1	30 500	4	22 000	6	6 500
2	25 600	4	23 000	7	6 400
3	20 000	4	19 200	7	8 500
3	24 300	5	16 000	8	4 200

We start our investigation of the association between price and age by constructing a scatterplot and using it to describe the association in terms of strength, direction and form.

In this analysis, *age* is the explanatory variable.



From the scatterplot, we see that there is a strong, negative, linear association between the price of the car and its age. There are no clear outliers. The correlation coefficient is $r = -0.964$.

The equation of the least squares regression line from these data is:

$$\text{price} = -3940 \times \text{age} + 35\,100$$

The values of the intercept and slope give us valuable information about how the price of secondhand cars varies over time.



Example 3 Interpreting the slope and intercept of a regression line

The equation of a regression line that enables the *price* of a secondhand car to be predicted from its *age* is:

$$\text{price} = -3940 \times \text{age} + 35\,100$$

- a** Interpret the slope in terms of the variables *price* and *age*.
- b** Interpret the intercept in terms of the variables *price* and *age*.

Solution

- a** On average, the price of these cars decreases by \$3940 each year.
- b** On average, the price of these cars when new was \$35 100.

Explanation

The *slope* predicts the average change (increase/decrease) in the *price* for each one-year increase in the *age*. Because the slope is negative, it will be a decrease.

The *intercept* predicts the value of the *price* of the car when *age* equals 0; that is, when the car was new.

Interpreting the slope and intercept of a regression line

For the regression line $y = mx + c$:

- the slope (m) estimates the average change (increase/decrease) in the *response variable* (y) for each one-unit increase in the *explanatory variable* (x)
- the intercept (c) estimates the average value of the *response variable* (y) when the *explanatory variable* (x) equals 0.

Using the equation of the least squares line to make predictions

Once the equation of the least squares line is known, we can use it to predict a likely value of the response variable based on a given value of the explanatory variable.



Example 4 Using the equation of a regression line to make predictions

The relationship between *eyesight test score* and *age*, based on a sample of adults aged from 20 to 60 years old, was found to be:

$$\text{eyesight test score} = -0.0387 \times \text{age} + 4.65$$

Use the equation to predict (to two decimal places) the eyesight test score for:

- a** a person aged 40
- b** a person aged 65

Solution

- a** Person aged 40 eyesight test score
 $= -0.0387 \times 40 + 4.65 = 3.10$
- b** Person aged 65 eyesight test score
 $= -0.0387 \times 65 + 4.65 = 2.13$

Explanation

Substitute in the formula and evaluate.

Substitute in the formula and evaluate.

Interpolation and extrapolation

When using the equation of the least squares line to make predictions, we must be aware that, strictly speaking, the equation we have found applies only to the range of data values used to derive the equation.

Consider again the equation:

$$\text{price} = -3940 \times \text{age} + 35\,100$$

Using this equation and rounding to the nearest dollar we would predict that:

- a car which is 2 years old would have a price of \$27 220 (price = $-3940 \times 2 + 35\,100$)
- a car which is 7 years old would have a price of \$7520 (price = $-3940 \times 7 + 35\,100$)
- a car which is 12 years old would have a price of -\$12 180 (price = $-3940 \times 12 + 35\,100$)

This last result, -\$12 180, points to one of the limitations of substituting into a regression equation without thinking carefully. Using this regression equation, we have predicted a negative price, which is clearly not correct.

The problem is using the least squares equation to make predictions well outside the range of values used to calculate this equation, which was for cars up to 8 years old. Without knowing that the model works equally well for cars older than this, we are venturing into unknown territory and can have little faith in our predictions.

As a general rule, a least squares equation only applies to the range of values of the explanatory variables used to determine the equation. Thus, we are reasonably safe using the line to make predictions that lie roughly within this data range, from 1 to 8 years. The process of making a prediction within the range of values of the explanatory variable used to derive the least squares equation is called **interpolation**, and we can have some confidence in these predictions.

However, we must be extremely careful about how much confidence we have in predictions made outside the range of values of the explanatory variable. Making predictions outside the data range is called **extrapolation**.

Interpolation and extrapolation

Predicting within the range of values of the explanatory variable is called **interpolation**. Interpolation is generally considered to give a **reliable** prediction.

Predicting outside range of values of the explanatory variable is called **extrapolation**. Extrapolation is generally considered to give an **unreliable** prediction.

Constructing and interpreting a residual plot

So far all of our analysis has been based on the assumption that there is a linear relationship between the two variables. This is why it has been essential to examine the scatterplot before proceeding with any further analyses. However, sometimes the scatterplot is not sensitive enough to reveal the non-linear structure of a relationship. To gain more information about the validity of the linearity assumption, a **residual plot** can be constructed.

Residuals are vertical distances between the least squares line and the actual data value.

Residual plot

A residual plot is a plot of the **residuals** (plotted on the vertical axis) against the **explanatory variable** (plotted on the horizontal axis), where:

$$\text{Residual value} = \text{actual data value} - \text{predicted data value}$$

Residuals can be positive, negative or zero:

- Data points above the least squares line have a positive residual.
- Data points below the least squares line have a negative residual.
- Data points on the least squares line have zero residual.

Consider again the age (in years) and price (in dollars) of a selection of secondhand cars which was fitted to the following data:

Age (years)	Price (dollars)	Age (years)	Price (dollars)	Age (years)	Price (dollars)
1	32 500	3	22 000	5	18 400
1	30 500	4	22 000	6	6 500
2	25 600	4	23 000	7	6 400
3	20 000	4	19 200	7	8 500
3	24 300	5	16 000	8	4 200

The equation of the least squares regression line from these data is:

$$\text{price} = -3940 \times \text{age} + 35\,100$$

To determine the appropriateness of fitting the least squares regression line to these data we will construct a residual plot. But first, we need to calculate the residual for each value of the explanatory variable, in this case *age*.



Example 5 Calculating a residual

The actual price of the 6-year-old car is \$6500. Calculate the residual when its price is predicted using the regression equation: $\text{price} = 35\,100 - 3940 \times \text{age}$

Solution

Actual price: \$6500

$$\begin{aligned}\text{Predicted price} &= 35\,100 - 3940 \times 6 \\ &= \$11\,460\end{aligned}$$

$$\begin{aligned}\text{Residual} &= \text{actual} - \text{predicted} \\ &= \$6500 - \$11\,460 \\ &= -\$4960\end{aligned}$$

Explanation

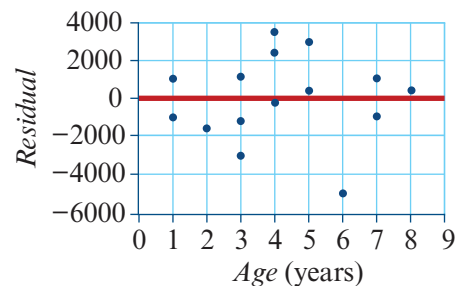
Write down the actual price.

Determine the predicted price using the least squares regression equation:

$$\text{price} = -3940 \times \text{age} + 35\,100$$

Calculate the residual.

By completing this calculation for all data points, we can construct a residual plot. Because the mean of the residuals is always zero, we will construct the horizontal axis for the plot at zero (indicated by the red line) as shown.



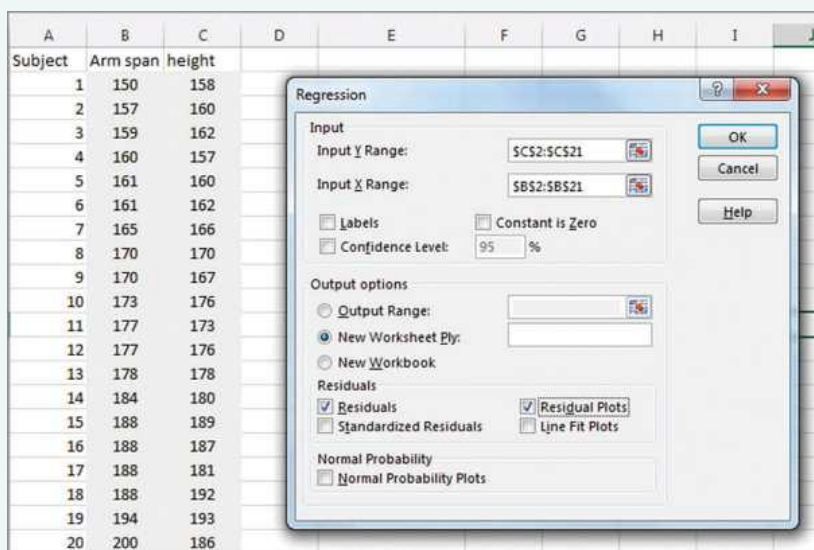


Example 6 Using Excel to construct a residual plot

Construct a residual plot for the least squares line fitted to the data relating height (y) to arm span (x) for the group of 20 adults from Example 2.

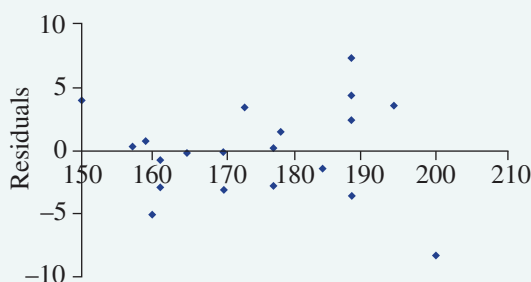
Solution

When setting up a spreadsheet to carry out the linear regression, check the boxes Residuals and Residual Plots.



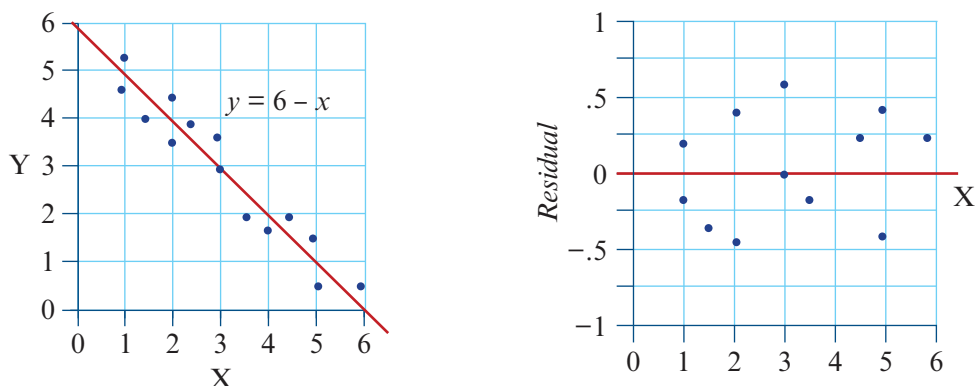
Select OK. In addition to the regression output columns containing the predicted values of the RV (in this case height), the associated residuals will be produced, together with the associated residual plot.

RESIDUAL OUTPUT		
Observation	Predicted $\hat{y}$	Residuals
1	153.968	4.032021
2	159.6145	0.38554
3	161.2277	0.772259
4	162.0344	-5.03438
5	162.841	-2.84102
6	162.841	-0.84102
7	166.0576	-0.05758
8	170.1008	-0.10078
9	170.1008	-3.10078
10	172.5207	3.479296
11	175.7473	-2.74726
12	175.7473	0.252735
13	176.5539	1.446095
14	181.3937	-1.39375
15	184.6203	4.379693
16	184.6203	2.379693
17	184.6203	-3.62031
18	184.6203	7.379693
19	189.4601	3.539852
20	194.3	-8.29999

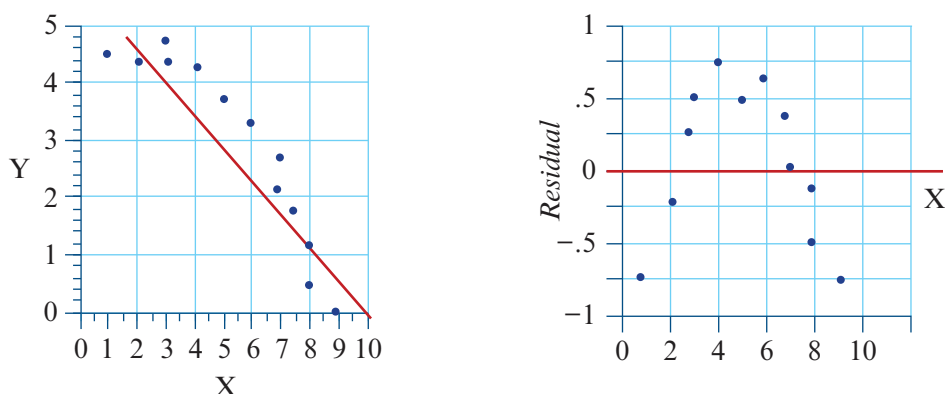


What are we looking for in a residual plot?

The residual plot is used to check the **linearity assumption** required for a least squares line. The below shows a relationship that is clearly linear. When a line is fitted to the data, the resultant residual plot appears to be a random collection of points roughly spread around zero.



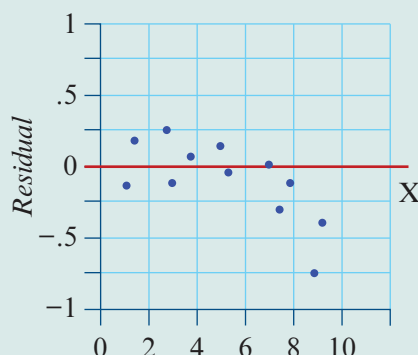
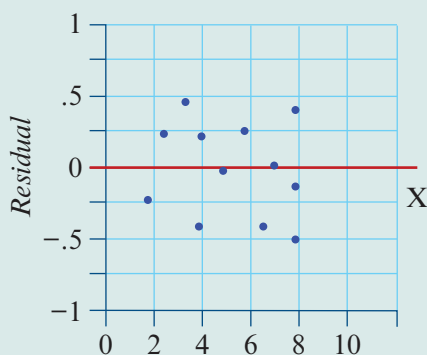
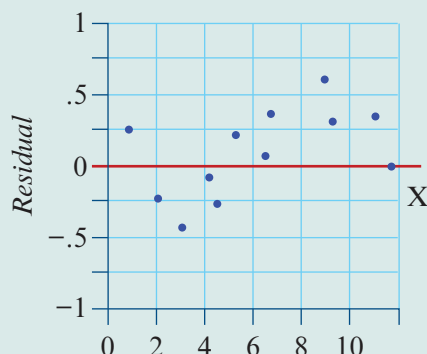
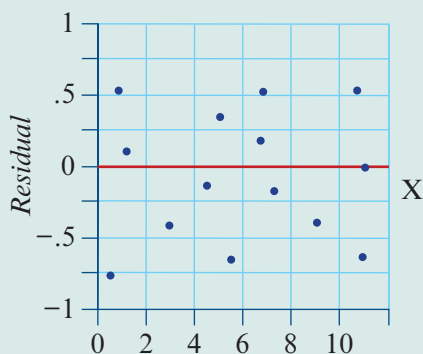
By contrast, the relationship shown in the following scatterplot is clearly non-linear. Fitting a straight line to the data results in the residual plot shown. While there is some random behaviour, there is also a clearly identifiable curve shown in the scatterplot.



In summary, if a residual plot shows evidence of some sort, then it is likely that the underlying relationship is non-linear. However, if the residual plot appears to be a random collection of points roughly spread around zero, then we can be happy that our original assumption of linearity was reasonable and that we have appropriately modelled the data. From a visual inspection, it is difficult to say with certainty that a residual plot is random. It is easier to see when it is not random. For present purposes, it is sufficient to say that a clear lack of a pattern in a residual plot is an indication of randomness.


Example 7 Interpreting a residual plot

Examine each of the following residual plots and determine if the assumption of linearity has been met.


Solution

Plot A – residuals look random, so the linearity assumption is met.

Plot B – there is a clear curve in the residuals, so the linearity assumption is not met.

Plot C – residuals look random, so the linearity assumption is met.

Plot D – there is a clear curve in the residuals, so the linearity assumption is not met.

Explanation

Examine each plot, looking for a pattern or structure in the residual.

The coefficient of determination

In the previous chapter we defined the coefficient of determination R^2 , which is numerically equal to the value of the correlation coefficient. The coefficient of determination can be considered a measure of the predictive power of a regression equation. While the association between the price of a second-hand car and its age does not explain all the variation in price, knowing the age of a car does give us some information about its likely price.

For a perfect relationship, the regression line explains 100% of the variation in prices. In this case, with $r = -0.964$ we have the:

$$\text{coefficient of determination} = r^2 = 0.964^2 = 0.930 \text{ or } 93.0\%$$

Thus, we can conclude that:

93% of the variation in price of the second-hand cars can be explained by the variation in the ages of the cars.

In this case, the regression equation has very good predictive power. As a guide, any relationship with a coefficient of determination greater than 30% can be regarded as having good predictive power. In practice, even much lower values of the coefficient of determination can be useful. We often determine that there are several explanatory variables that may help explain the value of the response variable, and we can use the value of the coefficient of determination to determine their relative importance.



Example 8 Using the coefficient of determination to compare explanatory variables

In a recent study across a number of countries, the correlation between educational attainment and the amount spent on education was found to be 0.26, whilst the correlation between educational attainment and the student : teacher ratio was found to be -0.38 .

- a** Determine the values of the coefficient of determination between *educational attainment* and the *amount spent on education*, and *student : teacher ratio* respectively.
- b** Which of the variables, *amount spent on education* or *student : teacher ratio* is a better predictor of educational attainment?

Solution

- a** *amount spent on education*: $R^2 = 0.26^2 = 6.8\%$
student : teacher ratio : $R^2 = (-0.38)^2 = 14.4\%$
- b** The variable *student : teacher ratio* explains 14.4% of the variation in *educational attainment*, making it a more important explanatory variable than the *amount spent on education*, which explains only 6.8%; hence, this variable is a better predictor.

Reporting the results

After the regression analysis has been carried out, a report can be written about the findings. Here is an example of a report that could be written to summarise the association between the price and age of secondhand cars.

Report

To investigate the association between the price and age of secondhand cars, data was collected from a sample of 15 cars. The scatterplot showed a strong, negative, linear relationship between the price and age of secondhand cars ($r = -0.964$), indicating that older cars tend to be lower in price. There were no obvious outliers, and the lack of a clear pattern in the residual plot confirmed the linearity assumption.

The equation of the least squares regression line is:

$$\text{price} = -3940 \times \text{age} + 35\,100$$

The intercept predicts that, on average, the price of the cars when new was \$35 100.

The slope predicts that, on average, the price of the cars decreases by \$3940 each year.

The coefficient of determination indicates that 93% of the variation in the price of these secondhand cars is explained by the variation in their age.

Performing a regression analysis

A complete analysis of the association between two numerical variables is often called a **regression analysis**, and it involves all of the following analyses, the results of which are collated in a report.

Performing a regression analysis

To carry out a **regression analysis** involves several processes, which include:

- constructing a scatterplot to investigate the nature of an association
- calculating the correlation coefficient to indicate the strength of the relationship
- determining the equation of the least squares regression line
- interpreting the coefficients of the y-intercept (c) and the slope (m) of the least squares regression line $y = mx + c$
- calculating residuals and using a residual plot to test the assumption of linearity
- calculating and interpreting the coefficient of determination
- using the regression line to make predictions
- writing a report on the findings.

Section Summary

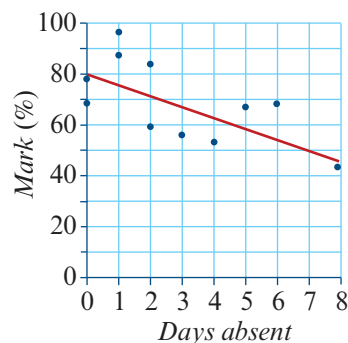
- ▶ For the regression line $y = mx + c$:
 - ▷ the slope (m) estimates the average change (increase/decrease) in the response variable (y) for each one-unit increase in the explanatory variable (x)
 - ▷ the intercept (c) estimates the average value of the response variable (y) when the explanatory variable (x) equals 0.
- ▶ A value of the explanatory variable can be substituted into the equation of the least squares line to predict the corresponding value of the response variable.
 - ▷ Predicting within the range of values of the explanatory variable is called **interpolation**.
 - ▷ Predicting outside the range of values of the explanatory variable is called **extrapolation**.
- ▶ A residual plot is used to investigate the linearity assumption. A random pattern suggests that the linear model is appropriate for the data.
- ▶ The coefficient of determination can tell us how well a statistical model predicts an outcome.



Exercise 2B

Some basics

- 1 Determine the equation of the least squares line shown on the scatterplot in terms of the variables *mark* and *days absent*. Determine the intercept correct to the nearest whole number and the slope correct to one decimal place.



Interpreting the intercept and slope of a least squares line

Example 3

- 2 The equation of a least squares line enabling hand span to be predicted from height is:

$$\text{hand span} = 0.33 \times \text{height} + 2.90$$

Complete the following sentences:

- a The slope tells us on average that *hand span* increases by cm for each cm increase in *height*.
- b The intercept tell us that on average people with a *height* of cm will have a *handspan* of cm.

- 3** The following least squares equation can be used to predict a company's weekly sales (\$) from their weekly online advertising expenditure (\$).

$$\text{sales} = 4.85 \times \text{expenditure} + 575$$

- a** Write down the value of the intercept, and interpret this value in the context of the variables in the equation.
- b** Write down the value of the slope, and interpret this value in the context of the variables in the equation.

Using the least squares line to make predictions

Example 4

- 4** For children between the ages of 36 and 60 months, the equation relating their *height* (in cm) to their *age* (in months) is:

$$\text{height} = 0.40 \times \text{age} + 72$$

Use this equation to predict the height (to the nearest cm) of a child with the following ages. In each case indicate whether you are interpolating or extrapolating.

- a** 20 months old **b** 50 months old **c** 65 months old

- 5** When preparing between 25 and 100 meals, a hospital's cost (in dollars) is given by the equation:

$$\text{cost} = 6.70 \times \text{meals} + 487.50$$

Use this equation to predict the cost (to the nearest dollar) of preparing the following meals. In each case indicate whether you are interpolating or extrapolating.

- a** 0 meals **b** 80 meals **c** 110 meals

- 6** For males of heights from 150 cm to 190 cm tall, the equation relating a *son's height* (in cm) to his *father's height* (in cm) is:

$$\text{son's height} = 0.525 \times \text{father's height} + 83.9$$

Use this equation to predict (to the nearest cm) the adult height of a male whose father is each of the following heights. State, with a reason, how reliable your predictions are in each case.

- a** 170 cm tall **b** 200 cm tall **c** 155 cm tall

Calculating a residual

Example 5

- 7** The equation of a regression line that enables hand span to be predicted from height is:

$$\text{hand span} = 0.33 \times \text{height} + 2.90$$

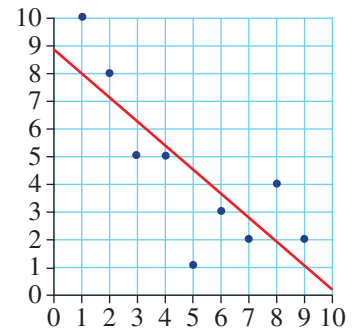
- a** Using this equation, show that the hand span of a person who is 160 cm is 55.7 cm.
- b** This person has an actual hand span of 58.5 cm. Show that the residual value for this person is 2.8 cm.

- 8 For a 100 km trip, the equation of a least squares line that enables the fuel consumption of a car (in litres) to be predicted from its weight (kg) is:

$$\text{fuel consumption} = 0.01 \times \text{weight} - 0.1$$

- a Use this equation to predict (to one decimal place) the fuel consumption of a car which weighs 980 kg.
- b This car has an actual fuel consumption of 8.9 litres. What is the residual value for this data point?
- 9 From the scatterplot shown, determine (to the nearest whole number) the residual values when the value of x is equal to:

- a 1
- b 3
- c 8



Constructing a residual plot

- 10 The table shows the number of sit-ups and push-ups performed by six students.

<i>Sit-ups</i>	52	15	22	42	34	37
<i>Push-ups</i>	37	26	23	51	31	45
<i>Residual</i>	-9.1		-6.0	10.6		7.5

Let the number of *sit-ups* be the explanatory (x) variable. The equation of the least squares line is:

$$\text{push-ups} = 0.57 \times \text{sit-ups} + 16.45$$

- a Complete the table of residuals.
- b Construct a residual plot.
- 11 The table shows average hours worked per week and university participation rate (%) in six countries.

<i>Hours</i>	35.0	43.0	38.2	39.8	35.6	34.8
<i>Rate</i>	26	20	36	25	37	55
<i>Residual</i>						

The equation of the least squares line that enables *hours* worked to be predicted from participation *rate* is:

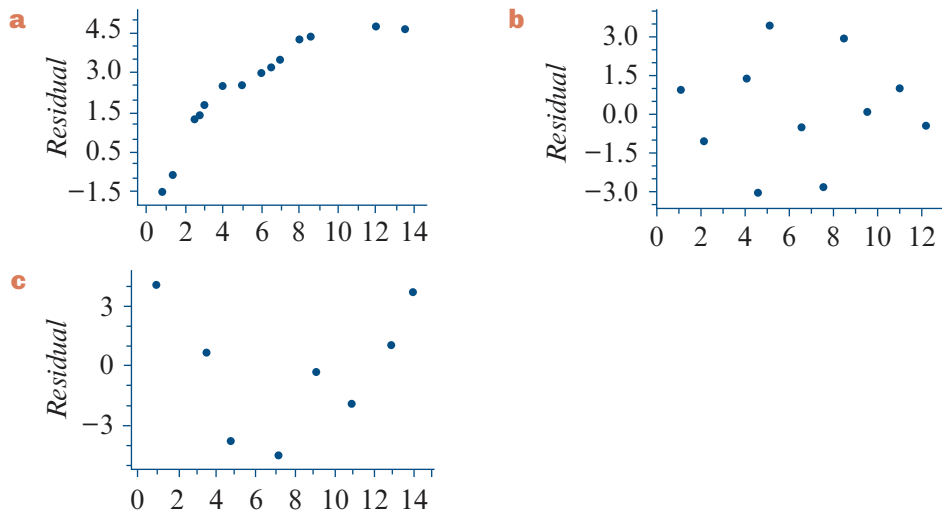
$$\text{hours} = -0.17 \times \text{rate} + 43.50$$

- a** Complete the table of residuals, giving the value of the residuals rounded to two decimal places.
- b** Construct a residual plot.

Interpreting a residual plot

Example 7

- 12** Each of the following residual plots has been constructed after a least squares line has been fitted to a scatterplot. Explain which of the residual plots suggest that the use of a linear model to fit the data was inappropriate.



Using the coefficient of determination to compare explanatory variables

Example 8

- 13** A teacher found that the correlation between her students' scores on an IQ test (*IQ*) and their final examination score in Year 12 (*exam score*) is 0.45, whilst the correlation between the average number of hours they spend each week studying mathematics (*hours*) and their final examination score in Year 12 (*exam score*) is 0.65.
- a** Determine the value of the coefficient of determination between *exam score* and *IQ*, expressed as a percentage rounded to one decimal place.
 - b** Determine the value of the coefficient of determination between *exam score* and *hours*, expressed as a percentage rounded to one decimal place.
 - c** Which of the explanatory variables, *IQ* or *hours*, is a better predictor of *exam score*?

Performing a regression analysis

- 14** In an investigation of the association between the food energy content (in calories) and the fat content (in g) in a standard-sized packet of chips, the least squares regression line was found to be:

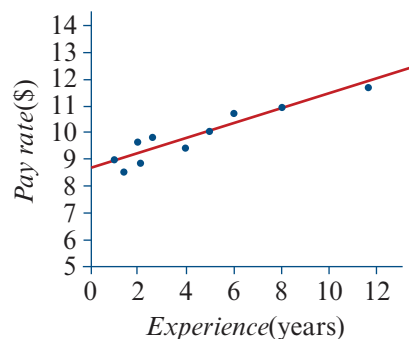
$$\text{energy content} = 14.7 \times \text{fat content} + 27.8 \quad R^2 = 0.7569$$

- a** Write down the value of the intercept, and interpret this value in the context of the variables in the equation.
 - b** Write down the value of the slope, and interpret this value in the context of the variables in the equation.
 - c** Interpret the value of the coefficient of determination in terms of the variables *energy content* and *fat content*.
 - d** Use this equation to predict the energy content of a packet of chips which contains 8 grams of fat.
 - e** If the actual energy content of a packet of chips containing 8 grams of fat is 132 calories, calculate the value of the residual.
- 15** In an investigation of the association between the success rate (%) of sinking a putt and the distance from the hole (in cm) of amateur golfers, the least squares regression line was found to be:

$$\text{success rate} = -0.278 \times \text{distance} + 98.5 \quad R^2 = 0.497$$

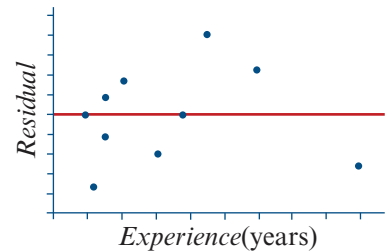
- a** Write down the slope of this regression equation and interpret.
 - b** Use the equation to predict the success rate when a golfer is 90 cm from the hole.
 - c** At what distance (in metres) from the hole does the regression equation predict an amateur golfer to have a 0% success rate of sinking the putt?
 - d** Calculate the value of r , rounded to three decimal places.
 - e** Write down the value of the coefficient of determination in percentage terms and interpret.
- 16** The scatterplot opposite shows the pay rate (dollars per hour) paid by a company to workers with different years of work experience. The least squares equation is:

$$y = 0.289x + 8.56 \quad \text{with } r = 0.967$$



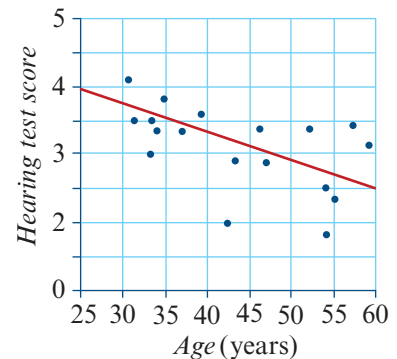
- a** Write down the equation of the least squares regression line in terms of the variables *pay rate* and *years of experience*.
- b** Determine percentage of the variation in a person's pay rate which can be explained by the variation in their work experience.
- c** Interpret the y -intercept in terms of the variables *pay rate* and *years of experience*.
- d** Interpret the slope in terms of the variables *pay rate* and *years of experience*.

- e** Use the least squares regression equation to:
- i** predict the hourly wage of a person with 8 years of experience
 - ii** determine the residual value if the actual hourly wage is \$11.20 per hour.
- f** The residual plot for this regression analysis is shown opposite. Use the residual plot to assess the assumption that the relationship between *pay rate* and *years of experience* is linear.



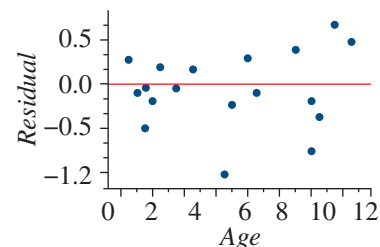
- 17** The scatterplot opposite shows scores on a hearing test against age. In analysing the data, a statistician produced the following statistics:

- coefficient of determination: $R^2 = 0.370$
- least squares line: $y = -0.043x + 4.9$



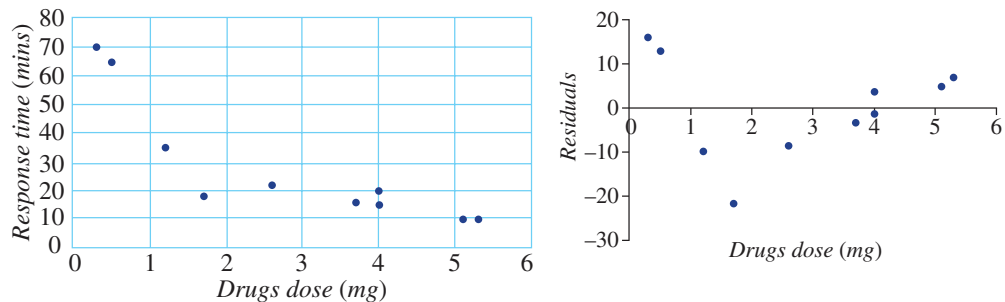
- a** Determine the value of correlation coefficient, r , for the data.
- b** Interpret the coefficient of determination in terms of the variables *hearing test score* and *age*.
- c** Write down the equation of the least squares line in terms of the variables *hearing test score* and *age*.
- d** Write down the slope and interpret.
- e** Use the least squares equation to:
 - i** predict the hearing test score of a person who is 20 years old
 - ii** determine the residual value if the person's actual hearing test score is 2.0.
- f** Use the graph to estimate the value of the residual for the person aged:
 - i** 35 years
 - ii** 55 years

- g** The residual plot for this regression analysis is shown opposite. Does the residual plot support the assumption that the relationship between hearing test score and age is essentially linear? Explain your answer.



Reporting the results of a regression analysis

- 18** In a study of the effectiveness of a pain relief drug, the response time (in minutes) was measured for different drug doses (in mg). A least squares regression analysis was conducted to enable response time to be predicted from drug dose. The results of the analysis are displayed.



Least squares equation: $y = mx + c$

$$m = -10.19 \quad c = 57.05 \quad r = -0.855 \quad R^2 = 0.731$$

Use this information to complete the following report.

Report

From the scatterplot we see that there is a strong relationship between response time and : $r =$. There are no obvious outliers.

The equation of the least squares line is:

response time = $\times$ drug dose +

The slope of the regression line predicts that, on average, response time increases/decreases by minutes for each 1-milligram increase in drug dose.

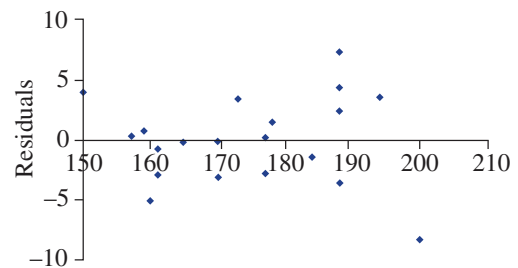
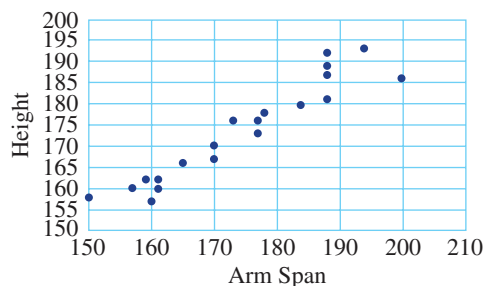
The y-intercept of the regression line predicts that, on average, the response time when no drug is administered is minutes.

The coefficient of determination indicates that, on average, % of the variation in is explained by the variation in .

The residual plot shows a , calling into question the use of a linear equation to describe the relationship between response time and drug dose.

- 19 The following information was generated using Excel to investigate the relationship between arm span (cm) and height (cm) for a group of 20 adults.

SUMMARY OUTPUT								
Regression	Statistics							
Multiple R	0.950668							
R Square	0.903769							
Adjusted R squared	0.898423							
Standard Error	3.788415							
Observations	20							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	1	2426.212	2426.212	169.0494	1.37E-10			
Residual	18	258.3376	14.35209					
Total	19	2684.55						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	32.97195	10.85292	3.038072	0.007074	10.17082	55.77308	10.17082	55.77308
X Variable 1	0.80664	0.06204	13.0019	1.37E-10	0.676299	0.936982	0.676299	0.936982



Use this information to complete the following report.

Report

From the scatterplot, we can see that there is a strong, relationship between height and arm span: $r =$. There are no obvious outliers. The equation of the least squares regression line is:

$$\text{height} = \text{} \times \text{arm span} + \text{}$$

The slope of the regression line predicts an increase/decrease of cm in height for each 1 cm increase in .

The coefficient of determination indicates that for this sample % of the variation in is explained by the variation in .

- 20** In a study of the relationship between height and weight for females, the following data were collected.

Subject	Height (cm)	Weight (kg)
1	169	55
2	155	54
3	175	64
4	168	56
5	170	59
6	168	60
7	160	47
8	153	45
9	166	60
10	165	52
11	160	49
12	183	63
13	170	57
14	173	56
15	154	48

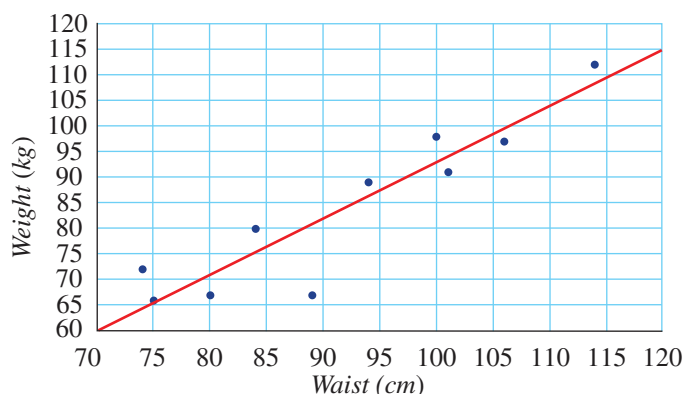
- a** Construct a scatterplot, calculate the value of the correlation coefficient, determine the equation of the least squares regression line, construct a residual plot, and determine the value of the coefficient of determination (use *height* as the explanatory variable and *weight* as the response variable).
- b** Use these analyses to construct a report, using the structure of the report from Question **16** as a guide.
- 21** A study of the association between the average score in an examination in each of 25 schools and the student : staff ratio in that school resulted in following information.

Variable	Mean	Stand dev
<i>student : staff ratio</i>	13.404	4.128
<i>score</i>	71.669	12.013
Correlation coefficient	$r = -0.651$	

Use this information to predict the average examination score in a school with a student : staff ratio of 15. Give your answer correct to one decimal place.

Paper 1-style multiple-choice questions

- 22** The scatterplot shows the weight (in kg) and waist measurement (in cm) for a group of people. A least squares line has been fitted to the scatterplot with *waist* as the explanatory variable. The equation of the least squares line is closest to:



- A** $weight = 1.10 \times waist + 60.0$ **B** $waist = 0.91 \times wrist + 60.0$
C $weight = 1.10 \times waist + 70.0$ **D** $weight = 1.10 \times waist - 17.0$
- 23** The table below shows the *life expectancy* in years and the percentage of government expenditure which is spent on health (*health*) in 10 countries.
- | | | | | | | | | | | |
|--------------------------------|------|------|-----|-----|------|-----|------|-----|------|------|
| <i>Health</i> | 17.3 | 10.3 | 4.7 | 6.0 | 20.1 | 6.0 | 13.2 | 7.7 | 10.1 | 17.5 |
| <i>Life expectancy (years)</i> | 82 | 76 | 68 | 69 | 83 | 75 | 76 | 76 | 75 | 75 |
- A least squares line which enables a country's *life expectancy* to be predicted from their expenditure on *health* is fitted to the data. The number of times that a country's predicted expenditure on *health* is greater than their actual expenditure on *health* is:
- A** 3 **B** 4 **C** 5 **D** 6
- 24** In a study of the association between the *length* in centimetres and *weight* in grams of a certain species of fish, the following least squares line was obtained:

$$weight = 23.3 \times length - 329$$

Which one of the following is a conclusion that can be made from this least squares line?

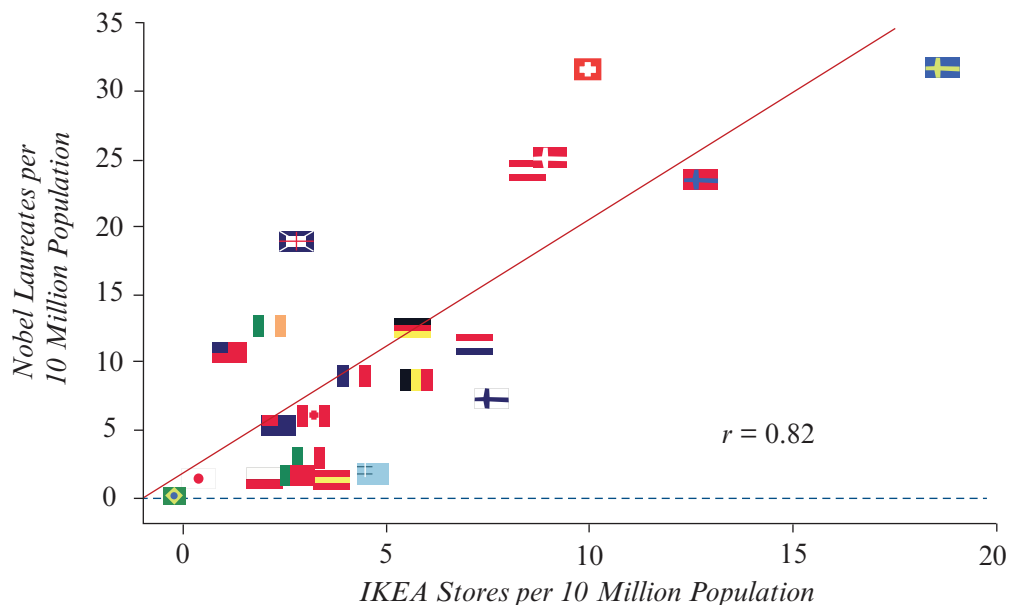
- A** On average, the *weight* of the fish increased by 23.3 g for each centimetre increase in *length*.
B On average, the *length* of the fish increased by 23.3 cm for each one gram increase in *weight*.
C On average, the *weight* of the fish decreased by 329 g for each centimetre increase in *length*.
D The equation cannot be correct as the *weight* of the fish can never be negative.

2C Association and causation

Learning intentions

- ▶ To define and differentiate the concepts of association and causation.
- ▶ To introduce the concepts of common response, confounding and coincidence.

Recently there has been interest in the strong association between the number of Nobel Prizes a country has won and the number of IKEA stores in that country ($r = 0.82$). This strong association is evident in the scatterplot below. Here, country flags are used to represent the data points.



Does this mean that one way to increase the number of Australian Nobel Prize winners is to build more IKEA stores?

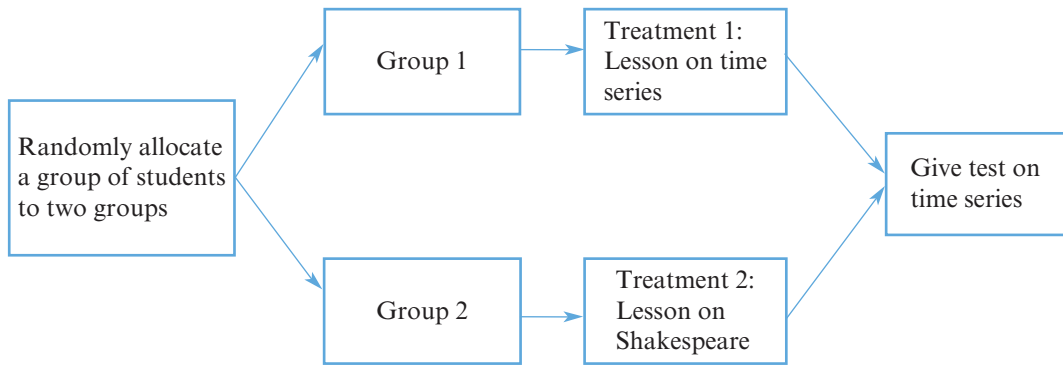
Almost certainly not, but this association highlights the problem of assuming that a strong *correlation* between two variables indicates the association between them is **causal**.

Correlation does not imply causality

A correlation tells you about the strength of the association between the variables, but no more. It tells you nothing about the source or cause of the association.

Establishing causality

To establish causality, you need to conduct an **experiment**. In an experiment, the value of the *explanatory variable* is deliberately manipulated, while all other possible explanatory variables are kept constant or controlled. A simplified version of an experiment is displayed here.



In this experiment, a class of students is randomly allocated into two groups. Random allocation ensures that both groups are as similar as possible.

Next, group 1 is given a lesson on time series (treatment 1), while group 2 is given a lesson on Shakespeare (treatment 2). Both lessons are given under the same classroom conditions. When both groups are given a test on time series the next day, group 1 does better than group 2.

We then conclude that this was because the students in group 1 were given a lesson on time series.

Is this conclusion justified?

In this experiment, the students' test score is the response variable and the type of lesson they were given is the explanatory variable. We randomly allocated the students to each group while ensuring that all other possible explanatory variables were controlled by giving the lessons under the same classroom conditions. In these circumstances, the observed difference in the response variable (*test score*) can reasonably be attributed to the explanatory variable (*lesson type*).

Unfortunately, it is extremely difficult to conduct properly controlled experiments, particularly when the people involved are going about their everyday lives.

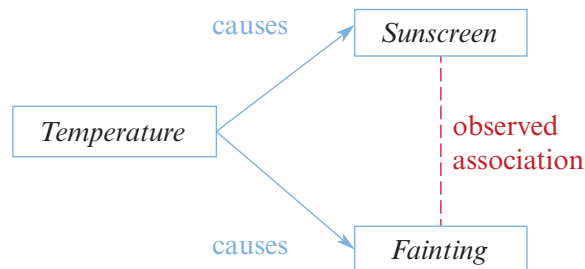
When data are collected through observation rather than experimentation, we must accept that a strong association between two variables is insufficient evidence by itself to conclude that an observed change in the response variable has been caused by an observed change in the explanatory variable. It may be, but unless all of the relevant variables are under our control, there will always be alternative non-causal explanations to which we can appeal. We will now consider the various ways this might occur.

Possible non-causal explanations for an association

Common response

Consider the following. There is a strong positive association between the number of people using sunscreen and the number of people fainting. Does this mean that applying sunscreen causes people to faint?

Almost certainly not. On hot and sunny days, more people apply sunscreen and more people faint due to heat exhaustion. The two variables are associated because they are both strongly associated with a common third variable, *temperature*. This phenomenon is called a **common response**. See the diagram below.

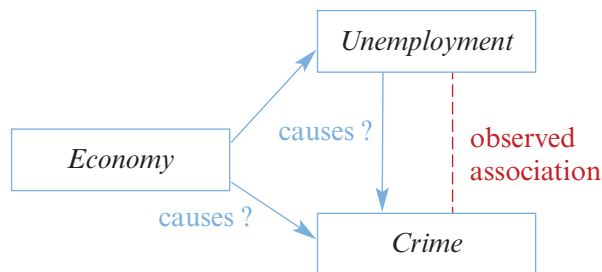


Unfortunately, being able to attribute an association to a single third variable is the exception rather than the rule. More often than not, the situation is more complex.

Confounding variables

Statistics show that *crime rates* and *unemployment rates* in a city are strongly correlated. Can you then conclude that a decrease in unemployment will lead to a decrease in crime rates?

It might, but other possible causal explanations could be found. For example, these data were collected during an economic downturn. Perhaps the state of the economy caused the problem. See the diagram below.



In this situation, we have at least two possible causal explanations for the observed association, but we have no way of disentangling their separate effects. When this happens, the effects of the two possible explanatory variables are said to be **confounded**, because we have no way of knowing which is the actual cause of the association.

Coincidence

It turns out that there is a strong correlation ($r = 0.99$) between the consumption of margarine and the divorce rate in the American state of Maine. Can we conclude that eating margarine causes people in Maine to divorce?

A better explanation is that this association is purely coincidental.

Occasionally, it is almost impossible to identify any feasible confounding variables to explain a particular association. In these cases we often conclude that the association is ‘spurious’ and it has just happened by chance. We call this **coincidence**.

Conclusion

However suggestive a strong association may be, this alone does not provide sufficient evidence for you to conclude that two variables are causally related. Unless the association is totally spurious and devoid of meaning, it will always be possible to determine at least one variable ‘lurking’ in the background that could explain the association.

Section Summary

- By itself, an observed association between two variables is never enough to justify the conclusion that two variables are causally related, no matter how obvious the causal explanation may appear to be.

Exercise 2C

- 1 A study of primary school children aged 5 to 11 years determines a strong positive correlation between height and score on a test of mathematics ability. Does this mean that taller people are better at mathematics? What common cause might counter this conclusion?
- 2 There is a clear positive correlation between the number of churches in a town and the amount of alcohol consumed by its inhabitants. Does this mean that religion is encouraging people to drink? What common cause might counter this conclusion?
- 3 There is a strong positive correlation between the total amount of ice cream consumed and the number of drownings each day. Does this mean that eating ice cream at the beach is dangerous? What common cause might explain this association?
- 4 The number of days a patient stays in hospital is positively correlated with the number of beds in the hospital. Can it be said that bigger hospitals encourage patients to stay longer than necessary just to keep their beds occupied? What common cause might counter this conclusion?
- 5 Suppose we found a high correlation between smoking rates and heart disease across a group of countries. Can we conclude that smoking causes heart disease? What confounding variable(s) could equally explain this correlation?
- 6 There is a strong correlation between cheese consumption and the number of people who died after becoming tangled in their bed sheets. What do you think is the most likely explanation for this correlation?

- 7** There is a strong positive correlation between the number of fire trucks attending a house fire and the amount of damage caused by the fire. What common cause might explain this association?
- 8** In a study of the relationship between years of experience (*years*) and annual salary (*salary*) for nurses, the Pearson's correlation coefficient between the variables was calculated to be 0.68.
- a** Describe the strength of the association between *years* and *salary*.
 - b** It is suggested that any additional qualifications achieved by the nurses above their basic training could be a confounding variable in this study.
 - i** Define confounding variable.
 - ii** Explain why the achievement of additional qualifications could be a confounding variable in this situation.
 - iii** Suggest an approach to this study which would address this problem.

Paper 1-style multiple-choice questions

- 9** There is a positive correlation between the Gross Domestic Product (GDP), a measure of a country's wealth, and the country's carbon dioxide emissions. From this information it can be concluded that:
- A** increasing a country's GDP will increase the carbon dioxide emissions of that country
 - B** decreasing a country's GDP will increase the carbon dioxide emissions of that country
 - C** countries with higher GDP also tend to have lower carbon dioxide emissions
 - D** countries with higher GDP also tend to have higher carbon dioxide emissions.
- 10** Which of the following best describes a confounding variable?
- A** A variable that gives unexpected results.
 - B** A variable that is difficult to measure.
 - C** A variable that may affect the response variable as well as, or instead of, the explanatory variable.
 - D** A variable that is made up only of categories.

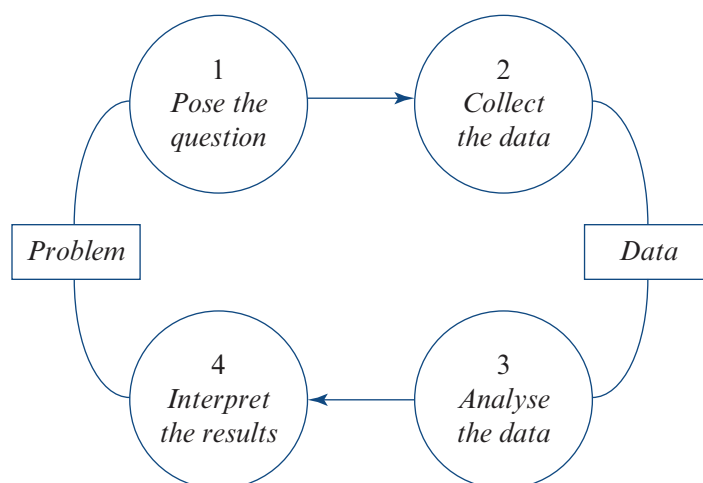
2D Solving practical problems by identifying, analysing and describing associations

Learning intentions

- To use the information in this chapter to solve practical problems involving the association between two categorical variables, or two numerical variables.

A systematic approach to solving practical problems is to follow these steps:

- 1 Pose the question – decide on the variables that allow you to address the question, and how you are going to measure these (what data you are going to collect).
- 2 Collect the data – collect or obtain the data.
- 3 Analyse the data – summarise and display the data to answer the question posed.
- 4 Interpret the results – use the results to address the question asked, and communicate what has been learned.



In this section we will focus on steps 2 and 3, and assume that the data has already be collected. It is important to identify which variable is the response variable (RV), and which is the explanatory variable (EV), and whether each of these variables is categorical or numerical. This will help to you choose the appropriate analyses, including statistical plots, to display and understand the data you have.

The following guidelines might help you make your decision. They are guidelines only, because in some instances there may be more than one suitable approach.

Type of variable		Analysis options
<i>Response</i>	<i>Explanatory</i>	
Categorical	Categorical	Two-way frequency table Percentaged two-way frequency table
Numerical	Categorical	Parallel boxplots Parallel dot plots Back-to-back stem plots Summary statistics (mean, median, standard deviation, interquartile range)
Numerical	Numerical	Scatterplot Pearson's correlation coefficient r Least squares regression line (intercept and slope) Residual plot Coefficient of determination R^2



Example 9 Identifying, analysing and describing the association between two categorical variables

Does money make us happy? Investigate this relationship. Are your conclusions the same for males and females?

Solution

One commonly used measure that we could use to measure the amount of money that people have is socioeconomic status (SES). For a measure of happiness, we could ask respondents if they were satisfied with their life overall (yes, no). Thus, we can pose the question: 'Is there a relationship between socioeconomic status and satisfaction with life overall?' Using a *data set* that recorded each respondent's socioeconomic status (low, mid, high) and answered the question 'Are you satisfied with your life overall?' (yes, no), the following analyses were carried out. Firstly, the following two-way percentage frequency table summarises the responses observed in the sample of 500 people.

<i>Are you satisfied with your life overall?</i>	<i>Socioeconomic status (SES)</i>		
	Low SES	Mid SES	High SES
Yes	84.6%	90.2%	79.5%
No	15.4%	9.8%	20.5%
Total	100.0%	100.0%	100.0%

The following two-way percentage frequency table summarises the responses for the females ($n = 250$) in the sample.

<i>Are you satisfied with your life overall?</i>	<i>Socioeconomic status (SES)</i>		
	Low SES	Mid SES	High SES
Yes	83.9%	85.2%	84.0%
No	16.1%	14.8%	16.0%
Total	100.0%	100.0%	100.0%

The following two-way percentage frequency table summarises the responses for the males ($n = 250$) in the sample.

<i>Are you satisfied with your life overall?</i>	<i>Socioeconomic status (SES)</i>		
	Low SES	Mid SES	High SES
Yes	85.2%	95.2%	74.6%
No	14.8%	4.8%	25.4%
Total	100.0%	100.0%	100.0%

Using these analyses, we are now in a position to *answer the question*.

Report

A study was conducted to investigate the relationship between socioeconomic status and satisfaction with life overall. We were particularly interested to know if those of High SES were more likely to be satisfied with their lives. Data were collected from a sample of 500 people, 250 males and 250 females. When the total group was examined, it appeared that there was a relationship between SES and satisfaction with life overall, but contrary to expectations it was the Mid SES group who were more likely to be satisfied (90.2%), followed by the Low SES group (84.6%) and then the High SES group, who were the least likely to be satisfied with their life overall (79.5%). However, further examination showed that this relationship did not hold for both males and females when each gender was examined separately. For the females, there was no relationship between SES and satisfaction with life, with each group showing similar percentages who were satisfied (Low SES: 83.9%, Mid SES: 85.2%, High SES: 84.0%). However, there was a clear relationship for males, with the Mid SES group more likely to be satisfied (95.2%), followed by the low SES group (85.2%) and then the High SES group, who were the least likely to be satisfied with their life overall (74.6%).


Example 10 Identifying, analysing and describing the association between two numerical variables

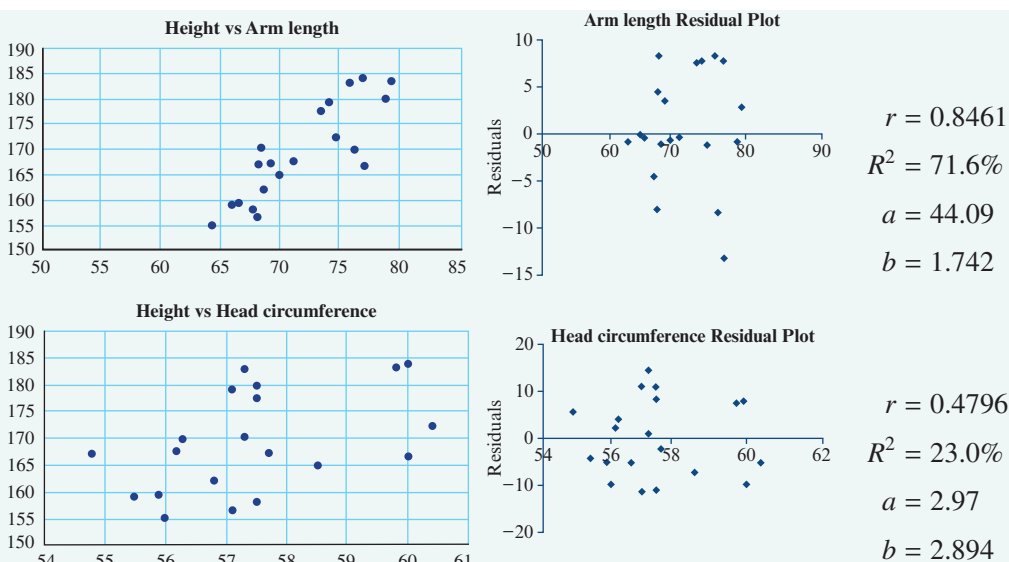
Which of our body measurements is the best predictor of height?

Solution

Let's look at two body measurements that are easy to collect: head circumference and arm length. Then we can *pose the question*: 'Which is the better predictor of height, head circumference or arm length?' The following data were *collected* from a group of 20 students (all measurements are in cm):

Subject	Height	Arm length	Head circumference
1	179.3	74.2	57.1
2	170.4	68.5	57.3
3	159.0	66.0	55.5
4	172.5	74.8	60.4
5	162.0	68.7	56.8
6	167.1	68.3	54.8
7	159.4	66.6	55.9
8	167.3	69.3	57.7
9	167.7	71.2	56.2
10	184.3	77.0	60.0
11	177.7	73.5	57.5
12	169.9	76.3	56.3
13	154.9	64.4	56.0
14	156.5	68.2	57.1
15	158.0	67.8	57.5
16	165.0	70.0	58.5
17	166.6	77.2	60.0
18	183.7	79.4	59.8
19	183.3	75.9	57.3
20	180.2	78.9	57.5

Using the data, the following *analyses* were able to be produced.



Based on these analyses the following report could be written to *answer the question*.

Report

A study was conducted to investigate which measure was a better predictor of a person's height, head circumference or arm length. Data were collected from a sample of 20 students. From the scatterplot of height versus arm length, we can see that there is a strong, positive, linear relationship between height and arm length: $r = 0.8461$. That is, those students with longer arms also tended to be taller. There are no obvious outliers, and the linearity assumption is confirmed by the residual plot. The equation of the least squares regression line is:

$$\text{height} = 1.742 \times \text{arm length} + 44.09$$

The slope of the regression line predicts an increase of 1.742 cm in height for each 1 cm increase in arm length.

From the scatterplot of height versus head circumference, we can see that there is a moderate, positive, linear relationship between height and arm length: $r = 0.480$. That is, those students with larger head circumference also tended to be taller. There are no obvious outliers, and the linearity assumption is confirmed by the residual plot. The equation of the least squares regression line is:

$$\text{height} = 2.894 \times \text{head circumference} + 2.97$$

The slope of the regression line predicts an increase of 2.894 cm in height for each 1 cm increase in head circumference. Comparing the values of the coefficient of determination for each variable, we can see that for this sample 71.6% of the variation in height is explained by the variation in arm length, while only 23% of the variation in height is explained by the variation in head circumference. Based on this comparison, we conclude that arm length is a much better predictor of height than head circumference.

Exercise 2D

- 1 Which display (parallel boxplots, parallel dot plots, back-to-back stem plot, a percentage two-way frequency table or a scatterplot) would be appropriate to display the relationships between the following? There may be more than one appropriate graph.
 - a vegetarian (yes, no) and gender (male, female)
 - b mark obtained on a statistics test and time spent studying (in hours)
 - c number of hours spent at the beach each year and state of residence
 - d number of CDs purchased per year and income (in dollars)
 - e runs scored in a cricket game and number of 'overs' faced
 - f attitude to compulsory sport in school (agree, disagree, no opinion) and school type (government, independent)
 - g income level (high, medium, low) and place of residence (urban, rural)
 - h number of cigarettes smoked per day and gender (male, female)

Example 9

- 2 Researchers were interested in the attitudes to women's role in society. They hypothesised that attitudes might differ based on ethnicity, and that this relationship might also differ for males and females. They also believe that attitudes may have changed substantially in the years between 1990 and 2010. Data were collected and on the basis of these data the following tables were created. Use these analyses to report on the researchers' hypothesis.

Males 1990	Ethnicity		
<i>A woman should devote her time to her family</i>	Australia	UK	Europe
Agree	139	174	185
Neither	166	140	124
Disagree	299	276	156

Females 1990	Ethnicity		
<i>A woman should devote her time to her family</i>	Australia	UK	Europe
Agree	147	160	145
Neither	126	121	113
Disagree	312	312	270

Males 2010	Ethnicity		
<i>A woman should devote her time to her family</i>	Australia	UK	Europe
Agree	120	151	146
Neither	144	151	126
Disagree	597	604	485

Females 2010	Ethnicity		
<i>A woman should devote her time to her family</i>	Australia	UK	Europe
Agree	58	63	74
Neither	116	121	112
Disagree	406	391	345

Example 10

- 3** The following table shows the results of a study of obesity for a sample of 12 women and 8 men. The lean body mass, in kilograms, and the resting metabolic rate for each subject in the sample are shown. The researchers hypothesised that lean body mass (a person's weight after allowing for all fat) would have a strong association with metabolic rate. Use the data below to investigate this hypothesis.

Subject	Gender	Mass (kg)	Rate
1	M	53.1	1586
2	M	52.0	1871
3	M	47.3	1363
4	F	40.8	1192
5	F	52.1	1373
6	F	42.1	1421
7	M	63.0	1669
8	F	33.4	921
9	F	34.4	1049
10	M	62.7	1812
11	F	39.8	1174
12	M	51.9	1465
13	F	35.9	989
14	M	47.0	1442
15	F	43.0	1286
16	F	54.4	1420
17	F	42.7	1132
18	M	48.5	1607
19	F	48.5	1405
20	F	49.9	1481

Key ideas and chapter summary

Linear regression	The process of fitting a line to data is known as linear regression .
Least squares method	The least squares method is one way of determining the equation of a regression line. It minimises the sum of the squares of the residuals. It works best when there are no outliers. The equation of the least squares regression line is given by $y = mx + c$, where c represents the y-intercept of the line and m the slope.
Residuals	The vertical distance from a data point to the straight line is called a residual : residual value = data value – predicted value.
Predicting using the least squares line	The least square line $y = mx + c$ enables the value of y to be determined for a given value of x. For example, a least squares line relating the cost of printing of a book to the number of pages in the book is $\text{cost} = 0.06 \times \text{number of pages} + 1.20$ which predicts that the cost of a 100-page book is: $\text{cost} = 0.06 \times 100 + 1.20 = \$7.20$
Slope and intercept	The slope of the regression line above predicts that the cost of printing the book increases by 6 cents (\$0.06) for each additional page. The intercept of the line predicts that a book with no pages costs \$1.20 (this might be the cost of the cover).
Residual plots	Residual plots can be used to test the linearity assumption by plotting the residuals against the EV. A residual plot that appears to be a random collection of points clustered around zero supports the linearity assumption. A residual plot that shows a clear pattern indicates that the association is not linear.
Interpolation and extrapolation	Predicting <i>within</i> the range of data is called interpolation. Predicting <i>outside</i> the range of data is called extrapolation.
Correlation and causation	A correlation between two variables does not automatically imply that the association is causal . Alternative non-causal explanations for the association include a common response to a common third variable, a confounding variable or simply coincidence .

Skills checklist



Download this checklist from the Interactive Textbook, then print it and fill it out to check your skills.



2A	1 I can determine the equation of the least squares line using the formula. ☐ See Example 1 and Exercise 2A Question 1
2A	2 I can determine the equation of the least squares line using technology ☐ See Calculator Activity 1, Example 2 and Exercise 2A Question 4
2B	3 I can interpret the slope and intercept of a least squares line. ☐ See Example 3 and Exercise 2B Question 2
2B	4 I can use the least squares line to make predictions. ☐ See Example 4 and Exercise 2B Question 4
2B	5 I can distinguish between interpolation and extrapolation and understand the potential dangers of extrapolation. ☐ See Exercise 2B Question 4
2B	6 I can calculate residual values. ☐ See Example 5 and Exercise 2B Question 7
2B	7 I can construct a residual plot. ☐ See Example 6 and Exercise 2B Question 10
2B	8 I can interpret a residual plot. ☐ See Example 7 and Exercise 2B Question 12
2B	9 I can use the coefficient of determination to compare explanatory variables. ☐ See Example 8 and Exercise 2B Question 13
2B	10 I can write a report based on a regression analysis. ☐ See Exercise 2B Question 18
2C	11 I understand that correlation does not imply causation. ☐ See Exercise 2C Questions 1–5
2D	12 I can solve practical problems by identifying, analysing and describing associations. ☐ See Example 9, Example 10 Exercise 2D and Questions 2 and 3

Multiple-choice questions

- 1 For the least squares regression line $y = 0.52x - 1.2$
- A** the y -intercept = -0.52 and slope = -1.2
B the y -intercept = 0.52 and slope = 1.2
C the y -intercept = -1.2 and slope = 0.52
D the y -intercept = 1.2 and slope = -0.52
- 2 If the equation of a least squares regression line is $y = -9x + 8$ and $R^2 = 0.25$:
- A** $r = -0.5$ **B** $r = -0.25$ **C** $r = -0.0625$ **D** $r = 0.50$
- 3 Given that $r = 0.357$, $s_x = 1.871$ and $s_y = 3.391$, the slope of the least squares regression line (m) is closest to:
- A** 0.660 **B** 0.647 **C** 0.197 **D** 0.773

- 4 The association between the number of *errors* made in a task, and the *time* spent practicing the task (in minutes) was found to be approximately linear, and the values of the following statistics were determined:

	<i>time</i>	<i>errors</i>
mean	8.00	34.5
standard deviation	2.40	12.5
correlation coefficient	$r = -0.236$	

The equation of the least squares line that enables *errors* to be predicted from *time* is given by

- A** $errors = -1.23 \times time + 52.2$ **B** $errors = -0.99 \times time + 10.1$
C $errors = 0.24 \times time + 34.6$ **D** $errors = -1.23 \times time + 44.3$
- 5 The *speed* at which a car is travelling (in km/hr), and the *distance* (in metres) taken by the car to come to a stop when the brakes are applied, were recorded over speeds from 60km/hr to 120km/hr.

	<i>speed</i>	<i>distance</i>
mean	90.5	52.7
standard deviation	1.124	1.349
correlation coefficient	$r = 0.948$	

The association was found to be approximately linear, and the values of the statistics shown were determined.

On average, for each additional km/hr of *speed*, the *distance* taken to come to a stop

- A** decreased by 1.14 metres **B** decreased by 0.79 metres
C increased by 1.14 metres **D** increased by 0.95 metres

- 6 The least squares regression line $y = -9x + 8$ predicts that, when $x = 5$, the value of y is:
- A** -45 **B** -37 **C** 37 **D** 45

- 7 A least squares regression line of the form $y = mx + c$ is fitted to the data set shown.

x	25	15	10	5
y	10	10	15	25

The equation of the line is:

- A** $y = -24.4x + 0.69$ **B** $y = -0.69x + 24.4$
C $y = 0.69x + 24.4$ **D** $y = -x + 28.7$

- 8 A least squares regression line of the form $y = a + bx$ is fitted to the data set shown.

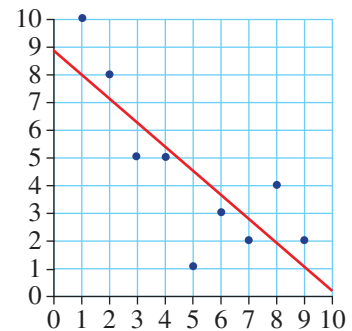
y	30	25	15	10
x	40	20	30	10

The equation of the line is:

- A** $y = 0.5x + 1$ **B** $y = x + 0.5$
C $y = 7.5x + 0.5$ **D** $y = 0.5x + 7.5$
- 9 Using a least squares regression line, the predicted value of a data point is 78.6. The residual value is -5.4 . The actual data value is:
- A** 73.2 **B** 84.0
C 88.6 **D** 94.6

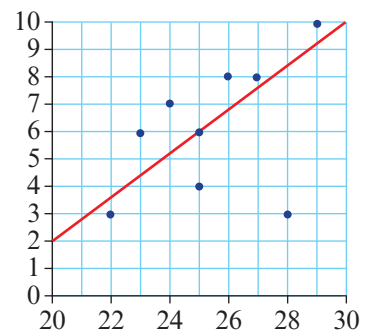
- 10 The equation of the least squares regression line plotted on the scatterplot opposite is closest to:

- A** $y = -0.9x + 8.7$
B $y = 0.9x + 8.7$
C $y = -8.7x + 0.9$
D $y = 8.7x + 0.9$



- 11 The equation of the least squares regression line plotted on the scatterplot opposite is closest to:

- A** $y = 0.8x - 14$
B $y = 14x + 0.8$
C $y = 0.8x + 2.5$
D $y = -0.8x + 14$



The following information relates to Questions 12–15.

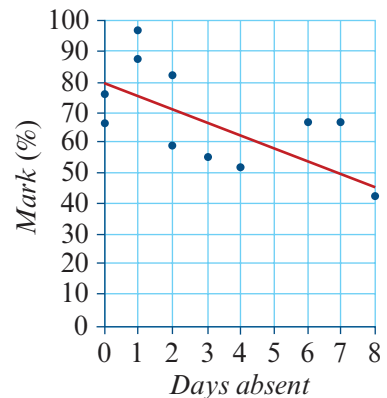
Weight (in kg) can be predicted from height (in cm) using the least squares line:

$$\text{weight} = 0.95 \times \text{height} - 96 \quad r = 0.79$$

- 12** Which of the following statements relating to the regression line is **false**?
- A** The slope of the regression line is 0.95.
 - B** The explanatory variable in the regression equation is *height*.
 - C** The least squares line does *not* pass through the origin.
 - D** The intercept is 96.
- 13** This regression line predicts that weight:
- A** decreases by 96 kg for each 1 centimetre increase in height
 - B** increases by 96 kg for each 1 centimetre increase in height
 - C** decreases by 0.95 kg for each 1 centimetre increase in height
 - D** increases by 0.95 kg for each 1 centimetre increase in height.
- 14** We can say that:
- A** 62% of the variation in weight can be explained by the variation in height
 - B** 79% of the variation in weight can be explained by the variation in height
 - C** 88% of the variation in weight can be explained by the variation in height
 - D** 79% of the variation in height can be explained by the variation in weight.
- 15** A person of height 179 cm weighs 71 kg. If the regression equation is used to predict their weight, then the residual will be closest to:
- A** –8 kg
 - B** 3 kg
 - C** –3 kg
 - D** 74 kg

The following information relates to Questions 16–21.

The scatterplot shows the association between a student's *mark* on a test, and the number of *days absent* during the term.



- 16** The coefficient of determination between *mark* and *days absent* is $R^2 = 0.5$. The correlation coefficient is closest to:
A -0.7 **B** 0.25 **C** 0.5 **D** 0.7
- 17** There were two students who were absent for two days that term. The values of the residuals for these students are closest to
A 10 **B** 60 and 80 **C** -10 and 10 **D** -10
- 18** Using the least squares line, we predict that a student who is absent for 4 days would receive a mark of about:
A 51 **B** 62 **C** 65 **D** 67
- 19** The table below shows the *weight* in grams and the *length* in cm for a certain species of fish.

<i>Length(cm)</i>	13.5	14.3	16.3	17.5	18.4	19.0	19.0	19.8	21.2	23.0
<i>Weight(gm)</i>	55	60	90	120	150	140	170	145	200	273

A least squares line which enables a fish's *weight* to be predicted from their *length* is fitted to the data. The number of times that a the fish's predicted *weight* is greater than their actual *weight* is:

- A** 3 **B** 4 **C** 5 **D** 6 **E** 7

The following information relates to Questions 20 and 21

The equation of a least squares line that enables weekly *amount* spent on entertainment (in dollars) to be predicted from weekly *income* is given by:

$$\text{amount} = 0.10 \times \text{income} + 40$$

- 20** Using this equation the amount spent on entertainment by an individual with a weekly income of \$600 is predicted to be:
A \$40 **B** \$46 **C** \$100 **D** \$240
- 21** From the equation of the regression line it can be concluded that, on average:
A the weekly *amount* spent on entertainment increases by 40 cents a week for each extra dollar of weekly income
B the weekly *amount* spent on entertainment increases by 10 cents a week for each extra dollar of weekly income
C the weekly *income* increases by 10 cents for each dollar increase in the amount spent on entertainment each week
D \$40.10 is spent on entertainment each week.

- 22** There is a strong, linear, positive correlation ($r = 0.85$) between the amount of *garbage recycled* and *salary level*.

From this information, we can conclude that:

- A** the amount of garbage recycled can be increased by increasing people's salaries
 - B** the amount of garbage recycled can be increased by decreasing people's salaries
 - C** people on high salaries tend to recycle less garbage
 - D** people on high salaries tend to recycle more garbage.
- 23** There is a strong, linear, positive correlation ($r = 0.95$) between the marriage rate in Kentucky and the number of people who drown falling out of a fishing boat.
- From this information, the most likely conclusion we can draw is:
- A** reducing the number of marriages in Kentucky will decrease the number of people who drown falling out of a fishing boat
 - B** increasing the number of marriages in Kentucky will increase the number of people who drown falling out of a fishing boat
 - C** this correlation is just coincidence
 - D** only married people in Kentucky drown falling out of a fishing boat.

Short-response questions

- 1** Determine the values of m and c in the least squares equation:

$$y = mx + c$$

where $r = 0.75$, $\bar{x} = 12.6$, $s_x = 2.4$, $\bar{y} = 124.8$, $s_y = 8.4$

- 2** A retailer recorded the number of ice creams sold and the day's maximum temperature over 8 consecutive Saturdays one summer. Use the data in the table to determine the equation of the least squares regression line for these data. Write your equation in terms of the variable in the table.

<i>Temperature (°C)</i>	22	25	36	34	21	28	41	31
<i>Ice creams</i>	145	155	200	198	150	179	230	180

- 3** The actual price of a 10-year-old car is \$15 600. Calculate the residual when its price is predicted using the regression equation: $price = -4250 \times age + 57\,500$
- 4** A regression equation that enables the price of a second-hand caravan (in dollars) to be predicted from its age (in years) is:

$$price = -5675 \times age + 87\,500$$

- a** Interpret the slope in terms of the variables *price* and *age*.
- b** Interpret the intercept in terms of the variables *price* and *age*.

- 5 The least squares equation between reaction time (in seconds) in a certain experiment and age (in years), based on a sample of a group of adults aged 40–70 years old, was determined to be:

$$\text{reaction time} = 0.036 \times \text{age} + 1.93$$

- a Use the equation to predict (to two decimal places) the reaction time for:
- i a person aged 55 years
 - ii a person aged 35 years.
- b Comment on the reliability of each of these predictions.
- 6 The equation of a least squares line that enables scores on a certain test to be predicted from IQ is:

$$\text{score} = 0.34 \times IQ + 32$$

Complete the following sentences:

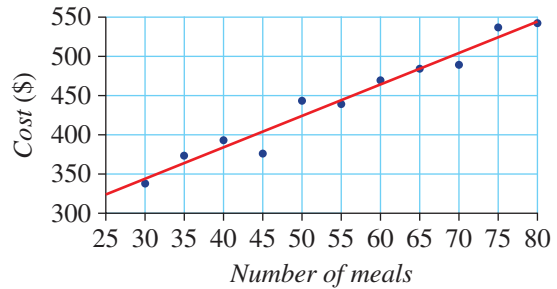
- a The response variable is .
- b The slope is and the intercept is .
- c A person has an IQ of 112. The equation predicts a test score of .
- d This person has an actual test score of 78. The residual value is .
- 7 In an investigation of the relationship between the number of hours of sunshine (per year) and the number of days of rain (per year) for 25 cities, the equation of the least squares line was found to be:

$$\text{hours of sunshine} = -6.88 \times \text{days of rain} + 2850, \text{ with } R^2 = 0.484$$

Use this information to complete the following sentences.

- a In this least squares equation, the explanatory variable is .
- b The slope is and the intercept is .
- c The least squares equation predicts that a city that has 120 days of rain per year will have hours of sunshine per year.
- d The slope of the least squares line predicts that the hours of sunshine per year will by hours for each additional day of rain.
- e $r = \text{}$, correct to three decimal places.
- f % of the variation in *hours of sunshine* can be explained by the variation in .
- g One city had 142 days of rain and 1390 hours of sunshine.
- i The least squares equation predicts that this city has hours of sunshine.
 - ii The residual value for this city is hours.
- h Using the least squares line to make predictions within the range of data used to determine the regression equation is called .

- 8 The cost of preparing meals, in dollars, and the number of meals prepared are plotted in the scatterplot shown. A least squares line has been fitted to the data which enables the cost of the meals prepared to be predicted from the number of meals prepared.



- a Write down the name of the response variable.
 b Describe the association in terms of strength, direction and form.

The equation of the least squares line that relates the cost of preparing meals to the number of meals produced is:

$$\text{cost} = 4.039 \times \text{number of meals} + 222.48$$

- c i Use the equation to predict the cost of preparing 21 meals. Round the answer to the nearest cent.
 ii In making this prediction, are you interpolating or extrapolating?
 d Write down:
 i the intercept of the regression line and interpret in terms of *cost* and the *number of meals* prepared
 ii the slope of the regression line and interpret in terms of *cost* and the *number of meals* prepared.
 e When the number of meals prepared was 50, the cost of preparation was \$444. Show that when the least squares line is used to predict the cost of preparing 50 meals, the residual is \$19.57, to the nearest cent.
- 9 We wish to determine the equation of the least squares line that will enable a person's height (in cm) to be predicted from femur (thigh bone) length (in cm).

- a Which is the response variable and which is the explanatory variable?
 b Use the following summary statistics to determine the equation of the least squares regression line that will enable height (y) to be predicted from femur length (x).

$$r = 0.9939 \quad \bar{x} = 24.246 \quad s_x = 1.873 \quad \bar{y} = 166.092 \quad s_y = 10.086$$

Give the equation in terms of *height* and *femur length*, with the values of the slope and intercept rounded to three decimal places.

- c Interpret the slope of the least squares equation in terms of *height* and *femur length*.
 d Determine the value of the coefficient of determination and interpret in terms of *height* and *femur length*.

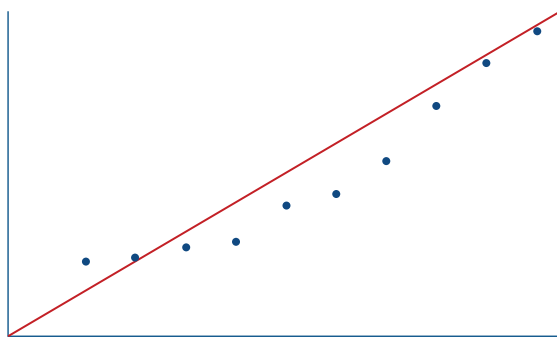
- 10** The data below shows the height (in cm) of a group of 10 children aged 2 to 11 years.

Height (cm)	86.5	95.5	103.0	109.8	116.4	122.4	128.2	133.8	139.6	145.0
Age (years)	2	3	4	5	6	7	8	9	10	11
Residual	-2.9	-0.2	0.9	1.3	1.6	1.2	0.6	-0.1	-0.7	-1.7

The equation of the least squares line used to predict height from age is

$$\text{height} = 6.366 \times \text{age} + 76.64$$

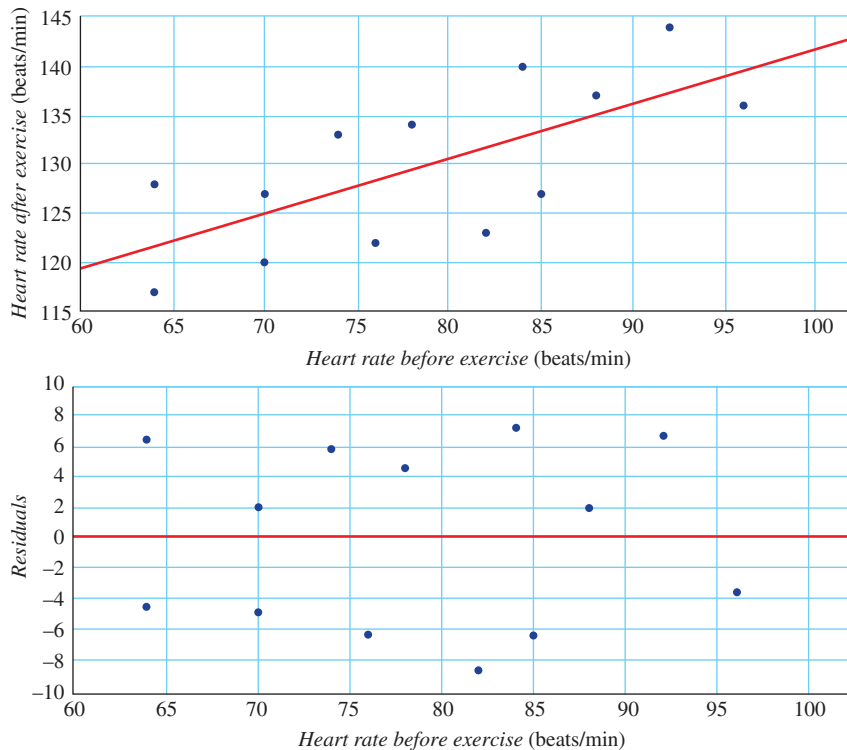
- a** Use the least squares equation to predict the height of a 1-year-old child to the nearest cm. Are you extrapolating or interpolating?
- b** The residuals for this analysis, rounded to the nearest cm, are also shown. Construct a residual plot.
- c** Is a linear equation the most appropriate model for this relationship? Explain.
- 11** In a recent study across a number of countries, the correlation between educational attainment and the amount spent on education was found to be 0.43, whilst the correlation between educational attainment and the student : teacher ratio was found to be -0.64.
- a** Interpret each of these correlation coefficients in terms of the variables in the study.
- b** Which of the variables, *amount spent on education* or *student : teacher ratio* is the better predictor of educational attainment? Explain your answer.
- 12** The relationship between two variables y and x is non-linear, as shown in the scatterplot below.



Sketch the residual plot which would result when the straight line shown is fitted to this data

- 13** Explain the difference between correlation and causation, including an example of each to illustrate your explanation.
- 14** There is a strong correlation between level of maturity and the number of children a person has. Can we assume from this that having children matures a person?

- 15** The heart rate (in beats/minute) was measured and recorded for a group of 13 students. The students then completed the same set of exercises and their heart rate was measured again immediately on completion. The scatterplot below shows the students' *heart rate after exercise* plotted against their *heart rate before exercise*, with a least squares regression line fitted. Also shown is the residual plot for this line.



The correlation between *heart rate before exercise* and *heart rate after exercise* is $r = 0.699$.

The equation of the least squares line is:

$$\text{heart rate after exercise} = 0.561 \times \text{heart rate before exercise} + 85.671$$

- a** **i** Use the equation to predict heart rate after exercise when heart rate before exercise is 100 beats/minute. Round to the nearest whole number.
- ii** Are you extrapolating or interpolating?
- b** The person with a heart rate of 122 beats/minute after exercise had a heartbeat of 76 beats/minute before exercise. If the least squares line is used to predict this person's heart rate after exercise, determine the error of prediction (residual).
- c** **i** What assumption can be tested using a residual plot?
- ii** Explain why this assumption is satisfied.
- d** Construct a report describing the association between *heart rate before exercise* and *heart rate after exercise*.

Chapter 3

Time series analysis

Chapter questions

UNIT 3 BIVARIATE DATA AND TIME SERIES ANALYSIS, SEQUENCES AND EARTH GEOMETRY

Topic 3: Time series analysis

- ▶ How do we recognise time series data?
- ▶ How do we construct a time series plot?
- ▶ How do we identify trends, seasonality and irregular fluctuations in a time series plot?
- ▶ How do we smooth a time series plot using the mean or median?
- ▶ How do we calculate and interpret seasonal indices?
- ▶ How do we deseasonalise data?
- ▶ How do we calculate and interpret a trend line?
- ▶ How do we make forecasts of future values?

In this chapter we will focus on a special case of numerical bivariate data, called **time series data**. In time series data the explanatory variable is a measure of time (for example hour, day, month or year), and we are concerned with understanding how the response variable is changing over time.

3A Time series data

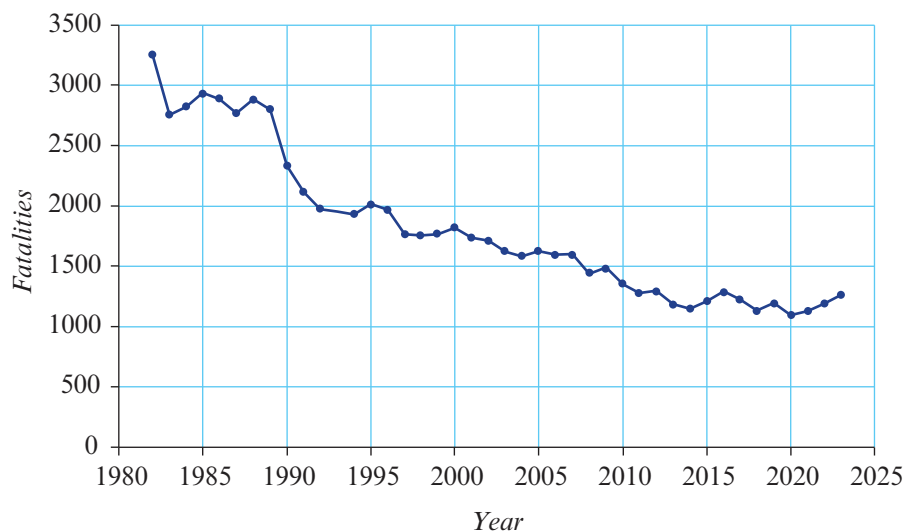
Learning intentions

- ▶ To recognise time series data.
- ▶ To construct a time series plot.
- ▶ To recognise features in the plot such as trend, seasonality and irregular fluctuations.

When data concerned with a variable is collected, observed or recorded at successive intervals of time, it is referred to as **time series data**. An example of time series data is Annual road accident fatalities for Australia, 1984–2023, given in the following table.

<i>Year</i>	<i>Fatalities</i>	<i>Year</i>	<i>Fatalities</i>	<i>Year</i>	<i>Fatalities</i>	<i>Year</i>	<i>Fatalities</i>
1984	2822	1994	1928	2004	1583	2014	1150
1985	2941	1995	2017	2005	1627	2015	1209
1986	2888	1996	1970	2006	1598	2016	1293
1987	2772	1997	1767	2007	1603	2017	1225
1988	2887	1998	1755	2008	1437	2018	1135
1989	2801	1999	1764	2009	1491	2019	1195
1990	2331	2000	1817	2010	1353	2020	1095
1991	2113	2001	1737	2011	1277	2021	1127
1992	1974	2002	1715	2012	1300	2022	1194
1993	1953	2003	1621	2013	1187	2023	1266

Time series data is just a special kind of bivariate data, where the explanatory variable is time. We will begin by drawing a scatterplot of the data. In this instance, the scatterplot is called a **time series plot**, with *time* always placed on the horizontal axis. A time series plot differs from a normal scatterplot in that, in general, the points will be joined by line segments in time order. A time series plot of the road accident fatality data is given below.



Looking at the time series plot, we can readily see a clear trend of decreasing road fatalities, which is good news for drivers. This provides some evidence that the many efforts being made to reduce the road toll across Australia have been effective, although the number of fatalities has started to increase again since 2020.

Constructing time series plots

As previously mentioned, time series data is a special case of bivariate numerical data, where the explanatory variable is time. Consider the variable *month*, which takes values such as January, February, March and so on. For the purpose of plotting and analysing time series data, we can consider the variable *month* as numerical, taking the values $\{1, 2, 3, \dots\}$. If we had monthly data for a two year period, then the variable *month* would take the values $\{1, 2, \dots, 24\}$. Whether the actual value of the variable is used in the plot (January, February, March, ...) or its numerical equivalent $(1, 2, 3, \dots)$ is used, both time series plots would be considered correct. We can use a similar approach for the variables *day*, or *quarter*.



Example 1 Constructing a time series plot

Maximum temperature was recorded each day for a week in a certain town. Construct a time series plot of the data.

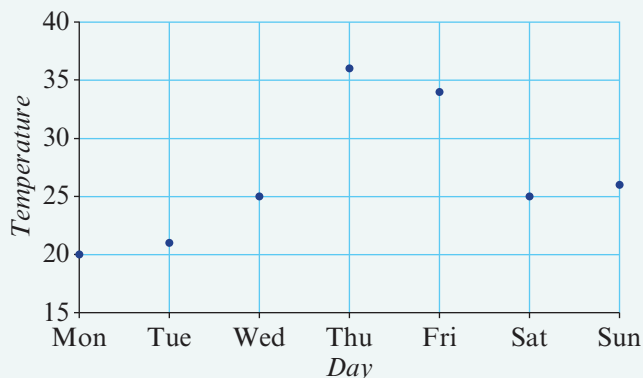
Day	Mon	Tues	Wed	Thur	Fri	Sat	Sun
Temperature ($^{\circ}\text{C}$)	20	21	25	36	34	25	26

Solution

Day is the EV – this will label the horizontal axis.

Temperature is the RV – this will label the vertical axis.

A horizontal scale from 0–7 with intervals of 1 for each day would be suitable. Temperature ranges from 20–36. A vertical scale from 15–40 with intervals of 5 would be suitable.

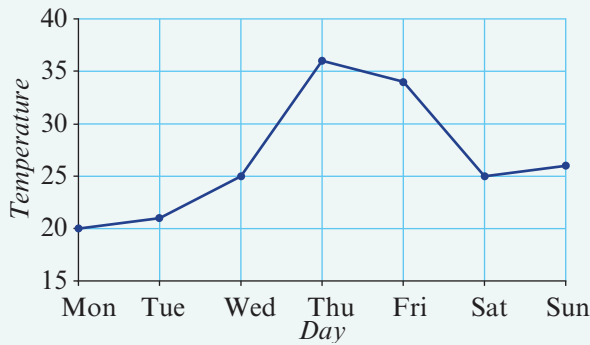


Explanation

In a time series plot, time (day in this case) is always the explanatory variable (EV) and is plotted on the horizontal axis.

Determine the scales for each axis.

Set up the axes, and then plot all seven data points as for a scatterplot.



Complete the graph by joining consecutive data points with straight lines.

Most real-world time series data come in the form of large datasets that are best plotted with the aid of a spreadsheet or statistical package. The availability of the data in electronic form via the internet greatly helps this process.



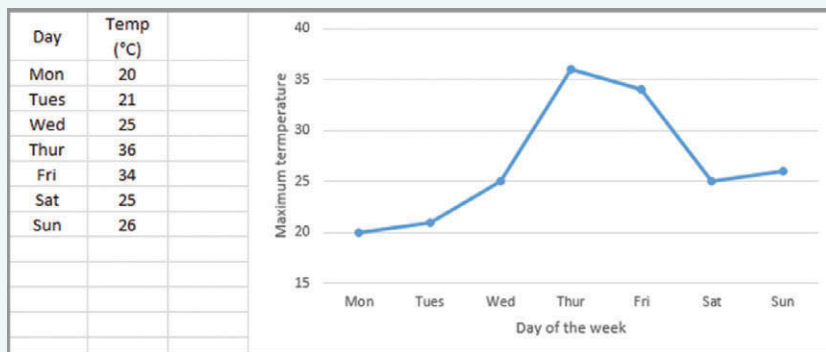
Example 2 Using Excel to construct a time series plot



Use the data from Example 1 to construct a time series plot using Excel.

Solution

Enter the data into two columns as shown below. Select both columns (including headings) and on the **Insert** tab, in the **Charts** group, click on **Scatter** and then the option **Scatter with Straight Lines and Markers**. Double click on each scale separately and edit to include the range of the data. Axis labels can be added using the **Add Chart Elements** option when editing the scale.



Looking for patterns in time series plots

The features we look for in a time series are:

- trend
- cycles
- seasonality
- structural change
- possible outliers
- irregular (random) fluctuations.

We would always expect to see irregular or random fluctuations in a time series, and it is common to see one or more of the other features as well.

Trend

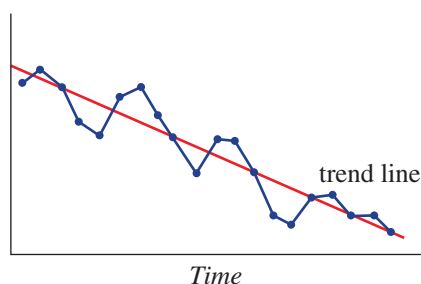
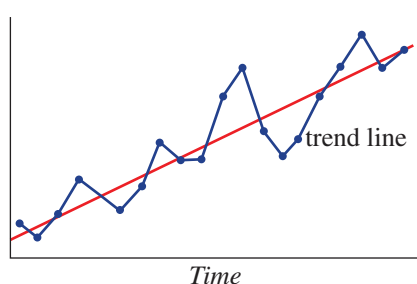
Examining a time series plot we can often see a general upward or downward movement over time. This indicates a long-term change over time that we call a trend.

Trend

The tendency for values in a time series to generally increase or decrease over a significant period of time is called a **trend**.

One way of identifying trends on a time series graph is to draw a line that ignores the fluctuations, but which reflects the overall increasing or decreasing nature of the plot. These lines are called **trend lines**.

Trend lines have been drawn on the time series plots below to indicate an **increasing** trend (line slopes upwards) and a **decreasing** trend (line slopes downwards).

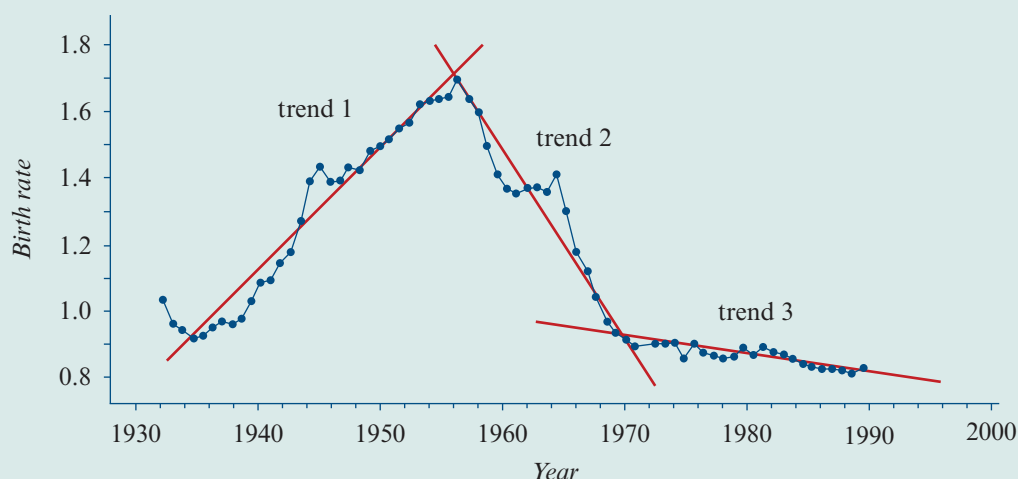


Sometimes, different trends are apparent in a time series over different time periods.



Example 3 Identifying trend

Consider the time series plot of the Australian annual birth rates over the years from 1931 to 1990, shown below. Comment on the trend shown in the plot.



Solution

There are three distinct trends, which can be seen by drawing trend lines on the plot. Each of these trends can be explained by changing socioeconomic circumstances.

Trend 1: Between 1940 and 1961 the birth rate in Australia grew quite dramatically. Those in the armed services came home from the Second World War, and the economy grew quickly. This rapid increase in the Australian birth rate during this period is known as the ‘Baby Boom’.

Trend 2: From about 1962 until 1980 the birth rate declined very rapidly. Birth control methods became more effective, and women started to think more about careers. This period is sometimes referred to as the ‘Baby Bust’.

Trend 3: During the 1980s, and beyond, the birth rate continued to decline slowly for a complex range of social and economic reasons.

Cycles

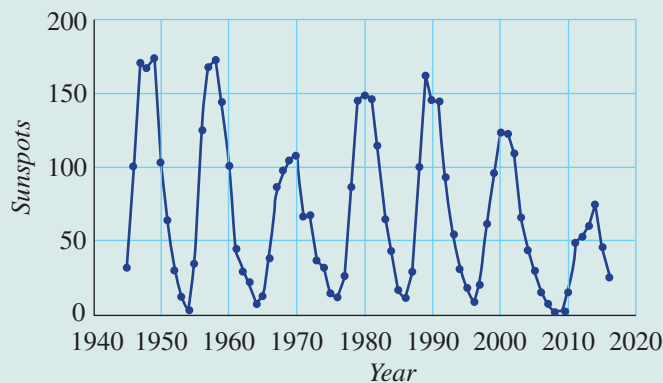
The term cycle refers to recurring variations in a time series that generally last longer than a year. These variations may not be of the same height and they may not repeat at regular intervals.

Cycles

Cycles are recurrent movements in a time series, generally over a period greater than one year.

**Example 4 Identifying cycles**

Sunspots are darker, cooler areas on the surface of the sun. The following plot shows the sunspot activity for the period 1945 to 2016. Comment on the cycles shown in the plot.

**Solution**

The recurrent pattern in the number of sunspots can be seen clearly from the time series plot. Looking at the plot, the years of lowest sunspot activity are at approximately 1954, 1964, 1975, 1986, 1996, 2008.

Many business indicators, such as interest rates or unemployment figures, also vary in cycles, but their periods are usually less regular.

Seasonality

Cycles with calendar-related periods of one year or less are of special interest and are referred to as **seasonality**.

Seasonality

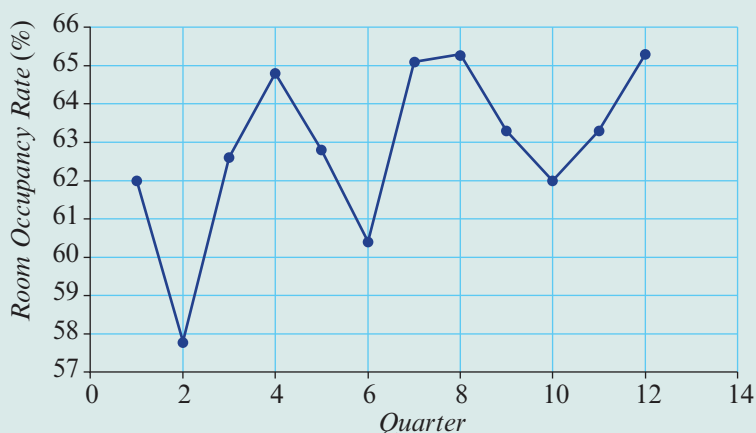
Seasonality is present when there is a periodic movement in a time series that is related to a calendar-related period, for example, a year, a month or a week.

Seasonal movements tend to be more predictable than trends, and occur because of variations in the weather, such as ice-cream sales, or institutional factors, like the increase in the number of unemployed people at the end of the school year.



Example 5 Identifying seasonality

The plot below shows the total percentage of hotel rooms occupied in Australia by quarter, over a three year period. Comment on the seasonality shown in the plot.



Solution

The regular peaks and troughs in the plot that occur at the same time each year signal the presence of **seasonality**. In this case, the demand for accommodation is at its lowest in Quarter 2 each year and highest in Quarter 4.

This time series plot reveals both seasonality and trend in the demand for hotel rooms. The upward sloping trend line signals the presence of a general increasing **trend**. This tells us that, even though demand for accommodation has fluctuated from quarter to quarter, demand for hotel accommodation has increased over time.

Structural change

A **structural change** in a time series is a sudden change in the pattern of the time series at a point in time.

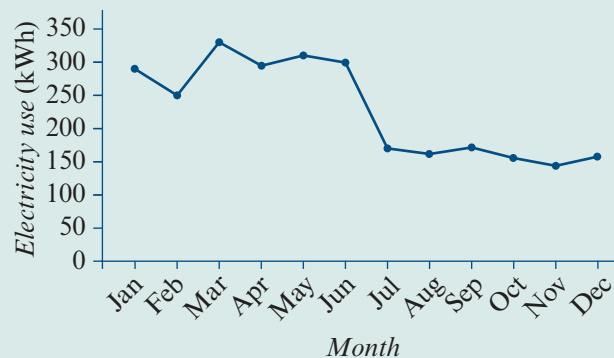
Structural change

Structural change is present when there is a sudden change in the established pattern of a time series plot.



Example 6 Identifying structural change

The time series plot below shows the power bill for a rental house (in kWh) for the 12 months of a year. Comment on any structural change in the plot.



Solution

The plot reveals an abrupt change in power usage in June to July. During this period, monthly power use suddenly decreases from around 300 kWh per month from January to June to around 175 kWh for the rest of the year. This is an example of structural change that can probably be explained by a change in circumstances, for example, from a family with children to a person living alone.

Structural change is also displayed in the birth rate time series plot we saw earlier. This revealed three quite distinct trends during the period 1900–2010. These reflect significant external events (like a war) or changes in social and economic circumstances.

One consequence of structural change is that we can no longer use a single mathematical model to describe the key features of a time series plot.

Outliers

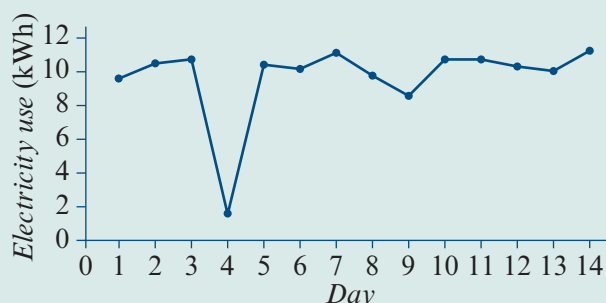
Outliers

Outliers are present when there are individual values that stand out from the general body of data.



Example 7 Identifying outliers

The time series plot below shows the daily power bill for a house (in kWh) for a fortnight. Comment on any outliers in the plot.



Solution

For this household, daily electricity use follows a regular pattern that, although fluctuating, averages about 10 kWh per day. In terms of daily power use, day 4 is a clear outlier, with less than 2 kWh of electricity used. A follow-up investigation found that, on this day, the house was without power for 18 hours due to a storm, so much less power was used than normal.

Irregular (random) fluctuations

Irregular (random) fluctuations

Irregular (random) fluctuations include all the variations in a time series that we cannot reasonably attribute to systematic changes like trend, cycles, seasonality and structural change or an outlier. There will **always** be irregular, random variation present in any real world time series data.

There can be many sources of irregular fluctuations, mostly unknown. A general characteristic of these fluctuations is that they are unpredictable.

One of the aims of time series analysis is to develop techniques to identify regular patterns in time series plots that are often obscured by irregular fluctuations. One of these techniques is smoothing, which you will meet in the next section.

Section Summary

The features we look for in a time series are:

■ trend	■ cycles	■ seasonality
■ structural change	■ possible outliers	■ irregular fluctuations

Trend is present when there is a long-term upward or downward movement in a time series.

Cycles are present when there is a periodic movement in a time series. The period is the time it takes for one complete up and down movement in the time series plot. In practice, this term is reserved for periods greater than 1 year.

Seasonality is present when there is a periodic movement in a time series that has a calendar related period, for example, a year, a month or a week.

Structural change is present when there is a sudden change in the established pattern of a time series plot.

Outliers are present when there are individual values that stand out from the general body of data.

Irregular (random) fluctuations are always present in any real-world time series plot. They include all the variations in a time series that we cannot reasonably attribute to systematic changes like trend, cycles, seasonality and structural change or an outlier.

Exercise 3A

Constructing a time series plot

Example 1

- 1 Construct a time series plot to display the following data.

Example 2

<i>Year</i>	2015	2016	2017	2018	2019	2020	2021	2022
<i>Sales</i>	2	23	35	31	45	23	67	70

- 2 Researchers recorded the number of penguins present on a remote island each month for 12 months. Construct a time series plot of the data.

<i>Month</i>	<i>Number of penguins</i>	<i>Month</i>	<i>Number of penguins</i>
January	449	July	180
February	214	August	241
March	170	September	311
April	265	October	499
May	434	November	598
June	102	December	674

- 3 The following table shows the minimum temperature in Brisbane during one week in January. Construct a time series plot of the data.

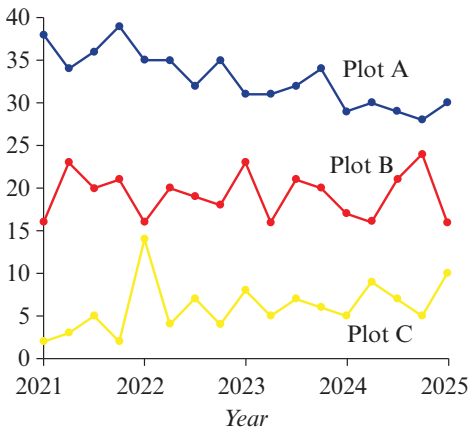
<i>Day</i>	Mon	Tues	Wed	Thur	Fri	Sat	Sun
<i>Temperature (°C)</i>	24.0	24.2	17.4	17.7	18.3	19.5	17.4

Identifying key features in a time series plot

Example 3

- 4 Complete the table below by indicating which of the listed features are present in each of the time series plots.

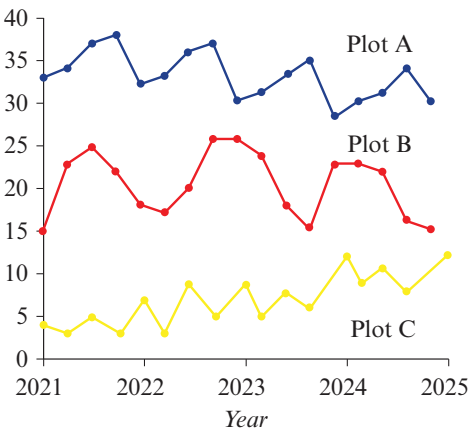
Feature	Plot		
	A	B	C
Irr. fluctuations			
Increasing trend			
Decreasing trend			



Example 4

- 5 Complete the table below by indicating which of the listed features are present in each of the time series plots.

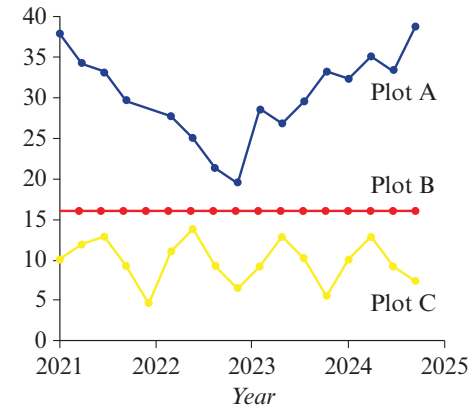
Feature	Plot		
	A	B	C
Irr. fluctuations			
Increasing trend			
Decreasing trend			
Cycles			
Seasonality			



Example 6

- 6 Complete the table below by indicating which of the listed features are present in each of the time series plots.

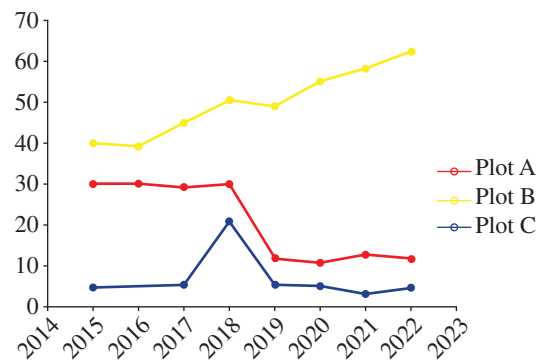
Feature	Plot		
	A	B	C
Irr. fluctuations			
Struct. change			
Increasing trend			
Decreasing trend			
Seasonality			



Example 7

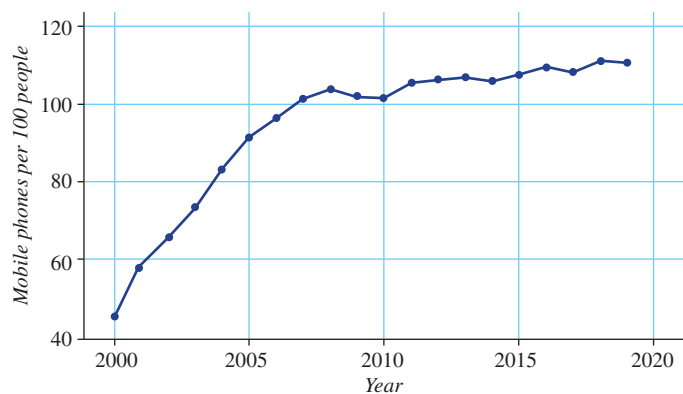
- 7 Complete the table below by indicating which of the listed features are present in each of the time series plots shown.

Feature	Plot		
	A	B	C
Irr. fluctuations			
Struct. change			
Increasing trend			
Decreasing trend			
Outliers			



Describing time series plots

- 8 The time series plot for the number of mobile phones per 100 people from 2000–2019 is shown opposite. Describe the features of the time series plot.



- 9 The data below shows the population (in millions) in Australia over the period 2012–2021.

Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Population	22.7	23.1	23.5	23.8	24.2	24.6	25.0	25.4	25.7	26.0

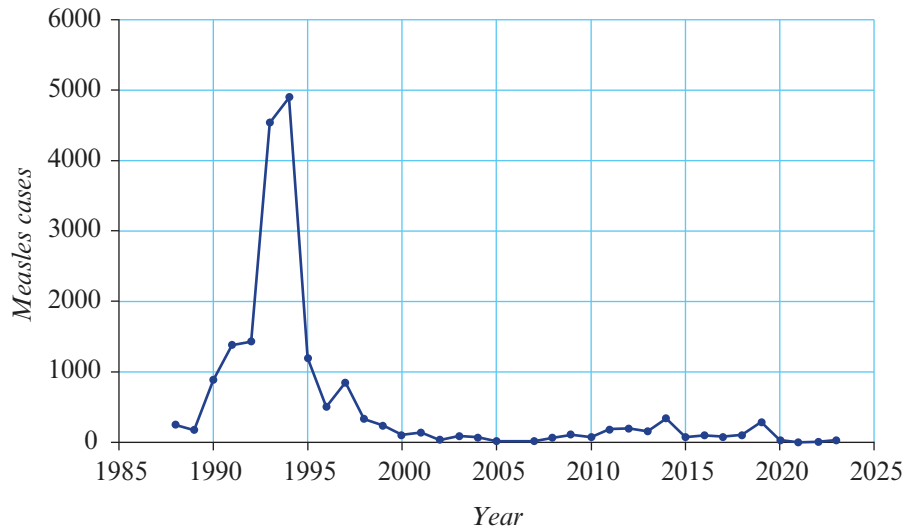
- a Construct a time series plot of the data.
b Describe the features of the plot.
- 10 The table below shows the motor vehicle theft rate per 100 000 cars in Australia from 2003 to 2018.

Year	2003	2004	2005	2006	2007	2008	2009	2010
Theft rate	500.9	442.4	398.3	367.2	337.6	320.0	274.2	214.8

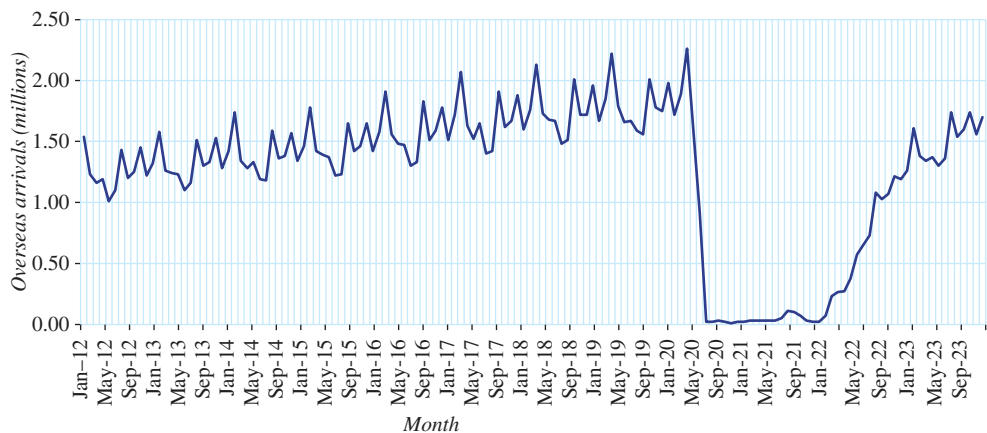
Year	2011	2012	2013	2014	2015	2016	2017	2018
Theft rate	220.0	228.4	204.2	191.0	194.5	231.0	213.3	214.1

- a Construct a time series plot of the data.
b Describe the features of the plot.

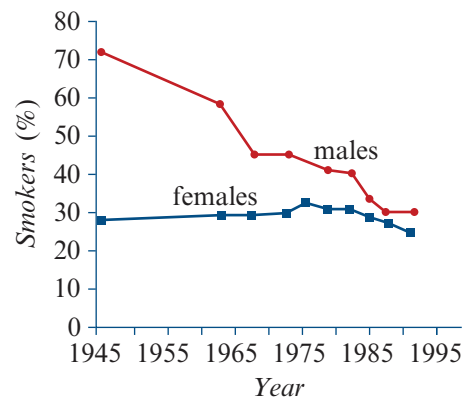
- 11** The time series plot below shows the number of measles cases reported in Australia from 1988 to 2023. Describe the features of the plot.



- 12** The time series below shows the number of overseas arrivals (millions people per month) in Australia from January 2012 until December 2023. Describe the features of the plot.



- 13 a** The time series plot shows the smoking rates (%) of Australian males and females over the period 1945–92.
- Describe the trends in the time series plot.
 - Compare the difference in the trends over the period 1945–92.



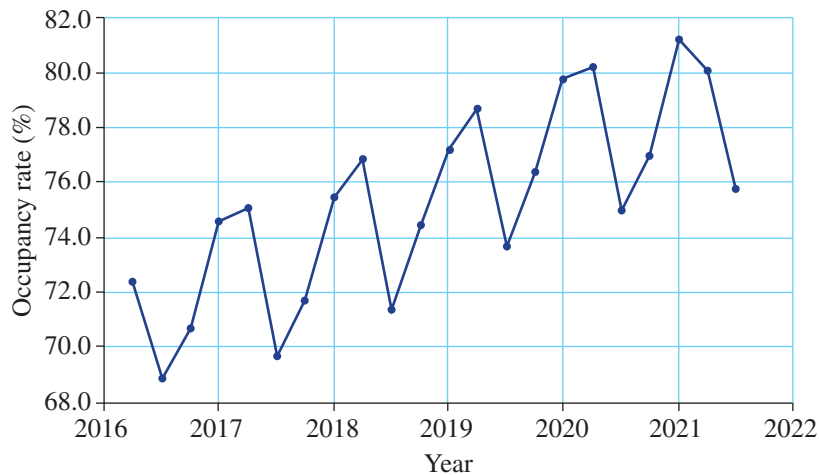
- b** The table below shows the smoking rates for females and males aged 15 years at several time points from 2000–2018 (smoking rate data is not collected every year).

Year	2000	2005	2007	2010	2011	2012	2013	2014	2015	2016	2018
Female	22.4	18.9	19.9	17.9	15.4	16.6	14.4	15.6	13.5	14.5	13.6
Male	26.7	22.9	24.9	22.9	19.1	21.9	18.0	20.7	17.0	19.7	18.7

- i** Construct time series plots of the smoking rates for males and females.
- ii** Describe any trends in the time series plot.
- iii** Compare the difference in the trends over the period 2000–2018.

Paper 1-style multiple-choice questions

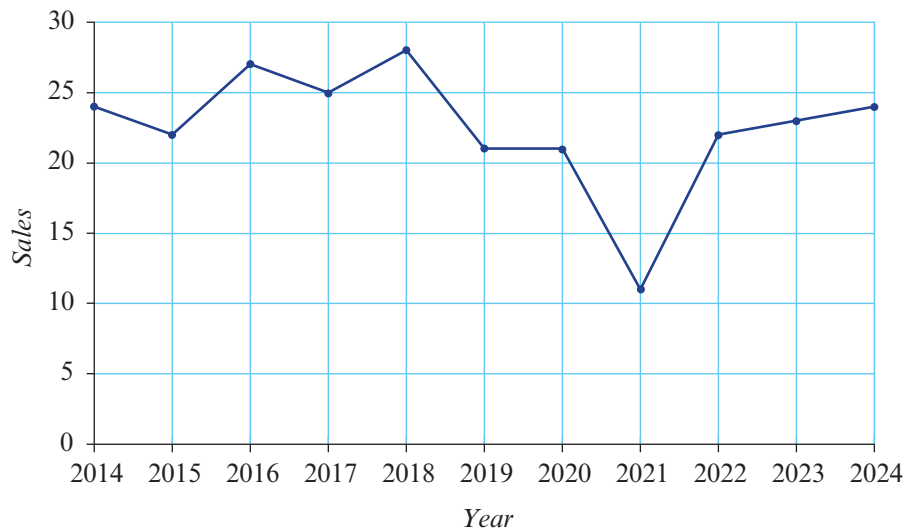
- 14** The time series plot below shows the quarterly room occupancy rate for a chain of hotels over the years 2016 to 2022.



The time series plot is best described as showing:

- A** seasonality only
- B** seasonality with irregular fluctuations
- C** an increasing trend with irregular fluctuations
- D** an increasing trend with seasonality and irregular fluctuations.

- 15** The time series plot below shows the annual sales (in \$ millions) for a car sales company.



The time series plot is best described as showing:

- A** irregular fluctuations only
- B** seasonality with irregular fluctuations
- C** irregular fluctuations with an outlier
- D** seasonality with an outlier.

3B Smoothing a time series using moving means

Learning intentions

- To smooth a time series plot using moving means.

A time series plot can incorporate many of the sources of variation previously mentioned: trend, cycles, seasonality, structural change, outliers and irregular fluctuations. One effect of the irregular fluctuations and seasonality can be to obscure an underlying trend. The technique of **smoothing** can sometimes be used to overcome this problem.

In this section we consider **moving mean smoothing**, which involves replacing individual data points in the time series with the mean of the data point and some adjacent points. The simplest method is to smooth over a small odd number of data points – for example, three or five, but any number of points can be used.

The three-moving mean

To use **three-moving mean smoothing**, replace each data value with the mean of that value and the values of its two neighbours, one on each side. That is, if y_1 , y_2 and y_3 are sequential data values, then:

$$\text{smoothed } y_2 = \frac{y_1 + y_2 + y_3}{3}$$

The first and last points do not have values on each side, so they are omitted.

The five-moving mean

To use **five-moving mean smoothing**, replace each data value with the mean of that value and the two values on each side. That is, if y_1 , y_2 , y_3 , y_4 , y_5 are sequential data values, then:

$$\text{smoothed } y_3 = \frac{y_1 + y_2 + y_3 + y_4 + y_5}{5}$$

The first two and last two points do not have two values on each side, so they are omitted.

These definitions can be readily extended for moving means involving more points.



Example 8 Three- and five-moving mean smoothing

The table below gives the temperature ($^{\circ}\text{C}$) recorded at a weather station at 9 a.m. each day for a week.

Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Temperature	18.1	24.8	26.4	13.9	12.7	14.2	24.9

- Calculate the three-mean smoothed temperature for Tuesday.
- Calculate the five-mean smoothed temperature for Thursday.

Solution

- a** 18.1, 24.8, 26.4

$$\text{Mean} = \frac{(18.1 + 24.8 + 26.4)}{3} = 23.1$$

The three-mean smoothed temperature for Tuesday is 23.1°C .

Explanation

Write down the three temperatures centred on Tuesday.

Calculate the mean and write down your answer.

b 24.8, 26.4, 13.9, 12.7, 14.2

$$\begin{aligned}\text{Mean} &= \frac{(24.8 + 26.4 + 13.9 + 12.7 + 14.2)}{5} \\ &= 18.4\end{aligned}$$

The five-mean smoothed temperature for Thursday is 18.4°C.

Write down the five temperatures centred on Thursday.

Calculate the mean and write down your answer.

The next step is to extend these computations to smooth all terms in the time series.



Example 9 Three- and five-moving mean smoothing of a time series

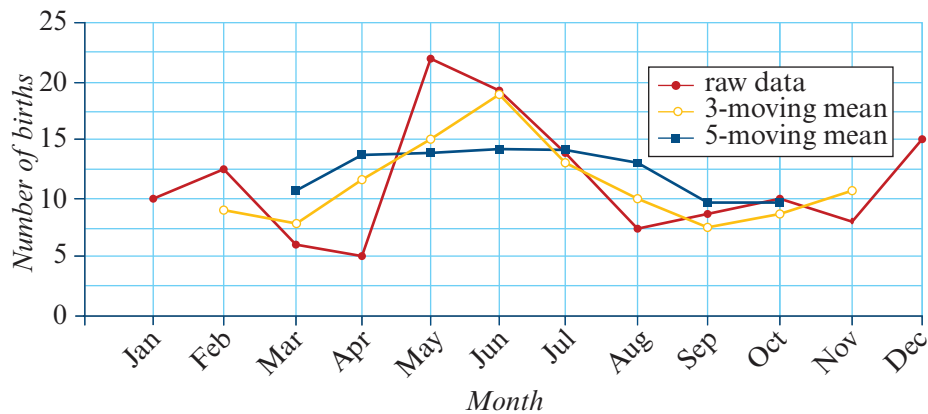
The following table gives the number of births per month over a calendar year in a country hospital. Use the three-moving mean and the five-moving mean methods, correct to one decimal place, to complete the table.

Solution

Complete the calculations as shown below.

Month	Number of births	3-moving mean	5-moving mean
January	10		
February	12	$\frac{10 + 12 + 6}{3} = 9.3$	
March	6	$\frac{12 + 6 + 5}{3} = 7.7$	$\frac{10 + 12 + 6 + 5 + 22}{5} = 11.0$
April	5	$\frac{6 + 5 + 22}{3} = 11.0$	$\frac{12 + 6 + 5 + 22 + 18}{5} = 12.6$
May	22	$\frac{5 + 22 + 18}{3} = 15.0$	$\frac{6 + 5 + 22 + 18 + 13}{5} = 12.8$
June	18	$\frac{22 + 18 + 13}{3} = 17.7$	$\frac{5 + 22 + 18 + 13 + 7}{5} = 13.0$
July	13	$\frac{18 + 13 + 7}{3} = 12.7$	$\frac{22 + 18 + 13 + 7 + 9}{5} = 13.8$
August	7	$\frac{13 + 7 + 9}{3} = 9.7$	$\frac{18 + 13 + 7 + 9 + 10}{5} = 11.4$
September	9	$\frac{7 + 9 + 10}{3} = 8.7$	$\frac{13 + 7 + 9 + 10 + 8}{5} = 9.4$
October	10	$\frac{9 + 10 + 8}{3} = 9.0$	$\frac{7 + 9 + 10 + 8 + 15}{5} = 9.8$
November	8	$\frac{10 + 8 + 15}{3} = 11.0$	
December	15		

The result of this smoothing can be seen in the following plot, which shows the raw data, the data smoothed with a three-moving means and the data smoothed with a five-moving means.



Note: In the process of smoothing, data points are lost at the beginning and end of the time series.

Two observations can be made from this plot:

- 1 Five-mean smoothing is more effective in reducing the irregular fluctuations than three-mean smoothing.
- 2 The five-mean smoothed plot shows that there is no clear trend although the raw data suggest that there might be an increasing trend.

There are many ways of smoothing a time series. Moving means of group size other than three and five are common and often very useful.

However, if we smooth over an even number of points, we run into a problem. The centre of the set of points is not at a time point belonging to the original series. Usually, we solve this problem by using a process called **centring**. Centring involves taking a two-moving mean of the already smoothed values so that they line up with the original time values. Smoothing with centring is beyond the scope of this course.

A spreadsheet is extremely useful when smoothing time series data.



Example 10 Smoothing a time series using Excel

Consider again the data in Example 8:

Day	Mon	Tues	Wed	Thur	Fri	Sat	Sun
Temperature (°C)	18.1	24.8	26.4	13.9	12.7	14.2	24.9

Use Excel to construct a plot showing the data, the three-mean smoothed plot, and the five-mean smoothed plot, all on the same axes.



Solution

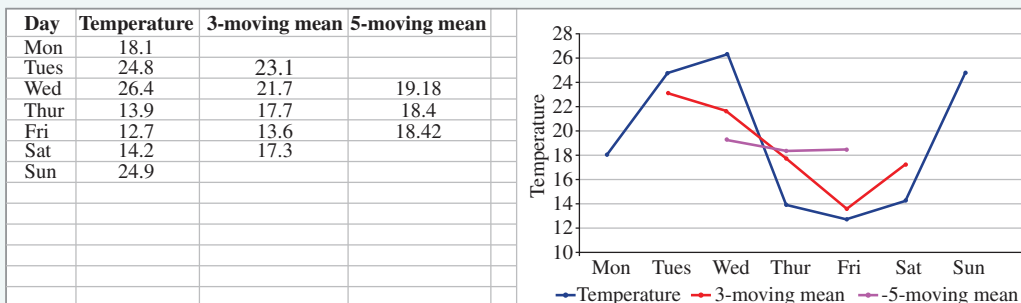
Enter the data into two columns. Place the cursor in cell D4, next to the temperature on Tuesday and enter the formula $= (C3 + C4 + C5)/3$ as shown below. Press Enter and the mean of these three cells will be calculated.

A	B	C	D
	Day	Temperature	3-moving mean
	Mon	18.1	
	Tues	24.8	$= (C3+C4+C5)/3$
	Wed	26.4	
	Thur	13.9	
	Fri	12.7	
	Sat	14.2	
	Sun	24.9	

Fill down the column with the formula, remembering to omit the first and last row, to calculate all three-mean smoothed values. Place the cursor in cell E5, next to the temperature on Wednesday and enter the formula $= (C3 + C4 + C5 + C6 + C7)/5$ as shown below. Press Enter and the mean of these five cells will be calculated.

	Day	Temperature	3-moving mean	5-moving mean
	Mon	18.1		
	Tues	24.8	23.1	
	Wed	26.4		$= (C3+C4+C5+C6+C7)/5$
	Thur	13.9	17.7	
	Fri	12.7	13.6	
	Sat	14.2	17.3	
	Sun	24.9		

Fill down the column with the formula, remembering to omit the first and last two rows, to calculate all five-mean smoothed values (see below). Select all four data columns (including headings) and on the **Insert** tab, in the **Charts** group, click on **2D-Line** and then the option **Lines with Markers**. Double click on each scale separately and edit as appropriate. Axis labels can be added using the **Add Chart Elements** option when editing the scale.



Section Summary

- ▶ To better identify an underlying trend in a time series, **moving mean smoothing** can be used.
- ▶ To use **three-moving mean smoothing**, replace each data value with the mean of that value and the values of its two neighbours, one on each side. The first and last points do not have values on each side, so they are omitted.
- ▶ To use **five-moving mean smoothing**, replace each data value with the mean of that value and the two values on each side. The first two and last two points do not have two values on each side, so they are omitted.
- ▶ These definitions can be readily extended for moving means involving more points.



Exercise 3B

Calculating the smoothed values of an odd number of individual points

Example 8

1

t	1	2	3	4	5	6	7	8	9
y	5	2	5	3	1	0	2	3	0

For the time series data in the table above, calculate:

a the three-mean smoothed y -value for

i $t = 4$

ii $t = 6$

iii $t = 2$

b the five-mean smoothed y -value for

i $t = 3$

ii $t = 7$

iii $t = 4$

c the seven-mean smoothed y -value for

i $t = 4$

ii $t = 6$

d the nine-mean smoothed y -value for $t = 5$

- 2** The table below gives the temperature ($^{\circ}\text{C}$) recorded at a weather station at 3.00 p.m. each day for a week.

Day	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Temperature ($^{\circ}\text{C}$)	28.9	33.5	21.6	18.1	16.2	17.9	26.4

- a** Calculate the three-mean smoothed temperature for Wednesday.
- b** Calculate the five-mean smoothed temperature for Friday.
- c** Calculate the seven-mean smoothed temperature for Thursday.

- 3 Complete the following table.

t	1	2	3	4	5	6	7	8	9
y	10	12	8	4	12	8	10	18	2
3-moving mean y	–								–
5-moving mean y	–	–						–	–

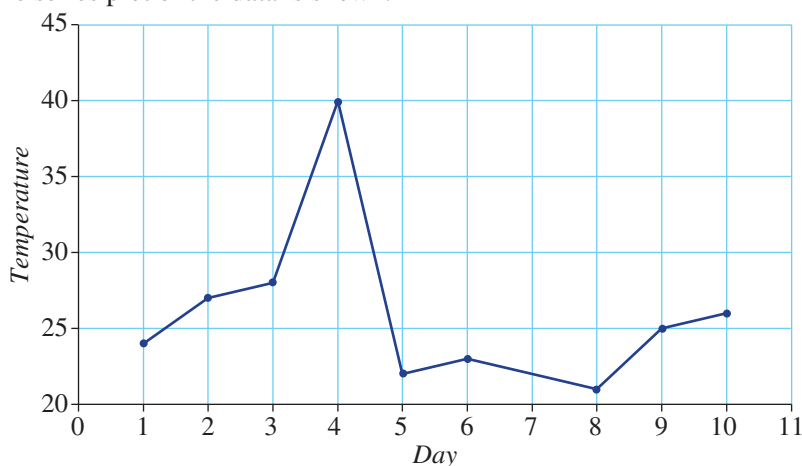
Smoothing and plotting a time series plot

Example 9

- 4 The maximum temperature of a city over a period of 10 days is given below.

Day	1	2	3	4	5	6	7	8	9	10
Temperature ($^{\circ}\text{C}$)	24	27	28	40	22	23	22	21	25	26
3-moving mean		26.3		30.0		22.3		22.7	24.0	
5-moving mean			28.2		27			23.4		

A time series plot of the data is shown.



- a Use the five-mean and seven-mean smoothing method to complete the table.
 b Plot the smoothed temperature data and compare the three plots.

Example 10

- 5 The value of the Australian dollar in US dollars (exchange rate) over 10 days is given below.

Day	1	2	3	4	5	6	7	8	9	10
Exchange rate	0.743	0.754	0.737	0.751	0.724	0.724	0.712	0.735	0.716	0.711
3-moving mean										
5-moving mean										

- a Construct a time series plot of the data.
 b Use the three-mean and five-mean smoothing method to complete the table.
 c Plot the smoothed exchange rate data and compare the plots.

Paper 1-style multiple-choice questions

- 6 The table shows the closing price of a company's shares on the stock market over a 10 day period.

Day	1	2	3	4	5	6	7	8	9	10
Price (\$)	0.99	1.05	1.10	1.25	1.29	1.37	2.42	1.95	2.05	2.35

The five-mean smoothed closing share *price* on Day 6 is closest to:

- A** \$1.56 **B** \$1.62 **C** \$1.66 **D** \$1.88
- 7 Hay Lam records the number of emails he receives over a one-week period.

Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Emails	85	65	77				10

The numbers of *emails* he received on Thursday, Friday and Saturday are not shown.

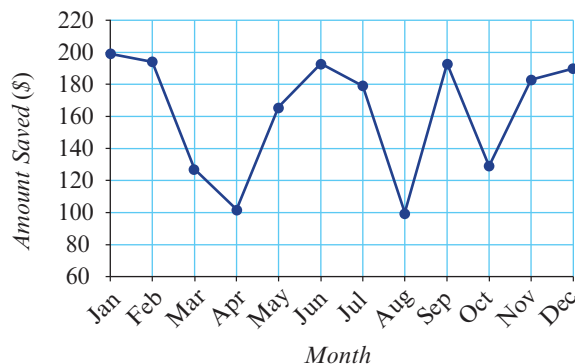
The five-mean smoothed number of *emails* he received on Friday is 39.

The three-mean smoothed number of *emails* he received on Friday is:

- A** 36 **B** 39 **C** 40 **D** 42

The following information relates to Questions 8 and 9

The time series plot below shows the amount that Arnold saved each month (in dollars) over a 12 month period.



- 8 If he saved a total of \$831 over the period from May to September, the five-mean smoothed amount that he saved in July is closest to:
- A** \$190 **B** \$182 **C** \$152 **D** \$166
- 9 If seven-mean smoothing is used to smooth this time series plot, the number of smoothed data points would be:
- A** 4 **B** 5 **C** 6 **D** 7

3C Smoothing a time series plot using moving medians

Learning intentions

- ▶ To locate the median of a data set graphically.
- ▶ To smooth a time series plot using moving medians.

Another simple and convenient method of smoothing a time series is to use **moving median smoothing**. The advantage of this method is that it can be done directly on the graph without needing to know the exact values of each data point. However, before smoothing a time series plot graphically using moving medians we will first need to know how to locate medians graphically.

Locating medians graphically

The graph opposite shows three data points plotted on set of coordinate axes. The task is to locate the median of these three points. The median will be a point somewhere on this set of coordinate axes. To locate this point we proceed as follows.

Step 1

Identify the middle data point moving in the x -direction. Draw a vertical line through this value as shown.

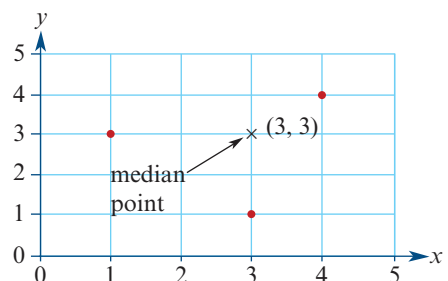
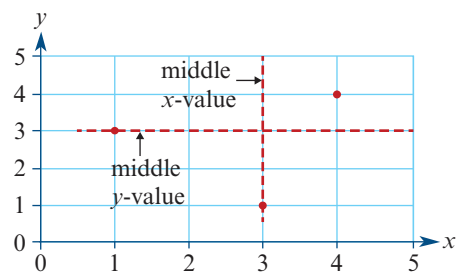
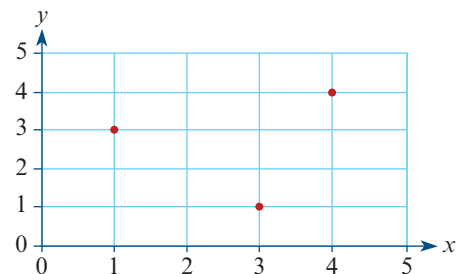
Step 2

Identify the middle data point moving in the y -direction. Draw a horizontal line through this value as shown.

Step 3

The median value is where the two lines intersect – in this case, at the point $(3, 3)$.

Mark this point with a cross ($\times$).



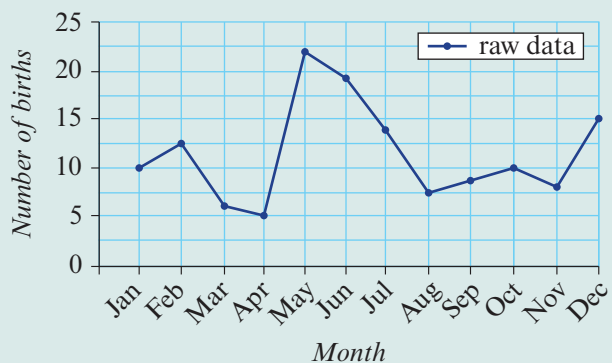
Smoothing a time series using moving median smoothing

The process of graphically smoothing a time series plot requires no more than repeating the above process for each group of three or five data points in the plot as required. The starting point for a median smoothing is a time series plot and we smooth directly onto the plot. The following worked examples demonstrate the process.

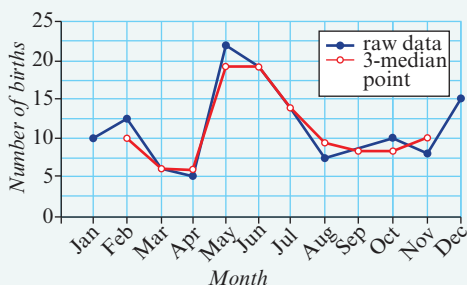
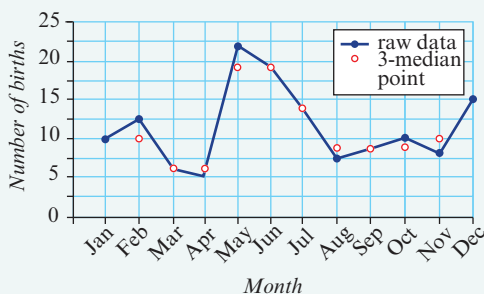
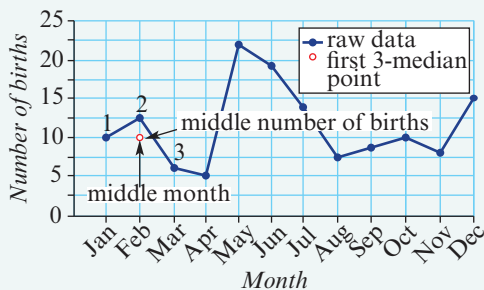


Example 11 Three-median smoothing using a graphical approach

Construct a three-median smoothed plot of the time series plot shown opposite.



Solution



Explanation

Locate on the time series plot the median of the *first* three points (Jan, Feb, Mar).

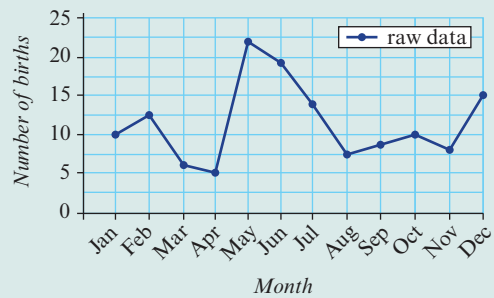
Continue this process by moving on to the next three points to be smoothed (Feb, Mar, Apr). Mark their medians on the graph, and continue the process until you run out of groups of three.

Join the median points with a line segment.

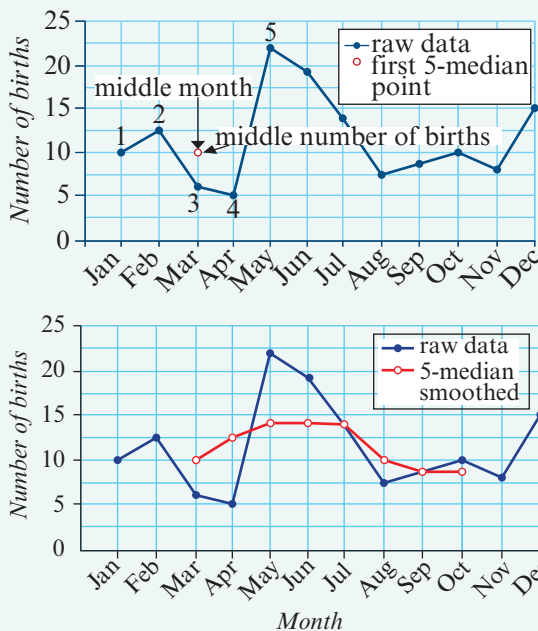


Example 12 Five-median smoothing using a graphical approach

Construct a five-median smoothed plot of the time series plot shown opposite.



Solution



Explanation

Locate on the time series plot the median of the first five points (Jan, Feb, Mar, Apr, May), as shown.

Then move on to the next five points to be smoothed (Feb, Mar, Apr, May, Jun). Repeat the process until you run out of groups of five points. The five-median points are then joined up with line segments to give the final smoothed plot, as shown.

Note: As expected, the five-median smoothed plot is smoother than the three-median smoothed plot.

Section Summary

- ▶ To better identify an underlying trend in a time series, **moving median smoothing** can be used.
- ▶ **Moving median smoothing** can be used directly on a times series plot.
- ▶ To use **three-moving median smoothing**, replace each data value with the median of that value and the values of its two neighbours, one on each side. The first and last points do not have values on each side, so they are omitted.
- ▶ To use **five-moving median smoothing**, replace each data value with the median of that value and the two values on each side. The first two and last two points do not have two values on each side, so they are omitted.

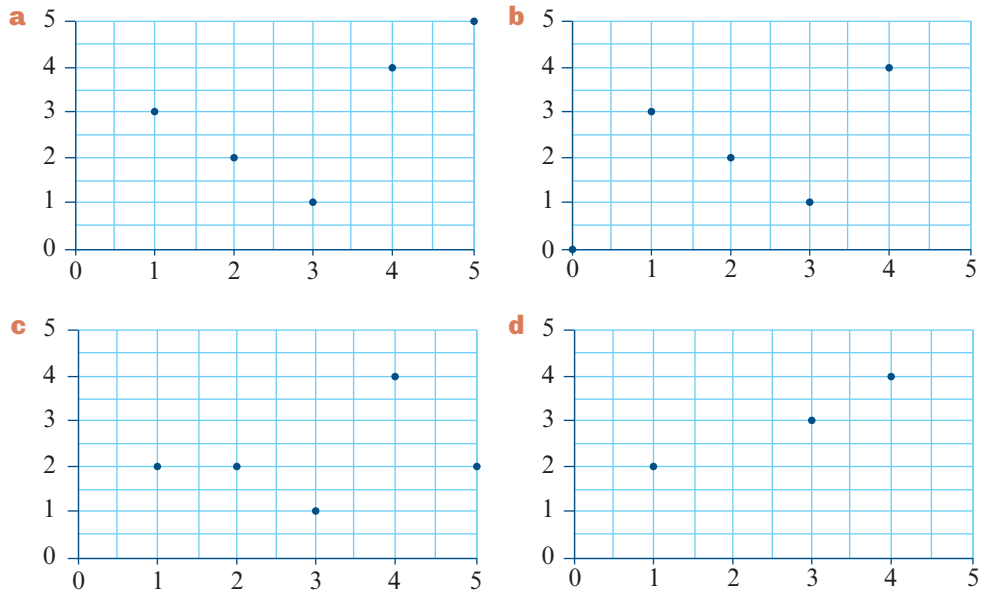


Exercise 3C

Note: Copies of all plots in this section can be accessed through the skillsheet icon in the Interactive Textbook.

Locating the median of a set of data points graphically

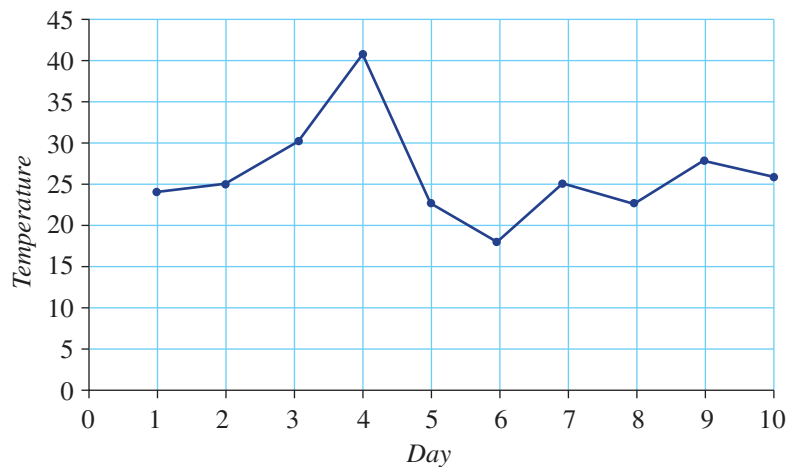
- 1 Mark the location of the median point for each of the sets of data points below.



Smoothing a time series graphically

Example 11

- 2 The time series plot below shows the maximum daily temperatures (in °C) in a city over a period of 10 consecutive days.

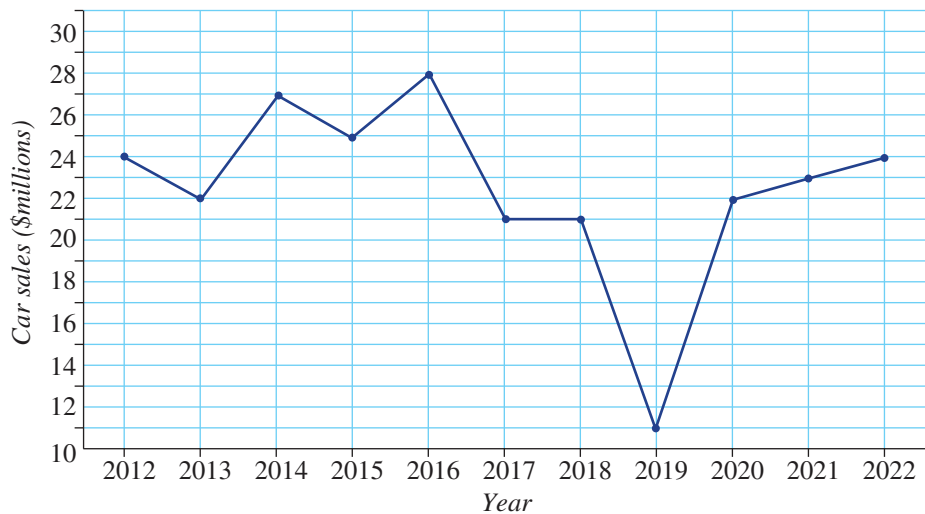


Use three-median smoothing to determine the smoothed temperature for:

a day 4

b day 8

- 3** The time series plot below shows the annual sales (in \$millions) for a car sales company. Use three-median smoothing to graphically smooth the plot and comment on the smoothed plot.

**Example 12**

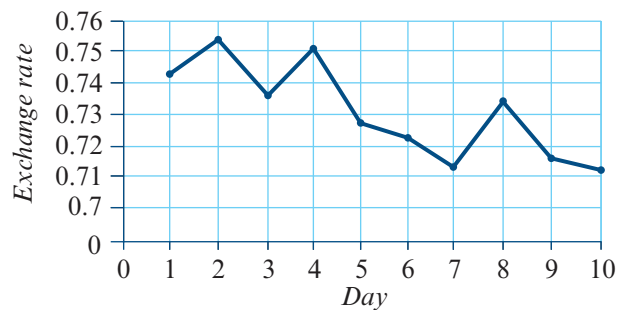
- 4** Using the time series plot shown in Question **2**, use five-median smoothing to determine the smoothed temperature for:

a day 4

b day 8

- 5** The time series plot opposite shows the value of the Australian dollar in US dollars (the exchange rate) over a period of 10 consecutive days in 2009.

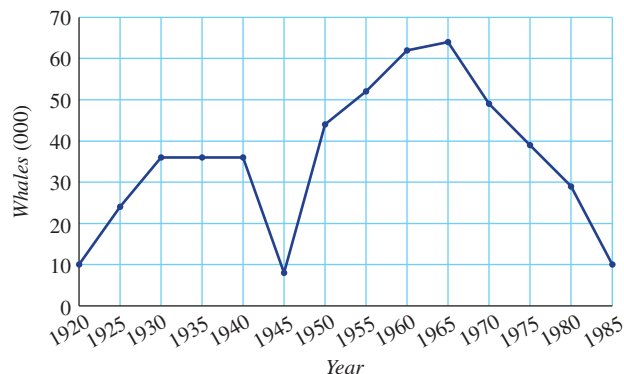
Use five-median smoothing to graphically smooth the plot and comment on the smoothed plot.



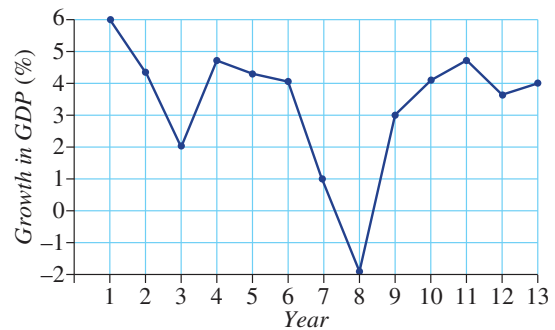
- 6** Use the graphical approach to smooth the time series plot below using:

a three-median smoothing

b five-median smoothing.



- 7** The time series plot opposite shows the percentage growth of GDP (gross domestic product) over a 13-year period.

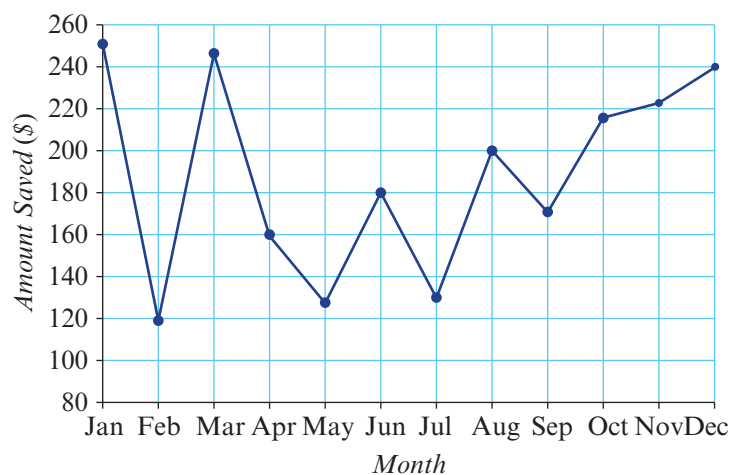


- a** Determine the median value of the percentage growth in GDP over the 13 year period.
- b** Smooth the times series graph:
- i** using three-median smoothing
 - ii** using five-median smoothing.
- c** What conclusions can be drawn about the variation in GDP growth from these smoothed time series plots?

Paper 1-style multiple-choice questions

Use the following information to answer Questions 8–10.

The time series plot below shows the amount that Lulu saved each month (to the nearest \$) over a 12 month period.



- 8** During the years shown in the time series plot, the median monthly amount Lulu saved is closest to:
- A** \$180 **B** \$155 **C** \$130 **D** \$190
- 9** The five-median smoothed amount saved by Lulu in July is closest to:
- A** \$130 **B** \$150 **C** \$170 **D** \$190
- 10** The nine-median smoothed amount saved by Lulu in August is closest to:
- A** \$132 **B** \$160 **C** \$168 **D** \$180

3D Seasonal indices

Learning intentions

- ▶ To interpret the meaning of seasonal indices.
- ▶ To seasonally adjust data using seasonal indices.
- ▶ To calculate seasonal indices from time series data.

When the data is considered to have a seasonal component, it is often necessary to remove this component so any underlying trend is clearer. The process of removing the seasonal component is called **deseasonalising** the data. To do this we need to calculate **seasonal indices**. Seasonal indices tell us how a particular season (generally a day, month or quarter) compares to the average season.

The concept of a seasonal index

Consider the (hypothetical) monthly seasonal indices for unemployment given in the table.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	Total
1.1	1.2	1.1	1.0	0.95	0.95	0.9	0.9	0.85	0.85	1.1	1.1	12.0

Key fact 1

Seasonal indices are calculated so that their **average** is 1. This means that the sum of the seasonal indices equals the number of seasons.

Thus, if the seasons are months, the seasonal indices add to 12. If the seasons are quarters, then the seasonal indices would add to 4, and so on.

Key fact 2

Seasonal indices tell us how a particular season (generally a day, month or quarter) compares to the average season.

For example, from the table above we can see:

- The seasonal index for unemployment for the month of February is 1.2. This means that February unemployment tends to be 1.2 times the average month, or equivalently 120% of the average month.
That is, February unemployment figures tend to be 20% **higher** than the monthly average. Remember, the average seasonal index is 1 or 100%.
- The seasonal index for unemployment for the month of August is 0.9. This tells us that the August unemployment figures tend to be 0.9 times the average month, or equivalently 90% of the average month.
That is, August unemployment figures tend to be 10% **lower** than the average month.


Example 13 Interpreting seasonal indices

Suppose that the seasonal indices (SI) for electricity usage in Esse's home are as shown in the table:

Summer	Autumn	Winter	Spring
1.16	0.94	1.26	0.64

- a** Interpret the seasonal index for Winter.
- b** Interpret the seasonal index for Spring.

Solution

- a** The seasonal index for Winter is 1.26. This tells us that Esse's electricity usage in Winter is typically 26% higher than the average season.
- b** The seasonal index for Spring is 0.64. This tells us that Esse's electricity usage in Spring is typically 36% lower than the average season.

Using seasonal indices to seasonally adjust a time series

We can use seasonal indices to remove the seasonal component (**deseasonalise**) from a time series, or to put it back in (**reseasonalise**). When we do this we are said to **seasonally adjust** the data.

To calculate deseasonalised figures, each entry is divided by its seasonal index as follows.

Deseasonalising data

Time series data are deseasonalised using the relationship:

$$\text{deseasonalised figure} = \frac{\text{actual figure}}{\text{seasonal index}}$$

The rule for determining deseasonalised data values can also be used to reseasonalise data – that is, convert a deseasonalised value into an actual data value.

Reseasonalising data

Time series data are reseasonalised using the rule:

$$\text{actual figure} = \text{deseasonalised figure} \times \text{seasonal index}$$


Example 14 Using seasonal indices

The seasonal indices (SI) for cold drink sales for Imogen's kiosk are as shown in the table:

Summer	Autumn	Winter	Spring
1.75	0.66	0.46	1.13

- a** If the actual cold drink sales last summer totalled \$21 653, what is the deseasonalised sales figure for that time period?
- b** If the deseasonalised cold drink sales last spring totalled \$10 870, what were the actual sales for that time period?

Solution

a Deseasonalised sales = $\frac{21\,653}{1.75}$
 $= \$12\,373.14$

b Actual sales = $10\,870 \times 1.13$
 $= \$12\,283.10$

Explanation

To deseasonalise we divide by the seasonal index for Summer (1.75).

To calculate the actual sales we multiply by the seasonal index for Spring (1.13).


Example 15 Using seasonal indices to determine percentage change required to correct for seasonality

Consider this table which gives the seasonal indices for heater sales at a discount store.

Summer	Autumn	Winter	Spring
0.65	1.25	1.35	0.075

- a** By what percentage should the sales in Summer be increased or decreased in order to deseasonalise the data? Give your answer as a percentage rounded to one decimal place.
- b** By what percentage should the sales in Winter be increased or decreased in order to deseasonalise the data? Give your answer as a percentage rounded to one decimal place.

Solution

a In general for Summer:

$$\begin{aligned}\text{deseseasonalised sales} &= \frac{\text{actual sales}}{0.65} \\ &= \frac{1}{0.65} \times \text{actual sales} \\ &= 1.538 \times \text{actual sales}\end{aligned}$$

Multiplying by the actual sales 1.538 is the equivalent of increasing the actual sales by 53.8%.

To correct for seasonality, the actual sales should be increased by 53.8%.

b In general for Winter:

$$\begin{aligned}\text{deseseasonalised sales} &= \frac{\text{actual sales}}{1.35} \\ &= \frac{1}{1.35} \times \text{actual sales} \\ &= 0.741 \times \text{actual sales}\end{aligned}$$

Multiplying the actual sales by 0.741 is the equivalent of decreasing the actual sales by $(100\% - 74.1\%) = 25.9\%$.

To correct for seasonality, the actual sales should be decreased by 25.9%.

Explanation

Insert the seasonal index for Summer into the rule

$$\text{deseseasonalised sales} = \frac{\text{actual sales}}{\text{seasonal index}}$$

Convert 1.538 into a percentage increase or decrease.

Write the answer in a sentence.

Insert the seasonal index for Winter into the rule

$$\text{deseseasonalised sales} = \frac{\text{actual sales}}{\text{seasonal index}}$$

Convert 0.741 into a percentage increase or decrease.

Write the answer in a sentence.

Calculating seasonal indices

To complete this section, you will now learn to calculate a seasonal index. We will start by using only one year's data to illustrate the basic ideas and then move onto a more realistic example where several years' data are involved.



Example 16 Calculating seasonal indices (1 year's data)

Mikki runs a shop and wishes to determine quarterly seasonal indices based on last year's sales (shown in table opposite).

Summer	Autumn	Winter	Spring
920	1085	1241	446

Solution

$$\text{Seasonal index} = \frac{\text{value for season}}{\text{seasonal average}}$$

$$\text{Quarterly average} = \frac{920 + 1085 + 1241 + 446}{4} = 923$$

$$SI_{\text{Summer}} = \frac{920}{923} = 0.997$$

$$SI_{\text{Autumn}} = \frac{1085}{923} = 1.176$$

$$SI_{\text{Winter}} = \frac{1241}{923} = 1.345$$

$$SI_{\text{Spring}} = \frac{446}{923} = 0.483$$

$$\text{Check: } 0.997 + 1.176 + 1.345 + 0.483 = 4.001$$

Seasonal indices

Summer	Autumn	Winter	Spring
0.997	1.176	1.345	0.483

Explanation

The seasons are quarters. Write the formula in terms of quarters.

Calculate the quarterly average for the year.

The seasonal index (SI) for each quarter is the ratio of that quarter's sales to the average quarter.

Check that the seasonal indices sum to 4 (the number of seasons). The slight difference is due to rounding error. Write out your answers as a table of the seasonal indices.

The next example illustrates how seasonal indices are calculated with three years' data. While the process looks more complicated, we just repeat what we did in Example 16 three times and average the results for each year at the end.

**Example 17** Calculating seasonal indices (several years' data)

Suppose that Mikki has three years of data, as shown. Use the data to calculate seasonal indices, correct to two decimal places.

Year	Summer	Autumn	Winter	Spring
1	920	1085	1241	446
2	1035	1180	1356	541
3	1299	1324	1450	659

Solution

- Calculate the seasonal indices for years 1, 2 and 3 separately. As we already have the seasonal indices for year 1 in the previous example we will save ourselves some time by simply quoting the result.
- Average the three sets of seasonal indices to obtain a single set of seasonal indices.

Solution

Year 1 seasonal indices:

Summer	Autumn	Winter	Spring
0.997	1.176	1.345	0.483

Year 2:

$$\text{Seasonal index} = \frac{\text{value for quarter}}{\text{quarterly average}}$$

$$\text{Quarterly average} = \frac{1035 + 1180 + 1356 + 541}{4} = 1028$$

$$SI_{\text{Summer}} = \frac{1035}{1028} = 1.007$$

$$SI_{\text{Autumn}} = \frac{1180}{1028} = 1.148$$

$$SI_{\text{Winter}} = \frac{1356}{1028} = 1.319$$

$$SI_{\text{Spring}} = \frac{541}{1028} = 0.526$$

$$\text{Check: } 1.007 + 1.148 + 1.319 + 0.526 = 4.000$$

Year 2 seasonal indices:

Summer	Autumn	Winter	Spring
1.007	1.148	1.319	0.526

Year 3:

$$\text{Quarterly average} = \frac{1299 + 1324 + 1450 + 659}{4} = 1183$$

$$SI_{\text{Summer}} = \frac{1299}{1183} = 1.098$$

$$SI_{\text{Autumn}} = \frac{1324}{1183} = 1.119$$

$$SI_{\text{Winter}} = \frac{1450}{1183} = 1.226$$

$$SI_{\text{Spring}} = \frac{659}{1183} = 0.557$$

$$\text{Check: } 1.098 + 1.119 + 1.226 + 0.557 = 4.000$$

Year 3 seasonal indices:

Summer	Autumn	Winter	Spring
1.098	1.119	1.226	0.557

Explanation

Write down the result for year 1.

Now calculate the seasonal indices for year 2.

The seasons are quarters. Write the formula in terms of quarters.

Calculate the quarterly average for the year.

Work out the seasonal index (SI) for each time period.

Check that the seasonal indices sum to 4.

Write out your answers as a table of the seasonal indices.

Now calculate the seasonal indices for year 3.

Determine the quarterly average for the year.

Work out the seasonal index (SI) for each time period.

Check that the seasonal indices sum to 4.

Write out your answers as a table of the seasonal indices.

Final seasonal indices:

$$SI_{\text{Summer}} = \frac{0.997 + 1.007 + 1.098}{3} = 1.03$$

$$SI_{\text{Autumn}} = \frac{1.176 + 1.148 + 1.119}{3} = 1.15$$

$$SI_{\text{Winter}} = \frac{1.345 + 1.319 + 1.226}{3} = 1.30$$

$$SI_{\text{Spring}} = \frac{0.483 + 0.526 + 0.557}{3} = 0.52$$

Check: $1.03 + 1.15 + 1.30 + 0.52 = 4.00$

Summer	Autumn	Winter	Spring
1.03	1.15	1.30	0.52

Determine the 3-year averaged seasonal indices by averaging the seasonal indices for each season.

Check that the seasonal indices sum to 4.

Write out your answers as a table of the seasonal indices.



Spread-sheet
↓

Example 18 Using Excel to calculate seasonal indices

Using the information in Example 17 repeat the calculation of seasonal indices for Mikki's shop using Excel.

Solution

Enter the data table as given in the question into Excel. In cell G3 enter the formula $= (C3 + D3 + E3 + F3) / 4$ to calculate the quarterly average for the first year. Press enter then **Fill down** to complete all quarterly averages.

	A	B	C	D	E	F	G
1							
2		Year	Summer	Autumn	Winter	Spring	Quarterly average
3		1	920	1085	1241	446	$=(C3+D3+E3+F3)/4$
4		2	1035	1180	1356	541	
5		3	1299	1324	1450	659	
6							

Now determine the seasonal indices for year 1, by entering the formulas as follows:

$C7 := C3/G3$ $D7 := D3/G3$ $E7 := E3/G3$ $F7 := F3/G3$

	A	B	C	D	E	F	G
1							
2		Year	Summer	Autumn	Winter	Spring	Quarterly average
3		1	920	1085	1241	446	923
4		2	1035	1180	1356	541	1028
5		3	1299	1324	1450	659	1183
6							
7			$=C3/G3$				
8							

	A	B	C	D	E	F	G
1							
2		Year	Summer	Autumn	Winter	Spring	Quarterly average
3		1	920	1085	1241	446	923
4		2	1035	1180	1356	541	1028
5		3	1299	1324	1450	659	1183
6							
7			0.99675	1.175515	1.344529	=F3/G3	
8							

Select cells C7, D7, E7, and F7 and **Fill down** two further rows, to determine the seasonal indices for each of the three years.

	A	B	C	D	E	F	G
1							
2		Year	Summer	Autumn	Winter	Spring	Quarterly average
3		1	920	1085	1241	446	923
4		2	1035	1180	1356	541	1028
5		3	1299	1324	1450	659	1183
6							
7			0.99675	1.175515	1.344529	0.483207	
8			1.006809	1.14786	1.319066	0.526265	
9			1.098056	1.119189	1.225697	0.557058	
10							

Finally, determine the average of the three years' seasonal indices by entering into C10 the formula = (C7 + C8 + C9)/3 as shown.

	A	B	C	D	E	F	G
1							
2		Year	Summer	Autumn	Winter	Spring	Quarterly average
3		1	920	1085	1241	446	923
4		2	1035	1180	1356	541	1028
5		3	1299	1324	1450	659	1183
6							
7			0.99675	1.175515	1.344529	0.483207	
8			1.006809	1.14786	1.319066	0.526265	
9			1.098056	1.119189	1.225697	0.557058	
10			= (C7+C8+C9)/3				

Select cell C10 and then **Fill right** to complete the seasonal index calculations. Round to two decimal places.

	A	B	C	D	E	F	G
1							
2		Year	Summer	Autumn	Winter	Spring	Quarterly average
3		1	920	1085	1241	446	923
4		2	1035	1180	1356	541	1028
5		3	1299	1324	1450	659	1183
6							
7			0.99675	1.175515	1.344529	0.483207	
8			1.006809	1.14786	1.319066	0.526265	
9			1.098056	1.119189	1.225697	0.557058	
10			1.03	1.15	1.30	0.52	
11							

Interpreting the seasonal indices

Having calculated these seasonal indices, what do they tell us about the previous situation?

The seasonal index of:

- 1.03 for summer tells us that summer sales are typically 3% above average
- 1.15 for autumn tells us that autumn sales are typically 15% above average
- 1.30 for winter tells us that winter sales are typically 30% above average
- 0.52 for spring tells us that spring sales are typically 48% below average.

Using seasonal indices to deseasonalise a time series

Once we have determined the seasonal indices, we can use the rule for deseasonalising the time series introduced earlier in this section to deseasonalise the data.

$$\text{deseasonalised figure} = \frac{\text{actual figure}}{\text{seasonal index}}$$



Example 19 Deseasonalising a time series

The quarterly sales figures for Mikki's shop over a 3-year period are given below.

Year	Summer	Autumn	Winter	Spring
1	920	1085	1241	446
2	1035	1180	1356	541
3	1299	1324	1450	659

Use the seasonal indices shown to deseasonalise these sales figures. Write the answers rounded to the nearest whole number.

Summer	Autumn	Winter	Spring
1.03	1.15	1.30	0.52

Solution

Deseasonalised Summer sales:

$$\text{Year 1: } \frac{920}{1.03} = 893$$

$$\text{Year 2: } \frac{1035}{1.03} = 1005$$

$$\text{Year 3: } \frac{1299}{1.03} = 1261$$

Deseasonalised sales figures

Year	Summer	Autumn	Winter	Spring
1	893	943	955	858
2	1005	1026	1043	1040
3	1261	1151	1115	1267

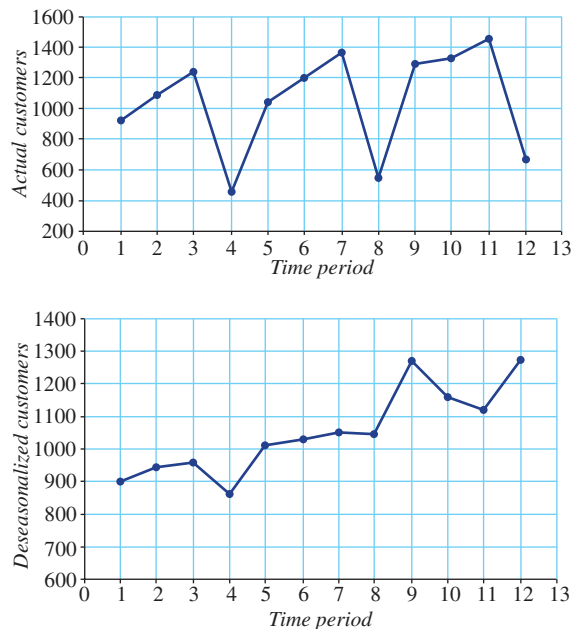
Explanation

To deseasonalise each sales figure in the table, divide by the appropriate seasonal index. For example, for summer, divide the figures in the 'Summer' column by 1.03. Round results to the nearest whole number.

Repeat for the other seasons.

Why deseasonalise?

The purpose of removing the seasonality component of a time series is generally so that any trend in the time series is clearer. Consider again the actual customer data, and the deseasonalised customer data from Example 18, both of which are shown in the following time series plots.



It is hard to see from the first plot whether there has been any growth in Mikki's business, but the deseasonalised plot reveals a clear underlying trend in the data.

It is common to deseasonalise time series data before you fit a trend line. We will consider this further in the next section.

Section Summary

- ▶ Seasonal indices tell us how a particular season (generally a day, month or quarter) compares to the average season.
- ▶ Seasonal indices are calculated so that their average is 1. This means that the sum of the seasonal indices equals the number of seasons.
- ▶ Time series data are deseasonalised using the rule:

$$\text{deseasonalised figure} = \frac{\text{actual figure}}{\text{seasonal index}}$$

- ▶ Time series data are reseasonalised using the rule:

$$\text{actual figure} = \text{deseasonalised figure} \times \text{seasonal index}$$
- ▶ Adjusting time series data for seasonality allows any trend in the data to be seen more clearly.



Exercise 3D

Interpreting and using seasonal indices

Use the following information to answer Questions 1–4.

The table below shows the monthly sales figures (in \$'000s) and seasonal indices (for January to November) for a product produced by the U-beaut company.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Sales (\$'000s)	9.6	10.5	8.6		7.1	6.0	5.4		6.4	7.2	8.3	7.4
Seasonal index	1.2	1.3	1.1	1.0	1.0	0.9	0.8	0.7	0.9	1.0	1.1	

Example 13

- 1 **a** Determine the seasonal index for December.
- b** Interpret the seasonal index for February.
- c** Interpret the seasonal index for September.

Example 14

- 2 **a** Determine the deseasonalised sales figure (in \$'000s) for March, giving your answer rounded to one decimal place.
- b** Determine the deseasonalised sales figure (in \$'000s) for June, giving your answer rounded to one decimal place.
- 3 **a** The deseasonalised sales figure (in \$'000s) for August is 5.6. Determine the actual sales (in \$'000s), giving your answer rounded to one decimal place.
- b** The deseasonalised sales figure (in \$'000s) for April is 6.9. Determine the actual sales (in \$'000s), giving your answer rounded to one decimal place.

Example 15

- 4 **a** By what percentage should the sales in August be increased or decreased in order to correct for seasonality? Give your answer as a percentage rounded to one decimal place.
- b** By what percentage should the sales in February be increased or decreased in order to correct for seasonality? Give your answer as a percentage rounded to one decimal place.
- 5 The table below shows the quarterly newspaper sales (in \$'000s) of a corner store. Also shown are the seasonal indices for newspaper sales for the first, second and third quarters.

	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Sales		1060	1868	1642
Seasonal index	0.8	0.7	1.3	

- a** Determine the seasonal index for quarter 4.
- b** Determine the deseasonalised sales (in \$'000s) for quarter 2.
- c** Determine the deseasonalised sales (in \$'000s) for quarter 3.
- d** Deseasonalised sales (in \$'000s) for quarter 1 are 1256. Determine the actual sales.

Deseasonalising a time series

- 6 The following table shows the number of students enrolled in a 3-month computer systems training course along with some seasonal indices that have been calculated from the previous year's enrolment figures. Complete the table by calculating the seasonal index for spring and the deseasonalised student numbers for each course.

	Summer	Autumn	Winter	Spring
<i>Number of students</i>	56	125	126	96
Deseasonalised numbers				
Seasonal index	0.5	1.0	1.3	

- 7 The number of waiters employed by a restaurant chain in each quarter of 1 year, along with some seasonal indices that have been calculated from the previous year's data, are given in the following table.

	Quarter 1	Quarter 2	Quarter 3	Quarter 4
<i>Number of waiters</i>	198	145	86	168
Seasonal index	1.30		0.58	1.10

- Determine the seasonal index for the second quarter.
- The seasonal index for quarter 1 is 1.30. Explain what this means in terms of the average quarterly number of waiters.
- Deseasonalise the data.

Calculating seasonal indices

Example 16

- 8 The table below records quarterly sales (in \$'000s) for a shop.

Quarter 1	Quarter 2	Quarter 3	Quarter 4
60	56	75	78

Use the data to determine the seasonal indices for the four quarters. Give your results rounded to two decimal places.

- 9 The table below records the monthly visitors (in '000s) to a museum over one year.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
12	13	14	17	18	15	9	10	8	11	15	20

Use the data to determine the seasonal indices for the 12 months. Give your results rounded to two decimal places.

- 10** The table below records the monthly sales (in \$'000s) for a shop over two years.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
22	19	25	23	20	18	20	15	14	11	23	30
21	20	23	25	22	17	19	17	16	11	25	31

Use the two years of data to determine monthly seasonal indices.

- 11** The daily number of cars carried on a car ferry service each day over a two-week period, together with the daily seasonal indices, are shown in the table below:

Week	Mon	Tues	Wed	Thur	Fri	Sat	Sun
1	124	110	45	67	230	134	330
2	120	108	57	74	215	150	345
Seasonal index	0.8	0.7	0.3	0.5	1.5	1.0	2.2

- a** Use the seasonal indices to deseasonalise the data, rounding answers to the nearest whole number.
b Construct a time series plot of the deseasonalised data.

Example 17

- 12** The number of retail job vacancies advertised on an online job board each quarter in each of three consecutive years are shown in the following table.

Example 19

Year	Quarter 1	Quarter 2	Quarter 3	Quarter 4
1	212	194	196	227
2	220	197	196	239
3	231	205	203	245

- a** Construct a time series plot of the data.
b Use the data to calculate seasonal indices, rounded to two decimal places.
c Use the seasonal indices to construct a table of the deseasonalised data.
d On the same axes, construct a time series plot of the deseasonalised data.
e Use the deseasonalised plot to describe any trend in the time series.
- 13** The table below shows the room occupancy rate (%) for a chain of hotels over the summer, autumn, winter and spring quarters for the years 2022–2024.

Season	2022	2023	2024
summer	72	73	75
autumn	62	61	67
winter	68	70	73
spring	75	85	88

- a** Construct a time series plot of the data.
- b** Use the data to calculate seasonal indices, rounded to two decimal places.
- c** Use the seasonal indices to construct a table of the deseasonalised data.
- d** On the same axes, construct a time series plot of the deseasonalised data.
- e** Use the deseasonalised plot to describe any trend in the time series.

Paper 1-style multiple-choice questions

Use the following information to answer Questions 14–16.

The table below shows the number of customers each month at a restaurant together with the seasonal indices for the number of customers each month of the year. The number of customers for August is missing.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Seasonal index	1.36	1.19	1.05	1.01	0.93	0.82	0.75	0.68	0.87	0.9	1.18	1.26
Number of customers	934	836	736	716	649	554	541		598	626	826	873

- 14** To adjust the number of customers in May for seasonality, the actual number of customers should be:
- A** increased by 93.0% **B** decreased by 7.0%
C decreased 7.5% **D** increased by 7.5%
- 15** To adjust the number of customers in November for seasonality, the actual number of customers should be:
- A** increased by 18.0% **B** increased by 15.3%
C decreased 15.3% **D** decreased by 18.0%
- 16** If the deasonalised number of customers for August is 700, the actual number of customers in that month is closest to:
- A** 1029 **B** 768 **C** 607 **D** 476
- 17** The table below records the monthly average electricity cost (in dollars) for a home.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
223	190	253	236	201	189	203	153	143	111	235	307

The seasonal index for August is closest to:

- A** 1.00 **B** 0.75 **C** 1.25 **D** 0.87

3E Fitting a trend line and forecasting

Learning intentions

- ▶ To use the method of least squares to fit a trend line to a time series.
- ▶ To use the trend line to make predictions.
- ▶ To use seasonal indices to add seasonality to predicted values as appropriate.

Fitting a trend line

If there appears to be a linear trend, we can use the least squares method to fit a line to the data to model the trend. The following example demonstrates fitting a trend line to time series data which shows no seasonal component.

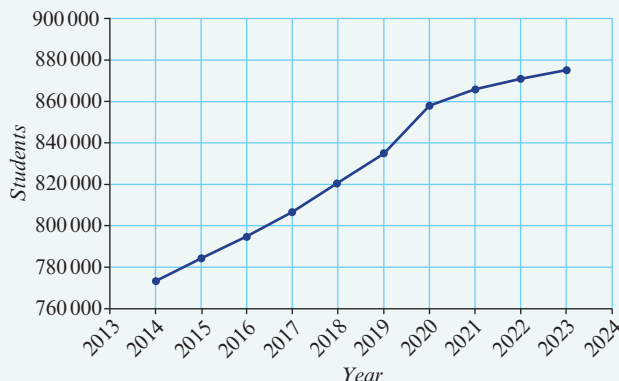


Example 20 Fitting a trend line

Fit a trend line to the data in the following table, which shows the total number of school students in Queensland over the years 2014–2023. Interpret the slope.

Year	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Students	773 309	784 224	794 815	806 555	820 700	834 818	857 920	865 888	870 819	875 232

Solution



$$\text{intercept} = -24\,213\,817$$

$$\text{slope} = 12\,406$$

$$\text{Number of students} = 12\,406 \times \text{year} - 24\,213\,817$$

The number of students at school in Queensland increased on average by 12 406 students per year over this time period.

Explanation

Construct a time series plot of the data to confirm that the trend is linear.

Fit a least squares regression trend line to the data with Year as the EV. Write down its equation. Interpret the slope.

Forecasting

Using a trend line fitted to a time series plot to make predictions about future values is known as **trend line forecasting**. When forecasting, we are using a value of the explanatory variable which is outside of the range of values used to fit the least squares line. That is, we are extrapolating; as such, our predictions become more unreliable the further into the future we predict.



Example 21 Using a trend line to forecast a future value

Use the data and least squares regression trend line from Example 20. How many students do we predict will be attending school in Queensland in 2030 if the same increasing trend continues? Give your answer rounded to the nearest 10 000 students.

Solution

$$\begin{aligned}\text{Number of students} &= 12\,406 \times 2030 - 24\,213\,817 \\ &= 970\,363 \\ &\approx 970\,000\end{aligned}$$

Explanation

Substitute the appropriate value for year in the equation determined using the least squares regression trend line. Round the answer to the nearest 10 000 students.

Forecasting taking seasonality into account

When time series data is seasonal, it is usual to deseasonalise the data before fitting the trend line.



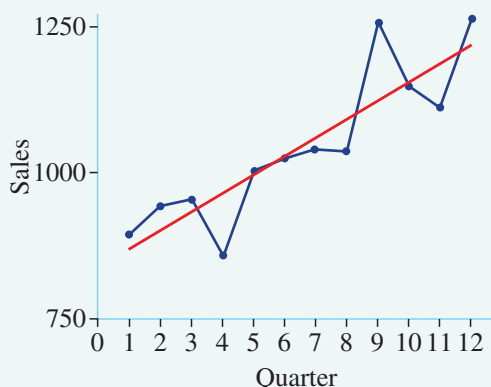
Example 22 Fitting a trend line (seasonality)

The *deseasonalised* quarterly sales data from Mikki's shop are shown below.

Quarter	1	2	3	4	5	6	7	8	9	10	11	12
Sales	893	943	955	858	1005	1026	1043	1040	1261	1151	1115	1267

Fit a trend line and interpret the slope.

Solution



Explanation

Plot the time series. Using your calculator (with *Quarter* as the EV and *Sales* as the RV), determine the equation of the least squares regression trend line. Plot it on the time series.

Solution

$$\text{Sales} = 32.1 \times \text{quarter} + 838.0$$

Over the 3-year period, sales at Mikki's shop increased at an average rate of 32 sales per quarter.

Explanation

Write down the equation of the least squares regression trend line.

Interpret the slope in terms of the variables involved.

Forecasting with seasonality

When using deseasonalised data to fit a trend line, you must remember that the result of any prediction is a deseasonalised value. To be meaningful, this result must then be reseasonalised by multiplying by the appropriate seasonal index.



Example 23 Forecasting (seasonality)

What sales do we predict for Mikki's shop in the winter of the 4th year? (Because many items have to be ordered well in advance, retailers often need to make such decisions.)

Solution

$$\begin{aligned}\text{Sales} &= 838.0 + 32.1 \times \text{quarter} \\ &= 32.1 \times 15 + 838.0 \\ &= 1319.5\end{aligned}$$

Deseasonalised sales prediction for winter of year 4 = 1319.5

$$\begin{aligned}\text{Seasonalised sales prediction for winter of year 4} &= 1319.5 \times 1.30 \\ &\approx 1715\end{aligned}$$

Explanation

Substitute the appropriate value for the time period in the equation for the trend line. Since summer year 1 was designated as quarter '1' in Example 19, then winter year 4 is quarter '15'.

The value calculated is the deseasonalised sales figure for the quarter in question. To obtain the *actual* predicted sales figure we need to reseasonalise this predicted value. To do this, we multiply this value by the seasonal index for winter, which was found to be 1.30 in Example 17.

Section Summary

- ▶ Least squares regression can be used to fit a trend line to time series data.
- ▶ If the time series data has a seasonal component then this should be removed before the trend line is fitted.
- ▶ The trend line can be used for forecasting future values of the time series.
- ▶ If the trend line is fitted to deseasonalised data then any forecast values need to be reseasonalised using the relevant seasonal index.

Exercise 3E

Fitting a least squares line to a time series plot (no seasonality)

Example 20

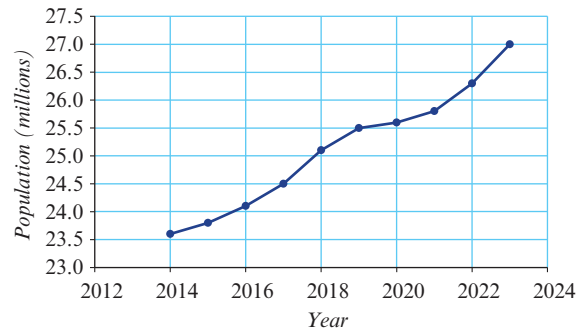
Example 21

- 1 Consider the population of Australia (in millions) over the period 2014–2023.

Year	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Number	23.6	23.8	24.1	24.5	25.1	25.5	25.6	25.8	26.3	27.0

The time series plot of the data is shown below.

- a Comment on the plot.
- b Fit a least squares regression trend line to the data, giving the values of the coefficients to three decimal places, and interpret the slope.
- c Use this equation to predict the population of Australia in 2030 to the nearest 100 000 people.

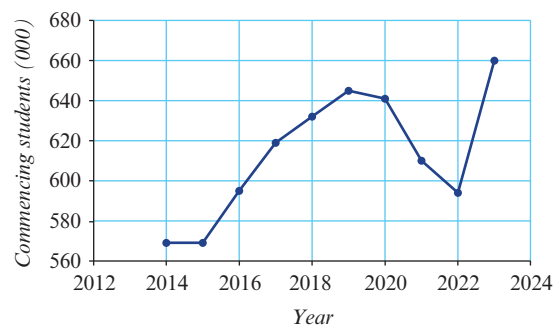


- 2 The data show the number of commencing university students (in thousands) in Australia for the period 2014–2023.

Year	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Number	569	569	595	619	632	645	641	610	594	660

The time series plot of the data is shown below.

- a Comment on the plot.
- b Fit a least squares regression trend line to the data, giving the values of the coefficients to three decimal places.
- c Use this equation to predict the number of students expected to commence university in Australia in 2030 to the nearest 10 000 students.



- 3 The table below shows the percentage of total retail sales that were made in department stores over an 11-year period:

Year	1	2	3	4	5	6	7	8	9	10	11
Sales (%)	12.3	12.0	11.7	11.5	11.0	10.5	10.6	10.7	10.4	10.0	9.4

- a Construct a time series plot.
 - b Comment on the time series plot in terms of trend.
 - c Fit a least squares regression trend line to the time series plot, rounding coefficients to three significant figures, and interpret the slope.
 - d Draw the trend line on your time series plot.
 - e Use the least squares equation to forecast the percentage of retail sales which will be made by department stores in year 15. Give your answer as a percentage rounded to one decimal place.
- 4 The median ages of first-time mothers in Australia over the years 2010–2020 are shown below.

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Age	30.7	30.7	30.7	30.8	30.8	30.9	31.1	31.2	31.3	31.4	31.5

- a Fit a least squares regression trend line to the data and interpret the slope.
 - b Use the trend line to forecast the average ages of mothers having their first child in Australia in 2030. Explain why this prediction is not likely to be reliable.
- 5 The average weekly earnings (in dollars) in Australia during the period 2014–2021 are given in the following table.

Year	2014	2015	2016	2017	2018	2019	2020	2021
Earnings	1454.10	1483.10	1516.00	1543.20	1585.30	1634.80	1713.90	1737.10

- a Fit a least squares regression trend line to the data, rounding coefficients to four significant figures, and interpret the slope.
- b Use this trend relationship to forecast average weekly earnings in 2030. Explain why this prediction is not likely to be reliable.

Fitting a least squares regression trend line to a time series with seasonality

Example 22

- 6 The table below shows the deseasonalised quarterly washing-machine sales of a company over 3 years.

Example 23

	Year 1				Year 2				Year 3			
Quarter	1	2	3	4	5	6	7	8	9	10	11	12
Deseasonalised sales	53	51	54	55	64	64	61	63	67	69	68	66

- a Use least squares regression to fit a trend line to the data.
- b Use this trend equation for washing-machine sales, with the seasonal indices below, to forecast the sales of washing machines in the fourth quarter of year 4, rounding your answer to the nearest whole number.

Quarter	1	2	3	4
Seasonal index	0.90	0.81	1.11	1.18

- 7** The quarterly seasonal indices for the sales of boogie boards in a surf shop are as follows.

Seasonal index	1.13	0.47	0.62	1.77

The actual sales of the boogie boards over a 2-year period are given in the table.

Year	Quarter 1	Quarter 2	Quarter 3	Quarter 4
1	138	60	73	230
2	283	115	163	417

- Use the seasonal indices to calculate the deseasonalised sales figures for this period to the nearest whole number.
- Plot the actual sales figures and the deseasonalised sales figures for this period and comment on the plot.
- Fit a trend line to the deseasonalised sales data. Write the slope and intercept rounded to three significant figures.
- Use the relationship calculated in **c**, together with the seasonal indices, to forecast the sales for the first quarter of year 4, rounding your answer to the nearest whole number.

Paper 1-style multiple-choice questions

- 8** The number of visitors to an adventure park is seasonal. A least squares regression line has been fitted to the data, and the equation is:

$$\text{deseasonalised number of visitors} = 286.5 \times \text{quarter} + 38\,345$$

where *quarter* number one is January–March 2022.

The quarterly seasonal indices for visitors to the adventure park are shown in the table below.

Quarter	Jan–Mar	Apr–Jun	Jul–Sept	Oct–Dec
Seasonal index	1.17	0.91	0.78	1.14

The predicted number of actual visitors for the April–June quarter in 2025 is closest to:

- A** 42 356 **B** 38 544 **C** 46 545 **D** 37 501
- 9** An electrical goods retailer knows that the sales of air conditioners are seasonal. A least squares regression line has been fitted to the data collected by the retailer in 2021 and 2022, and the equation is:

$$\text{deseasonalised number of air conditioners} = 1.2 \times \text{month} + 197$$

where *month* number one is January 2021.

The monthly seasonal indices for air conditioner sales are shown in the table below.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Seasonal index	1.36	1.19	1.05	1.01	0.93	0.82	0.75	0.68	0.87	0.90	1.18	1.26

The predicted number of actual sales for November 2025 is closest to:

- A** 316 **B** 227 **C** 299 **D** 333

Key ideas and chapter summary

Time series data	Time series data are a special case of bivariate data, where the explanatory variable is the time at which the values of the response variable were recorded.	
Time series plot	A time series plot is a bivariate plot where the values of the response variable are plotted in time order. Points in a time series plot are joined by line segments.	
Features to look for in a time series plot	■ Trend ■ Seasonality ■ Possible outliers	■ Cycles ■ Structural change ■ Irregular (random) fluctuations
Trend	A general increase or decrease over a significant period of time in a times series plot is called a trend .	
Cycles	Cycles are present when there is a periodic movement in a time series. The period is the time it takes for one complete up and down movement in the time series plot. This term is generally reserved for periodic movements with a period greater than one year.	
Seasonality	Seasonality is present when there is a periodic movement in a time series that has a calendar related period – for example, a year, a month, a week.	
Structural change	Structural change is present when there is a sudden change in the established pattern of a time series plot.	
Outliers	Outliers are present when there are individual values that stand out from the general body of data.	
Irregular (random) fluctuations	Irregular (random) fluctuations are always present in any real-world time series plot. They include all of the variations in a time series that we cannot reasonably attribute to systematic changes like trend, cycles, seasonality, structural change or the presence of outliers.	
Smoothing	Smoothing is a technique used to eliminate some of the irregular fluctuations in a time series plot so that features such as trend are more easily seen.	
Moving mean smoothing	In moving mean smoothing , each original data value is replaced by the mean of itself and a number of data values on either side.	

Moving median smoothing	Moving median smoothing is a graphical technique for smoothing a time series plot using moving medians rather than means.
Seasonal indices	Seasonal indices are used to quantify the seasonal variation in a time series.
Deseasonalise	The process of removing for the effects of seasonality in a time series is called deseasonalisation .
Reseasonalise	The process of converting seasonal data back into its original form is called reseasonalisation .
Trend line forecasting	Trend line forecasting uses the equation of a trend line to make predictions about the future.

Skills checklist



Download this checklist from the Interactive Textbook, then print it and fill it out to check your skills.


3A
1 I can construct a time series plot.


See Example 1 and Exercise 3A Question 1

3A
2 I can use technology to construct a time series plot.


See Example 2 and Exercise 3A Question 1

3A
3 I can identify trends in a time series plot.


See Example 3 and Exercise 3A Question 4

3A
4 I can identify cycles in a time series plot.


See Example 4 and Exercise 3A Question 5

3A
5 I can identify seasonality in a time series plot.


See Example 5 and Exercise 3A Question 5

3A
6 I can identify structural change in a time series plot.


See Example 6 and Exercise 3A Question 6

3A
7 I can identify outliers in a time series plot.


See Example 7 and Exercise 3A Question 7

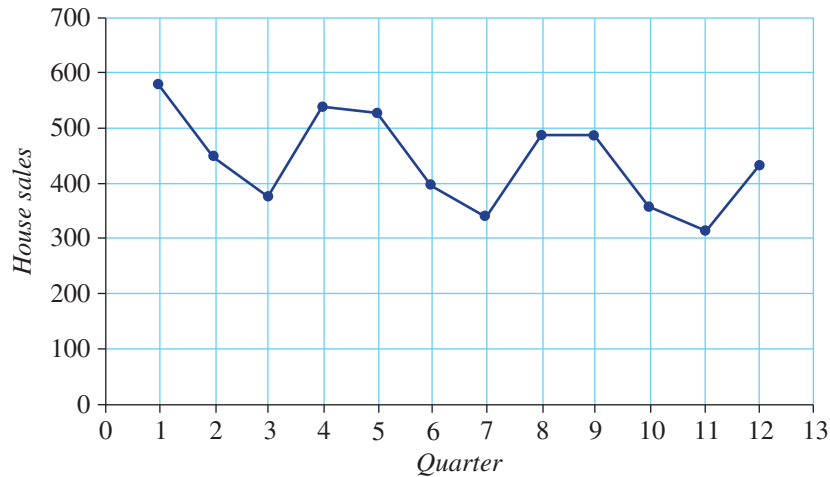
3B
8 I can smooth a time series using moving mean smoothing.


See Example 8, Example 9 and Exercise 3B Questions 1 and 4

3C	9 I can smooth a time series using moving median smoothing.	☐
	See Example 11, Example 12 and Exercise 3C Questions 2 and 4	
3D	10 I can interpret seasonal indices.	☐
	See Example 13 and Exercise 3D Question 1	
3D	11 I can use seasonal indices to deseasonalise and reseasonalise data.	☐
	See Example 14 and Exercise 3D Question 2	
3D	12 I can use seasonal indices to determine the percentage change required to correct for seasonality.	☐
	See Example 15 and Exercise 3D Question 4	
3D	13 I can calculate seasonal indices from 1 year of data.	☐
	See Example 16 and Exercise 3D Question 8	
3D	14 I can calculate seasonal indices from several years of data.	☐
	See Example 17 and Exercise 3D Question 12	
3D	15 I can use seasonal indices to deseasonalise a time series.	☐
	See Example 19 and Exercise 3D Question 12	
3E	16 I can fit a trend line to a time series plot.	☐
	See Example 20 and Exercise 3E Question 1	
3E	17 I can use a trend line to forecast a future value (no seasonality).	☐
	See Example 21 and Exercise 3E Question 1	
3E	18 I can fit a trend line to a time series plot with seasonality.	☐
	See Example 22 and Exercise 3E Question 6	
3E	19 I can use a trend line to forecast a future value (with seasonality).	☐
	See Example 23 and Exercise 3E Question 6	

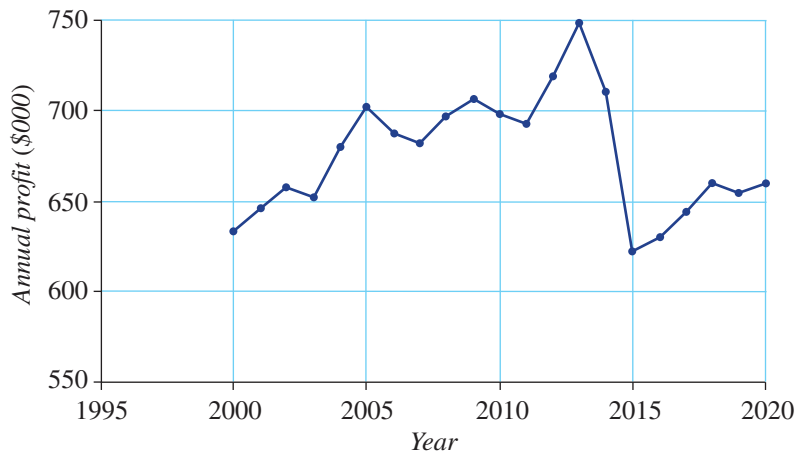
Multiple-choice questions

- 1 The time series plot below shows quarterly house sales for a real estate agency over a three year period.



The time series plot is best described as showing:

- A** seasonality only
 - B** seasonality with irregular fluctuations
 - C** an increasing trend with seasonality and irregular fluctuations
 - D** a decreasing trend with seasonality and irregular fluctuations.
- 2 The time series plot below shows the annual profit (in \$000) for a manufacturing company.



The time series plot is best described as showing:

- A** increasing trend
- B** seasonality with irregular fluctuations
- C** increasing trend with an outlier
- D** increasing trend with a structural change.

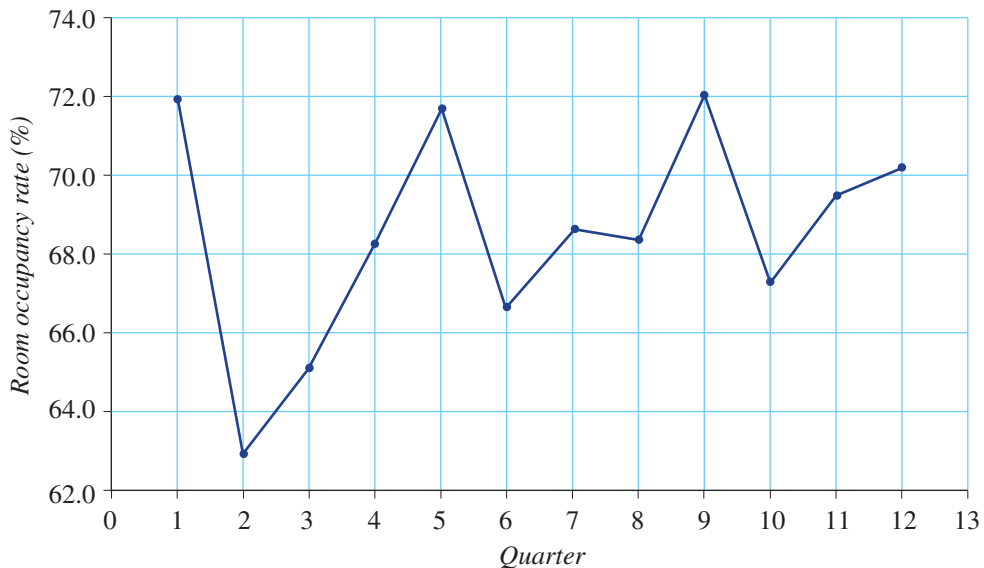
Use the following table to answer Questions 3–6.

Time period	1	2	3	4	5	6	7
Data value	2.3	3.4	4.4	2.7	5.1	3.7	4.2

- 3** The three-moving mean for time period 2 is closest to:
A 3.4 **B** 3.6 **C** 3.9 **D** 4.0
- 4** The five-moving mean for time period 3 is closest to:
A 3.6 **B** 3.9 **C** 4.1 **D** 4.2
- 5** The three-moving mean for time period 5 is closest to:
A 2.7 **B** 3.8 **C** 3.9 **D** 4.0
- 6** The seven-moving mean for time period 4 is closest to:
A 2.7 **B** 3.6 **C** 3.7 **D** 4.1

Use the following information to answer Questions 7 and 8.

The time series plot for hotel room occupancy rate (%) in a large city over a three year period is shown below.



- 7** The three-median smoothed value for Quarter 2 is closest to:
A 62 **B** 63 **C** 64 **D** 65
- 8** The five-median smoothed value for Quarter 3 is closest to:
A 64 **B** 65 **C** 68 **D** 69

- 9 The seasonal indices for the number of customers at a restaurant are as follows:

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1.0	p	1.1	0.9	1.0	1.0	1.2	1.1	1.1	1.1	1.0	0.7

The value of p is:

- A** 0.5 **B** 0.7 **C** 0.8 **D** 1.0

The following information relates to Questions 10 and 11.

- 10 The table shows the closing price (*price*) of a company's shares on the stock market over a 10 day period.

Day	1	2	3	4	5	6	7	8	9	10
Price (\$)	2.85	2.80	2.78	2.40	2.80	3.15	3.42	3.95	4.05	3.35

The five-mean smoothed *price* on Day 5 is closest to:

- A** \$2.80 **B** \$2.89 **C** \$2.91 **D** \$2.99

- 11 If five-mean smoothing was used to smooth this time series, the number of smoothed values would be:

- A** 5 **B** 6 **C** 7 **D** 8

- 12 Suppose that Lyn spent a total of \$427 on dining out over the period from January to March, and then another \$230 over the period April-May. The five-mean smoothed amount that she spent in March is closest to:

- A** \$115 **B** \$129 **C** \$131 **D** \$142

Use the following information to answer Questions 13–16.

The seasonal indices for the number of bathing suits sold at a surf shop are given in the table.

Quarter	Summer	Autumn	Winter	Spring
Seasonal index	1.8	0.4	0.3	1.5

- 13 The number of bathing suits sold one summer is 432. The deseasonalised number is closest to:

- A** 432 **B** 240 **C** 778 **D** 540

- 14 The deseasonalised number of bathing suits sold one winter was 380. The actual number was closest to:

- A** 114 **B** 152 **C** 380 **D** 1267

- 15** The seasonal index for spring tells us that, over time, the number of bathing suits sold in spring tends to be:
- A** 50% less than the seasonal average
 - B** 15% less than the seasonal average
 - C** 15% more than the seasonal average
 - D** 50% more than the seasonal average.
- 16** To correct for seasonality, the actual number of bathing suits sold in Autumn should be:
- A** reduced by 50%
 - B** reduced by 40%
 - C** increased by 40%
 - D** increased by 150%

- 17** The number of visitors to an information centre each quarter was recorded for one year. The results are tabulated below.

Quarter	Summer	Autumn	Winter	Spring
Visitors	1048	677	593	998

Using this data, the seasonal index for autumn is estimated to be closest to:

- A** 0.25
- B** 1.0
- C** 1.23
- D** 0.82

Use the following information to answer Questions 18 and 19.

A trend line is fitted to a time series plot displaying the percentage change in commencing international student enrolments in Australia each year compared to the previous year (*enrolments*) for the period 2012–2019.

The equation of this line is: $\% \text{ change in enrolments} = 1.73 \times \text{year} - 3480$

- 18** Using this trend line, the percentage change in enrolments from the previous year forecast for 2026 is:
- A** 24.98
 - B** -11.05
 - C** 1.73
 - D** 12.11
- 19** From the slope of the trend line it can be said that:
- A** on average, the number of commencing international student enrolments in Australia is increasing by 1.73 each year.
 - B** on average, the number of commencing international student enrolments in Australia is increasing by 1.73% each year.
 - C** on average, the number of commencing international student enrolments in Australia is decreasing by 1.73% each month.
 - D** on average, the number of commencing international student enrolments in Australia is decreasing by 1.73% each year.

20 Suppose that the seasonal indices for the wholesale price of petrol are:

Day	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
Index	1.2	1.0	0.9	0.8	0.7	1.2	1.2

The equation of the least squares regression line that could enable us to predict the *deseasonalised price* per litre in cents from the *day number* is

$$\text{deseasonalised price} = 0.23 \times \text{day number} + 189.9$$

where day number 1 is Sunday March 20. The predicted actual price for Sunday April 3 is closest to:

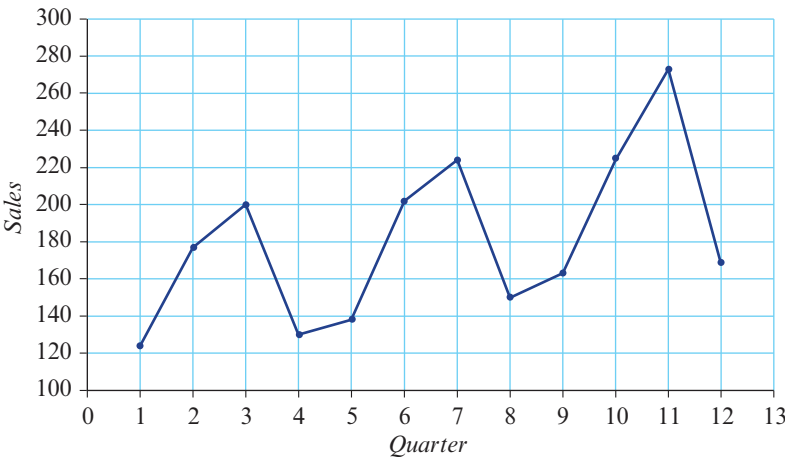
- A** 232.0 cents **B** 193.1 cents **C** 193.4 cents **D** 231.7 cents

Short-response questions

1 Construct a time series plot to display the following data:

Year	2017	2018	2019	2020	2021	2022	2023	2024
Sales	12	15	35	23	7	56	78	70

2 The time series plot below shows the number of houses sold per quarter in a certain town over a 3-year period. Describe the features of the plot.



3 The following table shows the maximum daily temperature in a certain town over a one-week period:

Day	Mon	Tues	Wed	Thu	Fri	Sat	Sun
Temp (°C)	20	21	26	36	34	42	26

- a** Determine the three-mean smoothed temperature for Tuesday.
b Determine the five-mean smoothed temperature for Friday.

- 4 The value of an Australian dollar in US dollars (exchange rate) over a 10-day period is given in the table.

Day	1	2	3	4	5	6	7	8	9	10
Exchange rate	0.718	0.715	0.709	0.710	0.708	0.704	0.705	0.701	0.704	0.707

- a Determine the three-median smoothed value for Day 5.
 b Determine the five-median smoothed value for Day 7.

- 5 The seasonal indices for the daily sales in an ice-cream shop are as follows:

Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Seasonal index	0.55	0.50	0.58	0.64	1.0	?	1.87

- a Determine the seasonal index for Saturday.
 b Interpret this index in terms of the average daily sales for this shop.
 c If the actual sales one Sunday was 456 ice-creams, determine the deseasonalised sales figure.

- 6 The number of staff employed by a restaurant chain in each quarter of one year is given in the following table.

Quarter	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Number of staff	35	98	65	23

Use these data to calculate quarterly seasonal indices for the number of staff employed by the restaurant chain.

- 7 The number of visitors to a website (in thousands) each year from 2018–2024 is given in the following table.

Year	2018	2019	2020	2021	2022	2023	2024
Visitors	6 490	6 970	6 771	7 032	7 382	7 868	8 244

- a Determine the equation of the least squares trend line for the data.
 b Use this to predict the number of visitors to the website in 2030, to the nearest 1000.
 c Comment on the reliability of this prediction.

- 8 The table below shows the average mortgage interest rates in Australia for the period 1987–97, a period of time when interest rates were very high. Also shown are the three-mean smoothed average rates but with one missing.

Year	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Av rate (%)	15.50	13.50	17.00	16.50	13.00	10.50	9.50	8.75	10.50	8.75	7.55
3-mean (%)		15.33	15.67	15.50	13.33		9.58	9.58	9.33	8.93	

- a Complete the table.
- b Construct a time series plot for the average interest rate during the period 1987–97.
- c Plot the smoothed interest rate data on the graph and comment on any trend.

- 9 The table below shows the carbon dioxide emissions in Australia (in millions of tonnes) for the period 2014 to 2023.

Year	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
CO ₂	396	401	408	411	409	408	393	385	375	374

- a Fit a least squares line to the time series plot that will enable *emissions* to be predicted from *year*.
- b Use the least squares equation to predict the carbon dioxide emissions in Australia in 2028. Round to the nearest whole number.
- c Comment on the reliability of your prediction.

- 10 The number of people arriving at a certain station each day is seasonal, with seasonal indices as shown below:

Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Seasonal index	1.32	1.20	0.95	0.95	1.48	0.65	0.45

It is also known that over a specific time period the number of travellers each day has generally been increasing, according to the following equation which was determined from deseasonalised data:

$$\text{number of travellers} = 22.5 \times \text{day} + 135$$

If a certain Saturday is designated as Day 1, determine the number of travellers predicted for the following Saturday.

- 11 The table below shows the number of dolphins spotted in a bay over each of the four seasons for the years 2022–2024.

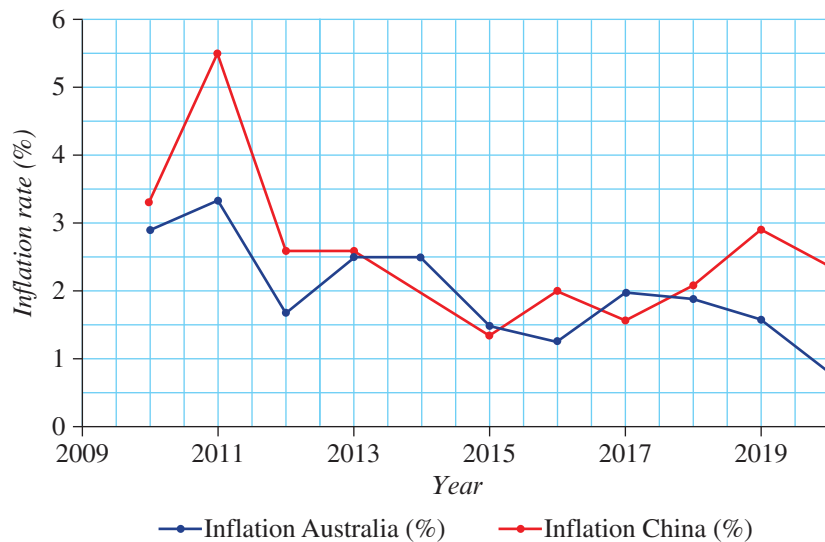
Year	Summer	Autumn	Winter	Spring
2022	97	112	480	678
2023	107	145	496	739
2024	78	86	350	540

- a Use the data in the table to determine seasonal indices. Give your answers rounded to two decimal places.
- b Fit a least squares line to the deseasonalised data.
- c Use this to predict the number of dolphins that will be observed each season in 2026. Round your answers to the nearest whole number.

- 12** The table below shows the annual inflation rates in Australia and China for the period 2010–20.

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Inflation Australia (%)	2.9	3.3	1.7	2.5	2.5	1.5	1.3	2.0	1.9	1.6	0.9
Inflation China (%)	3.3	5.4	2.6	2.6	2.0	1.4	2.0	1.6	2.1	2.9	2.4

These data are plotted in the time series plot shown.



- a** **i** Determine the equation of the least squares line which allows *inflation* to be predicted from *year* for China.
- ii** Draw the least squares line on the time series plot.
- b** **i** Determine the equation of the least squares line which allows *inflation* to be predicted from *year* for Australia.
- ii** Draw the least squares line on the time series plot.
- c** Explain why the equations of the least squares lines predict that the inflation rate for China is likely to remain higher than the inflation rate for Australia.

Chapter 4

Arithmetic and geometric sequences

Chapter questions

UNIT 3 BIVARIATE DATA AND TIME SERIES ANALYSIS, SEQUENCES AND EARTH GEOMETRY

Topic 4: Growth and Decay in Sequences

- ▶ How do we identify a sequence?
- ▶ How do we generate a sequence?
- ▶ How do we define a recurrence relation in a sequence?
- ▶ How do we calculate a certain term in a sequence?
- ▶ How can we use sequences to model and analyse practical situations involving arithmetic growth and decay and geometric growth and decay?
- ▶ How do we use sequences to determine the value of items depreciating?
- ▶ How do we use sequences to determine the value of loans and investments?

In this chapter we explore arithmetic sequences and how they can be used to model and analyse practical situations involving linear growth and decay. This includes simple interest loans and investments, calculating taxi fares based on a flag fall and a charge per kilometre, as well as calculating depreciation.

We also consider geometric sequences and how they can be utilised to study practical situations. Examples involving growth and decay, such as changes in bacterial populations, as well as calculating depreciation using a diminishing-value method are analysed.

4A Using recursion to generate an arithmetic sequence

Learning intentions

- ▶ To generate an arithmetic sequence.
- ▶ To generate an arithmetic sequence using a calculator.

Sequences

A list of numbers, written down in succession, is called a **sequence**. Each of the numbers in a sequence is called a *term*. The terms of a sequence are written as a list, separated by commas. If a sequence continues indefinitely, or if there are too many terms in the sequence to write them all, we use an *ellipsis* ‘...’ like this:

12, 22, 5, 6, 16, 43, ...

Sometimes there is no pattern or rule that allows the next number in the sequence to be predicted, so we have a random sequence. Some sequences of numbers display a pattern. For example:

1, 3, 5, 7, 9, ...

This sequence of numbers has a starting value, 1. We add 2 to this number to get 3. Then, add 2 again to get 5, and so on. The rule is ‘add 2 to each term’.



Recursion

Recursion is the process of generating a sequence of terms from a given starting point and a rule. Rules such as adding or subtracting a number, or multiplying or dividing by a number, or even squaring numbers allow us to generate sequences recursively. For now, we will focus on rules that involve adding or subtracting a number.



Example 1 Looking for a recursive rule

Find the pattern in the sequence, 4, 10, 16, 22, ..., and hence determine the next term.

Solution

6, 6, 6

$$22 + 6 = 28$$

Explanation

Write down the difference between each term.

Since the difference is 6 for each case, 6 must be added to find the next term.

In general, a sequence can be generated from a starting value and a rule that tells us how to find the next term in the sequence.

**Example 2** Generating a sequence by recursion

Generate the first five terms of the sequence with a starting value of 6 and the rule ‘add 4’.

Solution

6

$$6 + 4 = 10,$$

$$10 + 4 = 14,$$

$$14 + 4 = 18,$$

$$18 + 4 = 22$$

The sequence is 6, 10, 14, 18, 22

Explanation

Write down the starting value.

Apply the rule to generate the next term.

Calculate three more terms.

Write your answer.

Generating a sequence of numbers using a calculator

A calculator can perform recursive calculations very easily because it automatically stores the answer to the last calculation it performed, as well as the method of calculation.

**Example 3** Generating sequences of numbers from a rule using a calculator

Use a calculator to generate the first five terms of the sequence with a starting value of 5 and the rule ‘double and then subtract 3’.

Solution

5

Ans $\times$ 2 – 3

Pressing ‘=’ or enter 1 more time

Ans $\times$ 2 – 3

Pressing ‘=’ or enter another time

Ans $\times$ 2 – 3

The sequence is 5, 7, 11, 19, 35, ...

Explanation

Press **clear** to create a blank computation screen.

Type 5 and then press = (Casio) or **enter** (TI). This stores the starting value in the calculator memory.

Type $\times 2 - 3$ and then press = (Casio) or **enter** (TI). The second term will be calculated and displayed on the screen.

Press = (Casio) or **enter** (TI) repeatedly to apply the rule to the previous calculated term.

Note: The screen will show the calculation as ‘Ans $\times$ 2 – 3’ where ‘Ans’ is the previously calculated term value.

Write down the sequence terms.

Section Summary

- ▶ A **sequence** is a list of numbers or symbols in a particular order.
- ▶ Each number in a sequence is called a **term**.
- ▶ **Recursion** is the process of generating a sequence of terms from a given starting point and a rule to calculate subsequent terms.

Exercise 4A

Identifying rules and finding missing terms

Example 1

- 1** State the starting value and the rule for each of the following sequences.

a 5, 10, 15, 20, ...	b 2, 4, 6, 8, ...
c 10, 9, 8, 7, ...	d 22, 19, 16, 13, ...
e 101, 105, 109, 113, 117, ...	f 87, 92, 97, 102, ...

- 2** State the next term in each sequence.

a 15, 20, 25, 30, ...	b 10, 12, 14, 16, ...
c 100, 90, 80, 70, ...	d 27, 23, 19, 15, ...
e 111, 115, 119, 123, ...	f 132, 121, 110, 89, ...

- 3** State the next two terms in each sequence.

a 7, 13, 19, 25, ...	b 3, -3, -9, -15, ...
c 24, 29, 34, 39, ...	d 37, 33, 29, 25, ...
e 21, 25, 29, 33, ...	f 124, 113, 102, 91, ...

Generating a sequence by recursion

Example 2

- 4** Use the following starting values and rules to generate the first five terms of the following sequences recursively by hand.

a Starting value: 2, rule: add 6	b Starting value: 5, rule: subtract 3
c Starting value: 10, rule: add 12	d Starting value: 56, rule: subtract 11

Generating a sequence by recursion using a calculator

Example 3

- 5** Use the following starting values and rules to generate the first five terms of the following sequences recursively using a calculator.

a Starting value: 4, rule: add 2	b Starting value: 50, rule: divide by 5
c Starting value: 24, rule: subtract 4	d Starting value: 5, rule: multiply by 3
e Starting value: 2, rule: multiply by 5	

Harder arithmetic sequence questions

- 6** Given the following two terms of an arithmetic sequence, determine the starting term.
- a** The second term is 5 and the third term is 8.
 - b** The second term is 12 and the third term is 10.
 - c** The third term is 15 and the fourth term is 19.
 - d** The third term is 40 and the fourth term is 32.
 - e** The third term is 2 and the fourth term is 8.
 - f** The second term is 8 and the fourth term is 20.
- 7** Given the term and the rule provided, state the value of the first five terms of the sequence.
- a** The second term is 8 and the rule is 'Add 7'.
 - b** The second term is 12 and the rule is 'Subtract 2'.
 - c** The third term is 25 and the rule is 'Add 9'.
 - d** The third term is 47 and the rule is 'Subtract 11'.
 - e** The fourth term is 121 and the rule is 'Add 17'.
 - f** The fourth term is 18 and the rule is 'Subtract 3'.

Paper 1-style multiple-choice questions

- 8** Determine which of the following could be the first five terms of an arithmetic sequence.
- A** 1, 5, 1, 5, 1
 - B** 2, 4, 8, 16, 32
 - C** -7, -2, 3, 8, 13
 - D** 11, 17, 23, 29, 34
- 9** Determine which of the following is not an arithmetic sequence.
- A** 12, 3, -7, -18, ...
 - B** 3, 10, 17, 24, ...
 - C** 57, 51, 45, 39, ...
 - D** -4, -6, -8, -10, ...
- 10** The first term of a sequence is 3. Each subsequent term is 8 more than the previous term. The seventh term in the sequence is
- A** 3
 - B** 11
 - C** 51
 - D** 59

4B Defining an arithmetic sequence by recursion

Learning intentions

- ▶ To be able to define an arithmetic sequence and identify the common difference.
- ▶ To be able to generate a sequence from a recurrence relation.
- ▶ To be able to tabulate and graph an arithmetic sequence.

Numbering and naming terms

The symbols $t_1, t_2, t_3, \dots$ are used to label and name terms in a sequence. The term t_1 is the first term in the sequence, the term t_2 is the second term in the sequence and so on, where the term t_n is the n th term in the sequence. The letter t can be replaced by any letter in the alphabet.

This notation helps us to describe how sequences can be generated using recursion.



Example 4 Naming terms in a sequence

For the sequence, 3, 9, 15, 21, 27, 33, 39, 45 ..., state the value of

a t_1

b t_3

c t_7

Solution

3, 9, 15, 21, 27, 33, 39, 45
 t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8

$t_1 = 3, t_3 = 15, t_7 = 39$

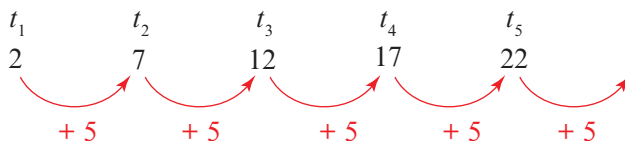
Explanation

Write the name for each term in the sequence under its value in the sequence.

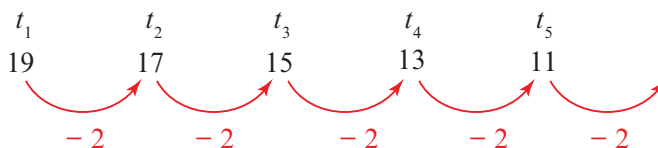
Read the value of each required term.

Arithmetic sequences and the common difference

A sequence that can be generated by adding the same number to the previous term is called an **arithmetic sequence**. For example, the sequence 2, 7, 12, 17, 22, ... is arithmetic because each successive term can be found by adding 5.



The sequence 19, 17, 15, 13, 11, ... is also arithmetic because each successive term can be found by adding -2 .



The fixed amount that we add or subtract to form an arithmetic sequence recursively is called the **common difference**. The symbol d is often used to represent the common difference.

If the sequence is known to be arithmetic, the common difference is found by calculating the difference between any pair of successive terms.

Common difference, d

In an **arithmetic sequence**, the fixed number added to (or subtracted from) each term to make the next term is called the **common difference**, d , where:

$$d = \text{any term} - \text{previous term}$$

For example, the common difference for the arithmetic sequence: 2, 7, 12, 17, ... is $d = 7 - 2 = 5$.



Example 5 Finding the common difference

Find the common difference in the following arithmetic sequences and use it to find the next term in each of the sequences below.

a 26, 24, 22, ...

b 1, 4, 7, ...

Solution

a $d = t_2 - t_1 = 24 - 26 = -2$
 $t_4 = t_3 + d = 22 + (-2) = 20$

b $d = t_2 - t_1 = 4 - 1 = 3$
 $t_4 = t_3 + d = 7 + 3 = 10$

Explanation

Because we know that the sequence is arithmetic, all we need to do is find the difference between t_1 and t_2 . To find the next term, t_4 , add the common difference to t_3 .

If the sequence is arithmetic, the difference between successive terms will be constant. We can use this to check if a sequence is arithmetic.


Example 6 Identifying an arithmetic sequence

Determine if either of the following sequences are arithmetic.

a 2, 8, 14, 22, ...

b 81, 74, 67, 60, ...

Solution

a 2, 8, 14, 22, ...

Differences:

$$8 - 2 = 6$$

$$14 - 8 = 6$$

$$22 - 14 = 8$$

As the difference between successive terms is not constant, the sequence is not arithmetic.

b 81, 74, 67, 60, ...

Differences:

$$74 - 81 = -7$$

$$67 - 74 = -7$$

$$60 - 67 = -7$$

As the difference between successive terms is constant, the sequence is arithmetic.

Explanation

Determine whether the difference between each successive pair of terms is constant. Write your conclusion.

Determine whether the difference between each successive pair of terms is constant. Write your conclusion.

Method of recursion to generate an arithmetic sequence

Method for using recursion to generate an arithmetic sequence

The method for using recursion to generate an arithmetic sequence has two parts:

- 1** A starting point: the value of the first term t_1 (or a) of the sequence.
- 2** A rule: add the common difference d to each term to obtain the next term

For example, in words, a rule for the recursion that can be used to generate the sequence:

$$10, 15, 20, 25, \dots$$

can be written as follows:

- 1** Start with 10.
- 2** To obtain the next term, add 5 to the current term.
- 3** Repeat the process.

This can be communicated in symbolic form using our notation developed earlier.

$$t_1 = 10, \quad t_{n+1} = t_n + 5$$

In general, a **recurrence relation** is written as:

$$t_1 = a \text{ and } t_{n+1} = t_n + d, \text{ where } a \text{ is our starting value and } d \text{ is the common difference.}$$

We can either generate the sequence by hand or use a calculator as explained in Exercise 4A.



Example 7 Generating an arithmetic sequence with a recurrence relation

An arithmetic sequence is defined by the recurrence relation:

$$t_1 = 9, t_{n+1} = t_n - 4$$

Find the first five terms.

Solution

$$t_1 = 9$$

$$t_2 = t_1 - 4 = 9 - 4 = 5$$

$$t_3 = t_2 - 4 = 5 - 4 = 1$$

$$t_4 = t_3 - 4 = 1 - 4 = -3$$

$$t_5 = t_4 - 4 = -3 - 4 = -7$$

Explanation

We start with the first term $t_1 = 9$.

Subtract 4 each time to generate the sequence.

Tables and graphs of arithmetic sequences

The terms of an arithmetic sequence can be tabulated, demonstrating that each input (n) has an associated value (t_n). This helps us to see that a sequence can be thought of as a **function**. The two cases below show that an arithmetic sequence can be increasing or decreasing.

Consider the arithmetic sequence defined as follows:

$$P_1 = 22, P_{n+1} = P_n + 2$$

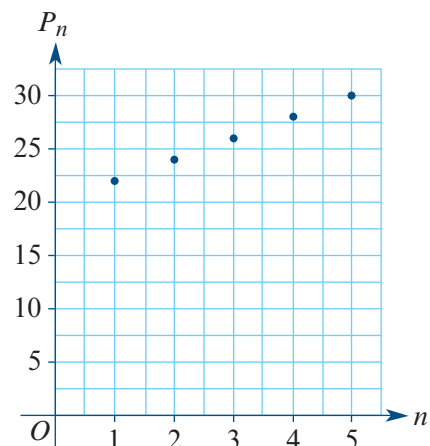
The first five values of the sequence can be shown in a table.

n	1	2	3	4	5
P_n	$P_1 = 22$	$P_2 = 24$	$P_3 = 26$	$P_4 = 28$	$P_5 = 30$

This gives the ordered pairs:

(1, 22), (2, 24), (3, 26), (4, 28), (5, 30).

These can be graphed, as shown opposite, illustrating that the sequence is increasing.



Consider the second arithmetic sequence, which is decreasing, defined as follows:

$$Q_1 = 18, Q_{n+1} = Q_n - 2$$

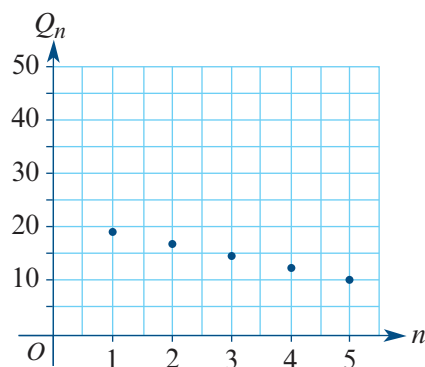
The first five values of the sequence can be shown in a table.

n	1	2	3	4	5
Q_n	$Q_1 = 18$	$Q_2 = 16$	$Q_3 = 14$	$Q_4 = 12$	$Q_5 = 10$

This gives the ordered pairs:

$(1, 18), (2, 16), (3, 14), (4, 12), (5, 10)$.

These can be graphed as shown opposite, illustrating that the sequence is increasing.



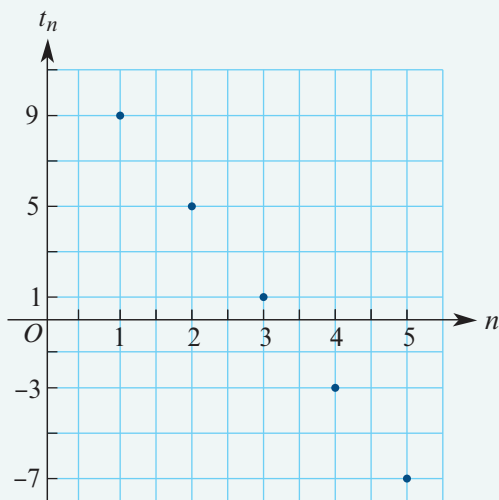
Example 8 Graphing arithmetic sequences

Prepare a table of values and plot a graph of the first five terms of the sequence defined by

$$t_1 = 9, t_{n+1} = t_n - 4$$

Solution

n	1	2	3	4	5
t_n	9	5	1	-3	-7



Explanation

Generate a table from the recurrence relation by subtracting 4 from a term to give the next term.

Plot the ordered pairs:

$(1, 9), (2, 5), (3, 1), (4, -3)$ and $(5, -7)$.

Linear growth and decay

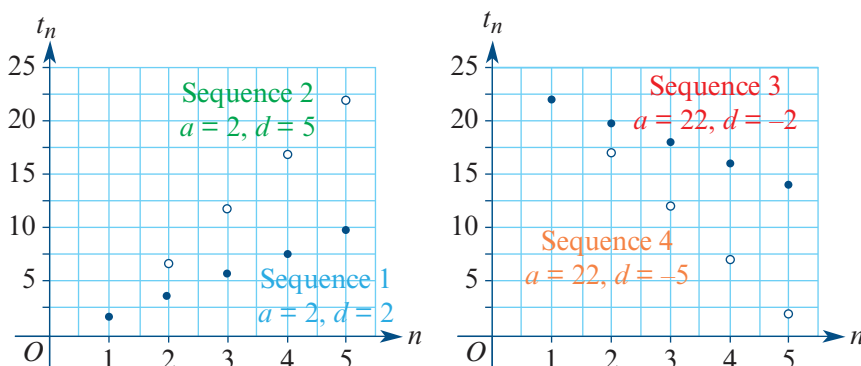
As well as considering whether graphs increase or decrease, we can also think about the rate at which they increase or decrease. The following sequences are graphed below to illustrate this point.

Sequence 1: $t_1 = a = 2, d = 2, n = 1, 2, 3, \dots$

Sequence 2: $t_1 = a = 2, d = 5, n = 1, 2, 3, \dots$

Sequence 3: $t_1 = a = 22, d = -2, n = 1, 2, 3, \dots$

Sequence 4: $t_1 = a = 22, d = -5, n = 1, 2, 3, \dots$



- The points in each graph are *collinear*; that is, the points lie on a *straight line*.
- If the common difference, d , is positive, the terms in the sequence increase. The bigger the value of d , the more rapid the increase. An arithmetic sequence with a positive common difference can be used to model **linear growth**.
- If the common difference, d , is negative, the terms in the sequence decrease. The bigger the absolute value of d , the more rapid the decrease. An arithmetic sequence with a negative common difference can be used to model **linear decay**.

Section Summary

- ▶ An **arithmetic sequence** is generated by adding the same number to the current term to generate the next term. This number is called the **common difference**, d . That is, the common difference is the difference between any two consecutive terms.
- ▶ Terms in a sequence are numbered as t_1, t_2, t_3 and so on.
- ▶ A **recurrence relation** for an arithmetic sequence is written as: $t_1 = a, t_{n+1} = t_n + d$ where a is the starting value.
- ▶ Sequences can be displayed in a table or a graph.
- ▶ An arithmetic sequence with a positive common difference exhibits **linear growth**.
- ▶ An arithmetic sequence with a negative common difference exhibits **linear decay**.

Exercise 4B

Naming terms in a sequence

Example 4

- 1** State the starting value of the sequence 18, 16, 14, 12, ...
- 2** Consider the following sequence: 8, 12, 16, 20, 24, ...
 - a** Copy the sequence into your book.
 - b** Write the name of each term below each term of the sequence.
- 3** Consider the sequence: 15, 18, 21, 24, 27, ... State the value of the following terms:

a t_1	b t_4	c t_5
- 4** Consider the sequence: 39, 33, 27, 21, 15, ... State the value of the following terms:

a t_3	b t_2	c t_6

Identifying arithmetic sequences and the common difference

Example 5

- 5** For each of the following sequences, find the difference between t_1 and t_2 and the difference between t_2 and t_3 .

a 6, 9, 12, 15, ...	b 19, 25, 31, 37, 43, ...
c 99, 95, 91, 87, 83, ...	d 101, 90, 79, 68, 57, ...

Example 6

- 6** Consider the following sequence:
4, 11, 18, 25, 32, ...
 - a** Calculate the common difference between each pair of consecutive terms, and decide if the sequence is arithmetic.
 - b** If the sequence is arithmetic, determine the value of t_6 and t_7 .
- 7** Consider the following sequence:
16, 8, 4, 2, 1, ...
 - a** Calculate the common difference between each pair of consecutive terms, and decide if the sequence is arithmetic.
 - b** If the sequence is arithmetic, determine the value of t_6 and t_7 .
- 8** For each of these arithmetic sequences, find the common difference and the value of t_5 .

a 3, 8, 13, 18, ...	b 19, 15, 11, 7, 3, ...
c 188, 181, 174, 167, ...	d 213, 311, 409, 507, ...

Generating an arithmetic sequence with a recurrence relation

Example 7

- 9** An arithmetic sequence is defined by:

$$t_1 = 3, t_{n+1} = t_n + 4$$

Find the first five terms.

- 10** An arithmetic sequence is defined by:

$$t_1 = 15, t_{n+1} = t_n - 4$$

Find the first five terms.

- 11** Consider the arithmetic sequence 20, 24, 28, 32, 36, ...

- a** State the common difference.
- b** State the next term in the sequence.
- c** Starting with 20, find the number of times the common difference must be added to obtain term 8. State the value of term 8.
- d** Starting with 20, find the number of times the common difference must be added to obtain term 13. State the value of term 13.

- 12** Consider the arithmetic sequence 5, 3, 1, -1, -3, ...

- a** State the common difference.
- b** State the next term in the sequence.
- c** Starting with 5, find the number of times the common difference must be added to obtain term 7. Find the value of term 7.
- d** Determine the value of term 10.
- e** Find the value of term 50.

Graphing arithmetic sequences

Example 8

- 13** Prepare a table of values and plot a graph to show the first five terms of the sequence defined by:

$$t_1 = 9, t_{n+1} = t_n - 4$$

- 14** Prepare a table of values and plot a graph to show the first five terms of the sequence defined by:

$$t_1 = 10, t_{n+1} = t_n + 5$$

- 15** Prepare a table of values and plot a graph to show the first five terms of the sequence defined by:

$$t_1 = 12, t_{n+1} = t_n - 2$$

Applying recurrence rules and understanding

- 16** Write the following as a recurrence relation in symbolic form, where t_n represents the value.

- a** The starting value is 7, and the rule is ‘add 3 to the current value and repeat the process.’
- b** The starting value is 19, and the rule is ‘add 15 to the current value and repeat the process.’
- c** The starting value is 62, and the rule is ‘subtract 8 from the current value and repeat the process.’

- 17** State the recurrence relation in symbolic form for the following sequences.

a $19, 23, 27, 31, 35, \dots$

b $26, 20, 14, 8, 2, \dots$

c $11, 20, 29, 38, 47, \dots$

d $4, -7, -18, -29, \dots$

- 18** The following recurrence relation can generate a sequence of numbers.

$$t_1 = 20, \quad t_{n+1} = t_n + 3$$

State the term name for the term that has a value of 44.

- 19** A restaurant starts with 370 wine glasses. After one week, they only had 358 wine glasses. If the restaurant has the same breakage rate every week, construct a recurrence relation in symbolic form and work out how long it will be until all wine glasses are broken.

- 20** Consider the following sequence:

$$-2, -5, -8, -11, \dots$$

a Determine t_{100} .

b Determine t_{200} .

Paper 1-style multiple-choice questions

- 21** The tenth term of the arithmetic sequence $44, 41, 38, \dots$ is:

A 8

B 17

C 20

D 23

- 22** The sequence generated by the recurrence relation $t_0 = 65, t_{n+1} = t_n - 9$ is:

A $65, 74, 83, 92, 101, \dots$

B $65, 54, 42, 31, 20, \dots$

C $65, 56, 47, 38, 29, \dots$

D $65, 76, 88, 99, 110, \dots$

- 23** Ollie has 10 collector's cards. Every month, he buys 4 more to add to his collection. A recurrence relation model, C_n , for the number of cards in Ollie's collection at the start of month n , is:

A $C_1 = 10, C_{n+1} = 4T_n$

B $C_1 = 10, C_{n+1} = 4T_n + 4$

C $C_1 = 10, C_{n+1} = T_n + 4$

D $C_1 = 10, C_{n+1} = T_n - 4$

4C A general rule for finding the n th term of an arithmetic sequence

Learning intentions

- ▶ To use a rule for an arithmetic sequence with a positive or negative difference.
- ▶ To determine an arithmetic sequence given two terms.
- ▶ To determine the number of terms required to reach a certain value.
- ▶ To graph an arithmetic sequence using a rule.

The rule for the n th term of an arithmetic sequence

Consider an arithmetic sequence with first term t_1 and common difference d . Then:

$$t_1 = t_1$$

$$t_2 = t_1 + d$$

$$t_3 = t_2 + d = t_1 + 2d$$

$$t_4 = t_3 + d = t_1 + 3d \text{ and so on.}$$

The rule for the n th term of an arithmetic sequence

The n th term of an arithmetic sequence in terms of the first term, t_1 , and the common difference, d , is given as: $t_n = t_1 + (n - 1)d$



Example 9 Using the rule for an arithmetic sequence (positive difference)

The first term of an arithmetic sequence is $t_1 = 6$ and the common difference is $d = 2$. Use a rule to determine the value of the 11th term.

Solution

$$\begin{aligned} t_{11} &= 6 + (11 - 1) \times 2 \\ &= 6 + 10 \times 2 \\ &= 26 \end{aligned}$$

Explanation

Substitute $t_1 = 6$, $d = 2$ and $n = 11$ in the rule $t_n = t_1 + (n - 1)d$.



Example 10 Using the rule for an arithmetic sequence (negative difference)

Use a rule to determine the value of the 15th term of the sequence 18, 15, 12, 9, ...

Solution

$$\begin{aligned} t_{15} &= 18 + (15 - 1)(-3) \\ &= 18 + 14 \times (-3) \\ &= -24 \end{aligned}$$

Explanation

Substitute $t_1 = 18$, $d = -3$ and $n = 15$ in the rule $t_n = t_1 + (n - 1)d$.


Example 11 Determining an arithmetic sequence given two terms

In an arithmetic sequence, the fifth term is 10 and the ninth term is 18.

Write down the first three terms of the sequence.

Solution

$$10 = t_1 + (5 - 1)d$$

$$18 = t_1 + (9 - 1)d$$

$$10 = t_1 + 4d \quad (1)$$

$$18 = t_1 + 8d \quad (2)$$

$$8 = 4d$$

$$\text{so } d = 2$$

$$10 = t_1 + 4 \times 2$$

$$= t_1 + 8$$

$$\therefore t_1 = 2$$

$$t_1 = 2$$

$$t_2 = t_1 + d = 2 + 2 = 4$$

$$t_3 = t_2 + 2d = 2 + 2 \times 2 = 6$$

Thus, the first three terms of the sequence are 2, 4, 6.

Explanation

Substitute $t_5 = 10$ and $n = 5$ into

$$t_n = t_1 + (n - 1)d.$$

Substitute $t_9 = 18$ and $n = 9$ into

$$t_n = t_1 + (n - 1)d.$$

We can now find t_1 and d by subtracting equation (1) from equation (2).

Substitute $d = 2$ in equation (1).

Use t_1 and d to compute the first three terms.


Example 12 Determining how many terms of an arithmetic sequence are required to reach a particular number

Consider the following sequence: 10, 14, 18, 22, ...

Determine the number of terms in the sequence until a term greater than 50 is reached.

Solution

Solving for n we have:

$$10 + (n - 1) \times 4 > 50$$

$$(n - 1) \times 4 > 40$$

$$(n - 1) > 10$$

$$n > 11$$

The first term to exceed 50 is t_{12} .

Thus, we need 12 terms in the sequence before a term greater than 50 is reached.

Explanation

We want to find n so that: $t_n = 10 + (n - 1) \times 4 > 50$.

Since $t_{11} = 50$, the first term to exceed 50 is t_{12} .

Graphing an arithmetic sequence

Using the rule for finding the n th term of an arithmetic sequence helps us understand a little more about the graphs of arithmetic sequences.

When $t_1 = 4$ and $d = -3$,

$$\begin{aligned}t_n &= 4 + (n - 1)(-3) \\&= -3n + 7\end{aligned}$$

When we plot the graph of t_n against n , the points lie on a straight line. This line has gradient -3 . The points on the line are not joined together as only whole number values of n make sense.



Example 13 Graphing an arithmetic sequence using a rule

Consider the arithmetic sequence with $t_1 = 2$ and $d = 3$.

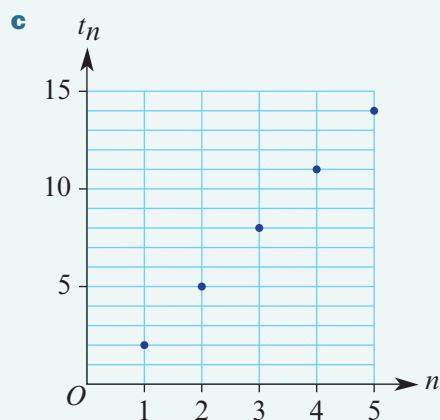
- Find the rule for this arithmetic sequence.
- Prepare a table of values for the sequence for $n = 1$ to $n = 5$.
- Plot a graph from the table of values.

Solution

a $t_n = 2 + (n - 1) \times 3$ for $n = 1, 2, 3 \dots$
 $= 3n - 1$

b

n	1	2	3	4	5
t_n	2	5	8	11	14



Explanation

Substitute $t_1 = 2$ and $d = 3$ into
 $t_n = t_1 + (n - 1)d$ and simplify.

Use the rule to generate the values of the table.

For example, $t_4 = 3 \times 4 - 1 = 11$

Plot the ordered pairs $(1, 2)$, $(2, 5)$, $\dots$

Section Summary

- ▶ The rule for the n th term of an arithmetic sequence is given as $t_n = t_1 + (n - 1)d$.
- ▶ The gradient of the line generated by the arithmetic sequence is the common difference of the sequence.

Exercise 4C

Using the rule for an arithmetic sequence

Example 9

- 1** The first term of an arithmetic sequence is $t_1 = 200$ and the common difference is $d = 10$. Use this rule to determine the value of:
- a** the third term
 - b** the 10th term
 - c** the 151st term.

Example 10

- 2** Use the rule for the n th term of an arithmetic sequence to determine the value of:
- a** the 11th term of the arithmetic sequence 5, 10, 15, 20, 25, ...
 - b** the 8th term of the arithmetic sequence 12, 8, 4, 0, ...
 - c** the 27th term of the arithmetic sequence 0.1, 0.11, 0.12, 0.13, ...
 - d** the 13th term of the arithmetic sequence $-55, -42, -29, -16, \dots$
 - e** the 10th term of the arithmetic sequence $-1.0, -1.5, -2.0, -2.5, \dots$
 - f** the 95th term of the arithmetic sequence 130, 123, 116, 109, ...
 - g** the 7th term of the arithmetic sequence $\frac{1}{2}, \frac{1}{4}, 0, -\frac{1}{4}, \dots$
- 3** Write down a rule for the n th term for each of these recurrence relations.
- | | |
|--|---|
| a $t_1 = 6, t_{n+1} = t_n + 7$ | b $t_1 = 48, t_{n+1} = t_n - 11$ |
| c $t_1 = 8, t_{n+1} = t_n + 11$ | d $t_1 = 1000, t_{n+1} = t_n - 20$ |

Determining an arithmetic sequence given two terms

Example 11

- 4** The first term of an arithmetic sequence is 6 and $t_8 = 34$. Determine the common difference.
- 5** In an arithmetic sequence, the common difference is 20 and $t_{10} = 188$. Determine t_1 .
- 6** In an arithmetic sequence, $t_2 = 15$ and $t_3 = 19$. Determine the first term.
- 7** The common difference for an arithmetic sequence is -8 and $t_6 = 30$. Determine t_1 .
- 8** Write down the first three terms (starting with t_1) of the arithmetic sequences in which:
- a** the seventh term is 37 and the ninth term is 47
 - b** the 11th term is 31 and the 15th term is 43
 - c** the 6th term is 0 and the 11th term is -20
 - d** the 8th term is 134 and the 13th term is 159
 - e** the 6th term is 0 and the 10th term is -16
 - f** the 7th term is 60 and the 12th term is 10
 - g** the 10th term is 20 and the 21st term is 75.

Determining how many terms of an arithmetic sequence are required to reach a particular number

Example 12

9 How many terms would we have to write down in the arithmetic sequence:

- a** 5, 7, 9, 11, ... before we found a term greater than 25?
- b** 132, 182, 232, 282, ... before we found a term greater than 1000?
- c** 100, 96, 92, 88, ... before we found a term less than 71?
- d** 10, 16, 22, 28, ... before we found a term equal to 52?
- e** 0.33, 0.66, 0.99, 1.32, ... before we found a term greater than 2?
- f** -17, -15, -13, -11, ... before we found a positive term?
- g** 127, 122, 117, 112, ... before we found a term less than zero?

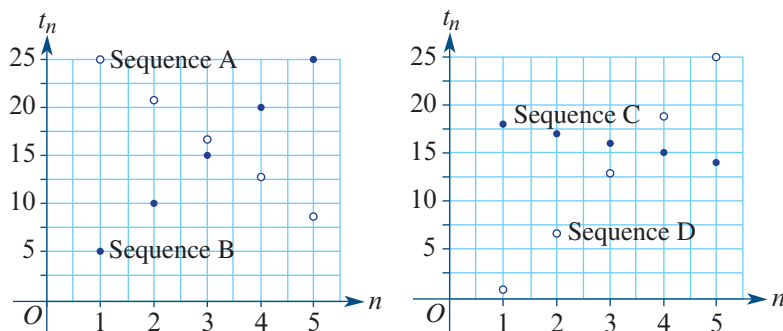
Graphing an arithmetic sequence using a rule

Example 13

10 Plot the first five terms of each of the following arithmetic sequences:

- a** Sequence A: $t_1 = 3, d = 2, n = 1, 2, 3, \dots$
- b** Sequence B: $t_1 = 12, d = -2, n = 1, 2, 3, \dots$
- c** Sequence C: $t_1 = 0, d = 4, n = 1, 2, 3, \dots$
- d** Sequence D: $t_1 = -6, d = 3, n = 1, 2, 3, \dots$
- e** Sequence E: $t_1 = 10, d = -5, n = 1, 2, 3, \dots$

11 Four sequences are displayed in the following graphs. For each sequence determine the value of the first term, t_1 , the common difference, d , and an expression for the n th term, t_n , in terms of n .



12 In an arithmetic sequence, the fifth term is 10 and the ninth term is 18.

- a** Write down the rule for the arithmetic sequence.
- b** Write down the first 4 terms of the sequence.

13 In an arithmetic sequence, the second term is 17 and the ninth term is 59.

- a** Write down the rule for the arithmetic sequence.
- b** Write down the first 4 terms of the sequence.

14 In an arithmetic sequence, the eighth term is 35 and the twelfth term is 23.

- a** Write down the rule for the arithmetic sequence.
- b** Write down the first 4 terms of the sequence.

- 15** Consider the arithmetic sequence with $t_1 = 5$ and $d = 2$.
- Find the rule for this arithmetic sequence.
 - Prepare a table of values for the sequence from $n = 1$ to $n = 5$.
 - Plot a graph of the sequence from the table of values.
- 16** Consider the arithmetic sequence with $t_1 = 20$ and $d = -4$.
- Find the rule for this arithmetic sequence.
 - Prepare a table of values for the sequence from $n = 1$ to $n = 5$.
 - Plot a graph of the sequence from the table of values.
- 17** Consider the arithmetic sequence with $t_1 = 10$ and $d = 2$.
- Find the rule for this arithmetic sequence.
 - Prepare a table of values for the sequence from $n = 1$ to $n = 5$.
 - Plot a graph of the sequence from the table of values.
- 18** An arithmetic sequence has rule $t_n = 5n + 2$. Find t_1 and the common difference d .
- 19** An arithmetic sequence has rule $t_n = 4 - 2n$. Find t_1 and the common difference d .
- 20** For each of the following arithmetic sequences, find t_1 and the common difference, then describe the arithmetic sequence using a recurrence relation.
- | | |
|--------------------------|--------------------------|
| a $t_n = 7n + 11$ | b $t_n = 8 - 5n$ |
| c $t_n = 12 - 3n$ | d $t_n = 41 + 6n$ |

Paper 1-style multiple-choice questions

- 21** A sequence is generated from the recurrence relation $V_1 = 20$, $V_{n+1} = V_n + 6$. The rule for the value of the term V_n is:
- A** $V_n = 20n + 6$ **B** $V_n = 14 + 6n$ **C** $V_n = 20n$ **D** $V_n = 20 + 6n$
- 22** The second and fifth terms of a sequence are 8 and 26, respectively. The rule for the n th term is:
- A** $t_n = 8 + 6n$ **B** $t_n = 2 + 8n$
C $t_n = 6n - 4$ **D** $t_n = 2 + 6n$
- 23** The 15th and 19th terms of an arithmetic sequence are 71 and 91, respectively. The rule for the n th term is:
- A** $t_n = 6 + 5(n - 2)$ **B** $t_n = 6 + 5(n - 1)$
C $t_n = 1 + 5(n + 1)$ **D** $t_n = 1 + 5(n - 1)$

4D Application of arithmetic sequences

Learning intentions

- ▶ To model an investment or loan with simple interest using an arithmetic sequence.
- ▶ To model a taxi fare using an arithmetic sequence.
- ▶ To model flat-rate and unit-cost depreciation using an arithmetic sequence.
- ▶ To model general situations with arithmetic sequences.

Many practical situations involving linear growth and decay can be modelled using arithmetic sequences.

Simple interest loans and investments

Simple interest is an interest charge that borrowers pay for a loan (or investors earn from an investment) based on the initial amount (the principal) and the interest rate. An arithmetic sequence is formed using the principal (starting value) and the common difference (interest).

Calculation of simple interest

Let P be the amount borrowed or invested (the principal), A be the value of the investment after n years and i be the percentage interest rate.

Then $d = i \times P$ is the amount of interest paid per year and is the common difference.

The value, A , of the simple interest investment after n years is

$$A = P + nd$$



Example 14 Modelling and analysing an investment with simple interest using arithmetic sequences

David invests \$20 000 into a bank account. He will be paid simple interest at the rate of 5% of the investment per annum (per year).

- a Find the expression for A , the value of the investment after n years.
- b Find the value of the investment after 5 years.
- c If David leaves the money in the account, after how many years will the investment be worth \$30 000?

Solution

a $d = 5\%$ of 20 000 = 1000

$$A = 20\,000 + n \times 1000$$

b $A = 20\,000 + 5 \times 1000$

$$= 25\,000$$

The investment will be worth \$25 000 after five years.

Explanation

Calculate the interest, $d = i \times P$. Substitute d and P into the rule.

Substitute $n = 5$ into the rule to find the value after 5 years and evaluate. Write your answer in a sentence.

$$\text{c } 20\,000 + n \times 1000 = 30\,000$$

$$n \times 1000 = 10\,000$$

$$n = 10$$

The investment will be worth \$30 000 after 10 years.

Form an equation and solve for n . Write your answer in a sentence.

Simple interest is an example of linear growth. Note that the formula we are using starts at $n = 0$ rather than $n = 1$. We should interpret the initial value being where $n = 0$ and that the value of n refers to the number of times that simple interest is added.

Modelling a taxi fare using arithmetic sequences

The amount that a taxi charges for a trip can be broken down into a flag fall (flat fee) and a charge based on the distance travelled. An arithmetic sequence is formed using the flag fall (initial amount) and the price per kilometre (common difference).



Example 15 Modelling a taxi fare

Trevor's taxi service charges a flag fall of \$4.30 and then \$2.17 per kilometre travelled.

a Find the expression for C , the cost of a taxi trip covering x kilometres.

b Find the cost of travelling 8 kilometres.

Solution

a $C = 4.30 + 2.17x$

b $C = 4.30 + 2.17 \times 8$
 $= \$21.66$

It costs \$21.66 to travel 8 kilometres.

Explanation

Substitute $P = 4.30$ and $d = 2.17$ into the rule for C .

Substitute $x = 8$ into the rule and evaluate.

Write your answer in a sentence.

In the example above, we have defined C as the cost of a taxi trip **after** x kilometres, so the rule takes the form $C = \text{flag fall} + \text{cost per kilometre} \times x$.

Depreciation

Over time, the value of large items will gradually decrease. For example, a car bought new this year will not be worth the same amount of money in five years. **Flat-rate depreciation** and **unit-cost depreciation** are two methods that can be modelled using arithmetic sequences. They are both examples of linear decay.

Flat-rate depreciation

Flat-rate depreciation is very similar to simple interest, but instead of adding a constant amount of interest, a constant amount is subtracted to reduce the value of the asset after every time period. This constant amount is called the **depreciation** amount, and, like simple interest, it is often given as a percentage of the initial purchase price of the asset.

Calculation of flat-rate depreciation

Let P be the initial value of the asset, A be the value of the asset after n years and i be the percentage depreciation rate.

Then $d = i \times P$ is the amount of depreciation per year and is the common difference.

The value, A , of the asset after n years is

$$A = P - n \times d$$

**Example 16** Modelling flat-rate depreciation using arithmetic sequences

The value of a machine is flat-rate depreciated in value by 4% of its initial value every year. Initially it was valued at \$100 000.

- a** Find an expression for A , the value of the machine after n years.
- b** Find the value of the machine after 5 years.

Solution

$$\begin{aligned}\mathbf{a} \quad d &= 4\% \text{ of } 100\,000 = 4000 \\ A &= 100\,000 - n \times 4000\end{aligned}$$

$$\begin{aligned}\mathbf{b} \quad A &= 100\,000 - 5 \times 4000 \\ &= 80\,000\end{aligned}$$

The machine will be worth \$80 000 after five years.

Explanation

Calculate the depreciation, $d = i \times P$.
Substitute into the rule for P and d .

Substitute $n = 5$ into the rule and evaluate.

Write your answer in a sentence.

**Example 17** Modelling flat-rate depreciation using arithmetic sequences

A new car was purchased for \$26 000 in 2014 and depreciates by 4.5% of the original amount each year. The value of the car, A , after n years is given as $A = 26\,000 - n \times 1170$.

Find the year when the value of the car is first expected to be less than \$13 000.

Solution

$$\begin{aligned}26\,000 - n \times 1170 &< 13\,000 \\ 26\,000 - 13\,000 &< n \times 1170 \\ n \times 1170 &> 13\,000 \\ n &> 11.11 \dots\end{aligned}$$

After 12 years, the value of the car will first be less \$13 000.

Explanation

Set up the inequality to solve.
Solve the inequality.

Write your answer in a sentence.

Unit-cost depreciation

Some items lose value because of how often they are used, rather than because of their age. A photocopier that is two years old but has never been used could be considered to be in 'brand new' condition and therefore worth the same as, or close to, what it was two years ago. But, if the same two-year-old photocopier had printed thousands of copies over those two years, it would be worth much less than its original value.

Similarly, cars can also be depreciated according to their use rather than time. When buying a secondhand car, people often consider the number of kilometres that the car has travelled. A secondhand car that has travelled fewer kilometres could be considered a better buy than a newer car that has travelled a large distance.

When the future value of an item is determined based upon use rather than age, we use a unit-cost depreciation method. Unit-cost depreciation can be modelled using a linear decay recurrence relation.

Calculation of unit-cost depreciation

Let P be the initial value of the asset, A be the value of the asset after n uses and d be the depreciation per use.

The value of the asset after n uses is:

$$A = P - nd$$



Example 18 Modelling unit-cost depreciation using arithmetic sequences

A lawn mower was purchased for \$270. Every time it is used to mow a lawn, the owner estimates a depreciation in value of 50 cents in the mower's worth.

- a** Write down the rule for the value of the asset, A , after n mows.
- b** Find the value of the lawn mower after it has mowed 25 lawns.

Solution

a $A = 270 - n \times 0.5$

b $A = 270 - 25 \times 0.5$
 $= 257.50$

The value of the lawn mower is
 \$257.50 after 25 mows.

Explanation

Substitute $P = 270$, $d = 0.50$ into the rule for A .

Substitute $n = 25$ to find the value after 25 mows and evaluate. Write the answer in a sentence.

Other applications of arithmetic sequences



Example 19 Solving other problems using arithmetic sequences

Before starting on a weight-loss program, a man weighs 124 kg. Using a combination of diet and exercise, he plans to lose weight at a rate of 1.5 kg a week until he reaches his recommended weight of 94 kg.

How many weeks will he take to reduce his weight to 94 kg?

Solution

$$w = 124.0 - n \times 1.5$$

$$94 = 124.0 - n \times 1.5$$

$$-30 = -n \times 1.5$$

$$n = 20$$

The man can expect his weight to be down to 94 kg after 20 weeks.

Explanation

Write down a rule for the man's weight, w , after n weeks using his initial weight of 124 kg and his planned loss of 1.5 kg per week.

Use the rule to determine how many weeks he must diet and exercise to reach his recommended weight of 94 kg.

Write your answer as a sentence.

Section Summary

- ▶ **Simple interest loans** and **simple interest investments** can be modelled using arithmetic sequences. The initial amount that is borrowed or invested is called the **principal**, P . The amount of interest that is paid or earned is calculated as $d = i \times P$ where i is the interest rate. The value of the loan or investment, A , after n years is $A = P + n \times d$.
- ▶ A **taxi fare** can be calculated by viewing the flag fall as the starting value and the cost per kilometre as the common difference.
- ▶ **Depreciation** is the amount the value of an item decreases over a period of time.
- ▶ **Flat-rate depreciation** is where a constant amount is subtracted from the value of the item at a regular time interval. It is usually based on a percentage of the initial value and is an example of linear decay. The value of the asset, A , after n years is given as $A = P - n \times d$, where P is the initial value and d is amount of depreciation each year.
- ▶ **Unit-cost depreciation** is calculated based on units of use rather than time. It is an example of linear decay. The value of the asset, A , after n uses is given as $A = P - n \times d$, where P is the initial value and d is amount that the asset depreciates after each use.



Exercise 4D

Calculating simple interest

Example 14

- 1** Suppose you invest \$5000 with simple interest of 7.5% per annum.

- a** State the amount of interest that is paid to you each year.
b Complete the following table:

Start of year	1	2	3	4	5
Value (\$)					

- c** Write down an expression for the value of the investment, A , after n years.
d Determine the value of the investment:
i at the end of the 15th year
ii after 25 years.

- 2** Suppose you borrow \$50 000 with simple interest rate of 9% per annum, for a period of five years.

- a** State the amount of interest that is charged each year.
b Complete the following table:

Start of year	1	2	3	4	5
Amount (\$)					

- c** Write down an expression for the amount you owe at the end of the n th year.
d If you decided to extend the loan, find how much would you owe:
i at the end of the 15th year.
ii at the end of the 25th year.

- 3** Monica invests \$60 000 into a bank account. She will be paid simple interest at the rate of 4.5% per annum.

- a** Find the expression for A , the value of the investment after n years.
b Find the value of the investment after 5 years.
c If Monica continues to invest the money, determine how many years it will take for the investment be worth more than \$80 000.

- 4** Two thousand dollars is invested at an interest rate of 3.8% per annum.

- a** Write down the rule for the value A for the simple interest investment at the start of the n th year.
b Use the rule to find the value of the investment after 6 years.
c Determine how many years it takes for the value of the investment to be more than \$3000.

- 5** A simple interest investment of \$7000 has an interest rate of 7.4% per annum.
- a** Write down the rule for the value of the simple interest investment A_n in terms of n .
 - b** Use this rule to find the value of the investment after 6 years.
 - c** Determine how many years it takes for the value of the investment to be more than \$10 000.

Calculating a taxi fare

Example 15

- 6** Andy's airport transfer service charges a flag fall of \$2.90 and then \$1.94 per kilometre travelled.
- a** Find an expression for $\$C$, the cost of travelling k kilometres.
 - b** Find the cost of travelling 17 kilometres.

- 7** Red cabs charge a flag fall of \$3.10 and then \$1.86 per kilometre travelled while Green cabs charge a flag fall of \$3.65 and then \$1.72 per kilometre travelled.
- a** Find an expression for $\$R$, the cost of travelling k kilometres with Red cabs.
 - b** Find an expression for $\$G$, the cost of travelling k kilometres with Green cabs.
 - c** Determine the smallest number of whole kilometres that a customer must travel for Green cabs to be the cheaper option.

Calculating flat-rate depreciation

Example 16

- 8** The value of a machine is flat-rate depreciated in value by 4% of its initial value every year. Initially it was valued at \$100 000.
- a** Find an expression for $\$A$, the value of the machine after n years.
 - b** Find the value of the machine after 5 years.

Example 17

- 9** The value of a harvester is flat-rate depreciated in value by 5.1% of its initial value every year. Initially it was valued at \$235 000 in 2026.
- a** Find an expression for $\$A$, the value of the harvester after n years.
 - b** Find the year when the value of the harvester is first expected to be less than \$180 000.

Calculating unit-cost depreciation

Example 18

- 10** The value of a delivery van with purchase price \$48 000 is depreciated by \$200 for every 1000 kilometres travelled.
- a** Write a rule for the value of the delivery van after travelling n kilometres.
 - b** Determine the value of the van after 15 000 kilometres.
 - c** Determine how many kilometres it takes for the value of the van to reach \$43 000.

- 11** The value of a truck, with purchase price \$82 000, is depreciated by \$350 for every 1000 kilometres travelled.
- a** Write a rule for the value of the truck after travelling n thousand kilometres.
 - b** Determine the value of the truck after 14 000 kilometres.
 - c** Determine how many kilometres it takes for the value of the truck to be \$47 000.

Solving other problems using arithmetic sequences

Example 19

- 12** To make up a set of notes, a printer charges \$2.25 for the cover and binding and an additional 2 cents per page.
- a** Write down a rule for the cost, C_n , of making up a set of notes with n pages.
 - b** State the cost to make up a set of notes with:
 - i** 10 pages.
 - ii** 35 pages.
 - c** State the number of pages the notes contain if the printer charges:
 - i** \$3.85.
 - ii** \$6.25.
- 13** When a garbage truck starts collecting rubbish, it first stops at a corner store where it collects 86kg of rubbish. It then travels down a long suburban street where it picks up 40kg of rubbish at each house.
- a** Write down a rule for the amount of garbage collected by the truck, g_n , after n pick-ups from houses.
 - b** State the amount of garbage that would be carried by the truck after:
 - i** 15 pick-ups from houses
 - ii** 27 pick-ups from houses.
 - c** The maximum amount of garbage that can be carried by the truck is 1500kg. After picking up from the corner store, find the maximum number of houses it can pick up rubbish from before it is fully loaded.
- 14** A coffee urn contains 15 litres of coffee. Coffee is served in 200mL cups.
- a** Write down a rule for determining the amount of coffee in the urn, a_n , after n cups of coffee have been served from the urn. Assume that each cup is completely filled.
 - b** Determine the amount of coffee that will be left in the urn after:
 - i** 23 cups have been served
 - ii** 45 cups have been served.
 - c** Determine the number of cups of coffee can be served from the urn if it is necessary to keep 1.5 litres of coffee in the urn for latecomers.

- 15** Jasmine is offered a job with a starting salary of \$20 500 per year and yearly pay rises of \$450.
- a** Write down a rule for determining Jasmine's salary, s_n , at the start of her n th year on the job.
 - b** Find Jasmine's salary:
 - i** at the start of the fifth year on the job
 - ii** at the start of the eighth year on the job.
 - c** At this rate, determine the number of years it would take Jasmine to have a salary of \$50 000 per year.
- 16** You have \$430 to spend on food while on an overseas holiday. To make the money last as long as possible, you budget for \$25.75 per day.
- a** Write down a rule for determining the amount of spending money, m_n , you will have left at the start of the n th day of your holiday.
 - b** Determine the amount of spending money would you have left:
 - i** at the start of the seventh day
 - ii** at the start of the thirteenth day.
 - c** At this spending rate, determine the number of days can you afford to stay on holidays.

Paper 1-style multiple-choice questions

- 17** Sandra invests \$13 000 with a bank. She will be paid simple interest at the rate of 6.8% per annum. If V_n is the value of Sandra's investment at the start of the n th year, the recurrence relation model for Sandra's investment is:
- A** $V_1 = 13\,000, V_{n+1} = V_n + 6.8$
 - B** $V_1 = 13\,000, V_{n+1} = 6.8 \times V_n$
 - C** $V_1 = 13\,000, V_{n+1} = V_n + 884$
 - D** $V_1 = 13\,000, V_{n+1} = V_n - 884$
- 18** A printer is depreciated using a flat-rate depreciation method. It was purchased for \$1900 and depreciates at the rate of 6% per annum. The amount of depreciation after 4 years is:
- A** \$114
 - B** \$1786
 - C** \$1444
 - D** \$456
- 19** A car purchased for \$37 990 is depreciated using a unit-cost depreciation method. After travelling a total of 15 000 kilometres, it has an estimated value of \$25 990. The depreciation amount, per kilometre, is:
- A** \$0.15
 - B** \$0.80
 - C** \$12 000
 - D** \$15 000

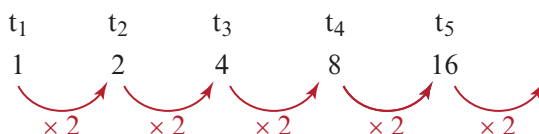
4E Geometric sequences

Learning intentions

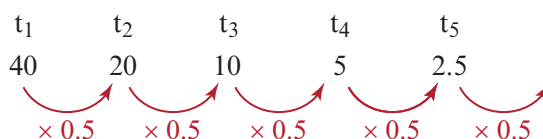
- ▶ To generate a geometric sequence with a recurrence relation.
- ▶ To graph a geometric sequence.

A sequence in which each successive term can be found by multiplying the previous term by a constant factor is called a **geometric sequence**.

For example, the sequence 1, 2, 4, 8, 16, ... is geometric because each successive term can be found by multiplying the previous term by 2.



The sequence 40, 20, 10, 5, 2.5, ... is also geometric because each successive term can be found by multiplying the previous term by 0.5.



The common ratio

In geometric sequences, the ratio between successive terms is constant and called the **common ratio**. For example, in the first sequence above, the common ratio is 2, while in the second sequence above, the common ratio is 0.5.

Common ratios can also be negative; for example, the common ratio for the geometric sequence 1, -2, 4, -8, 16, ... is -2.

Method of recursion to generate a geometric sequence

The method for using recursion to generate a geometric sequence has two parts.

- 1 A starting point: the value of the first term, t_1 , of the sequence.
- 2 A rule: multiply each term by the common ratio r to obtain the next term.

A rule for recursion that can be used to generate the sequence 10, 20, 40, 80, ... is:

- 1 Start with 10.
- 2 To obtain the next term, multiply the current term by 2 and repeat the process.

This information can also be written more compactly in symbolic form.

t_1 = starting value and $t_{n+1} = r \times t_n$, where r is the constant ratio.

This is the general recurrence relation for a geometric sequence and also gives us a way to test whether a sequence is geometric or not. For all whole numbers $n > 1$, we must have:

$$\frac{t_{n+1}}{t_n} = r$$



Example 20 Generating a geometric sequence with a recurrence relation

A geometric sequence is defined by $t_1 = 5, t_{n+1} = 2t_n$.

Find the first five terms.

Solution

$$t_1 = 5$$

$$t_2 = 2 \times t_1 = 2 \times 5 = 10$$

$$t_3 = 2 \times t_2 = 2 \times 10 = 20$$

$$t_4 = 2 \times t_3 = 2 \times 20 = 40$$

$$t_5 = 2 \times t_4 = 2 \times 40 = 80$$

$$5, 10, 20, 40, 80$$

Explanation

Write down the starting term.

Apply the rule (multiply by 2) to generate the next term.

Calculate three more terms.

Write your answer.

Graphing a geometric sequence

Geometric sequences can either be increasing or decreasing and show either geometric growth or decay respectively.

Sequence 1

Consider the sequence $V_1 = 15, V_{n+1} = 3V_n$, which generates the terms 15, 45, 135, ...

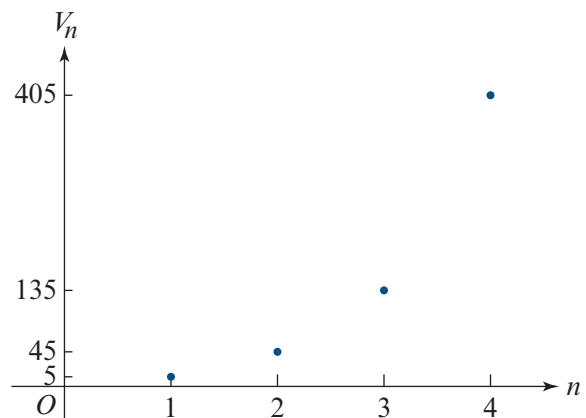
It is an **increasing sequence**, which can be used to model **geometric growth**. Values of a sequence can be shown in a table, with the first four values of Sequence 1 shown in the table below.

n	1	2	3	4
V_n	15	45	135	405

This gives ordered pairs

$(1, 15), (2, 45), (3, 135), (4, 405)$.

These can be graphed as shown opposite.



Sequence 2

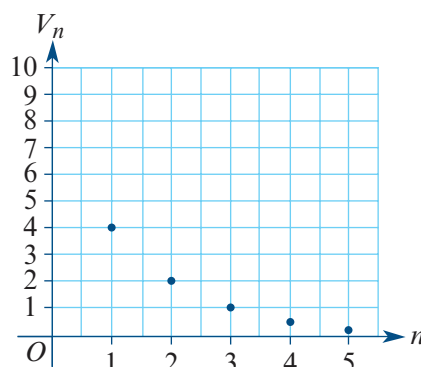
Consider the sequence $V_1 = 4$, $V_{n+1} = 0.5V_n$, which generates the terms 4, 2, 1, ... It is a **decreasing sequence**, which can be used to model **geometric decay**. The first five values of Sequence 2 are shown in the table below.

n	1	2	3	4	5
V_n	4	2	1	$\frac{1}{2}$	$\frac{1}{4}$

This gives ordered pairs

$(1, 4), (2, 2), (3, 1), (4, \frac{1}{2}), (5, \frac{1}{4})$.

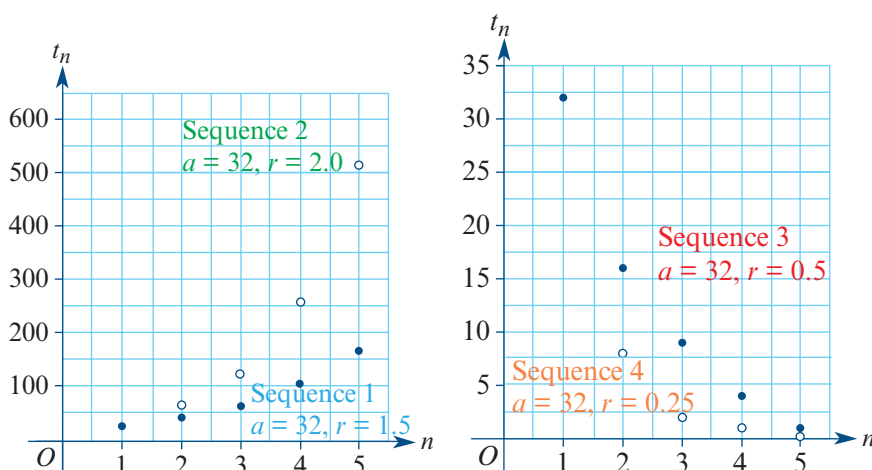
These can be graphed as shown opposite.



Graphs of geometric sequences

The graphs below display the terms in four different geometric sequences:

- Sequence 1: $t_1 = 32, r = 1.5, n = 1, 2, 3, \dots$
- Sequence 2: $t_1 = 32, r = 2.0, n = 1, 2, 3, \dots$
- Sequence 3: $t_1 = 32, r = 0.5, n = 1, 2, 3, \dots$
- Sequence 4: $t_1 = 32, r = 0.25, n = 1, 2, 3, \dots$



The key characteristics to note are that:

- the points in the graphs are not **collinear** but lie on what is called an **exponential curve**.
- if the common ratio, r , is **greater than 1**, the terms in the sequence **increase** in value. The bigger the value of r , the more rapid the increase.
- if the common ratio, r , is **less than 1**, the terms in the sequence **decrease** in value. The closer the value to 0, the more rapid the decrease.

In general:

- if $r > 1$, the recurrence relation $t_{n+1} = rt_n$ can be used to model geometric growth
- if $r < 1$, the recurrence relation $t_{n+1} = rt_n$ can be used to model geometric decay.

Note: When r is negative (not shown), terms oscillate between positive and negative values.



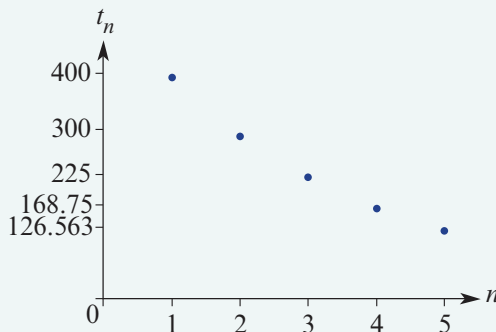
Example 21 Graphing geometric sequences

Prepare a table of values and plot the first five terms to illustrate the sequence defined by:

$$t_1 = 400, t_{n+1} = 0.75t_n$$

Solution

n	1	2	3	4	5
t_n	400	300	225	168.75	126.5625



Explanation

Start with $t_1 = 400$. Use the recursion relation $t_{n+1} = 0.75t_n$ to complete the table.

$$t_2 = 0.75t_1 = 0.75 \times 400 = 300$$

$$t_3 = 0.75t_2 = 0.75 \times 300 = 225$$

Use the values in the table to plot the ordered pairs $(1, 400)$, $(2, 300)$, $(3, 225)$, ...

Section Summary

- ▶ A **geometric sequence** is a sequence where each successive term is found by multiplying the previous term by a constant factor.
- ▶ The ratio, r , between successive terms in a geometric sequence is called the **common ratio**.
- ▶ A geometric sequence can be written as $t_1 = \text{starting value}$, $t_{n+1} = r \times t_n \dots$
- ▶ If $r > 1$, the sequence is **increasing** and exhibits **geometric growth**.
- ▶ If $0 < r < 1$, the sequence is **decreasing** and exhibits **geometric decay**.
- ▶ If $r < 0$, the sequence oscillates between positive and negative values.
- ▶ The terms in a geometric sequence can be displayed in both tabular and graphical form.

Exercise 4E

Generating a geometric sequence with a recurrence relation

Example 20

- 1** A geometric sequence is defined by:

$$t_1 = 3, t_{n+1} = 4t_n$$

Find the first five terms.

- 2** A geometric sequence is defined by:

$$t_1 = 15, t_{n+1} = 2t_n$$

Find the first five terms.

- 3** Consider the geometric sequence 2, 20, 200, 2000, ...

- a** State the common ratio.
- b** Determine the next term in the sequence.
- c** Starting with 2, determine the number of times you have to multiply by the common ratio to get to term 5.
- d** Starting with 2, determine the number of times you have to multiply by the common ratio to get to term 15.
- e** Determine the value of term 15. Give your answer using scientific notation correct to 2 decimal places.

- 4** Consider the geometric sequence 1024, 256, 64, 16, ...

- a** State the common ratio.
- b** Determine the next term in the sequence.
- c** Starting with 1024, determine the number of times you have to multiply by the common ratio to get to term 7.
- d** Determine the value of term 7.
- e** Determine the value of term 10.
- f** Determine the value of term 50. Give your answer in scientific notation, correct to two decimal places.

- 5** Consider the sequence -1, 5, -25, 125, ...

- a** State the common ratio.
- b** Determine the next term in the sequence.
- c** Starting with -1, determine the number of times you have to multiply by the common ratio to get to term 6.
- d** Determine the value of term 6.
- e** Starting with -1, determine the number of times you have to multiply by the common ratio to get to term 15.
- f** Determine the value of term 15.

Graphing a geometric sequence

Example 21

- 6** Prepare a table of values and plot a graph to show the first five terms of the sequence defined by:

$$t_1 = 9, t_{n+1} = 2t_n$$

- 7** Prepare a table of values and plot a graph to show the first five terms of the sequence defined by:

$$t_1 = 12, t_{n+1} = \frac{1}{2}t_n$$

- 8** Prepare a table of values and plot a graph to show the first five terms of the sequence defined by:

$$t_1 = 4, t_{n+1} = 4t_n$$

- 9** Write the following as a recurrence relation in symbolic form, where t_n represents the value.

a The starting value is 4, and the rule is ‘multiply the current term by 2 and repeat the process.’

b The starting value is 10, and the rule is ‘multiply the current term by 7 and repeat the process.’

c The starting value is 16, and the rule is ‘multiply the current term by $\frac{1}{2}$ and repeat the process.’

- 10** State the recurrence relation in symbolic form for the following sequences.

a 3, 9, 27, 81, ...

b 2, 8, 32, 128, ...

c 200, 100, 50, 25, ...

- 11** The following recurrence relation can generate a sequence of numbers.

$$t_1 = 6, \quad t_{n+1} = 2t_n$$

State the term name for the term that has a value of 1536.

- 12** Bobby starts a chain letter. On the first day, he sends the letter to two people. On the next day, each of the two people send the letter to two more people. On the third day, each of these people sends the chain letter to two more people.

a Write a recurrence relation to model the situation.

b Determine the day on which more than 5000 people receive the chain letter.

- 13** Bea decides to start a push-up challenge. On the first day, she does 4 push-ups. On the next day, she does twice as many push-ups as the day before.
- a** Write down the number of push-ups that Bea does each day for the first week.
 - b** Write a recurrence relation to model the situation.
 - c** Determine the day in which she first does more than 100 push-ups in one day.
- 14** Jackson has a population of goldfish in his garden pond. He realises that the population seems to be growing geometrically.
- On the first day of the third month, he counts 80 goldfish in the pond.
- On the first day of the fifth month, he counts 320 goldfish in the pond.
- Write a recurrence relation to model the situation and determine the original number of goldfish in the pond.

Paper 1-style multiple-choice questions

- 15** The first term of a sequence is 12. Each subsequent term is 0.8 times the previous term. The sixth term, correct to two decimal places, is
- A** 9.60 **B** 3.93 **C** 3.15 **D** 2.52
- 16** The following is a geometric sequence: 24, -48, 96, -192, ... The common ratio r is equal to:
- A** -3 **B** -2 **C** 1 **D** 2
- 17** A different type of recurrence relation is defined by $V_{n+1} = 2V_n + 1$, where $V_1 = 3$. The first four terms are
- A** 3, 7, 15, 31
B 2, 3, 7, 15
C 7, 15, 31, 63
D 3, 4, 5, 6
- 18** The first five terms of a recurrence relation are 972, 324, 108, 36, 12, 4. In symbolic form, this can be represented as
- A** $V_1 = 972, V_{n+1} = V_n - 648$
B $V_1 = 972, V_{n+1} = 3 \times V_n$
C $V_1 = 972, V_{n+1} = \frac{1}{3} \times V_n$
D $V_1 = 972, V_{n+1} = -\frac{1}{3} \times V_n$

4F A general rule for finding the n th term of a geometric sequence

Learning intentions

- ▶ To use a general rule for a geometric sequence.
- ▶ To determine a geometric sequence from two terms.
- ▶ To determine the number of terms required to reach a value in a geometric sequence.
- ▶ To graph a geometric sequence.

The rule for the n th term of a geometric sequence

Consider a geometric sequence with first term t_1 and common ratio r . Then:

$$t_1 = t_1$$

$$t_2 = t_1 \times r = t_1 r$$

$$t_3 = t_2 \times r = t_1 r^2$$

$$t_4 = t_3 \times r = t_1 r^3$$

The rule for the n th term of a geometric sequence

Thus, following the pattern, we can write:

$$t_n = t_1 r^{n-1}$$

which gives us a rule for finding the n th term of a geometric sequence in terms of the first term, t_1 , and the common ratio, r .



Example 22 Using the general rule for a geometric sequence

Given $t_1 = 6$ and $r = 2$, find t_7 .

Solution

$$\begin{aligned} t_7 &= 6 \times 2^{7-1} \\ &= 6 \times 2^6 \\ &= 384 \end{aligned}$$

Explanation

Substitute $t_1 = 6$ and $r = 2$ in the rule $t_n = t_1 r^{n-1}$.



Example 23 Using the general rule for a geometric sequence

Use the rule to determine t_8 of the geometric sequence 100, 50, 25, 12.5, ...

Solution

$$\begin{aligned} t_8 &= 100 \times 0.5^{8-1} \\ &= 100 \times 0.5^7 \\ &= 0.78125 \end{aligned}$$

Explanation

For this sequence $t_1 = 100$ and $r = 0.5$. Substitute into the rule $t_n = t_1 r^{n-1}$.


Example 24 Determining a geometric sequence given two terms

In a geometric sequence, the fourth term is 24 and the ninth term is 768. Write down the first three terms of the sequence.

Solution

$$24 = t_1 r^{4-1} = t_1 r^3 \quad (1)$$

$$768 = t_1 r^{9-1} = t_1 r^8 \quad (2)$$

$$\frac{768}{24} = \frac{t_1 r^8}{t_1 r^3} \quad 32 = r^5$$

$$r = 2$$

$24 = t_1 \times 2^3$ or $24 = 8t_1$ so $t_1 = 3$
The first three terms of the sequence are 3, 6, 12.

Explanation

Substitute $t_4 = 24$ in $t_n = t_1 r^{n-1}$ with $n = 4$, then substitute $t_9 = 768$ in $t_n = t_1 r^{n-1}$ with $n = 9$ to form two equations.

Find t_1 and r by solving the simultaneous equations (1) and (2). Divide (2) by (1).

Substitute $r = 2$ in equation (1).
Write the first three terms of the sequence.


Example 25 Determining how many terms of a geometric sequence are required to reach a particular number

Determine the number of terms in the geometric sequence 0.5, 5, 50, 500, ... before we reach a term greater than 1 000 000 is found.

Solution

$$0.5 \times 10^{n-1} > 1\,000\,000$$

$$10^{n-1} > 2\,000\,000$$

$$n - 1 > 6$$

$$n > 7$$

We need to write down 8 terms to find the first term that exceeds 1 000 000.

Explanation

Substitute $t_1 = 0.5$ and $r = 10$ in the general rule for a geometric sequence. We want to find n so that:
 $t_n = 0.5 \times 10^{n-1} > 1\,000\,000$

Alternate method Solve by trial and error with your calculator.

$$10^2 = 100,$$

$$10^3 = 1000, \dots,$$

$$10^6 = 1\,000\,000,$$

$$10^7 = 10\,000\,000$$

Graphing a geometric sequence

In investigating the properties of geometric sequences, particularly the rate at which the terms in the series either grow or decay, it is useful to have a graphical representation of the sequence.



Example 26 Graphing a geometric sequence

Consider the geometric sequence with $t_1 = 3$ and the common ratio $r = 2$.

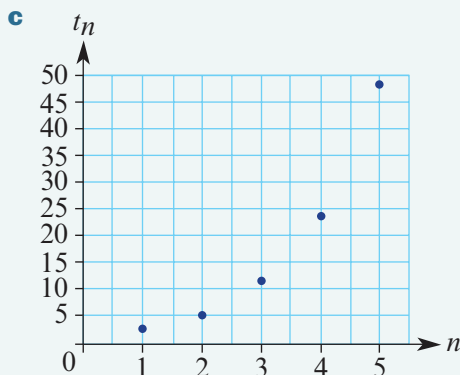
- Find the rule for this geometric sequence.
- Prepare a table of values for the sequence for $n = 1$ to $n = 5$.
- Plot a graph from the table of values.

Solution

a $t_n = 3 \times 2^{n-1}$ for $n = 1, 2, 3, \dots$

b

n	1	2	3	4	5
t_n	3	6	12	24	48



Explanation

Find an expression for the n th term, t_n , in terms of n .

Generate a table of values using the rule. For example, $t_3 = 3 \times 2^{3-1} = 12$.

Plot the points on a graph with n on the horizontal axis and t_n on the vertical axis. Do not join up the points, as the terms in the sequence are only defined for $n = 1, 2, 3, \dots$

Section Summary

- ▶ The rule for the n th term of a geometric sequence is given as $t_n = t_1 r^{n-1}$ where t_n is the n th term, t_1 is the first term and r is the common ratio.
- ▶ The rule for the n th term of a geometric sequence can be used to find a particular term by substituting in the starting value and the common ratio.

Exercise 4F

Using the rule to generate a geometric sequence

1 Write the rule for the n th term in a geometric sequence for the following.

a The starting value is 2 and the common ratio is 5.

b The starting value is 5000 and the common ratio is 0.8.

Example 22

2 The first term of a geometric sequence is $t_1 = 1$ and the common ratio is $r = 4$. Use the rule to determine the value of:

a the third term

b the seventh term

c the 15th term.

3 Write the rule for the n th term of each of the following recurrence relations.

a $t_1 = 8$ and $t_{n+1} = 5t_n$

b $t_1 = 5$ and $t_{n+1} = 0.6t_n$

c $t_1 = 32$ and $t_{n+1} = \frac{1}{4}t_n$

Example 23

4 Use the rule to determine the value of:

a the seventh term of the geometric sequence 1, 5, 25, 125, 625, ...

b the eighth term of the geometric sequence 10 000, 2000, 400, ...

c the 10th term of the geometric sequence $-1, 2, -4, 8, \dots$

d the ninth term of the geometric sequence $-20, -60, -180, \dots$

e the eighth term of the geometric sequence $2, 1, \frac{1}{2}, \dots$

f the fifth term of the geometric sequence 110, 121, 133.1, ...

g the seventh term of the geometric sequence $1, -\frac{1}{2}, \frac{1}{4}, -\frac{1}{8}, \dots$

Determining a geometric sequence given two terms

Example 24

5 Write down the first three terms of the geometric sequence in which:

a $t_5 = 81$ and $t_8 = 2187$

b $t_2 = 10\,000$ and $t_5 = 1250$

c $t_3 = 40$ and $t_6 = -320$

d $t_2 = 160$ and $t_4 = 250$ ($r > 0$)

Determining the required number of terms to reach a particular value

Example 25

6 Find how many terms we would have to write down in the following geometric sequences.

a 2, 4, 8, 16, ... before we found a term greater than 250.

b 1, 1.1, 1.21, ... before we found a term greater than 2.

c 100, 80, 64, ... before we found a term less than 10.

d $-8, -16, -32, \dots$ before we found a term equal to -4096 .

e 0.9, 0.81, 0.729, ... before we found a term less than 0.1.

f 2000, 2100, 2205, ... before we found a term greater than 4000.

g 6000, 5700, 5415, ... before we found a term less than 3000.

Graphing geometric sequences

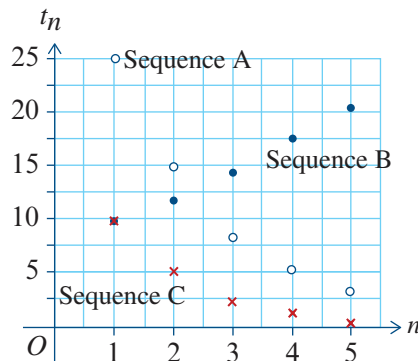
Example 26

7 Plot the first five terms of each of the following geometric sequences.

- a** Sequence A: $t_1 = 1, r = 2, n = 1, 2, 3, \dots$
- b** Sequence B: $t_1 = 100, r = 0.5, n = 1, 2, 3, \dots$
- c** Sequence C: $t_1 = 1024, r = 0.75, n = 1, 2, 3, \dots$
- d** Sequence D: $t_1 = 32, r = 1.5, n = 1, 2, 3, \dots$
- e** Sequence E: $t_1 = 1024, r = -0.25, n = 1, 2, 3, \dots$

8 Three sequences are displayed in the graphs below. For each sequence:

- a** determine the value of the first term, t_1
- b** determine from the trend of the points whether the value of r is greater than 1 or between 0 and 1
- c** use the first two points to estimate the value of r .



9 The sequence defined by the rule for the n th term, $t_n = 2^n$ is a geometric sequence.

- a** List the first four terms of the sequence.
- b** State the value of the common ratio.

Paper 1-style multiple-choice questions

10 The rule for the n th term in a geometric sequence where $t_1 = 5$ and $r = 2$ is

- A** $t_n = 5 + 2n$ **B** $t_n = 5n + 2$ **C** $t_n = 5 \times 2^n$ **D** $t_n = 2 \times 5^n$

11 The value of t_5 when $t_1 = 100$ and $r = \frac{1}{2}$ is

- A** 50 **B** 12.5 **C** 6.25 **D** 3.125

12 The value of $t_1 = 2$ and the value of $t_4 = 54$. If the sequence is geometric, the value of r is

- A** 2 **B** 3 **C** 4 **D** 13

4G Applications of geometric sequences

Learning intentions

- ▶ To model the growth of a bacterial population using geometric sequences.
- ▶ To model compound interest using geometric sequences.
- ▶ To calculate the value of an asset using reducing-balance depreciation.
- ▶ To apply geometric sequences to practical problems.

Population growth

Populations, whether bacteria or different animal species, often grow at the same rate but could also decline due to falling fertility rates or other negative impacts on the environment.



Example 27 Modelling growth of a bacterial population

A dish in a laboratory contains 150 000 bacteria. The population of bacteria, N , is expected to double in size every day.

- a** Using recursion, find the number of bacteria after 1, 2, 3 and 4 days.
- b** Write a rule for the number, N , of bacteria after n days.
- c** Use the rule to find the number of bacteria after 7 days.

Solution

- a** After day one: $2 \times 150\,000 = 300\,000$
 After day two: $2 \times 300\,000 = 600\,000$
 After day three: $2 \times 600\,000 = 1\,200\,000$
 After day four: $2 \times 1\,200\,000 = 2\,400\,000$
- b** $N = 150\,000 \times 2^n$
- c** $N = 150\,000 \times 2^7 = 19\,200\,000$
 After 7 days, there are 19 200 000 bacteria.

Explanation

The number of bacteria is doubled each day. Therefore, we multiply by 2 for each day.

$t_1 = 150\,000$, which is multiplied by 2, n times, or by 2^n .

Substitute $n = 7$ in rule.

Percentage increase and decrease

A population might increase by a percentage of its current amount. For example, a colony of koalas in a reserve increases by 5% each year. If there are 400 koalas this year, we can expect there to be an extra 5% next year, in addition to the 400 koalas that are already there.

$$\text{This year:} = 400$$

$$\text{Next year:} = 400 + 5\% \text{ of } 400$$

$$= 400 + \frac{5}{100} \times 400$$

$$= 400 + 20$$

$$= 420 \text{ koalas}$$

Alternatively, if we are adding on 5% to our initial 100% of koalas, we have 105% of our initial population, so we could calculate the number of koalas as follows:

$$\begin{aligned}\text{Next year:} &= 105\% \text{ of } 400 \\ &= \frac{105}{100} \times 400 \\ &= 1.05 \times 400 \\ &= 420 \text{ koalas}\end{aligned}$$

Alternatively, if the koala population decays by 5% each year, next year there would be $100\% - 5\%$ or 95% of this year's number of koalas.

The **rate** of geometric growth and decay is often given in the form of a percentage of the starting term. We use r for the common ratio and $i\%$ for the percentage growth.

Percentage increase and decrease

If the percentage **increase** for geometric growth is $i\%$, then $r = 1 + i\%$.

If the percentage **decrease** for geometric decay is $i\%$, then $r = 1 - i\%$.

Percentage increase and decrease can then be extended to cover multiple years, as we will see in the following sections and examples.

Compound interest

Most interest calculations involve **compound interest**, where any interest earned after one time period is added to the principal, with the interest in the following period being calculated on this new value. This means that the value of the investment grows in increasing amounts, or grows geometrically, and can be modelled with a recurrence relation.

Consider an investment of \$5000 that pays interest of 8% per annum. The interest will be paid into the account after each year, and this interest is re-invested and will earn interest in the next year. Thus, the interest is **compounding**. With a starting value of \$5000 and a **rate** of 8% each year, the multiplying factor will be:

$$\begin{aligned}r &= 1 + i\% \\ r &= 1.08\end{aligned}$$

The value of the investment is calculated as follows:

Initial amount: \$5000

After one year: $\$(5000 \times 1.08) = \$5400$

After two years: $\$(5000 \times 1.08 \times 1.08) = \$5832$

After three years: $\$(5000 \times 1.08 \times 1.08 \times 1.08) = \$6298.56$

After n years: $\$5000 \times 1.08^n$

Calculation of compound interest loans and investments

Let A be the value of a compound interest loan or investment after n years.

Let P be the amount borrowed or invested (principal).

Let i be the annual interest rate of the loan or investment.

$$A = P \times r^n \text{ where } r = 1 + i$$



Example 28 Modelling compound interest

An amount of \$2000 is invested with compound interest at 7.5% per annum.

- Find the rule for the value of the investment after n years.
- Find the value of the investment after 4 years.
- Determine when the value of the investment will first exceed \$3000.

Solution

$$\begin{aligned} \text{a } P &= 2000 \text{ and } r = 1 + \frac{7.5}{100} = 1.075 \\ A &= 2000 \times 1.075^n \end{aligned}$$

$$\begin{aligned} \text{b } A &= 2000 \times 1.075^4 \\ &= 2670.938 \dots \\ \text{The value of the investment is } \$2670.94 \\ &\text{after 4 years.} \end{aligned}$$

$$\begin{aligned} \text{c After 5 years,} \\ A &= 2000 \times 1.075^5 = 2871.258 \dots \\ \text{After 6 years,} \\ A &= 2000 \times 1.075^6 = 3086.60 \dots \\ \text{After 6 years the investment will} \\ &\text{exceed } \$3000. \end{aligned}$$

Explanation

Substitute the values for P and r into the formula $A = P \times r^n$, with $r = 1 + i$.

Substitute $n = 4$.

Work with trial and error to determine when the investment will exceed \$3000.

Note: When we are working with money, all **final values** are rounded to two decimal places unless we are told otherwise. You should not round intermediate values, only the final value.

Reducing-balance depreciation

Earlier in the chapter, we studied two different methods for depreciating the value of an asset, both of which were examples of linear decay. **Reducing-balance depreciation** or **diminishing-value depreciation** is another method of depreciation where the value of an asset decays geometrically. Each year, the value will be reduced by a percentage, $i\%$, of the previous year's value. The calculations are very similar to compounding interest, but with decay in value, rather than growth.

Calculation of reducing-balance depreciation

Let A be the value of the asset after n years.

Let P be the initial cost of the asset.

Let i be the annual depreciation rate of the asset.

$$A = P \times r^n \text{ where } r = 1 - i\%$$

**Example 29** Calculating reducing-balance depreciation

An item of office furniture has a purchase price of \$6900. It can be considered to be depreciating at a reducing-balance rate of 8.4% per annum.

- a** Find the rule for the value of the office furniture after n years.
- b** Find the value of the office furniture after 4 years.
- c** Determine when the value of the office furniture will be less than \$3000.

Solution

a $P = 6900$

$$r = 1 - \frac{8.4}{100} = 0.916,$$

$$A = 6900 \times 0.916^n$$

b $A = 6900 \times 0.916^4$

$$= 4857.7033 \dots$$

The value of the office furniture after 4 years is \$4857.70.

c $n = 8:$

$$A = 6900 \times 0.916^8 = 3419.895 \dots$$

$n = 9:$

$$A = 6900 \times 0.916^9 = 3132.624 \dots$$

$n = 10:$

$$A = 6900 \times 0.916^{10} = 2869.484 \dots$$

After 10 years the value is less than \$3000.

Explanation

Write down the purchase price of the furniture, P . Calculate the value of r where $i = 8.4\%$. Use the rule $A = P \times r^n$.

Substitute $n = 4$ into the rule.
Answer to the nearest cent.

Use trial and error to determine when the value of the furniture will be less than \$3000 by substituting values of n into the rule $A = 6900 \times 0.916^n$.

Note: Be careful to note if the question asks for the 'value' at the beginning or end of the year.

The same idea can be applied to other situations where an amount is decreasing by a constant percentage for each period of time.



Example 30 Applying geometric sequences

A swimming pool is filled with 200 000 litres of water.

Under certain conditions, 2% of the water in the pool will evaporate each day.

- a** Determine the rule for the amount of water in the pool at the end of the n th day.
- b** Find the amount of water in the pool at the end of the eighth day, to the nearest litre.

Solution

- a** Let V_n be the volume of water in the pool after n days. The water volume decays by 2% every day.

$$r = 1 - \frac{2}{100} = 0.98$$

$$V_n = 200\,000 \times 0.98^n$$

- b** $V_8 = 200\,000 \times 0.98^8$
 $= 170\,153$ litres to the nearest litre

The amount of water in the pool at the end of the eighth day is 170 153 litres.

Explanation

Determine the value of r .

Substitute r into the rule $V_n = 200\,000 \times r^n$.

Substitute $n = 8$.

Interest rates over different periods

Compound interest rates are usually quoted as annual rates, or interest rate per annum. The time period after which compound interest is calculated and paid is called the compounding period. For example, the compounding period could be months or days.

An annual interest rate can be converted to a compounding interest rate for a shorter period by **dividing** this interest rate per annum by the appropriate number. For example, an interest rate of 3.6% per annum gives a monthly interest rate of $(3.6 \div 12) = 0.3\%$ per month. This is explored more in Chapter 7.

Section Summary

- **Compound interest** is an example of geometric growth where the interest accumulates on both the principal and the interest from previous periods. It is calculated using the rule $A = P \times r^n$, $r = 1 + i$, where A is the value, P is the initial value and i is the annual interest rate.
- When the value of an item decreases as a percentage of its value each time period, it is said to be depreciating using a **reducing-balance method**. It is calculated using the rule $A = P \times r^n$, where $r = 1 - i$, A is the value, P is the initial cost of the asset and i is the annual depreciation rate.



Exercise 4G

Applying geometric sequences

Example 27

- 1 A dish on a counter contains 80 000 bacteria. The population of bacteria, N , is expected to double in size every day.
 - a Using recursion, find the number of bacteria after 1, 2, 3 and 4 days.
 - b Write a rule for the number, N , of bacteria after n days.
 - c Use the rule to find the number of bacteria after 5 days.

- 2 A dish on a counter contains 100 000 bacteria. The population of bacteria, N , is expected to double in size every 10 minutes. Each time period lasts 10 minutes.
 - a Using recursion, find the number of bacteria after 10 minutes and after 20 minutes.
 - b Write a rule for the number, N , of bacteria after n time periods (10 minutes).
 - c Use the rule to find the number of bacteria after 1 hour.

- 3 The population of kangaroos in a national park is increasing by 5% every year. There are currently 2700 kangaroos in the national park.
 - a Write down the rule for the number of kangaroos after n years.
 - b Find the number of kangaroos after 5 years.
 - c Determine how many years it takes for the kangaroo population to double.

There are also 830 wombats in the park. Their numbers are increasing by 4% every year.

 - d Find how many wombats there are in the park after 4 years.
 - e Determine how long it takes for the number of wombats in the park to double.

- 4 Suppose a newly discovered virulent bacteria replicates itself every five minutes.
 - a If we start off with 10 bacteria, determine the number of bacteria after 5 minutes.
 - b Complete the following table, where time period 1 is after the first 5 minutes have passed, time period 2 is after the second 5 minutes have passed, etc.

Time period	0	1	2	3
Number of bacteria	10			

 - c Write down an expression for the number of bacteria at the **end** of the n th time period.
 - d Determine the number of bacteria:
 - i at the end of the fourth time period
 - ii after 30 minutes
 - iii after 1 hour.

- 5** A fish population has an initial population of 1000 and increases its size by 40% every six months, provided the conditions are ideal.

- a** Determine the number of fish after six months.
b Complete the following table, where time period 1 is after the first six months have passed, time period 2 is after the second six months have passed, etc. Round to the nearest whole number.

Time period	0	1	2	3	4
Number of fish	1000				

- c** Write down an expression for the number of fish in the population at the end of the n th time period.
d Determine the number of fish:
 i at the end of the sixth time period
 ii after 5 years
 iii after 10 years.

Calculating compound interest

Example 28

- 6** An investment of \$5000 earns compounding interest at a rate of 3% per annum. Write down the value of the investment after 1, 2 and 3 years.
7 An investment of \$6000 earns compounding interest at a rate of 4.2% per annum.
a Write down the value of the investment after n years.
b State the value of the investment after 4 years.
8 An investment of \$15 000 earns compounding interest at the rate of 5.4% per annum.
a Write down the value of the investment after 1, 2 and 3 years.
b Write down the value of the investment after n years.
c Determine how many years it takes for the value of the investment to first exceed \$24 000.
9 A loan of \$20 000 is charged compounding interest at the rate of 6.3% per annum.
a Write down the value of the loan after n years.
b Write down the value of the loan after 5 years.
c Determine how many years it takes for the value of the loan to first exceed \$30 000.
d Write down the rule for the value of a loan of \$18 000 at a compounding interest rate of 9.4% per annum after n years.

- 10** An investment of \$8000 earns 12.5% compound interest each year.
- Write down a rule for the value of the investment after n years.
 - Use the rule to find the value of the investment after three years.
 - State the amount of interest earned over three years.
 - Determine how much interest was earned in the third year of the investment.
- 11** A loan of \$3300 is charged 7.5% compound interest each year.
- Write down a rule for the amount owed after n years.
 - Use the rule to find the value of the loan after 10 years.
 - State how much interest was charged over 10 years.
 - Determine how much interest was charged in the 10th year of the investment.

Calculating reducing-balance depreciation

Example 29

- 12** A motorcycle was purchased new for \$9800 and is depreciated using a reducing-balance depreciation method with an annual depreciation rate of 3.5%.
- Determine the rule for the value of the motorcycle after n years.
 - Determine the value of the motorcycle after 5 years.
 - State the amount of depreciation of the motorcycle in the third year.
- 13** A stereo system, initially valued at \$1200, is depreciated using reducing-balance depreciation of 12%.
- Write down a rule for the value of the stereo system after n years.
 - Use the rule to find the value of the stereo system after seven years.
 - Determine when the stereo is first worth less than half of its original value.
- 14** Suppose a car costs \$20 000 when new. Assume that it loses 7.5% of its value each year.
- State the value of the car after one year.
 - Complete the following table where time period 1 corresponds to the value after 1 year, time period 2 corresponds to after 2 years, etc. Give your answers correct to the nearest dollar.
- | | | | | | |
|-------------------|--------|---|---|---|---|
| Time period | 0 | 1 | 2 | 3 | 4 |
| Value of car (\$) | 20 000 | | | | |
- Write down an expression for the value of the car at the end of the n th year.
 - Determine the value of the car, to the nearest dollar:
 - at the end of the sixth year of its life
 - after 8 years
 - after 15 years.

- 15** When a man purchases an antique table for \$12 000, he is told its value will increase by 125% every 25 years.
- a** Making this assumption, determine how much you expect the table to be worth after 25 years.
- b** Complete the following table where the first time period is 25 years, time period 2 corresponds to after another 25 years, etc.

Time period	0	1	2	3
Value of antique table (\$)	12 000			

- c** Write down an expression for the value of the table at the end of the n th time period.
- d** Determine the value of the table:
- i** at the end of the 4th time period
 - ii** after 150 years
 - iii** after 250 years.

Applying geometric sequences

Example 30

- 16** The following rule can be used to model the number of shares an investor owns after n months, if the investor sells 4% of the shares he owns after every month and the investor initially owns P shares.

$$A = P \times 0.96^n$$

where A is the number of shares owned by the investor after n months.

- a** Determine the number of shares the investor has after 2 years if he originally owned 10 000 shares.
- b** Write down the rule for the number of shares owned after n months if the investor owned 30 000 shares and sold 3.5% of the shares owned after every month.
- 17** Geoff has \$570 000 in his superannuation account when he retires and it grows at a rate of 6.15% per annum.
- a** Write down the rule for the value of the superannuation account, S , after n years. Assume that Geoff does not add or subtract any money from the account.
- b** After three years, Geoff withdraws \$30 000 for a holiday. He also notices that the interest rate on his account has changed to 4.94%. Write down a new rule for the value of his superannuation account, N , after n years since this withdrawal.

- 18** A ball is dropped vertically from a tower 3.6 metres high and the height of its rebound is recorded for four successive bounces. The results are shown in the table below:

Bounce number	0	1	2	3
Height (centimetres)	360.00	270.00	202.50	151.875

- a** Determine whether the heights of the bouncing ball given in the table form a geometric sequence. Explain.
- b** Assuming that the height of the bouncing ball follows a geometric sequence:
- i** predict the height of the fourth bounce. Give your answer correct to two decimal places.
 - ii** write down an expression for the height of the n th bounce.
 - iii** predict the height of the 15th bounce. Give your answer correct to two decimal places.

Paper 1-style multiple-choice questions

- 19** A population of lizards is decreasing by 4% every year. There are currently 5200 lizards in the population.
The number of lizards in the population after n years, L , is
- A** 5200×1.4^n
B 5200×1.04^n
C 5200×0.96^n
D $5200 + 1.04^n$
- 20** Royce invests \$2000 in an account that pays compounding interest at the rate of 5.12% per annum.
The number of years it takes for the investment to first exceed \$4000 is
- A** 13 **B** 14 **C** 15 **D** 16
- 21** A tractor was purchased new for \$100 000 and is depreciated using a reducing-balance depreciation method with an annual depreciation rate of 3.8%.
The amount of depreciation of the tractor in the third year is
- A** \$96 200
B \$89 027.71
C \$3516.68
D \$3800

Key ideas and chapter summary

Sequence	A sequence is a list of numbers or symbols written in succession. For example, 5, 15, 25, ... Sequences often have patterns that mean we can write rules and predict the terms that make up the sequence.
Term	Each number or symbol that makes up a sequence is called a term .
Recursion	Recursion involves repeating the same calculation over and over, using the previous result to calculate the next result.
Recurrence relation	Recurrence relations define the terms of a sequence using recursive calculations. The rule of the recurrence relation relies on one term in the sequence, t_n , to generate the next term in the sequence, t_{n+1} . The recurrence relation must show the starting value, t_1 , and the rule.
Modelling	Modelling is the use of a mathematical rule or formula to represent or model real-life situations. Recurrence relations can be used to model situations involving the growth (increase) or decay (decrease) in values of a quantity.
Arithmetic sequences	A sequence is arithmetic if it satisfies the recurrence relation: $t_1 = \text{starting value}$, $t_{n+1} = t_n + d$, where d is a real number. Arithmetic sequences are used to model linear growth and linear decay situations. The rule for the nth term of an arithmetic sequence is: $t_n = t_1 + (n - 1)d$, where t_1 is the starting value.
Linear growth	When a recurrence relation rule involves adding a constant amount, d, to each term, the terms of the sequence will increase uniformly through the sequence. The terms will grow linearly. Linear growth can be modelled by the recurrence relation, $t_1 = \text{starting value}$, $t_{n+1} = t_n + d$, where d is a positive real number.
Linear decay	When a recurrence relation rule involves subtracting a constant amount from each term, the terms of the sequence will decrease uniformly through the sequence. The terms will decay linearly. Linear decay can be modelled by the recurrence relation, $t_1 = \text{starting value}$, $t_{n+1} = t_n - d$, where d is a positive real number.

Linear growth and decay rule

The value of a quantity that grows or decays linearly can be found using the general rule $t_n = t_1 + (n - 1)d$, where t_1 is the starting value, t_n is the value of the quantity after n steps and d is a real number.

Geometric sequences

A sequence is **geometric** if it satisfies the recurrence relation:

t_1 = starting value, $t_{n+1} = r \times t_n$, where r is a real number.

Geometric sequences are used to model geometric growth and linear decay situations.

The rule for the n th term of a geometric sequence is:

t_1 = starting value, $t_n = t_1 r^{n-1}$.

Geometric growth

When a recurrence relation rule involves multiplying by a factor, r , that is larger than 1, the terms of the sequence increase through the sequence. The terms will grow geometrically.

Geometric growth can be modelled by the recurrence relation

t_1 = starting value, $t_{n+1} = r \times t_n$, where $r > 1$.

Geometric decay

When a recurrence relation rule involves multiplying by a factor that is smaller than 1, the terms of the sequence decrease through the sequence. The terms will decay geometrically.

Geometric decay can be modelled by the recurrence relation

t_1 = starting value, $t_{n+1} = r \times t_n$, where $r < 1$.

Geometric growth and decay rule

The value of a quantity that grows or decays geometrically can be found using the general rule

$t_n = t_1 \times r^{n-1}$ where t_n is the value of the quantity after n steps.

Principal

The **principal** is the initial amount that is invested or borrowed.

Balance

The value of a loan or investment at any time during the loan or investment period is the **balance**.

Interest

Interest is the fee that is added to a loan or the payment for investing money.

Simple interest

Simple interest is a fixed amount of interest that is paid at regular time intervals. Simple interest is an example of linear growth.

Depreciation

Depreciation is the amount by which the value of an item decreases after a period of time.

Flat-rate depreciation

A constant amount that is subtracted from the value of an item at regular time intervals. **Flat-rate depreciation** is an example of linear decay.

Unit-cost depreciation

Depreciation that is calculated based on units of use rather than time. **Unit-cost depreciation** is an example of linear decay.

Compounding period

Interest rates are often quoted as annual rates (per annum). Interest is sometimes calculated more regularly than each year, for example, each quarter, month, fortnight, week or day. The time period for the calculation of interest is called the **compounding period**.

Compound interest

When interest is added to a loan or investment and then contributes to earning more interest, the interest is said to compound. **Compound interest** is an example of geometric growth.

Reducing-balance depreciation

When the value of an item decreases as a percentage of its value after each time period, it is said to be depreciating using a reducing-balance method, or diminishing-value depreciation. **Reducing-balance depreciation** is an example of geometric decay.

Skills checklist

Download this checklist from the Interactive Textbook, then print it and fill it out to check your skills. ☒

4A**1** I can generate an arithmetic sequence. ☐

See Example 1 and 2 and Exercise 4A Question 1 and 4

4A**2** I can generate an arithmetic sequence using a calculator. ☐

See Example 3 and Exercise 4A Question 5

4B**3** I can define an arithmetic sequence and identify the common difference. ☐

See Example 4, 5 and 6 and Exercise 4B Question 1, 5 and 6

4B**4** I can generate a sequence from a recurrence relation. ☐

See Example 7 and Exercise 4B Question 10

4B**5** I can tabulate and graph an arithmetic sequence. ☐

See Example 8 and Exercise 4B Question 14

4C**6** I can use a rule for an arithmetic sequence with a positive or negative difference. ☐

See Example 9 and 10 and Exercise 4C Question 1 and 2

- 4C** **7** I can determine an arithmetic sequence given two terms. ☐
See Example 11 and Exercise 4C Question 7
- 4C** **8** I can determine the number of terms required to reach a certain value. ☐
See Example 12 and Exercise 4C Question 12
- 4C** **9** I can graph an arithmetic sequence using a rule. ☐
See Example 13 and Exercise 4C Question 13
- 4D** **10** I can model an investment or loan with simple interest using an arithmetic sequence. ☐
See Example 14 and Exercise 4D Question 1
- 4D** **11** I can model a taxi fare using an arithmetic sequence. ☐
See Example 15 and 10 and Exercise 4D Question 6
- 4D** **12** I can model flat-rate and unit-cost depreciation using an arithmetic sequence. ☐
See Example 16, 17 and 18 and Exercise 4D Question 8, 9 and 10
- 4D** **13** I can model general situations using an arithmetic sequence. ☐
See Example 19 and Exercise 4D Question 11
- 4E** **14** I can generate a geometric sequence with a recurrence relation. ☐
See Example 20 and Exercise 4E Question 1
- 4E** **15** I can graph a geometric sequence. ☐
See Example 21 and Exercise 4E Question 6
- 4F** **16** I can use a general rule for a geometric sequence. ☐
See Example 22 and 23 and Exercise 4F Question 2 and 6
- 4F** **17** I can determine a geometric sequence from two terms. ☐
See Example 24 and Exercise 4F Question 7
- 4F** **18** I can determine the number of terms required to reach a value in a geometric sequence. ☐
See Example 25 and Exercise 4F Question 8
- 4F** **19** I can graph a geometric sequence. ☐
See Example 26 and Exercise 4F Question 9

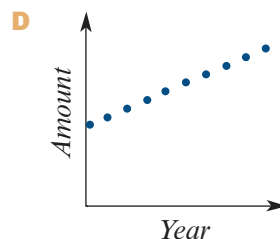
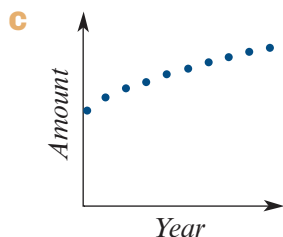
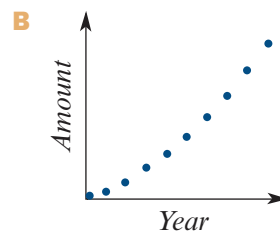
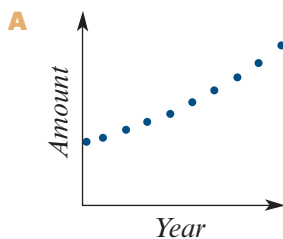
- 4G** **20** I can model the growth of a bacterial population using geometric sequences. ☐
See Example 27 and Exercise 4G Question 1
- 4G** **21** I can model compound interest using geometric sequences. ☐
See Example 28 and Exercise 4G Question 6
- 4G** **22** I can calculate the value of an asset using reducing-balance depreciation. ☐
See Example 29 and Exercise 4G Question 12
- 4G** **23** I can apply geometric sequences to practical problems. ☐
See Example 30 and Exercise 4G Question 16

Multiple-choice questions

- Determine which of the following could be the first five terms of an arithmetic sequence.
 - 2, 4, 2, 4, 2
 - 1, 10, 100, 1000, 10 000
 - 189, -89, 11, 111, 211
 - 1, 4, 9, 16, 25
- Determine which of the following is not an arithmetic sequence.
 - 11, 2, -8, -19, ...
 - 4, 7, 10, 13, ...
 - 57, 51, 45, 39, ...
 - 3, -5, -7, -9, ...
- The first term of a sequence is 3. Each subsequent term is 0.6 times the previous term. The sixth term, correct to two decimal places, is
 - 0.14
 - 0.23
 - 6.92
 - 7.15
- The n th term of the sequence defined by the recurrence relation $t_1 = 50$ and $t_{n+1} = 2t_n$ is:
 - $t_n = 50 \times 2^{-n}$
 - $t_n = 50 \times 2^{n-1}$
 - $t_n = 50 \times 2^{2n-1}$
 - $t_n = 50 \times 2^{2n}$
- The ninth term of the arithmetic sequence 44, 41, 38, ... is:
 - 8
 - 17
 - 20
 - 23

- 6** The following is a geometric sequence: $15, -45, 135, -405, \dots$. The common ratio r is equal to:
- A** -3 **B** -2 **C** 1 **D** 2
- 7** A recurrence relation is defined by $T_{n+1} = 4T_n + 5$, where $T_1 = 3$. T_4 is equal to:
- A** 73 **B** 75 **C** 85 **D** 297
- 8** The rungs of a ladder diminish uniformly in length from 30 cm at the bottom of the ladder to 22.5 cm at the top of the ladder. There are 16 rungs altogether. The length, in centimetres, of the 10 th rung up the ladder is:
- A** 24.5 **B** 25.0 **C** 25.3 **D** 25.5
- 9** The sequence generated by the recurrence relation $V_0 = 5, V_{n+1} = V_n - 3$ is:
- A** $5, 15, 45, 135, 405, \dots$
B $5, 8, 11, 14, 17, \dots$
C $5, 2, -1, -4, -7, \dots$
D $5, 15, 45, 135, 405, \dots$
- 10** Brian has two trees in his backyard. Every month, he will plant three more trees. A recurrence relation model, T_n , for the number of trees in Brian's backyard at the start of month n , is:
- A** $T_1 = 2, T_{n+1} = 3T_n$
B $T_1 = 2, T_{n+1} = 3T_n + 3$
C $T_1 = 2, T_{n+1} = T_n + 3$
D $T_1 = 2, T_{n+1} = T_n - 3$
- 11** Jennifer invests $\$2000$ with a bank. She will be paid simple interest at the rate of 5.1% per annum. If V_n is the value of Jennifer's investment at the start of the n th year, the recurrence relation model for Jennifer's investment is:
- A** $V_1 = 2000, V_{n+1} = V_n + 5.1$
B $V_1 = 2000, V_{n+1} = 5.1 \times V_n$
C $V_1 = 2000, V_{n+1} = 0.051V_n + 102$
D $V_1 = 2000, V_{n+1} = V_n + 102$
- 12** A sequence is generated from the recurrence relation $V_1 = 40, V_{n+1} = V_n - 16$. The rule for the value of the term V_n is:
- A** $V_n = 40n - 16$
B $V_n = 56 - 16n$
C $V_n = 40n$
D $V_n = 40 + 16n$

- 13** A computer is depreciated using a flat-rate depreciation method. It was purchased for \$2800 and depreciates at the rate of 8% per annum. The amount of depreciation after 4 years is:
- A** \$224 **B** \$448 **C** \$672 **D** \$794
- 14** A car purchased for \$18 990 is depreciated using a unit-cost depreciation method. After travelling a total of 20 000 kilometres, it has an estimated value of \$15 990. The depreciation amount, per kilometre, is:
- A** \$0.15 **B** \$0.80 **C** \$0.95 **D** \$6.67
- 15** A population of penguins is decreasing by 8% every year. There are currently 2700 penguins in the population. The number of penguins in the population after n years, A , is:
- A** 2700×1.8^n **B** 2700×1.08^n **C** 2700×0.92^n **D** $2700 + 1.08^n$
- 16** Sandra invests \$6000 in an account that pays compounding interest at the rate of 4.57% per annum. The number of years it takes for the investment to first exceed \$8000 is:
- A** 5 **B** 6 **C** 7 **D** 8
- 17** An investment of \$50 000 is made at a fixed rate of interest compounding annually over a number of years. Which graph best represents the value of the investment at the end of each year?



- 18** After 10 years, a compound interest investment of \$8000 earned a total of \$4000 in interest. The annual interest rate of this investment was closest to:
- A** 2.5% **B** 4.14%
C 5.03% **D** 7.2%
- 19** The second and fifth terms of a geometric sequence are -24 and 1536 , respectively. The rule for the n th term is:
- A** $t_n = 6 \times (-4)^{n-1}$ **B** $t_n = 6 \times 4^n$
C $t_n = 6 \times 4^{n-1}$ **D** $t_n = 4 \times 3^{n-1}$
- 20** The seventeenth and nineteenth terms of an arithmetic sequence are -28 and -102 , respectively. The rule for the n th term is:
- A** $t_n = 564 - 37(n - 2)$
B $t_n = 564 - 37(n + 1)$
C $t_n = 564 + 37(n + 1)$
D $t_n = 564 - 37(n - 1)$

Short-response questions

- 1** Consider the value of t_1 and d given for each arithmetic sequence and determine the required term.
- a** $t_1 = 6$ and $d = 5$. Determine t_9 . **b** $t_1 = 20$ and $d = 4$. Determine t_{10} .
- 2** Consider the value of t_1 and r given for each geometric sequence and determine the required term.
- a** $t_1 = 20$ and $r = 2$. Determine t_5 . **b** $t_1 = 2000$ and $r = 0.25$. Determine t_5 .
- 3** In an arithmetic sequence with $t_5 = 22$ and $t_{10} = 47$, determine t_{15} .
- 4** Write down the first three terms for the sequences defined by the following recurrence relations.
- a** $t_n = t_{n-1} + 6$, with $t_1 = 6$.
b $t_n = 0.5t_{n-1}$, with $t_1 = 1000$.
- 5** Write down a rule for the n th term of the sequence defined by the following recurrence relations.
- a** $t_n = t_{n-1} - 5$, with $t_1 = 8$
b $t_n = 3t_{n-1}$, with $t_1 = 1$

- 6** A sequence is generated from the recurrence relation $t_1 = 120, t_{n+1} = \frac{1}{4}t_n$
- Write down a rule for the value of the n th term of this sequence.
 - Use the rule to find t_3 .
 - Use the rule to find the value of t_{14} .
- 7** A car was purchased for \$38 500. It depreciates in value at a rate of 9.5% per year, using a reducing-balance depreciation method.
- Write down a rule for the value of the car after n years.
 - Use the rule to find the value of the car after five years.
 - Find the total depreciation of the car over five years.
- 8** Jack borrows \$20 000 from a bank and is charged simple interest at a rate of 9.4% per annum. Let t_n be the value of the loan after n years.
- Write down a rule for the value of the loan after n years.
 - Determine much Jack will need to pay the bank after 5 years.
 - Determine the number of years it takes for the value of Jack's loan to reach \$40 680.
- 9** A commercial cleaner bought a new vacuum cleaner for \$1650. The value of the vacuum cleaner decreases by \$10 for every 50 offices that it cleans.
- Determine the amount of depreciation for cleaning one office.
 - Write down a rule for the value of the vacuum cleaner after n offices are cleaned.
 - The cleaner has a contract to clean 10 offices per night, 5 nights a week for 40 weeks in a year. Determine the value of the vacuum cleaner after one year.
- 10** Kelly bought her current car five years ago for \$22 500. She is considering two different depreciation methods.
- First, she considers flat-rate depreciation of 12% per annum. Second, she considers reducing-balance depreciation of 16% per annum. Let A be the value of Kelly's car using flat-rate depreciation after n years and let B be the value of Kelly's car using reducing-value depreciation after n years.
- Write down an expression for Kelly's car using each method after n years and then sketch a graph of the value of Kelly's car for both methods on the same set of axes.
- 11** Meghan has \$5000 to invest. Company A offers her an account paying 6.3% per annum simple interest while Company B offers her an account paying 6.1% per annum compound interest, compounding monthly.
- Determine how much she would have after 5 years from each company. Determine, correct to one decimal place, the simple interest rate that Company A should offer if the two investments are to have equal value after 5 years.

12 An iron ore smelting works has a tall chimney stack from which a pollutant gas is emitted at a rate of 1500 kilograms per day. New technology has been developed that enables the emissions to be reduced in stages to a minimum of 200 kilograms per day. There are two methods of installing new equipment to reduce the emissions.

a Using the *first* installation method, the emissions will be reduced by a *constant amount* each day until the minimum emission of 200 kilograms per day is reached. Consider the case where the emissions are reduced by 130 kilograms each day. The installation will be completed by the end of the 10th day, and from the 11th day the emissions will be 200 kilograms per day. Use this information to complete the table below.

Day	1	2	3	4	5	6	7	8	9	10
Emission each day (kilograms)	1500	1370	1240	...	...	...	...	...	...	330

b Now suppose that the installation is to be completed by the end of the eighth day so that from the ninth day, the emission will be 200 kilograms per day. By what *constant amount* must the emission be reduced each day during the installation period?

c Using the *second* installation method, the emissions will be reduced by a *constant percentage* each day until the minimum emission of 200 kilograms per day is reached. Consider the case where the constant percentage is 25%. While the daily emissions are being reduced, the emissions each day will form a geometric sequence.

i Write down the common ratio of this geometric sequence.

ii Complete the table below, giving the entries correct to two decimal places.

Day	1	2	3	4	5	6	7	8	9	10
Emission each day (kilograms)	1500	1125	843.75	...	...					

d In the case described in part **c**, determine on which day the daily emission will first reach the minimum of 200 kilograms, within one kilogram.

e If the case described in part **d** is used rather than the case described in part **a**, find how much less, to the nearest kilogram, the total emission during days 1 to 10 (inclusive) is.

f Now suppose that the second installation method is used, but the minimum daily emission of 200 kilograms is not reached until the 10th day. Determine the *constant percentage* that emissions must be reduced by each day in this case. Give your answer correct to one decimal place.

Chapter 5

Earth geometry and time zones

Chapter questions

UNIT 3 BIVARIATE DATA AND TIME SERIES ANALYSIS, SEQUENCES AND EARTH GEOMETRY

Topic 5: Earth Geometry and time zones

- ▶ How do we define a great circle?
- ▶ How do we use latitude and longitude to describe a location on Earth?
- ▶ How do we find the latitude and longitude of a location on Earth?
- ▶ How do we calculate the distance between two places on Earth?
- ▶ How do we find time differences between two places on Earth using time zones?
- ▶ How do we solve problems in time planning associated with time differences between two places on Earth?

In this chapter we consider how to locate positions on Earth's surface given latitude and longitude, and how to calculate the distance between two points on the same meridian and between two points with the same latitude. We also consider the link between longitude and the time at a locality and use this to help calculate the time difference between two locations.

5A Angle measurement and arc length

Learning intentions

- ▶ To be able to convert angle measurements in decimal form to degrees and minutes and vice-versa.
- ▶ To be able to calculate the length of an arc of a circle.

We start with a brief section to remind you how to convert angle measurements given in decimal form to degrees and minutes and vice versa, and how to calculate the length of an arc. This section will help you to understand and make calculations relating to Earth geometry.

Conversion of angle measurements

There is a provision to further increase the accuracy of angle measurements with a third measurement of angle, which is seconds, but we will not do this here. Answers will be given to the nearest minute.

There are 60 minutes in a degree. We write 56 minutes as $56'$. To get a feeling for the conversion consider the following.

- $34.50^\circ = 34^\circ 30'$
- $34.25^\circ = 34^\circ 15'$
- $34.75^\circ = 34^\circ 45'$

Changing from decimal form to degrees and minutes

Multiply the decimal part of the number by 60.

- For 34.7° , multiply 0.7° by 60. The result is $42'$ and we have $34.7^\circ = 34^\circ 42'$.
- For 34.321° , multiply 0.321° by 60. The result is $19.26'$ and we have $34.321^\circ = 34^\circ 19'$ to the closest minute.

Changing from degrees and minutes to decimal form

Divide the minutes by 60.

- For $34^\circ 56'$, divide 56 by 60. The result is $0.9333 \dots$ and we have $34^\circ 56' = 34.93^\circ$, correct to two decimal places.
- For $54^\circ 19'$, divide 19 by 60. The result is $0.31666 \dots$ and we have $54^\circ 19' = 54.317^\circ$, correct to three decimal places.



Example 1 Converting angle measurements

- a Change 32.45° to degrees and minutes.
- b Change $44^\circ 32'$ to decimal form. Give your answer to two decimal places.

Solution

- a** $0.45^\circ = (0.45 \times 60) = 27'$ Therefore,
 $32.45^\circ = 32^\circ 27'$.
- b** $32' = (32 \div 60) = 0.533 \dots$ Therefore,
 $44^\circ 32' = 44.53^\circ$ correct to two decimal places.

Explanation

Multiply the decimal part of the number by 60.

Divide the minutes part of the angle by 60.

Calculator activity 5A Converting angle measurements on a calculator

Solve the following using a calculator.

- a** Change 32.45° to degrees and minutes.
b Change $44^\circ 32'$ to decimal form.

Casio fx82

- a** Type 32.45, then press $\boxed{\circ} \boxed{'} \boxed{''}$ $\>$ $=$ to get the degrees and minutes form.
 This gives the result $32.45^\circ = 32^\circ 27'$
- b** Type 44 $\boxed{\circ} \boxed{'} \boxed{''}$ $\>$ 32 $\>$ $\boxed{\circ} \boxed{'} \boxed{''}$ $\>$ $\boxed{\circ} \boxed{'} \boxed{''}$ $\>$ $=$ $\>$ to get the decimal form.
 This gives the result $44^\circ 32' = 44.533 \dots$

TI-30XB

- a** Type 32.45, then press 2nd $\>$ [Angle]. Scroll down the list to find $\boxed{D^\circ M' S''}$ $\>$ enter.
 This gives the result $32^\circ 27'$.
- b** Type 44 then press 2nd $\>$ [Angle] $\>$ 1 $\>$ type 32 $\>$ [Angle] $\>$ 2 $\>$ enter.
 This gives the result $44.533 \dots$

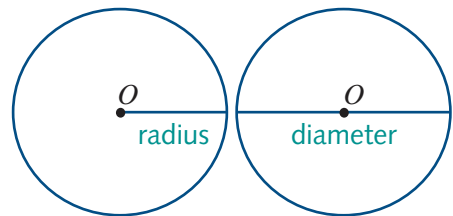
Sharp

- a** Type 32.45 press $\boxed{2^{nd}F}$ $\>$ $\boxed{D^\circ M' S''}$ $\>$ $\boxed{\leftrightarrow DEG}$.
 This gives the result $32^\circ 27'$.
- b** Type 44, then press $\boxed{D^\circ M' S''}$ $\>$ type 32 $\>$ $\boxed{D^\circ M' S''}$ $\>$ $\boxed{2^{nd}F}$ $\>$ $\boxed{D^\circ M' S''}$ $\>$ $\boxed{\leftrightarrow DEG}$.
 This gives the result $44.533 \dots$

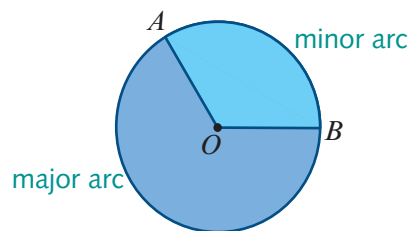
Arc length

Any line segment drawn from the centre of a given circle to any point on the circle is called a **radius** (plural radii).

Any line segment joining two points on the circle and passing through the centre is called the **diameter** of the circle.



Any two points on a circle divide the circle into **arcs**.
The shorter arc is called the *minor arc*; the longer is the *major arc*.



The arc ACB is said to subtend the angle $\angle AOB$ at the centre of the circle. If $\angle AOB = \theta$ and radius length is r units, then the length of arc ACB will be a fraction of the circumference.

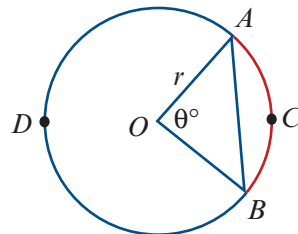
The fraction of the circumference will be $\frac{\theta}{360}$.

Recall that the circumference, C , of a circle of radius r is given by

$$C = 2\pi r.$$

Therefore, the length, s , of an arc that subtends an angle of θ at the centre is:

$$s = \frac{\theta}{360} \times 2\pi r$$



Length of an arc

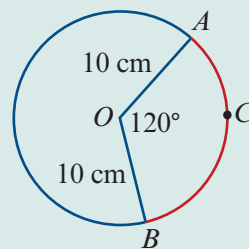
The length, s , of an arc of a circle of radius r that subtends an angle of θ at the centre is given by:

$$s = \frac{\pi r \theta}{180}$$



Example 2 Calculating the length of an arc

In this circle with centre O and radius length 10 cm, the angle subtended at O by arc ACB has magnitude 120° . Find the length of the arc ACB correct to one decimal place.



Solution

$$s = \frac{\pi r \theta}{180}$$

$$\begin{aligned} s &= \frac{\pi \times 10 \times 120}{180} \\ &= \frac{20\pi}{3} \end{aligned}$$

$$\approx 20.9 \text{ cm (correct to one decimal place)}$$

Explanation

Write down the formula for the length of an arc.

Substitute $\theta = 120^\circ$ and $r = 10$ into the equation.

Section Summary

- ▶ To change minutes to decimal form, divide the minutes by 60.
- ▶ To change decimal form of an angle to degrees and minutes, multiply the decimal part of the number by 60.
- ▶ Scientific calculators have the built-in facility to undertake the conversion from an angle expressed in decimal form to degrees and minutes.
- ▶ The length, s , of an arc of a circle of radius r that subtends an angle of θ at the centre is given by:

$$s = \frac{\pi r \theta}{180}$$

Exercise 5A

Converting angle measurements

Example 1

- 1** Convert each of the following angle measurements from decimal form to degrees and minutes.

a 32.45°

b 43.20°

c 122.46°

d 91.12°

e 0.75°

- 2** Convert each of the following angle measurements from degrees and minutes to decimal form, correct to two decimal places.

a $32^\circ 45'$

b $15^\circ 35'$

c $7^\circ 22'$

d $142^\circ 44'$

e $67^\circ 15'$

Calculating the length of an arc

Example 2

- 3** What is the circumference of each circle? Answer correct to two decimal places.

a Radius of 8 cm

b Radius of 14 m

c Radius of 45 mm

d Diameter of 12 mm

e Diameter of 14 m

- 4** What fraction of a circle is each sector?

a Angle at the centre is 90°

b Angle at the centre is 270°

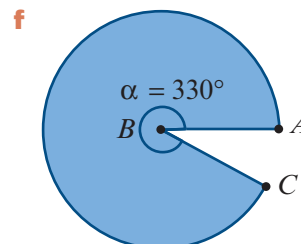
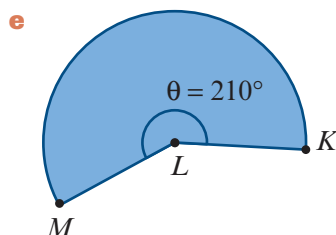
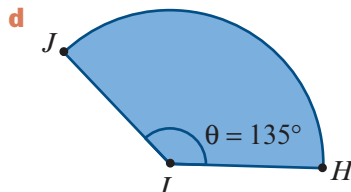
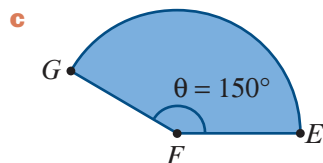
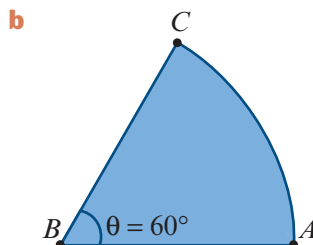
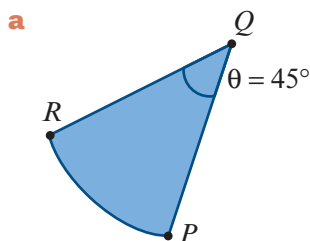
c Angle at the centre is 30°

d Angle at the centre is 120°

e Angle at the centre is 60°

f Angle at the centre is 150°

- 5 Determine the arc length of each sector. The radius is 10 cm. Answer correct to two decimal places.



- 6 Determine the arc length where the radius of the circle (given in cm) and the angle subtended at the centre are as given. Give your answer correct to two decimal places.

a $r = 15, \theta = 50^\circ$

b $r = 20, \theta = 15^\circ$

c $r = 30, \theta = 150^\circ$

d $r = 16, \theta = 135^\circ$

e $r = 40, \theta = 175^\circ$

f $r = 30, \theta = 210^\circ$

- 7 Determine the arc length that subtends an angle of magnitude 105° at the centre of a circle of radius 25 cm.

- 8 Determine the size of the angle subtended at the centre of a circle of radius length 30 cm by an arc length of:

a 50 cm

b 25 cm

- 9 A chord of length 6 cm is drawn in a circle of radius 7 cm. Determine the length of the minor arc cut off by the chord.

5B Latitude and longitude

Learning intentions

- ▶ To understand the meaning of great circles.
- ▶ To understand the meaning of latitude and longitude in relation to the equator and the prime meridian respectively.
- ▶ To locate positions on the Earth's surface given latitude and longitude.
- ▶ To be able to calculate the angular distance and distance between two places on earth on the same meridian.
- ▶ To be able to calculate the angular distance and distance between two places on Earth on the same parallel of latitude.

The Earth can be modelled by a sphere of radius 6371 km.

The radius at the equator is 6378.14 km, but the radius at the poles is only 6356.75 km.

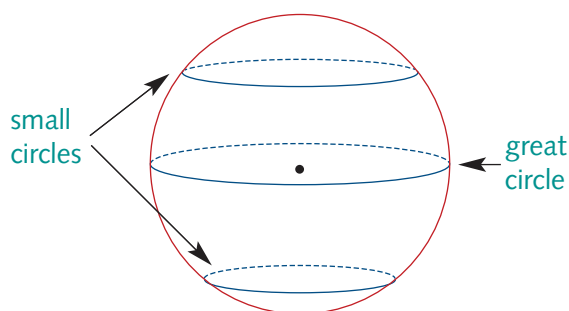
Here we will use the average radius for our calculations of distances on the Earth.

Elements of Earth geometry

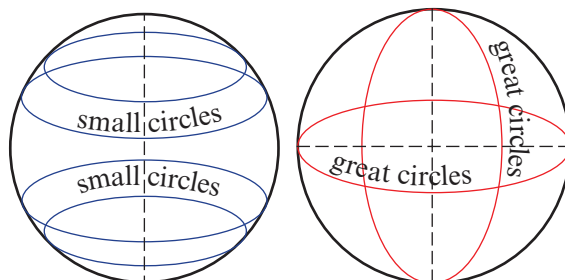
Great circles and small circles

A **great circle** is the circular boundary of a cross-section of a sphere that contains a diameter of the sphere. The cross-section contains the centre of the sphere.

The circular boundary of cross-sections of the sphere that do not contain a diameter of the sphere are called **small circles**. The cross-section for a small circle does not contain the centre of the sphere.



The great circle shown in the diagram above is in the *plane of the equator*.



The shortest distance between two points on the surface of the Earth is the distance along the great circle that passes through those two points.

The great circle path from New Delhi to New York is shown in the figure to the right.

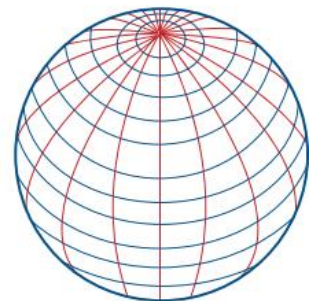


Below is a representation of the great circle route from Brisbane to London.



In this section, we are interested in describing the location of points on the surface of the Earth. We do this in a manner similar to how we described points in the plane with Cartesian coordinates/grid.

This is done using a grid of lines as shown here. They are used to give *coordinates* called **latitude** and **longitude**. The red lines will be used for longitude and the blue lines for latitude.



Meridians and parallels

Meridians of longitude are *semi-great circles* (arcs that go from pole to pole) that pass through the north and south poles.

The red lines on the sphere are meridians of longitude.

Parallels of latitude are small circles whose planes are parallel to that of the equator. The blue lines in the sphere are parallels of latitude. The equator is the only latitude that is a great circle.

Latitude

The blue lines in the diagram opposite are parallels of latitude. The latitude of a point on a sphere is the elevation of the point from the plane of the equator.

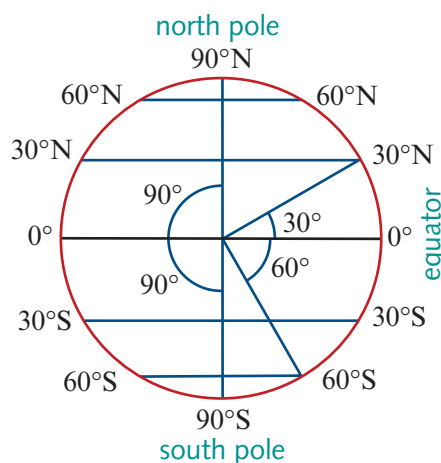
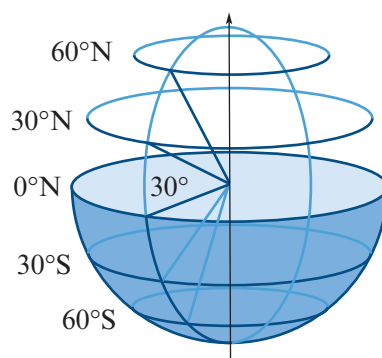
- The equator has latitude 0°N .
- The north pole has latitude 90°N .
- The south pole has latitude 90°S .

In the diagram the latitudes 60°N , 30°N , 60°S and 30°S are shown.

In the diagram opposite, the Earth has been sliced in half along a great circle. The vertical line through the poles is perpendicular, or at 90° , to the plane of the equator.

At the surface of the Earth, at a given latitude, draw a line from that location to the centre of the Earth. The angle between this line and the equator is the latitude measurement.

The diagram shows two examples: one for 30°N and one for 60°S .

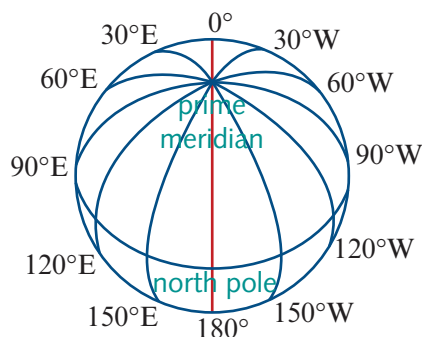
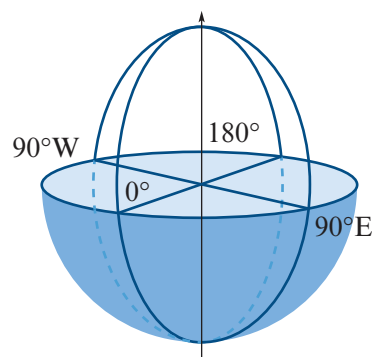


Longitude and the prime meridian

Lines of longitude are measured in degrees east or west of the **prime meridian** (0°). The lines of longitude shown in the diagram opposite are 0° , 90°E , 180° and 90°W . Note that you don't need to add E or W to the 0° or the 180° .

The prime meridian passes through Greenwich in England.

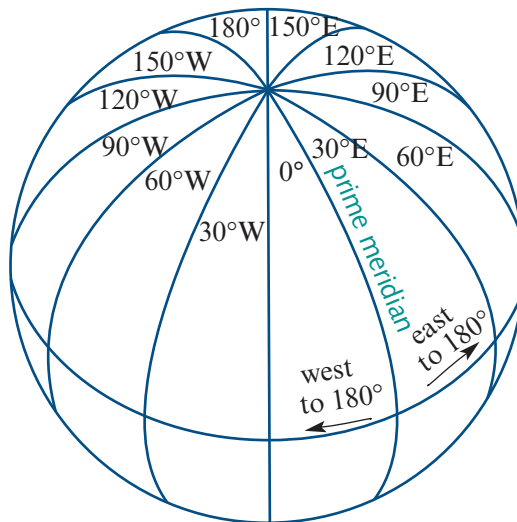
The diagram opposite is a diagram of the Earth looking down from the north pole. The evident plane in the diagram above is the plane of the equator. The angle formed between the prime meridian and the line from the centre of the Earth to the point where the meridian of longitude meets that plane is the longitude.



The meridian 120°W is on the same great circle as the meridian 60°E .

The meridian 30°W is on the same great circle as the meridian 150°E .

This is shown in a different way in the diagram below.



Latitude and longitude

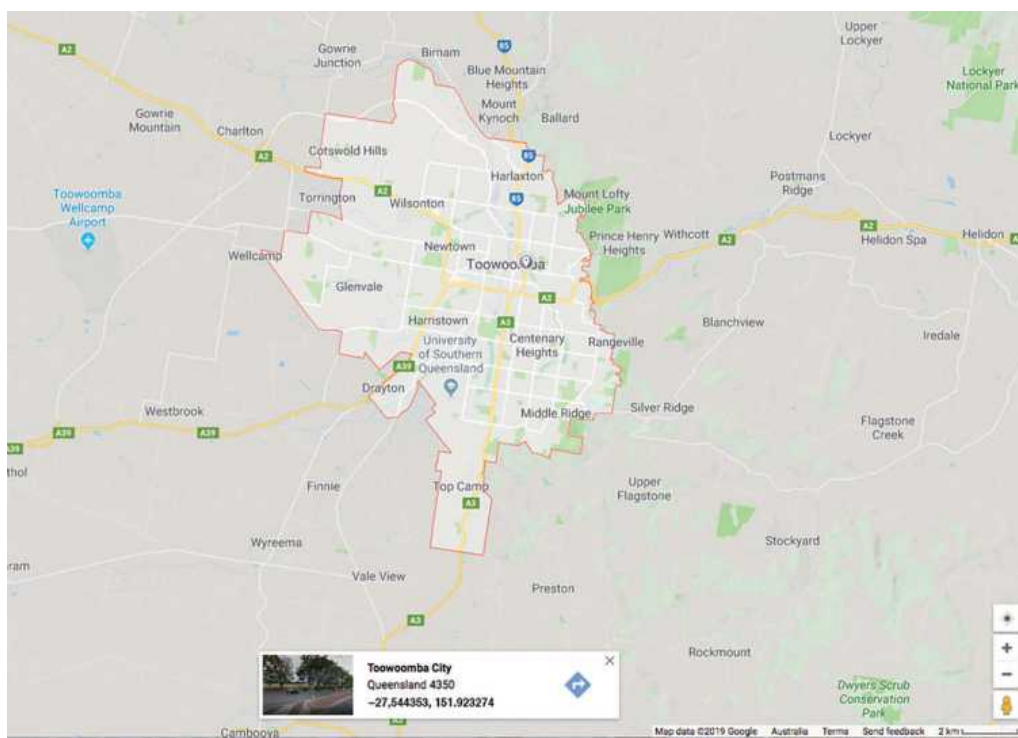
Any point on the Earth's surface can be described by giving its latitude and longitude. For example, Brisbane has latitude 27.4698°S and longitude 153.0251°E . Townsville has latitude 19.2590°S and longitude 146.8169°E . These are called the coordinates of the location.

Hemispheres

- Northern hemisphere, the half of the Earth that lies north of the equator. Locations in the northern hemisphere, such as London, have their latitude described using $^\circ\text{N}$. For London this is 51.5°N .
- Southern hemisphere, the half of the Earth that lies south of the equator. Locations in the southern hemisphere, such as Brisbane, have their latitude described using $^\circ\text{S}$. For Brisbane this is 27.5°S .
- Western and eastern hemispheres are defined through the prime meridian. Brisbane is in the eastern hemisphere and New York in the western hemisphere.

Finding the longitude and latitude of a location

Here is a map of a section of South East Queensland and northern NSW obtained from Google Maps. By clicking on a location, information is obtained including the latitude and longitude. (The precise method will depend on your device and browser or app; check Google Maps Help if needed.) In the diagram of the map below, the details for Toowoomba are shown.



Map Data © 2025 Google

The coordinates of Toowoomba are given as $(-27.544353, 151.923274)$. Changing this form to the hemisphere notations gives 27.5598°S , 151.9507°E and changing to degrees and minutes gives $27^{\circ}34'\text{S}$, $151^{\circ}57'\text{E}$. This may be done with any location. Other methods of finding a location include using a Global Positioning System (GPS) or using an atlas.

GPS

GPS is a system used for worldwide navigation and surveying. It is commonly used for determining an exact location anywhere on Earth by calculating the distances from multiple known satellite positions. This is made possible by the network of 24 man-made satellites, called GPS satellites.

GPS was originally used for military purposes, but was made available for wider public use in the 1990s. The system provides latitudes and longitudes given to a high degree of accuracy. For example, through Google Maps you could enter an address and the GPS coordinates will be given for the place, or vice-versa.

Distance along a meridian

On a flat surface, the shortest distance between two points is a straight line. Since the Earth's surface is curved, the shortest distance between A and B is the arc length AB of the great circle (the meridian) that passes through A and B . This is called the *great circle distance*.

We can calculate the distance between two points on Earth using the difference in their latitudes. Great circles of Earth have a radius of about 6371 km, so their circumference is $2 \times \pi \times 6371 \approx 40\,030$ km.

If two points subtend an angle of 1° at the centre of a great circle, the distance between them is:

$$\frac{1}{360} \times 40\,030 \approx 111.2 \text{ km}$$

Note: This result will be used throughout the chapter.

Angular distance with respect to a meridian

Beijing (China) and Perth (Australia) have coordinates $(40^\circ\text{N}, 116^\circ\text{E})$ and $(32^\circ\text{S}, 116^\circ\text{E})$, respectively. These two cities have the same longitude to the nearest degree. As the cities are in different hemispheres, north and south, we need to add the latitudes to determine the angular distance. The angular distance $= (40 + 32) = 72^\circ$.

For Brisbane and Coffs Harbour, the latitudes are 27°S and 30°S , respectively (to the nearest degree). As the cities are in the same hemisphere, south, we need to subtract the latitudes to determine the angular distance. The angular distance $= (30 - 27) = 3^\circ$. The cities are in the same hemisphere.

For locations on the same meridian in different hemispheres, we add the latitudes. For locations on the same meridian in the same hemisphere, we find the difference between the latitudes (always subtract smaller from larger).

The distance, D km, between two points on the same meridian is given by:

$$D = 111.2 \times \text{angular distance}$$



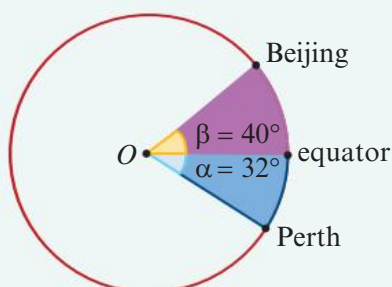
Example 3 Finding a distance along a meridian

Beijing (China) and Perth (Australia) have coordinates $(40^\circ\text{N}, 116^\circ\text{E})$ and $(32^\circ\text{S}, 116^\circ\text{E})$, respectively.

Calculate the shortest distance between Beijing and Perth, to the nearest kilometre.



Solution



$$\text{Angle} = (40 + 32) = 72^\circ$$

Therefore, the distance along the meridian
 $\approx 111.2 \times 72$
 $= 8006 \text{ km}$

Explanation

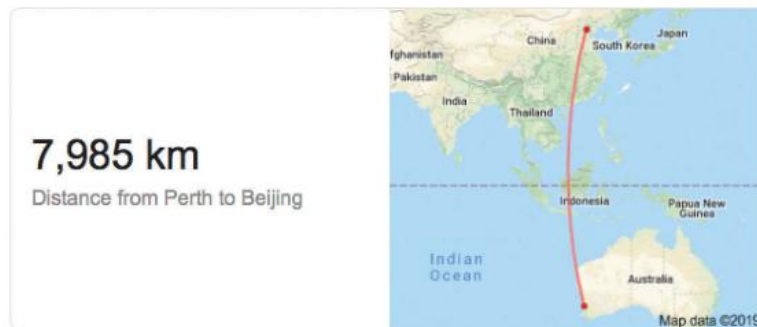
The two cities have the same longitude correct to the nearest degree. Therefore, they are on the great circle that is the meridian of longitude 116°E .

Add the latitudes of each city to find the angle subtended at the centre of the arc. As an angle of 1° at the centre of the great circle is subtended by 111.2 km on Earth, calculate the length of the arc by multiplying the angle measurement by 111.2 km.

Note: The calculated distance given on the internet is 7985 km, which is quite close to our approximate result.

Finding the distance along a meridian using a search engine

In a search engine, for example, Google, type 'Distance from Perth to Beijing'. The shortest arc along the great circle is shown.



Map Data © 2025 Google

Distance between two points on the equator

The equator is a great circle and therefore the distance between points on the equator can be found using our knowledge of arc length.



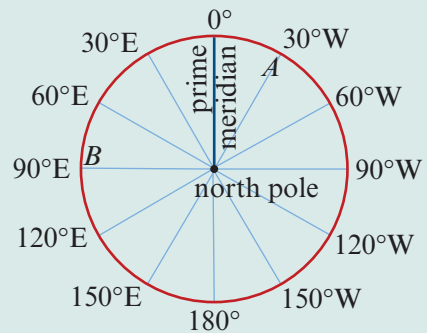
Example 4 Finding the distance between two points on the equator

Point A has longitude 30°W and latitude 0° .

Point B has longitude 90°E and latitude 0° .

Determine the distance between the two points:

- a** if you fly east from A to B
- b** if you fly west from A to B.



Solution

a $\theta = (30 + 90) = 120^\circ$

$$\begin{aligned} D &= 111.2 \times 120 \\ &= 13\,344 \text{ km} \end{aligned}$$

Flying east, the distance between A and B is 13 344 km.

b $\theta = (360 - 120) = 240^\circ$

$$\begin{aligned} D &= 111.2 \times 240 \\ &= 26\,688 \text{ km} \end{aligned}$$

Flying west, the distance between A and B is 26 688 km.

Explanation

Flying east, the angle between the locations is 120° .

We can use the formula obtained for locations on the same meridian because we are on a great circle.

$$D = 111.2 \times \text{angular distance}$$

Write the answer in a sentence.

Flying west, the angle is 240° , found by subtracting 120° from 360° .

Use the formula to calculate the distance between the points.

$$D = 111.2 \times \text{angular distance}$$

Write the answer in a sentence.


Example 5 Finding a distance between two places on the equator

The cities of Pontianak (Indonesia) and Quito (Ecuador) are on the equator to the nearest degree. The longitude of Pontianak is 109°E and the longitude of Quito is 78°W . Determine the distance between the two cities flying east from Pontianak to Quito, correct to the nearest whole number.


Solution

$$\theta = 78 + 109 = 187^{\circ}$$

Hence, the required angle is

$$\theta = 360 - 187 = 173^{\circ}$$

$$\begin{aligned} D &= 111.2 \times 173 \\ &= 19\,238 \text{ km} \end{aligned}$$

The approximate distance when flying east from Pontianak to Quito is 19 238 km.

Explanation

Add to find the total angle between the longitudes – we are considering the distance covered by flying west.

Determine the angle between the two longitudes if flying east.

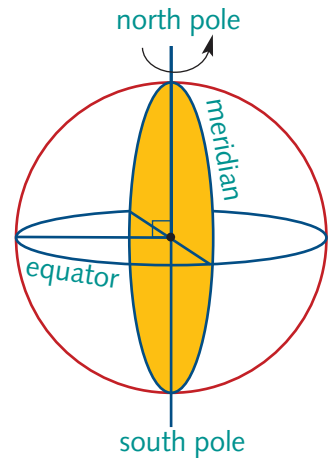
Use the formula to calculate the distance between Pontianak and Quito if flying east.

$$D = 111.2 \times \text{angular distance}$$

Write the answer in a sentence.

Distance from a pole or from the equator

All meridians pass through the poles and a point on the equator. Therefore, we can find the distance from any point on the surface of the Earth to a pole or the equator if we know its latitude.



Example 6 Finding the distance to the equator or a pole

Brisbane has latitude 27°S and longitude 153°E . Determine the distance of Brisbane to:

a the equator

b the south pole

c the north pole.

Solution

$$\begin{aligned}\mathbf{a} \quad D &= 111.2 \times 27 \\ &= 3002.4 \text{ km}\end{aligned}$$

The approximate distance between Brisbane and the equator is 3002.4 km.

$$\begin{aligned}\mathbf{b} \quad D &= 111.2 \times 63 \\ &= 7005.6 \text{ km}\end{aligned}$$

The approximate distance between Brisbane and the south pole is 7005.6 km.

$$\begin{aligned}\mathbf{c} \quad D &= 111.2 \times 117 \\ &= 13\,010.4 \text{ km}\end{aligned}$$

The approximate distance between Brisbane and the north pole is 13 010.4 km.

Explanation

Along the plane of the meridian 153°E , the difference between the equator and Brisbane is $(27 - 0 = 27^\circ)$. Use the formula $D = 111.2 \times \text{angular distance}$. Write the answer in a sentence.

Along the plane of the meridian 153°E , the difference between the south pole and Brisbane is $(90 - 27 = 63^\circ)$. Use the formula $D = 111.2 \times \text{angular distance}$. Write the answer in a sentence.

Along the plane of the meridian 153°E , the difference between the north pole and Brisbane is $(90 + 27 = 117^\circ)$. Use the formula $D = 111.2 \times \text{angular distance}$. Write the answer in a sentence.

Note: The distance from the north pole to the south pole along a great circle is approximately $\pi \times 6371 \approx 20015$ km

Distance along a parallel of latitude

In the following example, we find out how to find the distance between two points with the same latitude. We will illustrate this by finding the distance between Rockhampton and Alice Springs.



Example 7 Calculating distance along a parallel of latitude

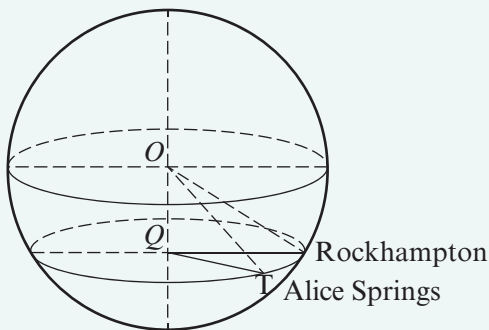
Rockhampton, Queensland, has latitude 23°S and longitude 150°E .

Alice Springs, Northern Territory, has latitude 23°S and longitude 134°E .

Determine the distance along the parallel of latitude 23°S from Rockhampton to Alice Springs, correct to the nearest kilometre.



Solution



Using right-angled triangle OTQ , the radius (QT) of the small circle of the latitude 23°S is

$$\begin{aligned}\cos 23^\circ &= \frac{QT}{6371} \\ QT &= 6371 \times \cos 23^\circ \\ &\approx 5864.54 \text{ km}\end{aligned}$$

Explanation

In the diagram of the Earth, Alice Springs and Rockhampton are shown. The circle passing through these two cities that is parallel to the equator is the parallel of latitude at 23°S .

Determine the radius of the small circle of latitude 23°S .

The Earth can be modelled by a sphere of radius 6371 km.

Continued on next page

$$\begin{aligned}\text{The required angle} &= 150^\circ - 134^\circ \\ &= 16^\circ\end{aligned}$$

$$\begin{aligned}\text{Distance} &= \frac{16}{360} \times 2 \times \pi \times 5864.54 \\ &= 1637.68 \\ &\approx 1638 \text{ km}\end{aligned}$$

These calculations are always the same for calculating distances around small circles, so we use $111.2 \cos 23 \times 16 \approx 1638$.

Determine the length of the arc connecting Alice Springs and Rockhampton.

The distance, D km, between two points that have the same parallel of latitude is given by

$$D = 111.2 \cos \theta \times \text{angular distance}$$

where the parallel of latitude is $\theta^\circ\text{N}$ or $\theta^\circ\text{S}$.



Example 8 Calculating distance along a parallel of latitude

The latitude of both Rockhampton in Queensland and Sao Paolo in Brazil is 23°S . Their longitudes are 151°E and 47°W , respectively. Determine the distance between the two cities:

- a** by flying west from Rockhampton to Sao Paolo
- b** by flying east from Rockhampton to Sao Paolo.

Solution

$$\begin{aligned}\text{a Angular distance} &= (151 + 47) = 198^\circ \\ \text{Latitude} &= 23^\circ\text{S}\end{aligned}$$

$$\begin{aligned}D &= 111.2 \cos 23^\circ \times 198 \\ &= 20\,267.31 \text{ km}\end{aligned}$$

Flying west, the distance is
20 267.31 km.

$$\text{b Angular distance} = (360 - 198) = 162^\circ$$

$$\begin{aligned}D &= 111.2 \cos 23^\circ \times 162 \\ &= 16\,582.34 \text{ km}\end{aligned}$$

Flying east, the distance is
16 582.34 km.

Explanation

For flying west, add the two angles as they are both measured from either side of the prime meridian.

Calculate the distance between the two locations using the formula.

Write the answer in a sentence.

For flying east, subtract the angle in part **a** from 360° .

Calculate the distance using the formula.

Write the answer in a sentence.


Example 9 Calculating distance using degrees and minutes

Yarraden in Queensland and Wyndham in Western Australian both have latitude 15°S . Yarraden has longitude $143^{\circ}18'\text{E}$ and Wyndham has longitude $128^{\circ}07'\text{E}$. Determine the distance along the small circle between the locations at the latitude 15°S (correct to two decimal places).

Solution

$$143^{\circ}18'\text{E} = 143.3^{\circ}\text{E}$$

$$128^{\circ}07'\text{E} = 128.12^{\circ}\text{E}$$

$$\text{Angular difference} = 15.18^{\circ}$$

$$\begin{aligned} D &= 111.2 \times \cos 15^{\circ} \times 15.18 \\ &= 1630.50 \text{ km} \end{aligned}$$

The distance between Yarraden and Wyndham is 1630.50 km.

Explanation

Convert the longitudes to decimal notation first.

The two locations are both east of the prime meridian. Find the difference between the two longitudes.

Calculate the distance using the formula.

Write the answer in a sentence.

Section Summary

- ▶ The shortest distance between two points on the surface of the Earth is the distance along the great circle that passes through those two points.
- ▶ **Meridians of longitude** are *semi-great circles* (arc that go from pole to pole) that pass through the north and south poles. **Parallels of latitude** are small circles whose planes are parallel to that of the equator. The equator is the only latitude that is a great circle.
- ▶ If two locations are both in same hemisphere (considering north or south hemispheres) and on the same meridian, the angular distance can be found by subtracting the smaller latitude of the two locations from the larger.
- ▶ If two locations are in different hemispheres (considering north or south hemispheres) and on the same meridian, the angular distance can be found by adding the two latitudes.
- ▶ To calculate the distance between two points on Earth on the same meridian, use the formula $D = 111.2 \times \text{angular distance}$ where D is the distance in kilometres.
- ▶ If two locations have the same latitude, the angular distance can be found by considering their longitudes. This will involve either adding or subtracting longitudes.
- ▶ To calculate the distance between two places on Earth on the same parallel of latitude, use the formula $D = 111.2 \cos \theta \times \text{angular distance}$ where D is distance in kilometres and where the parallel of latitude is $\theta^{\circ}\text{N}$ or $\theta^{\circ}\text{S}$.



Exercise 5B

Making latitude and longitude calculations

- 1 Use an atlas or Google Earth to name the place situated at the following coordinates.

- a** $44^{\circ}26'N, 26^{\circ}06'E$ ($44.43^{\circ}N, 26.10^{\circ}E$)
- b** $59^{\circ}54'N, 10^{\circ}45'E$ ($59.91^{\circ}N, 10.75^{\circ}E$)
- c** $29^{\circ}55'N, 95^{\circ}22'W$ ($29.76^{\circ}N, 95.36^{\circ}W$)
- d** $1^{\circ}21'N, 103^{\circ}49'E$ ($1.35^{\circ}N, 103.82^{\circ}E$)
- e** $33^{\circ}55'S, 18^{\circ}25'E$ ($33.92^{\circ}S, 18.42^{\circ}E$)

- 2 Use an atlas, Google Earth or another method to state the coordinates of:

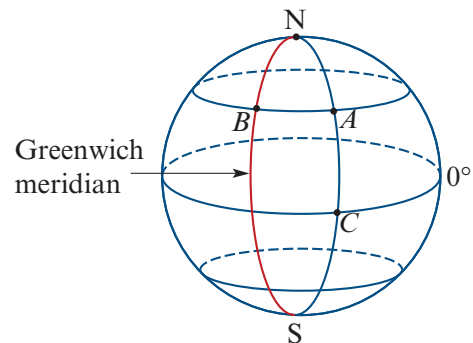
a Cairo	b Mexico City	c Buenos Aires
d Mumbai	e Lagos	f Harare

- 3 Use Google maps, an atlas or another method to find the coordinates of each of the following Queensland locations.

a St George	b Caloundra
c Mackay	d Weipa
e Bundaberg	f Charters Towers
g The Queensland and Northern Territory border (Longitude only)	

- 4 On the diagram to the right, the latitude and longitude of point A are ($65^{\circ}N, 75^{\circ}E$).

- a** What are the coordinates of point B?
- b** What are the coordinates of point C?



- 5 Athens has the coordinates ($38^{\circ}N, 24^{\circ}E$) and Sofia has the coordinates ($43^{\circ}N, 23^{\circ}E$).

- a** What is the latitude and longitude of a point 20° due south of Athens?
- b** What is the latitude and longitude of a point 20° due east of Athens?
- c** What is the latitude and longitude of a point 60° due south of Sofia?
- d** What is the latitude and longitude of a point 60° due west of Sofia?

- 6 The following table shows the latitude and longitude of cities around the world given to the nearest degree.

City	Latitude	Longitude
Brisbane	27°S	153°E
Melbourne	38°S	145°E
Cooktown	15°S	145°E
Townsville	19°S	147°E
Marseilles	43°N	5°E
Mexico City	19°N	99°W
Wellington	41°S	175°E
Zurich	47°N	8°E
London	52°N	0° (actually west of Greenwich)
Lima	12°S	77°W
Plymouth	50°N	4°W
Yangon	19°N	96°E

- a** Which city (or cities) is closest to the following latitudes?
- i** 15°N **ii** 28°S
- b** Which city (or cities) is closest to the following longitudes?
- i** 151°E **ii** 20°W
- iii** Greenwich meridian **iv** Longitude of Sofia (from Question 5)
- c** Which cities are in the northern hemisphere?
- d** Which cities are in the western hemisphere?
- e** Which cities have the same latitude?
- f** Which cities have the same longitude?
- g** Which city is closest to the north pole?
- h** Which city is closest to the south pole?

Calculating the distance between two points on the same meridian

Example 3

- 7 Two places on the same meridian have latitudes 22°N and 35°S. Determine the distance between the two places. Give your answer to the nearest kilometre.
- 8 Cairns 17°S and Melbourne 37.8°S are nearly on the same meridian. Assuming they are, determine the distance between them. Give your answer to the nearest kilometre.
- 9 How far apart are Esperance 34°S, 122°E and Broome 18°S, 122°E. Give your answer to the nearest kilometre.

- 10** Cairns and Griffith are 1920 km apart and both are on the same meridian. If the latitude of Cairns is 17°S , determine the latitude of Griffith. Griffith is south of Cairns.
- 11** Calculate the shortest distance along a meridian in each of the following cases. Give your answer to the nearest kilometre.
- a** Point X: latitude 10°N , longitude 18°W
Point Y: latitude 45°N , longitude 18°W
 - b** Point X: latitude 14°N , longitude 35°W
Point Y: latitude 13°S , longitude 35°W
 - c** Point X: latitude 23°S , longitude 140°W
Point Y: latitude 67°S , longitude 140°W
 - d** Point X: latitude 15°N , longitude 60°W
Point Y: latitude 25°S , longitude 60°W
 - e** Point X: latitude 15°N , longitude 70°W
Point Y: latitude 15°S , longitude 70°W

Calculating the distance between two points on the equator

Example 4

- 12** The difference of longitude between two points on the equator is 32° . Determine the distance between them in kilometres.

Example 5

- 13** There are places in Ecuador (South America) and Somalia (Africa) which are on the equator. In Ecuador, the longitude of a place X on the equator is 78°W and in Somalia the longitude of a place Y is 42°W . Determine the distance between them in kilometres.

- 14** Calculate the shortest distance along the equator between places A and B having the following longitudes:

a A 137°E	B 87°E	b A 57°E	B 13°W
c A 57°W	B 27°W	d A 140°E	B 160°W
e A 95°W	B 113°E		

- 15** The distance between two points on the equator is 600 km. What is the difference in their longitudes?

Finding the distance to the equator or a pole

Example 6

- 16** New Orleans has latitude 30°N and longitude 90°W . Determine the distance of New Orleans to:

- a** the equator **b** the north pole **c** the south pole.

- 17** Izmir has latitude 38°N and longitude 27°E . Determine the distance of Izmir to:

- a** the equator **b** the north pole **c** the south pole.

- 18** Ballarat has latitude 37.5500°S and longitude 143.8500°E . Determine the distance of Ballarat to:
- a** the equator **b** the north pole **c** the south pole.
- 19** Determine the distance from the given point to the equator and to each of the poles.
- a** Latitude 42°N , longitude 134°E **b** Latitude 55°N , longitude 45°W
c Latitude 15°S , longitude 35°E **d** Latitude 14°S , longitude 75°W

Calculating the distance around a parallel of latitude

- 20** Determine the radius of the small circle that is parallel to the latitude:
- a** 15°S **b** 30°S **c** 45°S **d** 60°S
- 21** For which parallels of latitude is the radius of the small circle half the radius of the equator?
- 22** Determine the distance around the parallel of latitude for the following locations.
- a** X: latitude 22°N , longitude 134°E ; Y: latitude 22°N , longitude 145°E
b X: latitude 32°S , longitude 50°E ; Y: latitude 32°S , longitude 80°E
c X: latitude 12°S , longitude 30°E ; Y: latitude 12°S , longitude 80°E
- 23** The position of Salzburg is 48°N , 13°E and the position of Seattle is 48°N , 122°W . Determine the distance around the 48°N parallel of latitude between Seattle and Salzburg.

Calculations using degrees and minutes

- 24** In the following, each pair of locations lie on the same meridian. Determine the approximate distance between them.
- a** X ($24^{\circ}15'\text{N}$), Y ($36^{\circ}35'\text{N}$) **b** X ($24^{\circ}15'\text{N}$), Y ($36^{\circ}45'\text{S}$)
c X ($15^{\circ}15'\text{S}$), Y ($26^{\circ}55'\text{N}$) **d** X ($58^{\circ}15'\text{S}$), Y ($36^{\circ}45'\text{S}$)
- 25** In the following, each pair of locations lie on the same latitude. Determine the approximate shortest distance between them travelling along the small circle.
- a** X ($124^{\circ}15'\text{E}$, $20^{\circ}15'\text{S}$), Y ($36^{\circ}45'\text{E}$, $20^{\circ}15'\text{S}$)
b X ($104^{\circ}15'\text{E}$, $32^{\circ}25'\text{S}$), Y ($28^{\circ}45'\text{W}$, $32^{\circ}25'\text{S}$)
c X ($120^{\circ}15'\text{E}$, $0^{\circ}25'\text{N}$), Y ($120^{\circ}45'\text{W}$, $0^{\circ}25'\text{N}$)
d X ($158^{\circ}15'\text{E}$, $40^{\circ}45'\text{N}$), Y ($26^{\circ}45'\text{E}$, $40^{\circ}45'\text{N}$)

Mixed exercises

- 26** An aircraft flies from Bairnsdale (38°S , 148°E) due north to Grenfell NSW (34°S , 148°E).
- a** How far is this (great circle) distance?
It then flies due west to Renmark SA (34°S , 141°E).
- b** How far is it from Grenfell to Renmark (distance around the parallel of latitude)?
- c** What is the total distance flown?

27 An aircraft flies from Warragul (38°S , 146°E) due north to Cairns, Qld (17°S , 146°E).

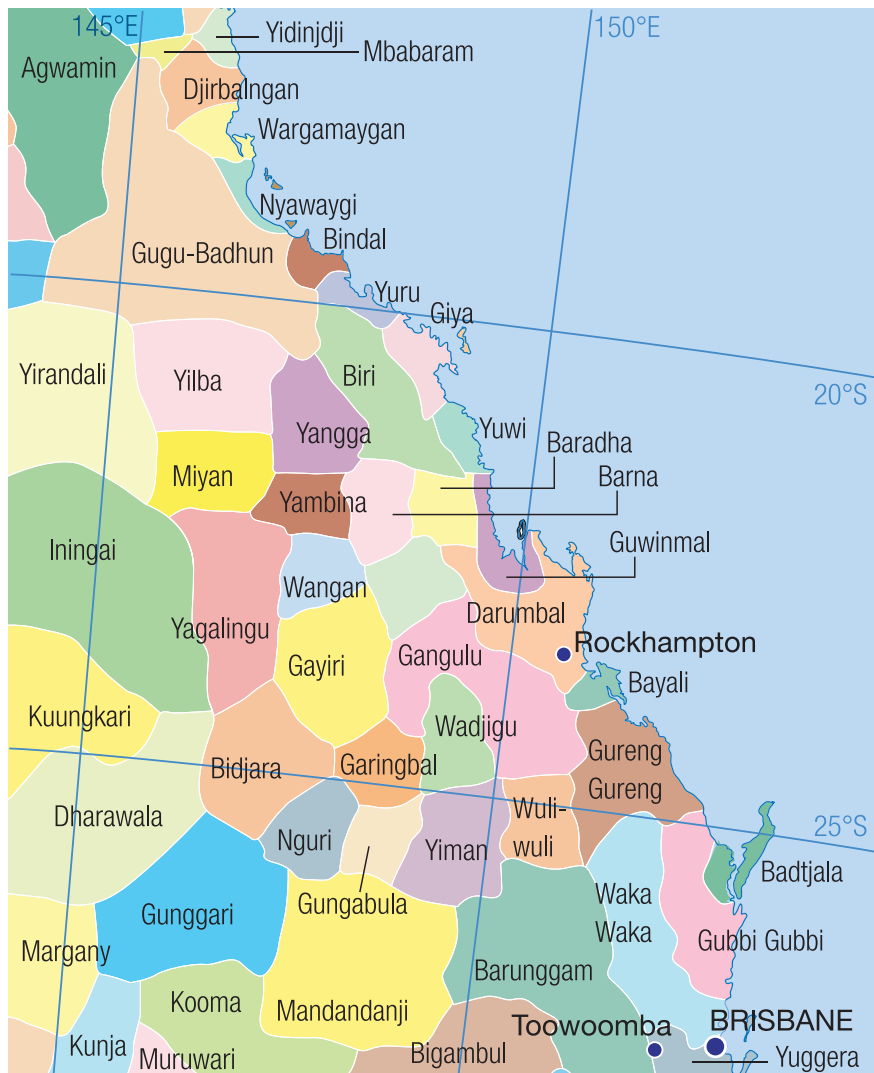
a How far is this (great circle) distance?

It then flies due west to Derby WA (17°S , 124°E).

b How far is it from Cairns to Derby (distance around the parallel of latitude)?

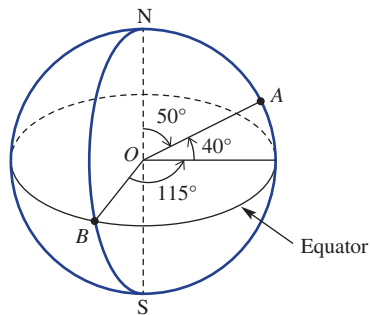
c What is the total distance flown?

28 The map shows various regions of Queensland based on language, social or nation groups of the Indigenous Australians. The lines of latitude show the parallels of latitude 20°S and 25°S and meridians 145°E and 150°E . Use this information, other resources and the map to help describe boundary positions of three chosen regions.



Paper 1-style multiple-choice questions

- 29** The actual distance between two locations has been correctly calculated as 890 km. The latitude and longitude respectively of these two cities could be
- A** 6°N 111°W and 2°S 111°W
 - B** 6°N 111°W and 2°N 111°W
 - C** 25°N 170°E and 30°S 170°E
 - D** 35°N 145°E and 43°S 145°E .
- 30** A location with coordinates $(36^{\circ}\text{S}, 26^{\circ}\text{E})$ is positioned
- A** 36° south of the prime meridian and 26° east of the equator
 - B** 36° south of the equator and 26° east of the prime meridian
 - C** 36° south of the 180° meridian and 26° east of the equator
 - D** 36° south of the equator and 26° east of the 180° meridian.
- 31** In this diagram of the Earth, O represents the centre and B lies on both the Equator and the Greenwich Meridian. What is the latitude and longitude of point A ?
- A** 40°N , 115°E
 - B** 40°N , 115°W
 - C** 50°N , 115°E
 - D** 50°N , 115°W



5C Time zones and time differences

Learning intentions

- ▶ To be able to understand the meaning of Greenwich Mean Time (GMT), International Date Line and Coordinated Universal Time (UTC).
- ▶ To be able to understand the link between longitude and time.
- ▶ To be able to determine the number of degrees of longitude for a given time difference.
- ▶ To be able to calculate time differences between two places on Earth.



The **time zones** are largely determined by the meridians of longitude. You can see from the map above that there are exceptions to this because of local requirements.

If it is 12 noon along a meridian, then on the other side of the world, along the meridian that makes up the other half of the great circle, it is midnight. For example, when it is noon in Victoria on the 145°E meridian, it is midnight in the far east of Brazil on the 35°W meridian.

Since the Earth turns 360° in 24 hours, it turns 15° in 1 hour. For every 15° of longitude, the time difference is 1 hour, and so for every 1° of longitude the time difference is 4 minutes.

15° longitude = 1 hour time difference

1° longitude = 4 minutes time difference

Local times around the world are given relative to the time along the prime meridian. The time at the prime meridian is taken as **Coordinated Universal Time (UTC)**. This is not a time zone but is the primary time standard by which the world regulates clocks and time. For most purposes, UTC is considered interchangeable with **Greenwich Mean Time (GMT)**, but GMT is no longer precisely defined by the scientific community.

In the map on the previous page, the time zones show the adjustments to UTC, which are taken in various locations in the world.

Places east of the prime meridian are ahead of GMT (UTC), while places west are behind GMT (UTC).



Example 10 Using time difference without time zones or summer time

Singapore is located at $1^{\circ}\text{N } 104^{\circ}\text{E}$ and Sydney is located at $34^{\circ}\text{S } 151^{\circ}\text{E}$. What is the time difference between Singapore and Sydney?

Solution

Difference in longitude = $151 - 104 = 47$

$47 \div 15 = 3\frac{2}{15}$ Therefore, the time difference is 3 hours. Sydney is three hours ahead.

Explanation

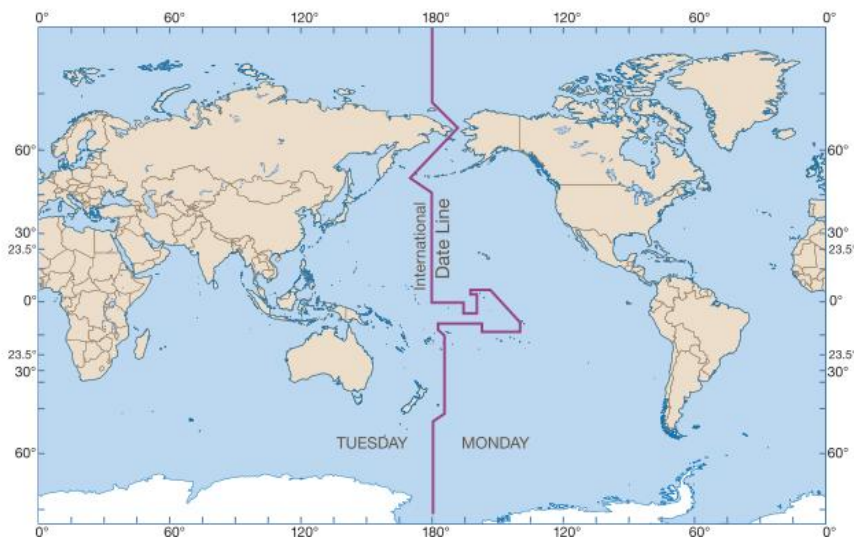
Calculate the difference in longitude.

Use 1 hour for each 15° of longitude.

International Date Line

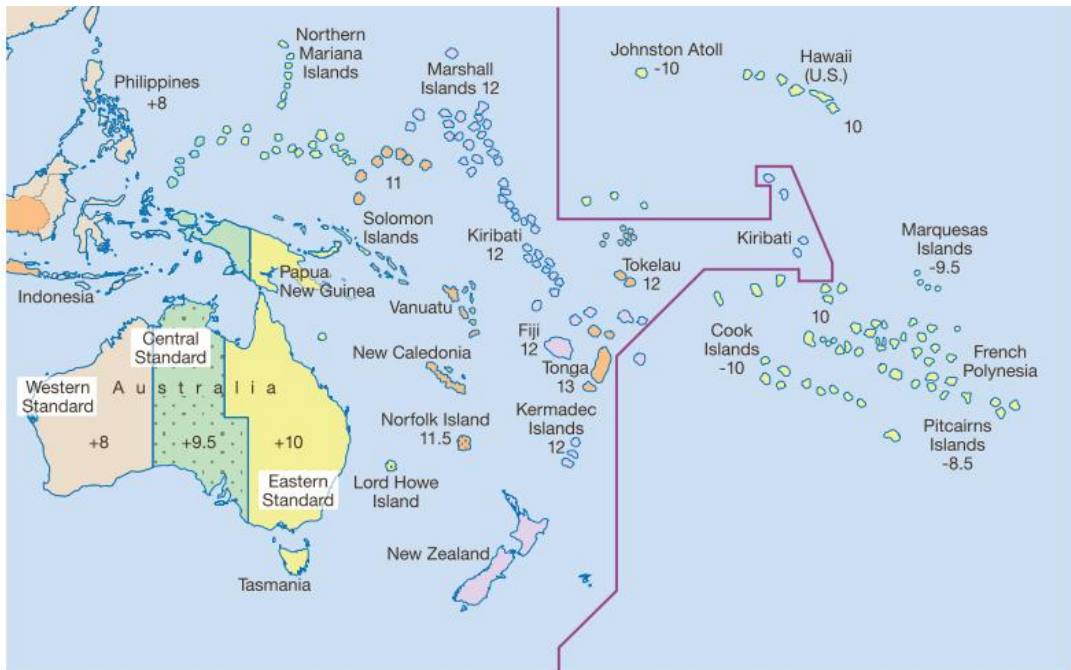
The International Date Line is an imaginary line on Earth's surface. **The International Date Line** is located halfway around the world from the prime meridian (0° longitude) at about the 180° meridian. The dateline runs from the north pole to the south pole. It is not straight but zigzags to avoid political and country borders and to not cut some countries in half. The location of the international dateline is shown in the map below. When you cross the International Date Line:

- from west to east, you subtract a day
- from east to west, you add a day.



Time zones for Australia and its neighbours

Australia is divided into three time zones: Eastern Standard, Central Standard and Western Standard. Some states change the clocks in summer to include daylight saving. Queensland time is 10 hours ahead of GMT. New Zealand time is ahead of eastern standard time by 2 hours and is 12 hours ahead of GMT.



Section Summary

- ▶ 15° longitude = 1 hour time difference.
- ▶ 1° longitude = 4 minutes time difference.
- ▶ Australia is divided into three time zones: Eastern Standard, Central Standard and Western Standard. Some states change the clocks in summer to include daylight saving.

Exercise 5C

Using time difference without time zones or summer time

Example 10

1 Give the time differences between the following places.

- a X longitude 150°E , Y longitude 120°E
- b X longitude 0° , Y longitude 75°W
- c X longitude 0° , Y longitude 75°E
- d X longitude 123°E , Y longitude 150°E

- 2** The following table shows the latitude and longitude of cities around the world.

City	Latitude	Longitude
Brisbane	27°S	153°E
Melbourne	38°S	145°E
Marseilles	43°N	5°E
Mexico City	19°N	99°W
Wellington	41°S	175°E
Zurich	47°N	8°E
London	52°N	0° (actually west of Greenwich)
Lima	12°S	77°W
Plymouth	50°N	4°W
Yangon	19°N	96°E

- a** Find the time difference between Brisbane and:
- i** Wellington
 - ii** Marseilles
 - iii** Yangon
 - iv** London
 - v** Mexico City
- b** If it is 6 a.m. in Brisbane what time is it in:
- i** Wellington?
 - ii** Marseilles?
 - iii** Yangon?
 - iv** London?
 - v** Mexico City?
- 3** The time in a Pacific island is 10 hours behind GMT.
- a** What is the longitude of the island?
 - b** What is the time in London when it is 4 a.m. on the island?
 - c** What is the time on the island when it is 5:30 a.m. in London?
- 4** The longitude of Hanoi is 105°E, while the longitude of Cape Howe is 150°E. When the time in Hanoi is 2:30 p.m., what is the time at:
- a** Cape Howe?
 - b** London?
- 5** Kalgoorlie has longitude 121°E while the Pacific island of Nauru has longitude 166°E.
- a** Calculate the difference in longitude between these two places.
 - b** Calculate the time difference between the two places.
 - c** What is the time in Nauru when it is 11:45 a.m. in Kalgoorlie?

- 6** You live in Broadbeach (28.0308°S, 153.4319°E) and want to telephone a friend in one of the places listed below at 9 a.m. on a Saturday (their time). For each city, at what time (your time) should you call?
- | | |
|------------------------------|--------------------------------|
| a Vancouver (123°W) | b Helsinki (25°E) |
| c Bologna (11°E) | d San Francisco (122°W) |
| e Vladivostok (132°E) | f Fiji (178°E) |

Using time zones to solve problems

- 7** Dimitri is in Athens, which is two hours ahead of Greenwich Mean Time. Allan is in New York, which is five hours behind Greenwich Mean Time.
- a** Dimitri is going to ring Allan at 10 a.m. on Wednesday, Athens time. What day and time will it be in New York when he rings?
- b** Allan is going to fly from New York to Athens. His flight will leave on Wednesday at 10 p.m., New York time, and will take 10 hours. What day and time will it be in Athens when he arrives?
- 8** Louise is in Dubai, which is three hours ahead of Greenwich Mean Time. David is in Sydney, which is ten hours ahead of Greenwich Mean Time.
- a** Louise is going to ring David at 10 a.m. on Wednesday, Sydney time. What day and time will it be in Dubai when she rings?
- b** David is going to fly from Sydney to Dubai. His flight will leave on Wednesday at 10 p.m., Sydney time, and will take 14 hours. What day and time will it be in Dubai when he arrives?
- 9** Los Angeles is 8 hours behind GMT and Brisbane is 10 hours ahead of GMT.
- a** If it is 10 a.m. in Brisbane on Tuesday, what time is it in Los Angeles?
- b** A plane leaves Brisbane at 10 a.m. on Tuesday and arrives in Los Angeles at 6 a.m. on Tuesday. What was the length of the flight? (Note that the plane crosses the International Date Line.)
- c** A plane leaves Los Angeles at 23:20 on Tuesday and arrives in Brisbane at 7:15 a.m. on Thursday. What was the length of the flight? (Note that the plane crosses the International Date Line.)

Using time zones of Australia and its neighbours

- 10** Use the map on page 262 to answer these questions using the time zones indicated on the map. Ignore summer time. You may need to find the location in an atlas or using the internet. If it is 12 midday in Brisbane, what time and day is it in:
- | | | |
|------------------------|-------------------------|---------------------------|
| a Melbourne? | b Alice Springs? | c Perth? |
| d Port Moresby? | e Auckland? | f Norfolk Island? |
| g Suva? | h Honolulu? | i Solomon Islands? |

Paper 1-style multiple-choice questions

- 11** The table gives the number of hours from UTC for each of the listed locations. Standard time is used and not Daylight time.

City	Number of hours on from UTC
Perth	+8
Eucla	+8:45
Adelaide	+9:30
Brisbane	+10
Lord Howe Island	+10:30

When it is 11:00 a.m. in Perth, what is the time in Lord Howe Island?

- A** 8:30 a.m.
B 12:30 p.m.
C 3:30 p.m.
D 1:30 p.m.
- 12** The local time in Delhi (29°N , 77°E) is two hours ahead of the local time in Tehran. What is the most likely longitude for Tehran?
- A** 36°E
B 65°E
C 51°E
D 120°E
- 13** Brisbane is 10 hours ahead of Greenwich in England. Nairobi in Kenya is 3 hours ahead of Greenwich. What is the time in Nairobi when it is 11 a.m. in Brisbane?
- A** 6:00 p.m.
B 4:00 a.m.
C 7:00 a.m.
D 2:00 a.m.

Key ideas and chapter summary

Length of an arc	The length, s , of an arc of a circle of radius r that subtends an angle of θ at the centre is given by $s = \frac{\pi r \theta}{180}$
Great circle	A great circle is a section of a sphere that contains a diameter of the sphere. The section contains the centre of the sphere.
Small circles	Sections of the sphere that do not contain a diameter are called small circles . A small circle does not contain the centre of the sphere.
Meridians of longitude	Meridians of longitude are semi-great circles that pass through the north and south poles.
Parallels of latitude	Parallels of latitude are small circles whose planes are parallel to that of the equator.
Australian time zones	Australia is divided into three time zones: Eastern Standard, Central Standard and Western Standard. Some states change the clocks in summer to include daylight saving.

Skills checklist



Download this checklist from the Interactive Textbook, then print it and fill it out to check your skills.


5A
1 I can change degrees and minutes to decimal form and vice versa.


See Example 1 and Exercise 5A Questions 1 and 2

5A
2 I can calculate the length of an arc of a circle.


See Example 2 and Exercise 5A Question 3

5B
3 I can calculate distances along a meridian.


See Example 3 and Exercise 5B Question 7

5B
4 I can find the distance between two points on the equator.


See Example 4, Example 5 and Exercise 5B Questions 12 and 13

5B
5 I can find the distance to the equator or a pole.


See Example 6 and Exercise 5B Question 16

- 5B** **6** I can find the distance between two points on a parallel of latitude. ☐
See Example 7, Example 8 and Exercise 5B Question 22
- 5B** **7** I can find the distance between two points on a parallel of latitude using degrees and minutes. ☐
See Example 9 and Exercise 5B Question 24
- 5C** **8** I can use time difference based on longitude to find the time difference between two locations. ☐
See Example 10 and Exercise 5C Question 1

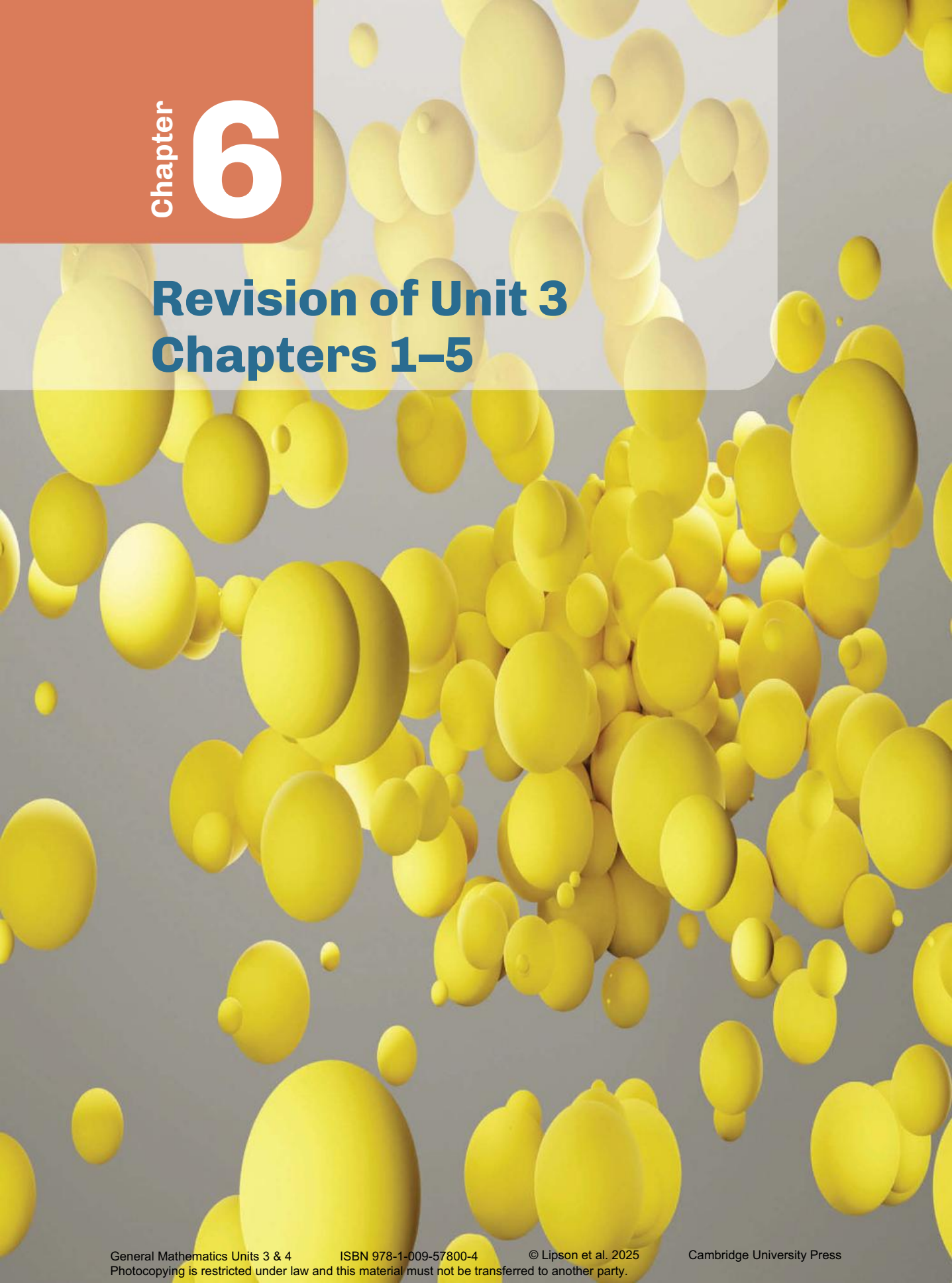
Multiple-choice questions

- A great circle on a newly found planet has a circumference of 11 000 km. The diameter of the planet is closest to:
A 320 km **B** 300 km
C 3500 km **D** 250 km
- The coordinates of two points M and N on the Earth's surface are (40°N , 40°E) and (25°S , 55°E). Which statement is most likely to be correct about the time difference?
A M is 5 hours behind N. **B** M is 1 hour behind N.
C N is 5 hours behind M. **D** N is 1 hour behind M.
- Point X on the Earth's surface has coordinates (29°S , 32°E), while point Y is at (8°S , 32°E). The distance between X and Y is closest to:
A 2335 km **B** 750 km
C 111 km **D** 1350 km
- X and Y are two towns on the equator. The longitude of X is 18°E and the longitude of Y is 48°W . Approximately how far apart are these two towns?
A 7400 km **B** 8500 km
C 10 100 km **D** 7340 km
- Trevor lives in Albany, which has a longitude of 118°E . He wants to watch a basketball game being played in Ottawa, which has a longitude of 76°W . The game starts at 10 p.m. on Wednesday Ottawa time. What is the time in Albany when the game starts? (Ignore time zones and daylight saving.)
A 11 a.m. on Wednesday **B** 1 a.m. on Thursday
C 9 a.m. on Thursday **D** 11 a.m. on Thursday

- ## Short response questions

- 

- 9** Arisa is from Japan and is studying in Brisbane. She plans to phone home on Sunday night at 8 p.m. What time is it in Tokyo? (Brisbane GMT +10, Tokyo GMT +9)
- 10** Terry is playing hockey at a tournament in Buenos Aires. After her team wins the semi-final at 5 p.m. on Friday, she phones her father in Toowoomba to tell him the news. What time is it in Toowoomba? (Toowoomba GMT +10, Buenos Aires GMT -3)
- 11** Elizabeth is in Rome, which is one hour ahead of Greenwich Mean Time. Leslie is in Boston, which is five hours behind Greenwich Mean Time.
- a** Elizabeth is going to ring Leslie at 10 p.m. on Tuesday, Rome time. What day and time will it be in Boston when she rings?
 - b** Elizabeth is going to fly from Boston to Rome. Her flight will leave on Wednesday at 10 a.m., Boston time, and will take 10 hours. What day and time will it be in Rome when she arrives?
- 12** Osaka is at 34°N , 135°E , and Dallas is at 33°N , 97°W .
- a** Find the time difference between the two cities. (Ignore time zones.)
 - b** Rex lives in Dallas and wants to ring a friend in Osaka. In Dallas it is 8 p.m. Monday. What time and day is it in Osaka?
 - c** Rex's friend in Osaka sent him a text message, which happened to take 14 hours to reach him. It was sent at 10 a.m. Thursday, Osaka time. What was the time and day in Dallas when Rex received the text?
- 13** Two locations X and Y have the same latitude 30°S . The longitude of X is 145°E and the longitude of Y is 130°E . (In this question take the radius of the Earth to be 6400 km.)
- a** Determine the radius of the small circle of the 30°S parallel of latitude.
 - b** Determine the distance between X and Y around the parallel of latitude 30°S .
- 14** The Tropic of Cancer is at latitude 23.5°N , while the Tropic of Capricorn is at latitude 23.5°S .
- a** Calculate the distance between these two tropics along the same great circle (correct to the nearest km).
 - b** Calculate the radius of the small circles of the Tropic of Capricorn and the Tropic of Cancer.
 - c** Rockhampton in Queensland and Sao Paulo in Brazil are both on the Tropic of Capricorn. Rockhampton has longitude 150.5°E and Sao Paulo has longitude 46.5°W . Find the distance around the Tropic of Capricorn from Sao Paulo to Rockhampton.



Chapter

6

Revision of Unit 3 Chapters 1–5

6A Topic 1: Bivariate data analysis 1

Multiple-choice questions

- 1 Researchers believe that reaction time might be lower in cold temperatures. They devise an experiment where *reaction time* in seconds is measured at three different *temperature* levels (1 = less than 8°C, 2 = from 8°C to 18°C, 3 = more than 18°C). The explanatory variable and its classification are:

- A** *reaction time*, categorical **B** *temperature*, categorical
C *reaction time*, numerical **D** *temperature*, numerical

Use the following information to answer Questions 2 to 4.

The data in the following table was collected to investigate the association between a person's age and their satisfaction with their career choice.

	Age group		
<i>Satisfied with career choice?</i>	Under 35	35 or more	Total
Yes	136	136	272
No	42	86	128
Total	178	222	400

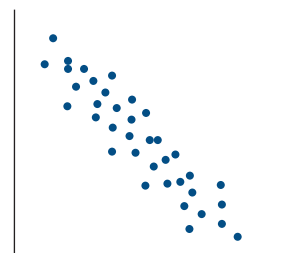
- 2 The percentage of participants in the study aged 35 or more is closest to:
A 55.5% **B** 50.0% **C** 44.5% **D** 68.0%
- 3 Of those people aged under 35, the percentage who are satisfied with their career choice is closest to:
A 34.0% **B** 50.0% **C** 61.3% **D** 76.4%
- 4 The data in the table supports the contention that there is an association between *age group* and *satisfaction with career choice* because:
A 68.0% of people are satisfied with their career choice, compared to 32.0% who are not.
B the number of people satisfied with their career choice aged under 35 is the same as the number aged 35 or more who are satisfied with their career choice.
C 76.4% of people aged under 35 are satisfied with their career choice, which is more than the 61.3% of those aged 35 or more who are satisfied with their career choice.
D 50.0% of people who are satisfied with their career choice are aged under 35.

- 5 For which one of the following pairs of variables would it be appropriate to construct a scatterplot?

A weight in kg and blood pressure in mmHg
B number of cups of coffee drunk each day and stress level (high, medium, low)
C age in years and football team
D time spent watching TV each week in hours and educational level (primary, secondary, tertiary)

- 6 The value of r for the scatterplot is closest to:

A 0.8
B 0.5
C -0.5
D -0.9



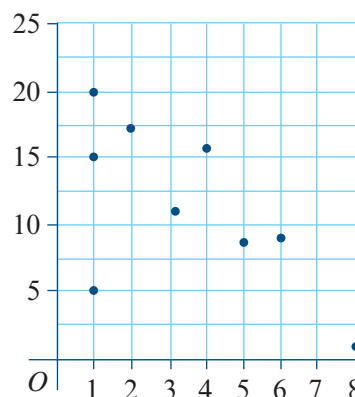
- 7 The association pictured in the scatterplot in the previous question is best described as:

A strong, positive, linear
B strong, negative, linear
C weak, negative, linear
D strong, negative, non-linear

- 8 When the correlation coefficient, r , was calculated for the data displayed in the scatterplot, it was found to be $r = -0.64$.

If the point $(1, 5)$ was replaced with the point $(6, 5)$ and the correlation coefficient, r , recalculated, then the value of r would be:

A unchanged
B positive but closer to 1
C positive but closer to 0
D negative but closer to -1 .



The following information relates to Questions 9 to 11.

There is a linear association between *computer ownership* (computers/1000 people) and *car ownership* (cars/1000 people), and the coefficient of determination is equal to 0.8464.

- 9 If the people who own more cars also tend to own more computers, then the value of the correlation coefficient, r (rounded to two decimal places), is closest to.

A 0.64 **B** 0.72 **C** 0.85 **D** 0.92

- 10** The percentage of variation in *computer ownership* explained by the variation in the *car ownership* is closest to:
- A** 71.6% **B** 84.6% **C** 92.0% **D** 15.4%
- 11** The percentage of variation in *computer ownership* NOT explained by the variation in the *car ownership* is closest to:
- A** 71.6% **B** 84.6% **C** 92.0% **D** 15.4%
- 12** The correlation between the score on a maths test and height for a group of primary school students is found to be 0.7. From this information, it is reasonable to conclude that:
- A** learning maths makes children grow taller
B there is no association between height and maths test scores
C a child's maths ability depends only on their height
D the children who obtained high maths test scores tended to be taller.

The following information relates to Questions 13 and 14

In a study of the association between left-handedness and dyslexia, the dominant hand (left or right) was recorded for two groups of students, a group of students who had been diagnosed with dyslexia (yes), and a control group (no). The results are summarised in the table below.

<i>Dominant hand</i>	<i>Dyslexia</i>	
	Yes	No
Left	10	10
Right	40	90

- 13** The percentage of students diagnosed with dyslexia who are left-handed is closest to:
- A** 10% **B** 20%
C 30% **D** 80%
- 14** The variables *dominant hand* and *dyslexia* appear to be associated because:
- A** 20% of students diagnosed with dyslexia are left-handed, compared to only 10% of the control group
B only 30% of the students are left-handed
C 80% of students diagnosed with dyslexia are left-handed, compared to 90% of the control group
D a higher percentage of the control group are left-handed compared to the students diagnosed with dyslexia.

- 15** The *age* in years and percentage *bodyfat* for a group of 8 people is given in the following table.

<i>age</i>	58	72	67	43	51	52	25	35
<i>bodyfat</i>	20.1	26.1	25.8	19.5	14.1	27.0	6.1	4.1

The value of the Pearson correlation coefficient, r , for these data is closest to:

- A** 0.6
- B** 0.7
- C** 0.8
- D** 0.9

Short-response questions

- 1** For the following pairs of variables, classify each variable as either categorical or numerical, and choose which of the following analysis techniques you would use to investigate the association between them:
- two-way frequency table
 - scatterplot
 - parallel boxplots
- a** *age* (years) and *reaction time* (seconds)
- b** *gender* (male, female) and *reaction time* (seconds)
- c** *gender* (male, female) and *reaction time* (fast, average, slow)
- 2** In a large university, students in some courses were asked if they would prefer to attend on-campus lectures and tutorials, or study online and not attend the campus. Data was collected for the following variables:
- *number* student number
 - *study mode* 1 = on-campus, 2 = online
 - *age* age in years
 - *course* 1 = Business, 2 = Health, 3 = Social Science
 - *gender* F = female, M = male
 - *distance* the distance the student lives from the campus, to the nearest km

The data collected for 12 students are shown in the following table:

<i>Number</i>	<i>Study mode</i>	<i>Age (years)</i>	<i>Course</i>	<i>Gender</i>	<i>Distance (kms)</i>
23455	1	18	1	M	2
13425	2	23	1	F	8
28445	1	18	2	M	4
19889	1	19	3	F	9
10340	2	25	2	F	13
23001	2	22	1	M	2
19968	1	19	1	F	7
20012	2	34	3	F	6
21980	1	18	3	M	12
17884	2	45	1	M	8
19456	2	22	2	F	9
21111	1	22	3	F	6

- a** Write down the names of the numerical variables in the table.
b Use the data in the table to complete the following two-way frequency table.

<i>Study mode</i>	<i>Gender</i>	
	Female	Male
On campus		
Online		
Total		

- 3** The number of hours spent studying for an examination by each member of a class, and the marks they received are given in the table:

<i>Student</i>	1	2	3	4	5	6	7	8	9	10
<i>Hours</i>	4	36	23	28	25	11	18	13	4	8
<i>Mark</i>	27	87	67	84	66	52	61	43	38	52

- a** Which of these variables is the explanatory variable and which is the response variable?
b Construct a scatterplot to display the data.
c From the scatterplot, describe the association between *hours* and *mark* in terms of direction, form and strength.

- 4 The following table gives the life expectancies, in years, for males and females across a group of countries.

	Male life expectancy (years)	Female life expectancy (years)
Australia	80.4	84.5
Canada	79.6	83.8
China	74.5	77.5
France	79.2	85.5
India	66.9	69.9
Indonesia	67.0	71.2
Italy	80.3	84.9
Japan	80.8	87.1
Mexico	72.3	77.7
Russia	65.9	76.7
South Africa	55.5	59.5
United Kingdom	79.2	82.8
United States	76.3	81.2

- a For these data, calculate the value of the correlation coefficient, r , and interpret.
- b What assumptions have you made about the relationship between male and female life expectancies in these countries for the calculation of the correlation coefficient to be valid?
- 5 The data in the table below is based on a study of dolphin behaviour. In this study, the main activities of dolphins observed in the wild were classified as ‘travelling’, ‘feeding’, and ‘socialising’. The time of day was also noted.

	Time of observation		
Activity	Morning	Afternoon	Evening
Travelling	11.4%	53.3%	16.5%
Feeding	38.0%	6.7%	70.9%
Socialising	50.6%	40.0%	12.6%
Total	100.0%	100.0%	100.0%

- a Which is the explanatory variable and which is the response variable?
- b Of the dolphins who were observed in the morning, what percentage were feeding?
- c Use the information in the table to write a brief report describing the association between behaviours of the dolphins and the time of day.

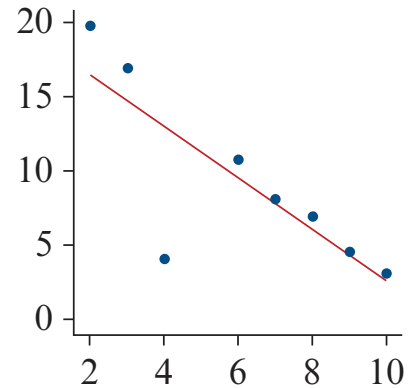
- 6** The following table gives the correlation coefficients between infant mortality rate and a range of possible explanatory variables for countries across the world.

	<i>Infant Mortality Rate</i> (per 1000 live births)
Birth rate per 1000 people	$r = 0.876$
Births attended by skilled health staff (% total)	$r = -0.714$
Exclusive breastfeeding (% of children under 6 months)	$r = 0.170$
Health expenditure per capita (current US\$)	$r = -0.495$
Literacy rate, adult female (% of females ages 15 and above)	$r = -0.800$
Literacy rate, adult male (% of males ages 15 and above)	$r = -0.849$
People using safely managed sanitation services (% of population)	$r = -0.636$
People using safely managed drinking water services (% of population)	$r = -0.817$

- a** Determine the values of the coefficient of determination for each of these variables.
- b** Discuss the relative importance of each of the explanatory variables in understanding infant mortality.
- 7** In a large university, students were offered the choice of attending traditional lectures and tutorials on campus or working independently using an online method of instruction. The study intentions were recorded for a random sample of full-time students from three different faculties (Arts, Business and Science). Researchers suggested that students enrolled in the Arts faculty would be more likely to select the traditional on campus study than students from other faculties. Write a report addressing the researchers' hypothesis.

		<i>Faculties</i>			
		Arts	Science	Business	Total
<i>Type of instruction</i>	On campus	102	128	58	288
	Online	58	72	82	212
		160	200	140	500

- 8 When the Pearson correlation coefficient was calculated for the data displayed in the following scatterplot, it was found to be -0.433 , and the slope of the least squares regression line is -1.8 . Discuss the effect on the value of the correlation coefficient, and the value of the slope, if the outlier is removed.



- 9 A sample of 608 males were asked whether Australia should retain the King or become a republic. Their answers, together with the political affiliation of the respondent, are summarised in the following two-way frequency table.

<i>King or republic?</i>	<i>Political affiliation</i>		Total
	Liberal	Labor	
Definitely keep King	74	28	102
Probably keep King	60	39	99
Probably become republic	81	78	159
Definitely become republic	47	201	248
	262	346	608

A sample of 560 females was also asked whether Australia should retain the King or become a republic. Their answers, together with the political affiliation of the respondent, are summarised as follows.

<i>King or republic?</i>	<i>Political affiliation</i>		Total
	Liberal	Labor	
Definitely keep King	93	51	144
Probably keep King	67	43	110
Probably become republic	49	72	121
Definitely become republic	38	147	185
	247	313	560

Discuss whether the relationship between political affiliation and attitude to the monarchy is the same for the females as for the males.

6B Topic 2: Bivariate data analysis 2

Multiple-choice questions

- 1 A teacher collected the following statistical information about her students' scores in their mathematics examination in Year 11 and their scores in Year 12:

	Year 11	Year 12
mean	75.1	64.8
standard deviation	2.567	4.983
correlation coefficient	$r = 0.675$	

The slope of the least squares regression line which would allow their score in Year 12 to be predicted from their score in Year 11 is closest to:

- A** 0.35 **B** 0.68 **C** 1.3 **D** 1.7
- 2 The statistical analysis of the set of bivariate data involving variables x and y resulted in the information displayed in the table below:

	x	y
mean	123.5	38.7
standard deviation	4.65	4.78
least squares equation	$y = -140 + 0.475x$	

Using this information, the value of the correlation coefficient, r , for this set of bivariate data is closest to:

- A** 0.73 **B** 0.34 **C** 0.46 **D** 0.49

The following data relate to Questions 3 and 4.

Number of hot dogs sold	190	168	146	155	150	170	185
Temperature ($^{\circ}\text{C}$)	10	15	20	15	17	12	10

We wish to determine the equation of the least squares regression line for the data that will enable the number of hot dogs sold to be predicted from temperature.

- 3 The equation of the least squares regression line fitted to the data is closest to:
- A** $\text{number of hot dogs sold} = -4.31 \times \text{temperature} + 227$
- B** $\text{number of hot dogs sold} = -0.206 \times \text{temperature} + 48.4$
- C** $\text{number of hot dogs sold} = -48.4 \times \text{temperature} + 0.206$
- D** $\text{number of hot dogs sold} = 4.31 \times \text{temperature} + 227$

- 4 The coefficient of determination will be closest to:
- A** -0.94 **B** -0.89 **C** 0.21 **D** 0.89

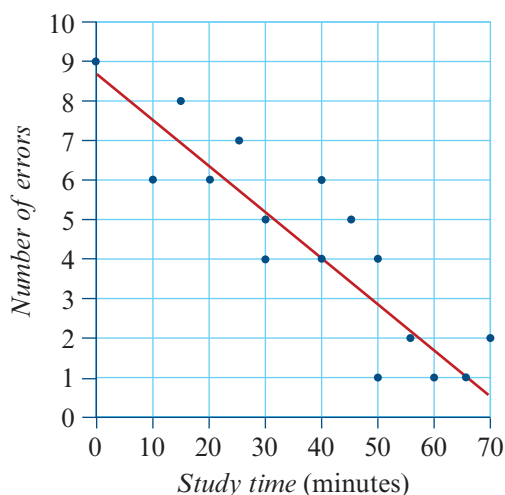
The following information relates to Questions 5 to 10.

Eighteen students sat for a 15-question multiple-choice test. In the scatterplot opposite, the number of errors made by each student on the test is plotted against the time they reported studying for the test.

A least squares regression line has been determined for the data and is also displayed on the scatterplot. The equation for the least squares regression line is:

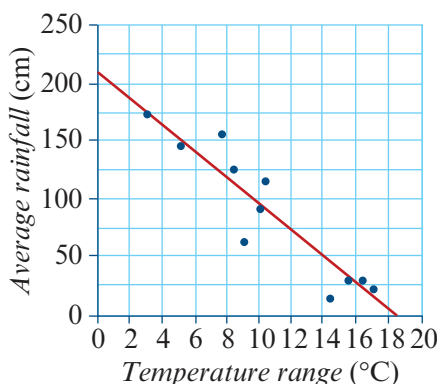
$$\text{number of errors} = -0.12 \times \text{study time} + 8.8$$

and the coefficient of determination is 0.8198.



- 5 The least squares regression line predicts that a student reporting a study time of 35 minutes would make:
- A** 4.3 errors **B** 4.6 errors **C** 4.8 errors **D** 5.0 errors
- 6 The value of Pearson's correlation coefficient, r , is closest to:
- A** -0.91 **B** -0.82 **C** 0.82 **D** 0.91
- 7 The student who reported a study time of 10 minutes made six errors. The predicted score for this student would have a residual of:
- A** -7.6 **B** -1.6 **C** 1.6 **D** 7.6
- 8 Which of the following statements that relate to the regression line is *not* true?
- A** The equation predicts that a student who spends 40 minutes studying will make approximately four errors.
- B** The least squares regression line does *not* pass through the origin.
- C** On average, a student who does not study for the test will make around 8.8 errors.
- D** The explanatory variable in the regression equation is *number of errors*.
- 9 This regression line predicts that, on average, the number of errors made:
- A** decreases by 0.82 for each extra minute spent studying.
- B** decreases by 0.12 for each extra minute spent studying.
- C** increases by 0.12 for each extra minute spent studying.
- D** increases by 8.8 for each extra minute spent studying.

- 10** Given that the coefficient of determination is 0.8198, we can say that close to:
- A** 18% of the variation in the number of errors made can be explained by the variation in the time spent studying
 - B** 67% of the variation in the number of errors made can be explained by the variation in the time spent studying
 - C** 82% of the variation in the number of errors made can be explained by the variation in the time spent studying
 - D** 95% of the variation in the number of errors made can be explained by the variation in the time spent studying.
- 11** The average rainfall and temperature range at several locations are displayed in the scatterplot below.



A least squares regression line has been fitted to the data, as shown. The equation of this line is closest to:

- A** $\text{average rainfall} = -11 \times \text{temperature range} + 210$
 - B** $\text{average rainfall} = 11 \times \text{temperature range} + 210$
 - C** $\text{average rainfall} = -0.08 \times \text{temperature range} + 18$
 - D** $\text{average rainfall} = 0.08 \times \text{temperature range} + 18$
- 12** In a certain state, the correlation coefficient between:
- university entrance score and score on a Mathematics aptitude test is $r = 0.462$
 - university entrance score and score on an English aptitude test is $r = 0.662$
- Given this information, which *one* of the following statements is *true*?
- A** Around 21.3% of the variation in score on the Mathematics aptitude test is explained by the variation in score on the English aptitude test.
 - B** Around 43.8% of the variation in score on the English aptitude test is explained by the variation in score on the Mathematics aptitude test.
 - C** Together the scores on the Mathematics and English aptitude tests explain 100% of the variation in university entrance score.
 - D** The score on English aptitude tests is a better predictor of the university entrance score than is the score on the Mathematics aptitude test.

Short-response questions

- 1** The following table gives the adult heights (in cm) of ten pairs of mothers and daughters:

<i>Mother</i>	170	163	157	165	175	160	164	168	152	173
<i>Daughter</i>	178	175	165	173	168	152	163	168	160	178

- Identify which variable is the explanatory variable and which is the response variable.
- Calculate the value of the correlation coefficient, r , and classify its strength.
- Determine the equation of the least squares regression line, and interpret the intercept and slope.
- Determine the value of the coefficient of determination, R^2 , and interpret in terms of the variables in this question.
- Estimate the adult height of a girl whose mother is 170 cm tall.
- Is your prediction in part **e** interpolation or extrapolation? Explain.

- 2** The following table shows the daily maximum temperature and the number of ice-creams sold at a kiosk on the beach over a nine-day period:
The equation of the least squares regression line that allows the number of ice-creams sold to be predicted from the temperature is:
 $sales = 10.3 \times temperature + 97.2$

<i>Temperature</i>	<i>Sales</i>
18	280
21	298
22	333
24	359
25	360
26	355
27	378
32	427
36	465

- a** Complete the following table of the residuals.

<i>Temperature</i>	18	21	22	24	25	26	27	32	36
<i>Residual</i>	-2.6		9.2	14.6	5.3	-10.0		0.2	-3.0

- b** Construct a residual plot, and comment on the linearity assumption.

- 3** A marketing firm wanted to investigate the relationship between the number of times a song was played on the radio (*played*) and the number of downloads sold the following week (*weekly sales*).

The following data was collected for a random sample of ten songs.

<i>Played</i>	47	34	40	34	33	50	28	53	25	46
<i>Weekly sales</i>	3950	2500	3700	2800	2900	3750	2300	4400	2200	3400

- Which is the explanatory variable and which is the response variable?
 - Construct a scatterplot of this data.
 - Determine the value of the Pearson correlation coefficient, r , for this data.
 - Describe the relationship between *weekly sales* and *played* in terms of direction, strength and form and outliers (if any).
 - Determine the equation for the least squares regression line and write it down in terms of the variables *weekly sales* and *played*.
 - Interpret the slope and intercept of the least squares regression line in the context of the problem.
 - Use your equation to predict the number of downloads of a song when it was played on the radio 100 times in the previous week.
 - Comment on the reliability of this prediction.
- 4** A caterer collected the following data on the cost to the company of the preparation of a differing number of meals.

<i>Number of meals (x)</i>	30	70	90	25	50	60	75	100
<i>Cost(\$)(y)</i>	345	595	720	300	485	530	585	750

- Using the method of least squares regression, find the equation of a straight line that relates the two variables.
- Determine is the caterer's base cost for the standard menu.
- Determine the cost of each meal, over and above this base cost.

- 5 A regression analysis was conducted to investigate the nature of the relationship between femur (thigh bone) length and radius (the shorter, thicker bone in the forearm) length in 18-year-old males. The bone lengths are measured in centimetres. The results of this analysis are reported below. In this investigation, femur length was treated as the explanatory variable.

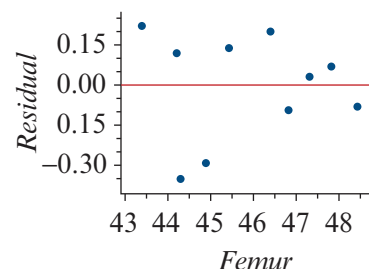
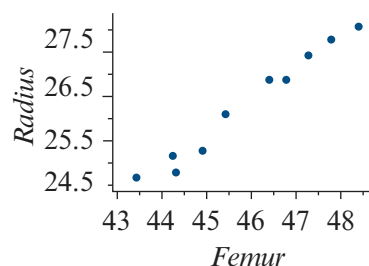
Regression equation $y = a + bx$

$$a = -7.24946$$

$$b = 0.739556$$

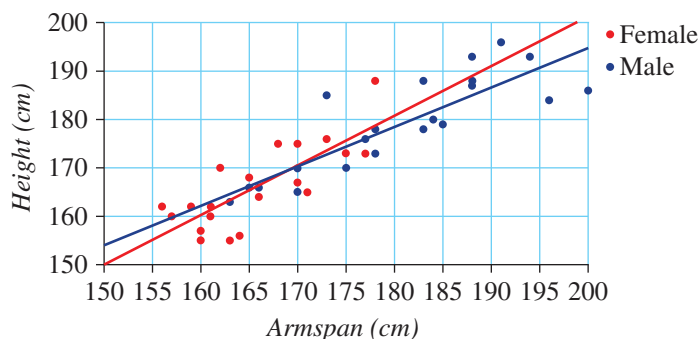
$$R^2 = 0.975291$$

$$r = 0.987568$$



Based on these analyses, write a report describing the association between femur length and radius length.

- 6 To address the question “Can we predict a person’s height from their armspan?” height (in cm) and armspan (in cm) measurements were collected from 60 students in Year 11 and 12. Of the 60 students 30 were male and 30 female. The following scatterplot shows the data collected, with least squares regression lines fitted for males and females.



The equation of the least squares regression line for females is:

$$\text{height} = -4.199 + 1.028 \times \text{armspan}$$

The equation of the least squares regression line for males is:

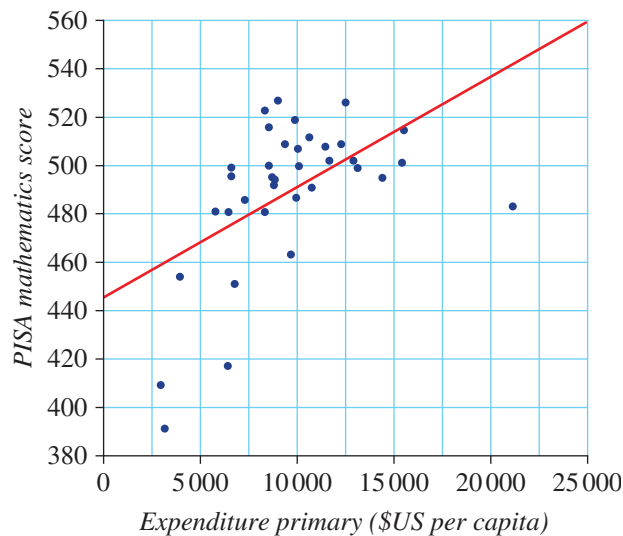
$$\text{height} = 31.705 + 0.815 \times \text{armspan}$$

- a Interpret the slope of the regression equation in terms of *height* and *armspan* for males.

In determining this equation, the summary statistics displayed in the table were also calculated.

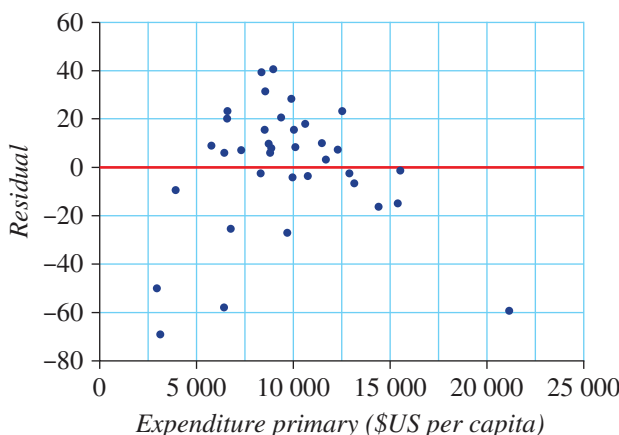
	Armspan	Height
mean (females)	164.0	164.5
standard deviation (females)	6.319	8.083
mean (males)	178.1	177.0
standard deviation (males)	9.832	9.583

- b**
- i** Determine the value of the coefficient of determination for females and interpret in terms of *height* and *armspan*. Give your answer as a percentage rounded to one decimal place.
 - ii** Determine the value of the coefficient of determination for males and interpret in terms of *height* and *armspan*. Give your answer as a percentage rounded to one decimal place.
 - iii** Explain why *armspan* is a better predictor of *height* for males than for females, quoting appropriate statistics.
- c**
- i** Use the least squares regression line to predict the difference in height between males and females who both have armspans of 160 cm.
 - ii** Use the least squares regression line to predict the difference in height between males and females who both have armspans of 190 cm.
 - iii** Are the prediction made in parts **ci** and **cii** reliable? Explain.
- 7** The average student PISA mathematics scores (*score*) for OECD countries, as well as the expenditure per primary school child in those countries in \$US per capita (*expenditure*), are shown in the following scatterplot.



- a** Describe the association between *score* and *expenditure* in terms of form and strength.

- b** A least squares regression line, with expenditure as the explanatory variable, was fitted to the data, and the following residual plot constructed.

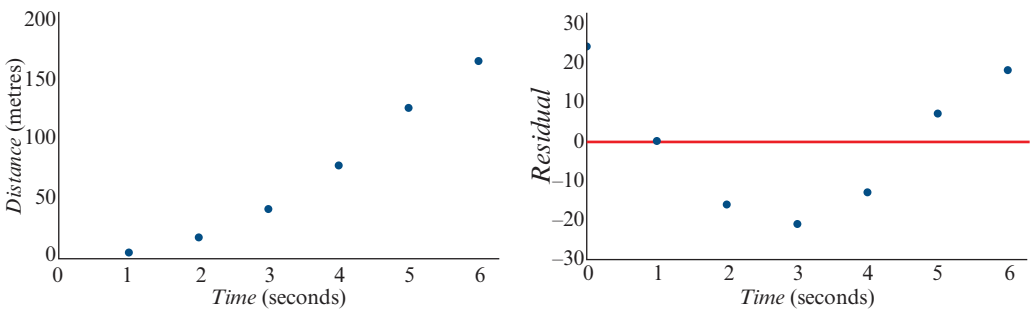


- i** A residual plot can be used to test an assumption about the nature of the association between two numerical variables. What is this assumption?
- ii** Does the residual plot support this assumption? Explain your answer.
- 8** Suppose that the manager of a company determines that the cost of manufacturing jeans is:
- $$\text{cost} = 22.90 \times \text{number of pairs of jeans produced} + 12\,500$$
- a** What is the marginal cost of manufacturing each pair of jeans?
- b** What is the cost per pair of jeans if 100 pairs of jeans are produced?
- c** For the manufacturer to make 75% profit on each pair of jeans, how much should he sell them each for:
- i** if 100 pairs of jeans are produced?
- ii** if 500 pairs of jeans are produced?

- 9 The following table displays the distance fallen by an object across one-second intervals.

<i>Time (seconds)</i>	0	1	2	3	4	5	6
<i>Distance (metres)</i>	0	5.2	18	42	79	128	168

A scatterplot of the data is shown below, together with the residual plot resulting when a least squares line is fitted to the data.



- a Comment on the linearity assumption.
- b The science teacher suggested the student consider $(\text{time})^2$ as the explanatory variable.

i Complete the following table:

<i>Time (seconds)</i>	$(\text{Time})^2$	<i>Distance (metres)</i>
0		0
1		5.2
2		18
3		42
4		79
5		128
6		168

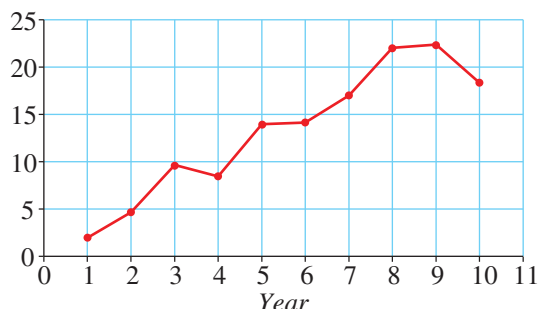
- ii Construct a scatterplot of the data, with $(\text{time})^2$ as the explanatory variable and distance as the response variable.
- iii Determine the equation of the least squares regression line for this data.
- iv Construct a residual plot, and comment on the linearity assumption.
- v Use the equation to predict the distance travelled after 7 seconds.

6C Topic 3: Time series analysis

Multiple-choice questions

- 1 The time series plot shown is best described as showing:

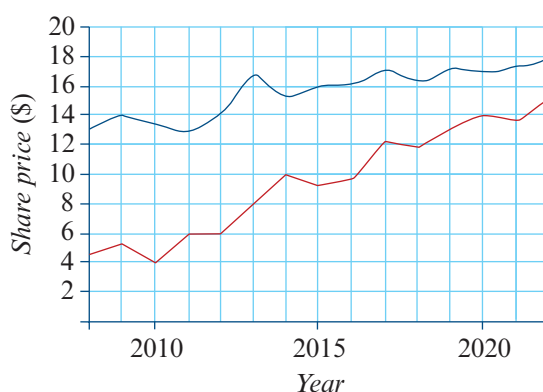
A trend only
B trend and irregular variation
C trend and seasonality
D irregular variation only.



- 2 The time series plot shows the share price of two companies over a period of time.

From the plot, it can be concluded that over the period 2010–2020, the *difference* in share price between the two companies has shown:

A a decreasing trend
B an increasing trend
C seasonal variation
D no trend.



Use the information in the table below to answer Questions 3 to 6.

t	1	2	3	4	5	6	7	8	9	10
y	4	5	4	4	8	6	9	10	9	12

- 3 The three-mean smoothed value for $t = 2$ is closest to:

A 4.3 **B** 6.2 **C** 6.4 **D** 6.5

- 4 The five-mean smoothed value for $t = 5$ is closest to:

A 4.3 **B** 6.2 **C** 6.4 **D** 6.5

- 5 The three-median smoothed value for $t = 6$ is closest to:

A 4.3 **B** 6 **C** 8 **D** 9

- 6 The five-median smoothed value for $t = 3$ is closest to:

A 4 **B** 4.5 **C** 5 **D** 8

- 7 To help work out her staffing roster, Fleur records the number of customers who come into her cafe between 7.00 a.m. and 8.00 a.m. each morning for a week.

Day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Customers	42	25				84	100

The numbers of *customers* on Wednesday, Thursday and Friday are not shown. The five-mean smoothed number of *customers* on Thursday is 38.

The three-mean smoothed number of *customers* on Thursday is:

A 27 **B** 29 **C** 30 **D** 38

Use the following information to answer Questions 8 and 9.

The table below records the monthly electricity cost (in dollars) for an apartment over one calendar year.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
123	90	153	136	101	129	153	143	95	61	85	107

- 8 Based on this information, the seasonal index for September is closest to:

A 1.00 **B** 0.78 **C** 1.25 **D** 0.83

- 9 Using data collected over several years, the seasonal index for December was determined to be 0.90. To correct the cost of electricity for seasonality in December, the actual cost should be:

A decreased by 11.1% **B** decreased by 10.0%
C increased by 11.1% **D** increased by 10.0%.

Use the information below to answer Questions 10 and 11.

The quarterly sales figures for a soft drink company and the seasonal indices are as shown.

Quarter	1	2	3	4
Sales (\$'000s)	1200	1000	800	1200
Seasonal index	1.1	0.90	0.8	

- 10 The deseasonalised figure (in \$'000s) for quarter 3 is:

A 640 **B** 667 **C** 800 **D** 1000

- 11 The seasonal index for quarter 4 is:

A 0.6 **B** 0.8 **C** 1.1 **D** 1.2

- 12** The deseasonalised sales (in dollars) for a company in June were \$91 564. The seasonal index for June is 1.45.

The *actual* sales for June were closest to:

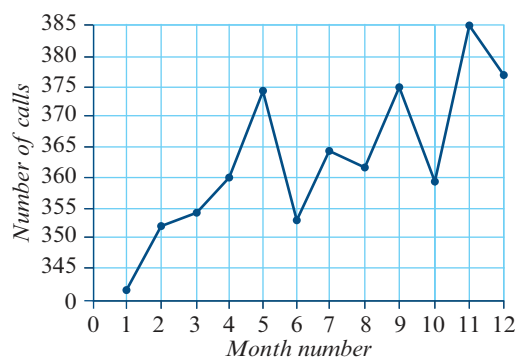
- A** \$41 204 **B** \$132 768 **C** \$63 148 **D** \$91 564

- 13** Sales for a major department store are reported quarterly. The seasonal index for the third quarter is 0.85. This means that sales for the third quarter are typically:

- A** 85% below the quarterly average for the year
B 15% below the quarterly average for the year
C 15% above the quarterly average for the year
D 18% above the quarterly average for the year

Use the information below to answer Questions 14 and 15.

The time series plot opposite shows the number of calls each month to a call-centre over a 12-month period.



- 14** The three-median smoothed number of calls for month 9 is closest to:

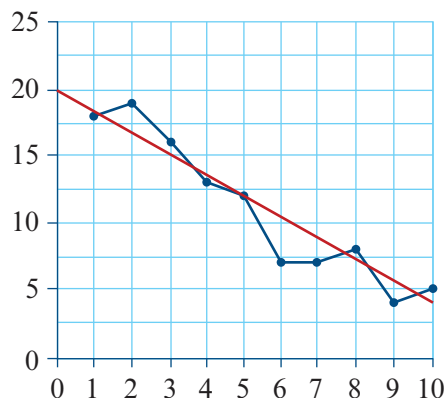
- A** 358 **B** 362 **C** 371 **D** 375

- 15** The five-median smoothed number of calls for month 10 is closest to:

- A** 358 **B** 362 **C** 371 **D** 375

- 16** A time series for y is shown in the graph, where t represents time. If a linear trend line is fitted to this data, as shown, then the equation of the line is closest to:

- A** $y = -1.6t + 20$
B $y = 1.6t + 20$
C $y = -0.6t + 20$
D $y = 0.6t + 20$

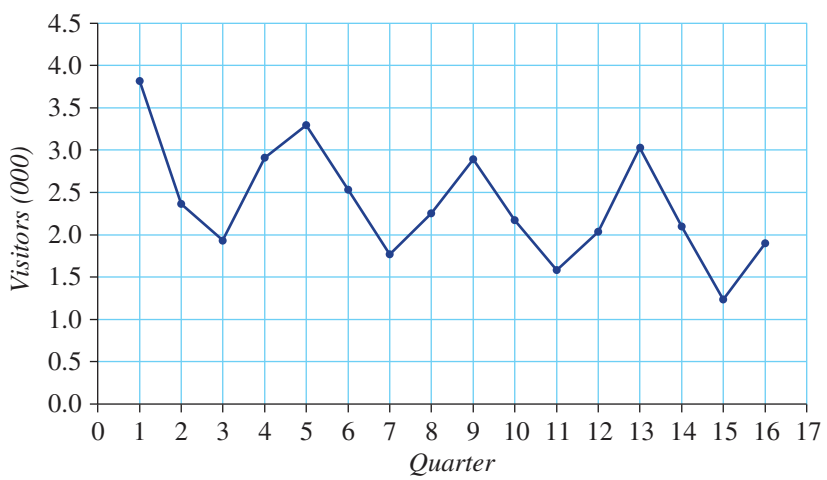


Short-response questions

- 1 The number of new vehicles (000's) purchased in Queensland in a particular year are given in the following table. Construct a time series plot and describe features of the plot.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
17.2	18.7	21.5	17.8	19.6	27.3	18.4	18.1	19.7	16.9	18.7	19.3

- 2 The time series plot below shows the number of visitors to a tourist attraction per quarter over a 4-year period. Quarter 1 is Summer 2024, Quarter 2 is Autumn, 2024, Quarter 3 is Winter, 2024, and Quarter 4 is Spring, 2024.
 - a In which season is the seasonal index highest?
 - b Describe the features of the time series plot.



- 3 The value of an Australian dollar in Euro (exchange rate) over a 10-day period is given in the table.

Day	1	2	3	4	5	6	7	8	9	10
Exchange rate	0.6035	0.6098	0.6134	0.6196	0.5996	0.6144	0.6208	0.6055	0.6166	0.5974

- a Determine the three-median smoothed value for Day 4.
 - b Determine the five-median smoothed value for Day 5.
- 4 The seasonal indices for the number of cars in a car park are as follows:

Time period	Morning	Afternoon	Evening	Overnight
Seasonal index	1.2	1.4	0.9	

- a Determine the seasonal index for Overnight.
- b Interpret this index in terms of the average number of cars in the car park each time period.

- 5** The number of caravans sold each quarter of one year by a certain company is given in the following table.

Quarter	Quarter 1	Quarter 2	Quarter 3	Quarter 4
Number of caravans	105	98	45	123

Use these data to calculate quarterly seasonal indices for the number caravans sold, rounded to two decimal places.

- 6** The monthly earnings for one particular person over a period of 12 months are as shown.

Month	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
Income \$000's	5.0	2.6	5.2	6.0	2.4	3.2	0.2	8.4	6.2	3.2	3.6	4.0

- a** Determine the three-mean smoothed value for March.
b Determine the five-mean smoothed value for October.
c The seasonal indices for 12 months are shown with the exception of December.

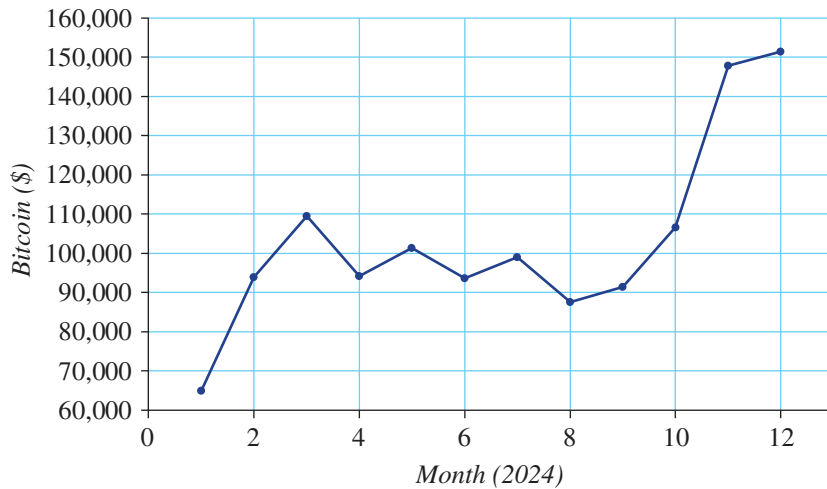
Jan	Feb	Mar	April	May	June	July	Aug	Sept	Oct	Nov	Dec
1.2	0.7	1.2	1.4	0.6	0.7	1.0	2.0	8	0.7	0.9	

- i** Determine the seasonal index for December.
ii In the following year the person earns \$4200 in January. Determine their deseasonalised earnings for that month.
- 7** The number of purchases made at a shop over the period of three years from 2022 to 2024 is recorded in the table below.

Number of purchases	Summer	Autumn	Winter	Spring
2022	1380	1627	1840	720
2023	1552	1770	2056	725
2024	1949	1986	2150	990

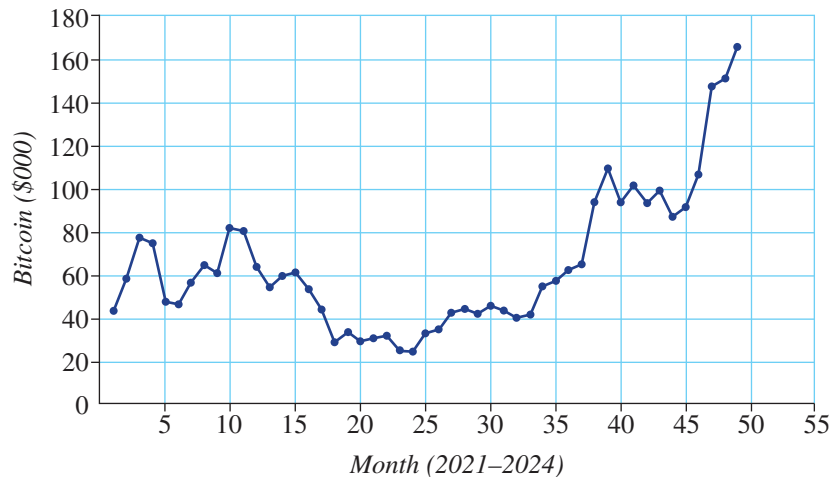
- a** Calculate the seasonal indices based on the three years of data.
b Construct a table to show the number of purchases after deseasonalisation.
c Determine the equation of the least squares regression line for this time series (deseasonalised data vs quarter number).
d Interpret the value of the gradient of the least squares regression line in this case.
e Determine the predicted actual number of purchases for summer 2026.

- 8** The value of one Bitcoin in Australian dollars at the beginning of each month in 2024 is shown in the following time series plot.



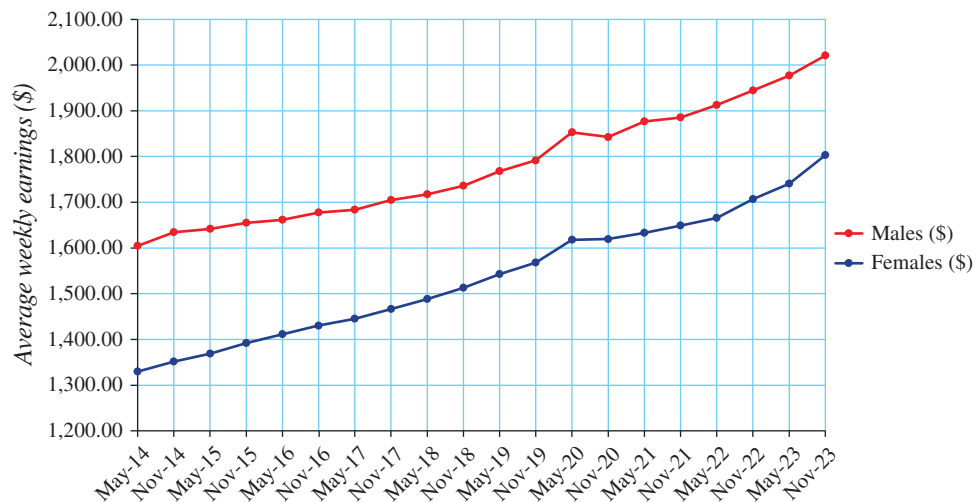
- a** Determine (to the nearest \$000):
- the 3-median smoothed value for Month 8
 - the 7-median smoothed value for Month 9.
- b** A least squares regression line fitted the monthly Bitcoin data for the year 2021 (where January 2021 is month 1), giving the following equation:
- $$\text{Bitcoin} = 1536.04 \times \text{month} + 53\,083.71$$
- Interpret the slope in terms of the variables in the question.
 - Use the equation to predict the value of Bitcoin in July 2025. Give your answer to the nearest dollar.
- c** A least squares regression line fitted the monthly Bitcoin data for the year 2024 (where January 2024 is month 1), giving the following equation:
- $$\text{Bitcoin} = 4776.66 \times \text{month} + 72\,401.57$$
- Interpret the slope in terms of the variables in the question.
 - Use the equation to predict the value of Bitcoin (to the nearest dollar) in July 2025.

- d** The value of one Bitcoin in Australian dollars at the beginning of each month from 2021 to 2024 is shown in the following time series plot.



Which of the two models given in parts **b** and **c** for predicting the future value of bitcoin do you prefer? Use mathematical reasoning to justify your answer.

- 9** Average weekly earnings in Australia are reported twice yearly, in May and November. The following time series plot shows this data for the years 2014–2023, plotted separately for males and females.



- a** Describe the time series plots.
- b** Least squares regression was used to fit lines to each data set separately, giving the following equations:

Males: $ave\ earning = 21.75 \times time\ period + 1553.25$

Females: $ave\ earnings = 24.08 \times time\ period + 1286.88$

Explain why, based in this data, the average earning for females will eventually exceed the average earnings for males, and determine when this will happen.

6D Topic 4: Growth and decay in sequences

Multiple-choice questions

- Determine which of the following could be the first five terms of an arithmetic sequence.

A 1, 4, 9, 16, 25	B 1, 3, 9, 27, 81
C 1, 4, 7, 10, 13	D 1, 3, 6, 9, 12
- The first term in a sequence is 3. Each subsequent term is 6 more than the previous term. The sixth term is:

A 3	B 6	C 36	D 33
- Jane has 11 different flavoured lip glosses in her collection. Each month, she purchases 3 more to add to her collection. A recurrence relation model, J_n , for the number of glosses in Jane's collection at the start of the month n , is:

A $J_1 = 11, J_{n+1} = 3J_n$	B $J_1 = 11, J_{n+1} = 3J_n + 3$
C $J_1 = 11, J_{n+1} = J_n + 3$	D $J_1 = 11, J_{n+1} = J_n - 3$
- A sequence is generated from the recurrence relation $V_1 = 12, V_{n+1} = V_n + 2$. The rule for the value of V_n is:

A $V_n = 12n + 2$	B $V_n = n + 14$
C $V_n = 2n + 10$	D $V_n = 12 - 2n$
- The third and seventh terms of an arithmetic sequence are 20 and 32 respectively. The rule for the n th term is:

A $V_n = 20n + 32$	B $V_n = 32 - 20n$
C $V_n = 12n + 20$	D $V_n = 3n + 11$
- Ezekiel purchased a car for \$22 500, which is depreciated using the flat-rate depreciation method at a rate of 4.5% per annum. The amount of depreciation after five years is:

A \$1012.50	B \$2250	C \$4500	D \$5062.50

- 7** Zara invests \$12 000 with a bank. She will be paid simple interest at the rate of 5.4% per annum. If V_n is the value of Zara's investment at the start of the n th year, the recurrence relation model for Zara's investment is:
- A** $Z_1 = 12\ 000, Z_{n+1} = Z_n + 648$
B $Z_1 = 12\ 000, Z_{n+1} = 5.4 \times Z_n$
C $Z_1 = 12\ 000, Z_{n+1} = Z_n + 5.4$
D $Z_1 = 12\ 000, Z_{n+1} = Z_n - 648$
- 8** The following is a geometric sequence 100 000, 90 000, 81 000, 72 900, ...
 The common ratio r is equal to:
- A** -10 000 **B** -9000 **C** 0.9 **D** 1.1
- 9** The value of t_8 when $t_1 = 3$ and $r = 2$ is:
- A** 3 **B** 6 **C** 19 **D** 384
- 10** A car purchased on 1 June 2017 loses value at a reducing-balance depreciation rate of 20% per year. The original purchase price was \$65 000. The value of the car on 1 June 2022 will be closest to:
- A** \$32 000
B \$22 000
C \$26 600
D \$21 300

Short-response questions

- 1** For each of the following sequences, find the required term.
- a** For an arithmetic sequence with $a = t_1 = 6$ and $d = 5$, determine t_{10} .
b For an arithmetic sequence with $a = t_1 = 400$ and $d = -10$, determine t_{10} .
c For a geometric sequence with $a = t_1 = 10\ 000$ and $r = 0.8$, determine t_5 .
d For a geometric sequence with $a = t_1 = 2$ and $r = 3$, determine t_5 .
- 2** Write down the values of t_1, t_2 and t_3 for each of the following sequences.
- a** An arithmetic sequence with rule $t_n = 8n - 3$.
b A geometric sequence with rule $t_n = 2 \times 5^{n-1}$.
- 3** Write down a rule for the n th term of the sequence for each of the following.
- a** A sequence is defined by the recurrence relation $t_n = t_{n-1} - 15$, with $t_1 = 800$.
b A sequence is defined by the recurrence relation $t_n = 6t_{n-1}$, with $t_1 = 2$.

- 4** Max has \$200 000 to invest and a bank offers him an interest rate of 2.8% per annum compounded annually for the duration of the investment. Find how much would his investment will be worth in eight years time, giving your answer correct to the nearest dollar.
- 5** Given the following information about two sequences, find the required term.
- a** For an arithmetic sequence $t_6 = 13$ and $t_{10} = 5$, find t_{20} .
 - b** For a geometric sequence $t_6 = 80$ and $t_{10} = 1280$, find t_4 .
- 6** A car was purchased for \$48 000. It depreciates in value at the rate of 8% per year, using a reducing-balance depreciation method.
- a** Write down a rule for the value of the car after n years.
 - b** Use this rule to find the value of the car after six years.
 - c** Find the total depreciation of the car over five years, giving your answer to the nearest dollar.
- 7** Philip borrows \$25 500 from a bank and is charged simple interest at the rate of 16% per annum. Let t_n be the value of the loan after n years.
- a** Write down a rule for the value of the loan after n years.
 - b** Find how much Philip needs to pay the bank after 3 years.
 - c** Determine how many years it takes for the value of the loan to reach \$50 000.
- 8** The cost of hiring a photocopier for a year involves a flat rate of \$5000 and then a cost of 2 cents per copy.
- a** Write down the cost of hiring the machine for a year where n copies are made.
 - b** Find how much it will cost if 120 000 copies are made for the year.
 - c** Determine how many copies can be made in a year if no more than \$10 000 can be spent on hiring the photocopier.
- 9** Wild deer are causing a problem in a nature reserve. Under normal conditions, the deer population grows at a rate of 22% per year. When counted at the start of the year, there were 1654 deer in the nature reserve.

Write down a mathematical model of the form:

$$N_{n+1} = rN_n \quad \text{where} \quad N_1 = a$$

that can be used to describe the growth of the deer population in the nature reserve under normal conditions. (N_n represents the number of deer in the nature reserve at the end of the n th year.)

Using your model, plot a graph of deer numbers against the year for the first five years. Determine whether the population of deer will exceed 5000 over this time period and whether the model is realistic.

- 10** When purchased new, a machine for manufacturing car components costs \$2 500 000.
- a** Let V_n be the value of the machine at the end of the n th year of its working life assuming that it depreciates in value by a constant amount of \$200 000 per year.
- i** Write down an expression for the value of the machine after n years.
 - ii** Calculate the value of the machine at the end of each of the first five years of its working life.
 - iii** Plot these values on a graph.
- b** Let U_n be the value of the machine at the end of the n th year of its working life assuming that it depreciates in value by 10% of its value each year.
- i** Write down an expression for the value of the machine after n years.
 - ii** Calculate the value of the machine at the end of each of the first five years of its working life.
 - iii** Plot these values on a graph.
- c** From some points of view, the best depreciation method is the one that gives you the lowest value of the machine at the time of disposal. Determine which depreciation method you should use if you plan to keep the machine for:
- i** two years
 - ii** 10 years.

6E Topic 5: Earth geometry and time zones

Multiple-choice questions

- 1** City M has latitude 5°N and longitude 5°E . City N has latitude 35°S and longitude 5°E . The shortest distance along the meridian between M and N , in kilometres, is closest to:
- A** 4431
 - B** 4448
 - C** 6200
 - D** 3336
- 2** Location A has latitude 25°N and longitude 5°E . Location B has latitude 25°N and longitude 50°W . The distance along the small circle of the 25°N latitude between A and B, in kilometres, is closest to:
- A** 1112
 - B** 5543
 - C** 3833
 - D** 6062

- 3** The coordinates of two points X and Y on the Earth's surface are $(25^\circ\text{N}, 15^\circ\text{E})$ and $(25^\circ\text{S}, 45^\circ\text{W})$. Which statement is most likely to be correct about the time difference?
- A** X is 2 hours behind Y .
 - B** Y is 4 hours ahead of X .
 - C** Y is 5 hours behind X .
 - D** X is 4 hours ahead of Y .
- 4** Catherine lives in Cooktown 145°E . She wants to watch a basketball game in Los Angeles 118°W starting at 8 p.m. on Friday. What is the closest time in Cooktown to when the game starts in Los Angeles? (Ignore time zones and daylight saving.)
- A** 11 p.m. on Friday
 - B** 2 a.m. on Saturday
 - C** 9 a.m. on Saturday
 - D** 2 p.m. on Saturday



The map is to be used in the Questions 5–8.

- 5** Which City has latitude 33°N and longitude 84°W ?
- A** Dallas
 - B** Memphis
 - C** Atlanta
 - D** St Louis
- 6** We can consider Denver and Philadelphia to lie on the 40°N latitude. Given this, the distance between the two cities is closest to:
- A** 2555 km
 - B** 1780 km
 - C** 2023 km
 - D** 1943 km
- 7** The longitude of Boston is 71°W and the longitude of Seattle is 122°W . Use this to calculate the time in Seattle if it is 11:00 p.m. in Boston.
- A** 2:00 a.m. the next day
 - B** 8:00 p.m. the same day
 - C** 6:00 p.m. the same day
 - D** 1:00 a.m. the next day
- 8** Pittsburgh and Miami lie on the 75° meridian. The latitude of Pittsburgh is 40°N and the latitude of Miami is 26°N . The distance between the two cities is closest to:
- A** 1455 km
 - B** 1780 km
 - C** 1560 km
 - D** 1943 km

Short-response questions

- 1 Two locations lie on the same meridian of longitude. One is 37° south of the other. What is the distance between the two locations, correct to the nearest kilometre?
- 2 Two locations lie on the same parallel of latitude 15°S . One is 140° west of the other. What is the distance around the small circle between the two locations, correct to the nearest kilometre?
- 3 Vancouver, Canada, has longitude 118°W , while Cooktown has longitude 146°E .
 - a Calculate the difference in longitude between these two places.
 - b Calculate the time difference between the two places. (Ignore time zones and daylight saving.)
 - c What is the time in Cooktown when it is 6 p.m. in Vancouver? (Ignore time zones and daylight saving.)
- 4 How far is London (51.5°N , 0.1°W) from the:
 - a the equator?
 - b the north pole?
 - c the south pole?
- 5 Two locations A and B are situated on the 140°E meridian. Location A is on the 5°S parallel of latitude and B is north of A. If the distance between the two locations is 3000 km, determine the approximate position of location B.
- 6 Two places are situated on latitude 25°S . If their difference of longitude is 48° , determine the distance between the two places measured along the parallel of latitude.
- 7 Determine the length of the parallel of latitude correct to the nearest 10 km.
 - a 27°S
 - b 51°N
- 8 A plane flies from M (40°N , 40°E) over the north pole to N (40°N , 140°W).
 - a How far does the plane fly?
 - b If the plane flies from M to N around the 40°N parallel of latitude, how far is this?
 - c What is the difference between the distances?
- 9 A plane flies from a point with longitude 5°E on the equator and flies around the equator in a westerly direction for 3500 km. Give the approximate location of the plane.

- 10** Complete the following questions relating to the distance between meridians.
- a** What is the distance between the 145°E and 150°E meridians at the equator?
 - b** What is the distance between the 145°E and 150°E meridians at the 30°S parallel?
 - c** At what latitude is the distance between the 145°E and 150°E meridians 500 km?
- 11** A plane leaves location A (5°N , 5°E).
- a** If the plane flies south along the 5°E meridian until it reaches the 10°S parallel of latitude and then flies west along the 10°S parallel of latitude until it reaches the 50°W meridian, what is the total distance flown? (Final location B (10°S , 50°W)).
 - b** If the plane flies west along the 5°N parallel of latitude until it reaches the 50°W meridian and then flies south along the 50°W meridian until it reaches the 10°S meridian parallel of latitude, what is the total distance flown? (Final location B (10°S , 50°W)).
 - c** Find the difference of the two total distances.
 - d** Find the time difference between points A and B. (Ignore time zones and summer-time.)

6F List of Unit 3 assessment and examination practice online items

These assessment practice items can be found in the interactive textbook and in the teacher resources of the online teaching suite.

Interactive Textbook

For student and teacher access:

- 1** IA1: A practice PSMT from Unit 3 Topics 1–3.
- 2** IA2: A practice internal exam on Unit 3.

Online Teaching Suite

For teacher access:

- 1** IA1: A PSMT from Unit 3 Topics 1–3.
- 2** IA2: An internal exam on Unit 3.

Assessment items for Unit 4, and for Units 3 and 4 together, are listed at the end of Chapter 13.

Chapter 7

Loans, investments and annuities 1

Chapter questions

Unit 4 INVESTING AND NETWORKING

Topic 1: Loans, investments and annuities 1

- ▶ How do we use and construct a recurrence relation model for a compound interest loan or investment?
- ▶ How do we use the compounding interest formula for a compound interest loan or investment?
- ▶ How do we compare loans and investments with and without technology?
- ▶ How do we find unknown values in compound interest problems?
- ▶ How do we use a recursive model and repayment schedule for an ordinary annuity?
- ▶ How do we use the present value annuity formula to model the present value of an ordinary annuity?
- ▶ How do we solve practical problems involving the present value of an ordinary annuity?

In this chapter we explore loans, investments and annuities with a focus on compound interest loans and investments. We use both recurrence models and formulas to model compound interest loans and investments to find unknown values.

In addition, we model ordinary annuities using a recurrence relation then use the present value annuity formula to analyse and solve practical problems such as reducing balance loans and retirement pensions.

7A Using a recurrence relation to model compound interest loans and investments

Learning intentions

- To use a recurrence relation to model a compound interest loan or investment.
- To convert annual interest rates for different compounding periods.
- To construct recurrence relation models for compound interest loans and investments.

Mathematical modelling

Mathematical **modelling** is the process of describing or explaining a real-life situation using mathematical terms and symbols. The mathematical *model* that is created can be used to help us understand the situation clearly, so we can explore and make predictions.

Recurrence relations can be used as a mathematical model for many different situations and they will be particularly useful to explain or understand financial situations.

Decimal interest rates

The interest rate for a compound interest investment or loan is usually given as a percentage annual rate of interest. When we perform calculations using this percentage rate of interest, it must be converted to a decimal rate of interest by dividing by 100.

Converting percentage interest rates to decimal interest rates

Let i be the decimal interest rate for the percentage interest rate $x\%$ per annum.

$$i = \frac{x}{100}$$

Modelling compound interest situations with recurrence relations

In Section 4G, we introduced the idea of using recurrence relations to model and analyse compound interest loans and investments. Recall that compound interest loans and investments are when any interest that is earned after one time period is added to the principal and then contributes to the earning of interest in the next time period.

Consider a compound interest investment of \$5000 with an interest rate of 8% paid every year, or *per annum*.

Every year the balance of the investment will increase by 8% of the previous year's balance.

So : $\text{balance next year} = \text{balance this year} + \text{interest earned}$

= balance this year + 8% of the balance this year
 = 100% of the balance this year + 8% of the balance this year
 = 108% of the balance this year
 = $1.08 \times$ the balance this year

Let A_n be the balance of the investment after n years.

The starting value of the recurrence relation is the principal value, $A_0 = 5000$.

In recurrence relation symbols, the rule is:

$$A_{n+1} = 1.08 \times A_n$$

We now have a recurrence relation that can be used to model the balance of a compound interest loan or investment.

A recurrence relation model for a compound interest loan or investment

Let A_n be the balance of a compound interest loan or investment after n years.

Let i be the annual decimal rate of interest.

A recurrence relation model for the balance of a compound interest investment or loan is:

$$A_0 = \text{principal of loan or investment}, \quad A_{n+1} = r \times A_n$$

where $r = 1 + i$.

The total interest earned or charged after n years $= A_n - A_0$.

The interest earned or charged in the n th year $= A_n - A_{n-1}$.



Example 1 Constructing a recurrence relation model for compound interest

Darren borrows \$4000 from a bank. The bank will charge him interest at the annual rate of 9.8%.

Construct a recurrence relation model for the value of Darren's loan after n years.

Solution

Let A_n be the value of the loan after n years.

$$A_0 = 4000$$

$$x = 9.8$$

$$i = \frac{9.8}{100} = 0.098$$

$$r = 1 + 0.098$$

$$r = 1.098$$

The recurrence relation is

$$A_0 = 4000, A_{n+1} = 1.098 \times A_n$$

Explanation

Define the variable of the recurrence relation.

The amount borrowed initially, the principal, is the starting value, A_0 .

Calculate the value of r , using the decimal rate of interest, i .

Write your answer.


Example 2 Using recurrence relation models for compound interest loans and investments

The following recurrence relation can be used to model a compound interest loan of \$2000 charged interest at the percentage rate of 7.5% per annum.

$$A_0 = 2000, \quad A_{n+1} = 1.075 \times A_n$$

In the recurrence relation, A_n is the balance of the loan after n years.

- Use the recurrence relation to find the balance of the loan after one, two and three years.
- Find how much interest in total has been charged after three years.
- Find how much interest has been charged after the third year.
- Determine when the balance of the loan will first exceed \$2500.

Solution

a $A_0 = 2000$

2000	2000
ans $\times$ 1.075	2150
ans $\times$ 1.075	2311.25
ans $\times$ 1.075	2484.59375

After one year, the balance is \$2150.00

After two years, the balance is

\$2311.25 After three years, the balance is \$2484.59

b Interest after three years
 $= \$2484.59 - \$2000$
 $= \$484.59$

After three years, a total of \$484.59 in interest has been charged.

Explanation

Write down the principal of the loan, A_0 .

Use calculator recursion to apply the recurrence relation rule and calculate A_1 , A_2 and A_3 .

Note: The value after three years must be rounded to the nearest cent.

Write your answer.

Subtract the principal from the balance to calculate the total interest charged.

Write your answer.

c Interest in the third year

$$= A_3 - A_2$$

$$= \$2484.59 - \$2311.25$$

$$= \$173.34$$

After the third year, \$173.34 in interest has been charged.

Subtract the balance after two years from the balance after three years.

Write your answer.

d

2000	2000

ans $\times$ 1.075	2150

ans $\times$ 1.075	2311.25

ans $\times$ 1.075	2484.59375

ans $\times$ 1.075	2670.938281

The balance of the loan will first exceed \$2500 after 4 years.

Use calculator recursion to count how many years are required to reach the balance of \$2500.

Press = (Casio), or **enter** (TI), repeatedly, counting the number of times before the balance first exceeds \$2500.

Write your answer.

Nominal and compounding interest rates

Compound interest rates are usually quoted as an annual rate, or an interest rate per annum. This rate is called the **nominal interest rate** for the investment or loan. Sometimes an annual rate might be quoted, but the interest can be calculated and paid according to a different time period, such as monthly. The time period after which compound interest is calculated and paid is called the **compounding period**.

Annual (nominal) interest rates can be converted to interest rates for other compounding periods using simple arithmetic.

It must be assumed that there are:

- 12 equal months in every year (even though some months have different numbers of days)
- 4 quarters in every year (a quarter is equal to 3 months)
- 26 fortnights in a year (even though there are slightly more than this)
- 52 weeks in a year (even though there are slightly more than this)
- 365 days in a year (ignore the existence of leap years).

Converting between annual percentage rates of interest and decimal rates of interest per compounding period

Let x be the nominal annual percentage rate of interest.

Let k be the number of compounds per year.

Let i be the decimal interest rate per compounding period for the loan or investment.

$$i = \frac{x}{k \times 100} \quad \text{or} \quad x = k \times 100 \times i$$



Example 3 Converting annual interest rates to decimal interest rates per compounding period

An investment account will pay interest at the rate of 6.24% per annum.

Convert this to a decimal interest rate, i , if the interest compounds are:

a monthly

b fortnightly

c quarterly

Solution

a $i = \frac{6.24}{12 \times 100} = 0.0052$

b $i = \frac{6.24}{26 \times 100} = 0.0024$

c $i = \frac{6.24}{4 \times 100} = 0.0156$

Explanation

Divide the interest rate by 12×100 .

Divide the interest rate by 26×100 .

Divide the interest rate by 4×100 .



Example 4 Converting decimal interest rates to annual interest rates per compounding period

Convert the following decimal interest rates to annual rates.

a 0.0024 (monthly)

b 0.015 (quarterly)

c 0.00312 (weekly)

Solution

a Annual interest rate = $0.0024 \times 12 \times 100$
= 2.88% per annum

b Annual interest rate = $0.015 \times 4 \times 100$
= 6% per annum

c Annual interest rate = $0.00312 \times 52 \times 100$
= 16.224% per annum

Explanation

Multiply the interest rate by 12×100 .

Multiply the interest rate by 4×100 .

Multiply the interest rate by 52×100 .

The recurrence relation model for compound interest with compounding periods other than one year can now be written.

A recurrence relation model for a compound interest loan or investment

Let A_n be the balance of a compound interest loan or investment after n compounding periods.

Let i be the decimal rate of interest, per compounding period.

A recurrence relation model for the balance of a compound interest investment or loan is:

$$A_0 = \text{principal of loan or investment}, \quad A_{n+1} = r \times A_n$$

where $r = 1 + i$.

The total interest earned or charged after n compounding periods $= A_n - A_0$.

The interest earned or charged in the n th compounding period $= A_n - A_{n-1}$.



Example 5 Constructing recurrence relation models for compound interest loans and investments

Diego will invest \$7500 and will earn compound interest at the nominal rate of 9.6% per annum.

Let A_n be the balance of the investment after n compounding periods.

Construct a recurrence relation to model the balance of Diego's investment if interest is compounded:

a yearly

b quarterly

c monthly

Solution

$$\mathbf{a} \quad i = \frac{9.6}{1 \times 100} = 0.096$$

$$r = 1 + 0.096 = 1.096$$

$$A_0 = 7500, A_{n+1} = 1.096 \times A_n$$

$$\mathbf{b} \quad i = \frac{9.6}{4 \times 100} = 0.024$$

$$r = 1 + 0.024 = 1.024$$

$$A_0 = 7500, A_{n+1} = 1.024 \times A_n$$

$$\mathbf{c} \quad i = \frac{9.6}{12 \times 100} = 0.008$$

$$r = 1 + 0.008 = 1.008$$

$$A_0 = 7500, A_{n+1} = 1.008 \times A_n$$

Explanation

Calculate the decimal rate of interest per year (one compounding period).

Calculate the value of r .

Write the recurrence relation.

Calculate the decimal rate of interest per quarter (four compounding periods).

Calculate the value of r .

Write the recurrence relation.

Calculate the decimal rate of interest per month (twelve compounding periods).

Calculate the value of r .

Write the recurrence relation.

Graphs of compound interest

Consider an investment of \$5000 earning interest at the rate of 10% per annum.

The table on the right shows the rounded balances (where necessary) for both simple interest and compound interest, after each year for a period of six years.

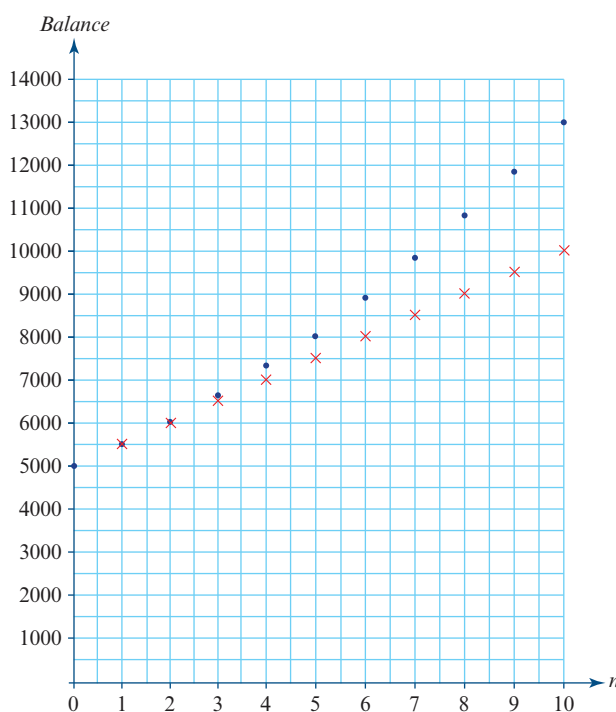
Both investments grow in value over time. After one year, the balance is the same, but after each subsequent year, the compound interest investment balance is higher than that of the simple interest investment.

The graph shows the year number, n , on the horizontal axis and the balance of the investment on the vertical axis.

Crosses are used to represent the balances of the simple interest investment and dots are used to represent the balances of the compound interest investment.

The graph of the simple interest investment is a straight line because it grows by a constant amount each year, while the graph of the compound interest investment curves upwards because it grows by an amount that increases each year.

Year (n)	Simple interest	Compound interest
0	5000.00	5000.00
1	5500.00	5500.00
2	6000.00	6050.00
3	6500.00	6655.00
4	7000.00	7320.50
5	7500.00	8052.55
6	8000.00	8857.81



As the number of years for the investments increases, the balance of the compound interest investment grows much larger than the simple interest balance. After ten years, the compound interest investment is almost \$3000 more than the simple interest investment.



Desmos activity 7A: To investigate and compare the compound interest growth rates for different loans and investments

The effect of the compounding period on total interest

Changing the compounding period of a compound interest loan or investment has an effect on the total interest that is earned or charged.

You can use recurrence relations to investigate the effect of compounding periods by using the investigation below.

Investigation 7A-1: Investigating the effect of compounding period using a recurrence relation.



Investigation 7A-2: Investigating the effect of compounding period using a spreadsheet.

Changing the compounding period impacts the amount of interest that is applied to the balance of the loan or investment.

The effect of changing the compounding period on compound interest

- Increasing the number of compounding periods per year will mean more interest is earned or charged over the same period of time.
- For investments, from the viewpoint of the investor, interest should compound as often as possible to maximise total interest.
- For loans, from the viewpoint of the borrower, interest should compound as infrequently as possible to minimise total interest.



Example 6 Using a spreadsheet or recurrence relation to analyse compounding periods for loans and investments

- a** Use the compound interest spreadsheet or a recurrence relation to complete the table below.

Principal: \$4000	
Annual interest rate: 4.2%	
Compounds per year	Balance after 1 year
1	
2	
4	
12	
52	

- b** If this was an investment that was closed after one year, find how much extra interest is earned by choosing monthly compounds instead of yearly compounds.
- c** If this was a loan that was paid out after one year, find how much interest is saved by choosing monthly compounds instead of weekly compounds.

Solution

a Using a spreadsheet:

Principal: \$4000	
Annual interest rate: 4.2%	
Compounds per year	Balance after 1 year
1	\$4168.00
2	\$4169.76
4	\$4170.66
12	\$4171.27
52	\$4171.51

b Extra interest = \$4171.27 – \$4168.00
= \$3.27

Compared to yearly compounds, monthly compounds would earn an extra \$3.27 interest.

c Interest saved = \$4171.51 – \$4171.27
= \$0.24

Compared to weekly compounds, monthly compounds would save \$0.24 interest.

Explanation

Use the spreadsheet to complete the table, or use the recurrence relations:

$$A_0 = 4000, A_{n+1} = 1.042 \times A_n$$

$$A_0 = 4000, A_{n+1} = 1.021 \times A_n$$

$$A_0 = 4000, A_{n+1} = 1.0105 \times A_n$$

$$A_0 = 4000, A_{n+1} = 1.0035 \times A_n$$

$$A_0 = 4000, A_{n+1} = 1.000807692 \times A_n$$

Extra interest is the difference between the monthly value and the annual value

Write your answer.

Interest saved is the difference between the weekly value and the monthly value.

Write your answer.

Section Summary

- ▶ A recurrence relation for a compound interest loan or investment is $A_0 =$ principal of loan or investment, $A_{n+1} = r \times A_n$ where $r = 1 + i$. The total interest earned or charged in n years = $A_n - A_0$. The interest earned or charged after the n th year = $A_n - A_{n-1}$.
- ▶ If x is the nominal annual interest rate and k is the number of compounds per year, then $i = \frac{x}{k \times 100}$ is the decimal interest rate per compounding period for the loan or investment.
- ▶ Increasing the number of compounding periods per year will mean more interest is earned or charged over the same period of time. For investments, more compounding periods is better while for loans, fewer compounding periods is better.

Exercise 7A

Constructing recurrence relation models

Example 1

- 1 Franco borrows \$5000 from a bank. The bank will charge him interest at the annual rate of 8.4%, compounding annually. Construct a recurrence relation model for the value of Franco's loan after n years.
- 2 Grace invests \$8500 and will earn interest at the rate of 4.2%, compounding annually. Construct a recurrence relation model for the value of Grace's investment after n years.

Example 2

- 3 The following recurrence relation can be used to model a compound interest loan of \$2000 earning interest at the annual rate of 2.5% per annum.

$$A_0 = 2000, \quad A_{n+1} = 1.025 \times A_n$$

In the recurrence relation, A_n is the balance of the loan after n years.

- a Use the recurrence relation to find the balance of the loan after one, two and three years.
 - b Find how much interest in total has been charged after three years.
 - c Find how much interest has been charged in the third year.
 - d Determine when the balance of the loan will first exceed \$2500.
- 4 The following recurrence relation can be used to model a compound interest investment of \$25 000 earning interest at the annual rate of 6.4% per annum.

$$A_0 = 25\,000, \quad A_{n+1} = 1.064 \times A_n$$

In the recurrence relation, A_n is the balance of the investment after n years.

- a Use the recurrence relation to find the balance of the investment after one, two and three years.
- b Find how much interest in total has been earned after three years.
- c Find how much interest has been earned in the third year.
- d Determine when the balance of the investment will first exceed \$40 000.

Converting between annual percentage rates of interest and decimal rates of interest

Example 3

- 5 Convert the following annual percentage rates of interest to decimal rates of interest for the given compounding periods.
- a 7.2% per annum, compounding monthly
 - b 11.16% per annum, compounding monthly
 - c 8.06% per annum, compounding fortnightly
 - d 13.52% per annum, compounding fortnightly
 - e 7.6% per annum, compounding quarterly
 - f 10.44% per annum, compounding quarterly

Example 4

- 6** Convert the following decimal rates of interest to annual percentage rates of interest for the given compounding periods.
- a** 0.0375 (compounding yearly)
 - b** 0.064 (compounding yearly)
 - c** 0.008 (compounding monthly)
 - d** 0.0094 (compounding monthly)
 - e** 0.0153 (compounding quarterly)
 - f** 0.0042 (compounding fortnightly)

Constructing recurrence relation models for compound interest

Example 5

- 7** Jackson borrows \$2000 and will be charged interest at the nominal rate of 8.4% per annum, compounding monthly. Let J_n be the balance of the loan after n compounding periods. Construct a recurrence relation to model the balance of Jackson's loan.
- 8** Rupert invests \$34 000 and will earn interest at the nominal rate of 9.36% per annum, compounding weekly. Let R_n be the balance of the investment after n compounding periods. Construct a recurrence relation to model the balance of Rupert's investment.

Example 6

- 9** Use a spreadsheet or recurrence relation to complete the tables below. If you use a recurrence relation, do not fill the last row of the table (52 compounds per year).

Principal: \$12 000 Annual interest rate: 5.2%	
Compounds per year	Balance after 1 year
a 1	
2	
4	
12	
52	

Principal: \$25 000 Annual interest rate: 9.4%	
Compounds per year	Balance after 1 year
b 1	
2	
4	
12	
52	

- 10** An investment of \$6000 earns compounding interest at the rate of 5.76% per annum, compounding monthly. A recurrence relation that can be used to model the balance of the investment, A_n over n months is

$$A_0 = 6000, A_{n+1} = 1.0048 \times A_n$$

- a** Apply the recurrence relation to find the balance after one, two and three months.
- b** Find how much interest is earned in total after three months.
- c** Find how much interest is earned in the second month.
- d** Determine how many months it will take for the total interest earned to exceed \$200.

- 11** The following recurrence relation can be used to model a simple interest investment of \$2000 earning interest at the rate of 3.8% per annum.

$$A_0 = 2000, A_{n+1} = A_n + 76$$

In the recurrence relation, A_n is the value of the investment after n years.

- a** Apply the recurrence relation to find the value of the investment after one, two and three years.
- b** Use your calculator to determine how many years it takes for the value of the investment to first be worth more than \$3000.
- 12** The following recurrence relation can be used to model a simple interest loan of \$7000 being charged interest at the rate of 7.4% per annum.

$$A_0 = 7000, A_{n+1} = A_n + 518$$

In the recurrence relation, A_n is the value of the loan after n years.

- a** Apply the recurrence relation to find the value of the loan after one, two and three years.
- b** Use your calculator to determine how many years it takes for the value of the loan to first be worth more than \$10 000.
- 13** A loan of \$8400 is charged compounding interest at the rate of 12.6% per annum, compounding monthly. A recurrence relation that can be used to model the balance of the loan, A_n , after n months is:

$$A_0 = 8400, A_{n+1} = 1.0105 \times A_n$$

- a** Apply the recurrence relation to find the balance of the loan after one, two and three months.
- b** Find how much interest is charged after the first month.
- c** Find how much interest is charged in the second month.
- d** Find how much interest is charged in total after three months.
- e** Determine how many months it takes for the total interest charged to exceed \$1000.
- 14** Tom has invested \$7600 and will earn compound interest at the rate of 6% per annum, compounding monthly. Let the balance of Tom's investment be A_n after n months.
- a** State the monthly interest rate for Tom's investment.
- b** Construct a recurrence relation model for the balance of Tom's investment.
- c** Find the balance of Tom's investment after six months.
- 15** Jun has borrowed \$3500 and will be charged compound interest at the rate of 8% per annum, compounding quarterly. Let the balance of Jun's loan be A_n after n quarters.
- a** State the quarterly interest rate for Jun's loan.
- b** Construct a recurrence relation model for the balance of Jun's loan.
- c** Jun fully repays his loan after one year. Find how much he will need to repay.
- d** Determine how much extra interest Jun will be charged if he waits another year to repay the loan.

- 16** A sum of \$12 800 is invested into an account earning compound interest at the rate of 5.8% per annum.
- a** If there is a choice, determine whether an investor should choose weekly or monthly compounds.
 - b** Use a spreadsheet or recurrence relation to calculate the difference in interest earned during the first year of investment by monthly and weekly compounds.
- 17** A sum of \$3500 is borrowed from a money lender that charges compound interest at the rate of 14.8% per annum.
- a** If there is a choice, determine whether a borrower should choose quarterly or fortnightly compounds.
 - b** Use a spreadsheet or recurrence relation to calculate the difference in interest charged during the first year of the loan by quarterly and fortnightly compounds.

Paper 1-style multiple-choice questions

- 18** A loan of \$9200 is charged compounding interest at the rate of 7.2% per annum, compounding annually. The recurrence relation that can be used to model the balance of the loan, A_n , after n years is $A_0 = 9200$, $A_{n+1} = 1.072 \times A_n$. The amount of interest charged over three years is
- A** \$662.40
 - B** \$761.22
 - C** \$2133.71
 - D** \$11 333.71
- 19** An investment that compounds quarterly is modelled by a recurrence relation $A_0 = 6000$, $A_{n+1} = 1.0145 \times A_n$. The principal and the annual interest rate are
- A** \$1500, 1.45%
 - B** \$6000, 4.058%
 - C** \$6000, 1.45%
 - D** \$6000, 5.8%
- 20** A loan of \$12 800 is taken out at an interest rate of 6.63% per annum. The loan compounds monthly. If A_n is the value of the loan after n months, the recurrence relation can be written as
- A** $A_0 = 12\,800$, $A_{n+1} = 1.0663 \times A_n$
 - B** $A_0 = 12\,800$, $A_{n+1} = 1.00633 \times A_n$
 - C** $A_0 = 12\,800$, $A_{n+1} = 1.005525 \times A_n$
 - D** $A_0 = 12\,800$, $A_{n+1} = 1.5525 \times A_n$

7B Using a compound interest formula to model a compound interest loan or investment

Learning intentions

- To use the compound interest formula to model a compound interest loan or investment.

A rule for the balance of compound interest loans and investments

Consider an investment of \$2000 that earns compound interest at the rate of 5% per annum, compounding yearly. If we let A_n be the balance of this investment after n years, the following recurrence relation can be used to model this investment:

$$A_0 = 2000, A_{n+1} = 1.05 \times A_n$$

Using this recurrence relation, we can write out the sequence of terms it generates as follows:

$$A_0 = 2000$$

$$A_1 = 1.05 \times A_0$$

$$A_2 = 1.05 \times A_1 = 1.05 (1.05 \times A_0) = 1.05^2 \times A_0$$

$$A_3 = 1.05 \times A_2 = 1.05 (1.05^2 \times A_0) = 1.05^3 \times A_0$$

$$A_4 = 1.05 \times A_3 = 1.05 (1.05^3 \times A_0) = 1.05^4 \times A_0$$

and so on.

Following this pattern, after n years, the balance of the investment will be $A_n = 1.05^n \times A_0$.

Instead of using recurrence relation symbols in this rule, we can use P to represent the principal amount of the loan (A_0) or investment and A to represent the *future value* of the loan or investment after n compounding periods (A_n).

This rule allows the balance of a compound interest loan or investment after any number of compounding periods to be calculated.

A rule for the future value of compound interest loans and investments

Let A be the future value of a compound interest loan or investment.

Let n be the total number of compounding periods.

Let i be the decimal rate of interest per compounding period.

Let P be the principal of the loan or investment.

The future value of the compound interest loan or investment after n compounding periods is

$$A = P(1 + i)^n$$

**Example 7** Using the rule for the future value of compound interest loans and investments

Tyson would like to borrow \$8000 and will pay compound interest at the rate of 5.6% per annum, compounding annually.

Find the balance of Tyson's loan after 5 years.

Solution

$$P = 8000$$

$$i = 0.056$$

$$n = 5$$

$$A = P(1 + i)^n$$

$$\begin{aligned} A &= 8000 \times (1 + 0.056)^5 \\ &= 10\,505.3270657 \dots \end{aligned}$$

The balance of Tyson's loan is \$10 505.33 after 5 years.

Explanation

Write down the value of P , i and n .

Apply the rule to find the future value, A .

Write your answer, rounded to the nearest cent.

**Example 8** Using the rule for the future value of compound interest loans and investments

Bongile would like to invest \$25 000 into an account that will pay her compound interest at the rate of 4.2% per annum, compounding monthly.

- a** Find the balance of Bongile's investment after 10 years.
- b** Find how much interest Bongile's investment earned in total.
- c** Find how much Bongile's investment earned in the 10th year.

Solution	Explanation
a $P = 25\,000$ $i = \frac{4.2}{12 \times 100}$ $= 0.0035$ $n = 10 \times 12 \text{ months}$ $= 120 \text{ months}$ $A = P \times (1 + i)^n$ $A = 25\,000 \times (1 + 0.0035)^{120}$ $= 38\,021.14816$ The balance of Bongile's investment is \$38 021.15 after 10 years of investment.	Write down the value of P and i. Calculate the number of compounding periods in the time of the loan, n. Apply the rule to find the future value, A. Write your answer, rounded to the nearest cent.
b interest = $38\,021.15 - 25\,000$ $= 13\,021.15$	Calculate the difference between the principal and the balance after 10 years.
c $P = 25\,000$ $i = \frac{4.2}{12 \times 100}$ $= 0.0035$ $n = 9 \times 12 \text{ months}$ $= 108 \text{ months}$ $A = P \times (1 + i)^n$ $A = 25\,000 \times (1 + 0.0035)^{108}$ $= 36\,460.0034665$ interest = $38\,021.15 - 36\,460$ $= 1561.146533$ The interest earned in the 10th year is \$1561.15	Write down the value of P, i and n. Apply the rule to find the future value, A after 9 years. Calculate the increase in the value after 9 years. Write your answer, rounded to the nearest cent.

Section Summary

- ▶ The future value of the compound interest loan or investment with a principal of P and an interest rate of i after n compounding periods is $A = P \times (1 + i)^n$.
- ▶ To find the amount of interest earned or paid, calculate the difference between the value of the loan or investment after a different number of compounding periods.

Exercise 7B

Using the rule to determine the future value of compound interest loans and investments

Example 7

- 1** Hugh borrowed \$2500 with compound interest of 3.5% per annum, compounding yearly.
 - a** Write down a rule for the value of Hugh's loan, H_n , in terms of n .
 - b** Use the compound interest rule to determine the balance of Hugh's loan after 3 years.
- 2** Griffin invested \$15 000 with compound interest of 2.8% per annum, compounding yearly.
 - a** Write down a rule for the value of Griffin's investment, G_n , in terms of n .
 - b** Use the compound interest rule to determine the balance of Griffin's investment after 4 years.

Example 8

- 3** Taylor borrowed \$5800 at a compound interest rate of 9.6% per annum, compounding quarterly. Use the compound interest rule to determine the balance of Taylor's loan for one year.
- 4** Levi invested \$26 000 with compound interest of 4.6% per annum, compounding monthly. Use the compound interest rule to determine the balance of Levi's investment after 60 months.
- 5** Will borrowed \$6400 at a compound interest rate of 8.5% per annum, compounding weekly. Use the compound interest rule to determine the balance of Will's loan for five years.
- 6** Rhiannon invested \$12 500 with compound interest of 4.7% per annum, compounding monthly. Use the compound interest rule to determine the balance of Rhiannon's investment after 6 years.

Analysing compound interest loans and investments using a rule

- 7** Bhavna invested \$10 000 into an account that will pay her compound interest at the rate of 5.2% per annum, compounding annually.
 - a** Write down a rule for the value of Bhavna's investment.
 - b** Calculate the amount of interest Bhavna earned on the investment over 10 years.
- 8** Damon borrowed \$14 500 for 5 years at a compound interest rate of 7.7% per annum, compounding annually. Calculate the amount of interest Damon must pay.

- 9** Indi invested \$13 800 into an account for five years that will pay her compound interest at the rate of 4.8% per annum, compounding quarterly. Calculate the amount of interest Indi earned in the fifth year of her investment.
- 10** Rob borrowed \$22 100 at a compound interest rate of 4.8% per annum, compounding monthly, to be paid at the end of the loan. Calculate the amount of interest Rob paid in the fourth year of his loan.
- 11** Nathan invests \$5000 at a compound interest rate of 7.2% per annum, compounding monthly. Calculate the amount of interest that Nathan earns on his investment in each of the first 5 years of the investment.
- 12** Sarah wants to borrow \$14 500 to purchase a new car. Bank A offers her a loan of 6.5% per annum, compounding annually for 3 years. Bank B offers her a loan of 6.4% per annum, compounding quarterly for 4 years. Bank C offers her a loan of 6% per annum, compounding monthly for 5 years. Determine which loan would see Sarah pay the least amount of interest for the whole loan.
- 13** A simple interest investment offers an annual interest rate of 5.2% while a compounding interest rate is advertised at 4.8% per annum, compounding monthly. Christian wishes to invest \$12 080. Determine the fewest number of years he would need to invest for the compound interest rate to pay more interest than the simple interest investment.

Paper 1-style multiple-choice questions

- 14** Bryan borrowed \$3200 with compound interest of 3.6% per annum, compounding quarterly. The rule for the balance of the loan, A , in terms of n is
- A** $A = 2000 + 3.6n$
B $A = 2000 \times 3.6n$
C $A = 2000 \times 1.036^n$
D $A = 2000 \times 1.009^n$
- 15** Salma invests \$6800 for 8 years at an interest rate of 7.2% per annum, compounding monthly. The rule for the balance of the investment, A , after 8 years is
- A** $A = 6800 \times 1.072^8$
B $A = 6800 \times 1.72^8$
C $A = 6800 \times 1.006^8$
D $A = 6800 \times 1.006^{96}$
- 16** Tristan invested \$10 000 for 9 years at an interest rate of 8.2% per annum, compounding quarterly. The amount of interest earned in the 9th year was
- A** \$1540.39 **B** \$1618.68 **C** \$20 325.69 **D** \$20 761.96

7C Effective annual rate of interest

Learning intentions

- ▶ To compare loans and investments with effective annual rates of interest rates.
- ▶ To use spreadsheet calculators to compare loans or investments using effective interest rates.

Defining the effective annual rate of interest

As we discussed in Section 7A, the more frequently interest is calculated and compounded, the more rapidly the value of the investment or loan increases.

The table below compares the value of a \$5000 investment earning interest at the nominal rate of 4.8% per annum with the value of the investment with interest calculated on a quarterly and monthly basis. When interest is added monthly, the investment earned \$245.35 in interest, which is greater than the \$240.00 earned with interest added monthly.

Principal of investment: \$5000			
Nominal annual interest rate: 4.8%			
	Value of investment for interest earned at the rate of:		
Month	4.8% per annum	1.2% per quarter	0.4% per month
0	5000.00	5000.00	5000.00
1			5020.00
2			5040.08
3		5060.00	5060.24
4			5080.48
5			5100.80
6		5120.72	5121.21
7			5141.69
8			5162.26
9		5182.17	5182.91
10			5203.64
11			5224.45
12	5240.00	5244.35	5245.35
Total interest earned*	240.00	244.35	245.35
Effective annual interest rate	4.80%	4.89%	4.91%

***Note:** The total interest earned is the value of the investment at the end of the year less the principal.

The **effective annual rate of interest** of a loan or investment is the annual interest rate that would generate the same amount of interest with one single compound per year as that generated by the original loan or investment.

For example, the effective annual rate of interest for the investment above compounding monthly is the interest rate that would earn \$245.25 interest in one annual compound. This can be calculated by writing the interest amount as a percentage of the principal value.

$$\text{effective interest rate for 1 compound per year} = \frac{240}{5000} \times 100\% = 4.8\%$$

$$\text{effective interest rate for 4 compounds per year} = \frac{244.35}{5000} \times 100\% = 4.887\%$$

$$\text{effective interest rate for 12 compounds per year} = \frac{245.35}{5000} \times 100\% = 4.907\%$$

An investment with a nominal interest rate of 4.907% per annum with one yearly compound will earn the same interest (\$245.35) as an investment with a nominal interest rate of 4.8% but with monthly compounds.

A rule for effective annual rate of interest

The calculations above required the amount of interest earned or charged to be known before the effective annual rate of interest was calculated. It is possible to use a rule to calculate the effective annual rate of interest for a loan or investment given the nominal interest rate and the number of compounding periods per year.

Effective annual rate of interest

The effective annual rate of interest of a loan or investment is the nominal interest rate with one annual compound that generates the same amount of interest over the course of one year.

Let n be the number of compounding periods in one year and i be the annual decimal rate of interest.

Then the effective annual decimal rate of interest, $i_{\text{effective}}$ for the loan is

$$i_{\text{effective}} = \left(1 + \frac{i}{n}\right)^n - 1$$

Note: To convert to an annual percentage rate of interest, multiply by 100%.



Example 9 Comparing loans and investments with effective annual rates of interest

Brooke would like to borrow \$20 000. She is deciding between two loan options:

- Option A: 5.95% per annum compounding weekly
- Option B: 6% per annum compounding quarterly.
 - a Calculate the effective annual rate of interest for each loan.
 - b Determine which loan option is the best and why.

Solution

a	Option A	Option B
	$n = 52$	$n = 4$
	$i = \frac{5.95}{100} = 0.0595$	$i = \frac{6}{100} = 0.06$
	$i_{\text{effective}} = \left(1 + \frac{0.0595}{52}\right)^{52} - 1$	$i_{\text{effective}} = \left(1 + \frac{0.06}{4}\right)^4 - 1$
	$= 0.06126 \dots$	$= 0.06136 \dots$
	Annual percentage effective rate	Annual percentage effective rate
	$= 0.06126 \dots \times 100\% = 6.127\%$	$= 0.06136 \dots \times 100\% = 6.136\%$
	$= 6.13\%$ (to two decimal places)	$= 6.14\%$ (to two decimal places)

- b Brooke is borrowing money, so the best option is the one with the lowest effective interest rate. She will pay less interest with option A.

Note: Either the effective decimal interest rate or effective percentage interest rate can be compared.

Explanation

- a Write down the values of n for each loan and calculate the annual decimal rate of interest for each loan. Apply the effective interest rate formula, then convert the effective rates to annual percentage rates.
- b Compare the effective interest rates.

If the previous example instead modelled Brooke **investing** \$20 000 (instead of borrowing), she would prefer a **higher** effective interest rate. Thus, she would prefer option B.

The amount of interest charged or earned over a particular time period depends on the number of compounds within that time period. In a short period of time, the number of compounds per year has little effect on the total interest charged or earned. Over a long period of time, however, the number of compounds per year can have a significant effect on the total interest charged or earned. For example, most mortgages in Australia are for 30 years and compound monthly. Thus, there is a large effect.



Spreadsheet activity 7C: A spreadsheet calculator can be used to find the effective annual rate of interest with ease.



Example 10 Using a spreadsheet calculator to compare loans or investments using effective interest

Ronnie has \$30 000 to invest. She has the choice of two investment accounts:

- Account 1 pays compound interest at 6.2% per annum, compounding monthly.
- Account 2 pays compound interest at 6.05% per annum, compounding weekly.

Determine which investment Ronnie should choose.

Solution

	A	B	C	D	E
1	Effective annual rate of interest calculator				
2	Enter the nominal percentage interest rate and compounds per year into the boxes. Click 'Calculate' to find the effective annual rate of interest. Click 'Clear' to remove all numbers and start again.				Clear
3					
4		nominal percentage interest rate		6.2	
5			compounds per year	12	
6			effective annual rate of interest	6.379253169	Calculate

	A	B	C	D	E
1	Effective annual rate of interest calculator				
2	Enter the nominal percentage interest rate and compounds per year into the boxes. Click 'Calculate' to find the effective annual rate of interest. Click 'Clear' to remove all numbers and start again.				Clear
3					
4		nominal percentage interest rate		6.05	
5			compounds per year	52	
6			effective annual rate of interest	6.233023748	Calculate

Account 1 has the higher effective annual rate of interest (6.379%) compared to Account 2 (6.233%) and so Ronnie will earn more interest with Account 1.

Explanation

Use the effective annual rate of interest spreadsheet calculator (or formula) to find the effective annual rate of interest for Account 1.

Use the effective annual rate of interest spreadsheet calculator (or formula) to find the effective annual rate of interest for Account 2.

Ronnie is investing and so should choose the account that has the highest effective annual rate of interest.

If the previous example had instead modelled Ronnie **borrowing** \$30 000 (instead of investing), then Ronnie would want a **smaller** effective annual rate of interest and so would have preferred Account 2.

Section Summary

- The effective annual rate of interest of a loan or investment is the annual interest rate that would generate the same amount of interest with one single compound per year as that generated by the original loan or investment. It is given by

$$i_{\text{effective}} = \left(1 + \frac{i}{n}\right)^n - 1$$

where i is the annual decimal rate of interest and n is the number of compound periods in one year.

- When borrowing, a **smaller** effective annual rate of interest is preferred, but if investing, a **larger** effective annual rate of interest is preferred.

Exercise 7C**Calculating effective annual rates of interest****Example 9**

- 1** Calculate the effective annual rate of interest for the following nominal annual interest rates and compounding periods using the rule. Round your answer to two decimal places.
- a** 5.4% per annum compounding monthly
 - b** 8.4% per annum compounding daily
 - c** 4.8% per annum compounding weekly
 - d** 12.5% per annum compounding quarterly
 - e** 7.5% per annum compounding every six months

Example 10

- 2** Calculate the effective annual rate of interest for the following nominal annual interest rates and compounding periods using a spreadsheet. Round your answer to two decimal places.
- a** 6.8% per annum compounding monthly
 - b** 5.6% per annum compounding daily
 - c** 11.2% per annum compounding weekly
 - d** 4.5% per annum compounding quarterly
 - e** 2.3% per annum compounding every six months

Problem-solving and modelling

- 3** Brenda invests \$15 000 in an account earning nominal compound interest of 4.60% per annum, compounding quarterly.
- a** Explain why Brenda would be better off with *more frequent* compounds per year.
 - b** Calculate the effective annual rate of interest for the current investment with quarterly compounds, correct to two decimal places.
 - c** Calculate the effective annual rate of interest for this investment with monthly compounds, correct to two decimal places.
 - d** Explain how these effective annual rates of interest support your answer to part **a**.
- 4** Stella borrows \$25 000 from a bank and pays nominal compound interest of 7.94% per annum, compounding fortnightly.
- a** Explain why Stella would be better off with *less frequent* compounds per year.
 - b** Calculate the effective annual rate of interest for the current loan with fortnightly compounds, correct to two decimal places.
 - c** Calculate the effective annual rate of interest for this loan with monthly compounds, correct to two decimal places.
 - d** Explain how these effective annual rates of interest support your answer to part **a**.

- 5** Luke is considering a loan of \$35 000. His bank has two compound interest rate options:
- A: 8.3% per annum, compounding monthly
B: 7.8% per annum, compounding weekly.
- a** Calculate the effective annual rate of interest for each of the loan options.
b Calculate the amount of interest Luke would pay in the first year for each of the loan options.
c Determine which loan Luke should choose and why.
- 6** Sharon is considering investing \$140 000. Her bank has two compound interest investment options:
- A: 5.3% per annum, compounding monthly
B: 5.5% per annum, compounding quarterly.
- a** Calculate the effective annual rate of interest for each of the investment options.
b Calculate the amount of interest Sharon would earn in the first year for each of the investment options.
c Determine which investment option Sharon should choose and why.

Paper 1-style multiple-choice questions

- 7** Isla is considering investing some money. The investment option that gives the best return is
- A** 5.92% p.a. compounding daily
B 5.94% p.a. compounding monthly
C 5.96% p.a. compounding quarterly
D 5.98% p.a. compounding six-monthly
- 8** Timothy wishes to take out a loan of \$6000 for 3 years. The option that he should take is
- A** 6.13% p.a. compounding daily
B 6.15% p.a. compounding monthly
C 6.17% p.a. compounding quarterly
D 6.18% p.a. compounding six-monthly
- 9** Jane has a two year loan of \$5000 that compounds monthly and has an interest rate of 6.3% per annum. The effective percentage interest rate is closest to
- A** 5.25% **B** 0.00525% **C** 1.06485% **D** 6.485%

7D Practical problems involving compound interest loans and investments

Learning intentions

- ▶ To use the rule for compounding interest loans and investments to find the principal or interest rate.
- ▶ To use technology, such as a spreadsheet, to find unknown values in compound interest problems.

Using the rule for compounding loans and investments

The compound interest rule used in Section 7B can be rearranged into alternate forms. These forms allow the calculation of the principal amount of the loan or investment and the annual percentage interest rate of the loan or investment.

Rule for the principal and interest rates of compounding interest loans and investments

Let A be the future value of a compound interest loan or investment.

Let n be the total number of compounding periods.

Let i be the decimal rate of interest per compounding period.

Let P be the principal of the loan or investment.

The principal of the compound interest loan or investment is:

$$P = \frac{A}{(1 + i)^n}$$

The decimal interest rate per compounding period for the loan or investment is:

$$i = \left(\frac{A}{P}\right)^{\frac{1}{n}} - 1$$

Note: The decimal interest rate per compounding period, i , can be converted to a percentage rate per compounding period by multiplying by 100 and then to an annual percentage rate by multiplying by the number of compounds per year.



Example 11 Using the rule for the principal of compound interest loans and investments

Lowanna has been offered the opportunity to invest some money. She will earn interest at the annual percentage interest rate of 9% per annum, compounding quarterly. Lowanna would like to have at least \$15 000 in her investment after 4 years.

Find how much Lowanna should invest in order to achieve this goal. Round your answer to the nearest dollar.

Solution

$$n = 4 \times 4 \text{ quarters} = 16 \text{ quarters}$$

$$i = \frac{9}{4 \times 100} = 0.0225$$

$$A = 15\,000$$

$$P = \frac{15\,000}{(1 + 0.0225)^{16}}$$

$$P = 10\,506.98693$$

To have a balance of \$15 000 after 4 years,
Lowanna should invest \$10 507 now.

Explanation

Write down the values of n and i .

The future value is the amount of
Lowanna's savings goal.

Apply the rule to find the value of P .

Write your answer, rounding to the nearest
dollar.

**Example 12****Using the rule for the annual percentage interest rate of compound interest loans and investments**

Daaruk has \$35 000 to invest now and would like this investment to grow to at least \$45 000 over a period of six years.

If Daaruk's investment earns interest that compounds monthly, determine the minimum annual percentage interest rate that he would require in order to achieve his savings goal.

Round your answer to one decimal place.

Solution

$$P = \$35\,000$$

$$n = 6 \times 12 \text{ months} = 72 \text{ months}$$

$$A = \$45\,000$$

$$i = \left(\frac{A}{P}\right)^{\frac{1}{n}} - 1$$

$$i = \left(\frac{45\,000}{35\,000}\right)^{\frac{1}{72}} - 1$$

$$i = 0.003496\dots$$

Annual percentage interest rate

$$= 0.003496 \times 12 \times 100$$

$$= 4.19589\dots$$

Daaruk would need an annual percentage
interest rate of 4.2% to achieve his savings
goal.

Explanation

Write down the value of P , n and A .

Apply the rule to find the annual
percentage rate of interest.

Convert i to an annual percentage interest
rate.

Write your answer rounded to one decimal
place.

Solving compound interest problems using technology

As seen above, compound interest calculations involve five different values: Principal (P), Annual interest rate (r), Number of compounds per year, Future value after n compounds and n , the total number of compounding periods. If any four of these values are known, the fifth can be calculated either using the formula or by using technology such as a spreadsheet.



Spreadsheet activity 7D: A spreadsheet compound interest calculator



Example 13 Solving compound interest problems using a spreadsheet

The balance of Ahmet's investment account is \$15 480.03 after a period of 2 years. His initial investment was \$12 000. If compound interest is calculated and added to the account monthly, find the annual percentage interest rate for Ahmet's investment. Round your answer to one decimal place.

Solution

Principal = \$12 000

Balance = \$15 480.03

Compounds per year = 12 (monthly)

Number of compounds = $n = 2 \times 12 = 24$

	A	B	C	D	E
1	Compound Interest Calculator				
2	Click the clear button to begin. Enter the four compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				
3					Clear
4		number of compounds	24		Calculate
5		annual percentage interest rate			Calculate
6		compounds per year	12		Calculate
7		future value	\$ 15,480.03		Calculate
8		principal	\$ 12,000.00		Calculate

	A	B	C	D	E
1	Compound Interest Calculator				
2	Click the clear button to begin. Enter the four compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				
3					Clear
4		number of compounds	24		Calculate
5		annual percentage interest rate	12.79999281		Calculate
6		compounds per year	12		Calculate
7		future value	\$ 15,480.03		Calculate
8		principal	\$ 12,000.00		Calculate

The annual percentage interest rate for Ahmet's investment is 12.8%.

Explanation

Identify the known quantities, noting that the investment is monthly over 2 years.

Click 'Clear' on the compound interest spreadsheet and enter the known values.

Note: You do not need to type the dollar sign or thousands comma. Cells for dollar values have currency formatting applied which includes them automatically.

Click the 'Calculate' button next to annual interest rate. The annual percentage interest rate will be calculated and entered into the box for you.

Write your answer, rounding to one decimal place.

Section Summary

- ▶ Compound interest calculations involve five different values: Principal, annual interest rate, number of compounds per year, future value after n compounds and the total number of compounding periods. If any four of these are known, the fifth can be found using technology such as a spreadsheet.
- ▶ The principal of the compound interest loan or investment is:

$$P = \frac{A}{(1 + i)^n}$$

The decimal interest rate per compounding period for the loan or investment is:

$$i = \left(\frac{A}{P}\right)^{\frac{1}{n}} - 1$$

Exercise 7D



Using the rule to determine the principal of compound interest loans and investments

Example 11

- 1 An investment earning compound interest at the rate of 6.9% per annum, compounding quarterly, has a future value of \$14 692.82 after 8 years. Use the rule for the principal of compound interest loans and investments to determine the principal value. Round your answer to the nearest cent.
- 2 A loan charging compound interest at the rate of 12.6% per annum, compounding quarterly, has a future value of \$34 821.06 after 3 years. Use the rule for the principal of compound interest loans and investments to determine the principal value. Round your answer to the nearest cent.
- 3 An investment earning compound interest at the rate of 4.2% per annum, compounding monthly, has a future value of \$43 162.90 after 5 years. Use the rule for the principal of compound interest loans and investments to determine the principal value. Round your answer to the nearest cent.
- 4 A loan charging compound interest at the rate of 14.5% per annum, compounding monthly, has a future value of \$7944.62 after 18 months. Use the rule for the principal of compound interest loans and investments to determine the principal value. Round your answer to the nearest cent.
- 5 An investment earning compound interest at the rate of 3.8% per annum, compounding weekly, has a future value of \$33 446.91 after 2 years. Use the rule for the principal of compound interest loans and investments to determine the principal value. Round your answer to the nearest cent.

Using the rule to determine the interest rate of compound interest loans and investments

Example 12

- 6** An investment of \$2000.00 earning compound interest that compounds quarterly has a future value of \$2560.67 after 6 years. Use the rule for the annual percentage interest rate of compound interest loans and investments to determine the interest rate after the given period of time. Round your answer to two decimal places.
- 7** An investment of \$8500.00 earning compound interest that compounds quarterly has a future value of \$10 198.86 after 3 years. Use the rule for the annual percentage interest rate of compound interest loans and investments to determine the interest rate after the given period of time. Round your answer to two decimal places.
- 8** An investment of \$50 000.00 earning compound interest that compounds monthly has a future value of \$63 828.57 after 4 years. Use the rule for the annual percentage interest rate of compound interest loans and investments to determine the interest rate after the given period of time. Round your answer to two decimal places.
- 9** An investment of \$15 000 earning compound interest that compounds monthly has a future value of \$33 059.63 after 15 years. Use the rule for the annual percentage interest rate of compound interest loans and investments to determine the interest rate after the given period of time. Round your answer to two decimal places.
- 10** An investment of \$45 000 earning compound interest that compounds weekly has a future value of \$52 153.57 after 3 years. Use the rule for the annual percentage interest rate of compound interest loans and investments to determine the interest rate after the given period of time. Round your answer to two decimal places.

Problem-solving and modelling with the aid of technology

Example 13

- 11** Tenile invested \$20 000. After 18 months, her investment had grown to a balance of \$21 522.15, with interest compounding monthly. Determine the percentage annual interest rate for Tenile's investment. Round your answer to two decimal places.
- 12** Charlie's investment of \$14 000 has grown to a balance of \$14 863.49 after 12 months, with interest compounding monthly. Determine the percentage annual interest rate for Charlie's investment. Round your answer to two decimal places.
- 13** Sarah invested \$3500 at 6.75% per annum, compounding annually. Determine the number of years that it took for the value of Sarah's investment to first exceed \$5000.
- 14** Determine how long it will take for \$2000 to first exceed \$20 000 if it was invested at a compound interest rate of 4.75% per annum, compounding annually. Give your answer in years to the nearest year.

- 15** If \$45 000 was invested in a compound interest account earning interest at the rate of 6.8% per annum, compounding quarterly, find the number of quarters that it would take for the balance of the investment to exceed \$100 000.
- 16** Suppose that an investment of \$1000 has grown to \$1601.03 after 12 years. If this investment earned compound interest at the rate of $i\%$ per annum compounding yearly, determine the value of i . Round your answer to two decimal places.
- 17** Jannie invested \$25 000 in an account earning compound interest at the rate of $i\%$ per annum, compounding monthly. Jannie's investment had a balance of \$29 216.11 after 30 months. Calculate the value of i . Round your answer to two decimal places.

Paper 1-style multiple-choice questions

- 18** If \$62 000 was invested in a compound interest account earning interest at the rate of 8.4% per annum, compounded quarterly, how many quarters will it take for the balance of the investment to exceed \$100 000?
- A** 12 quarters
B 16 quarters
C 20 quarters
D 24 quarters
- 19** A loan charging compound interest at the rate of 7.2% per annum, compounding monthly, has a future value of \$8000 after 2 years. The principal value of the loan is closest to
- A** \$6930
B \$7000
C \$7200
D \$7450
- 20** An investment earns compound interest at the rate of 5.4% per annum, compounding monthly for three years. It has a future value of \$28 892. The principal value of the investment is closest to
- A** \$18 207
B \$24 675
C \$24 580
D \$4350
- 21** A loan charged compound interest at the rate of 8.4% per annum, compounding quarterly. The future value of the loan was \$18 941.96 after 5 years. Using the rule for the principal of compound interest loans and investments, the principal value of the loan is closest to
- A** \$12 500
B \$12 655
C \$17 072
D \$28 704

7E Using a recurrence relation to model the present value of an ordinary annuity

Learning intentions

- ▶ To model an ordinary annuity, such as a reducing-balance loan, with a recurrence relation.
- ▶ To calculate the total interest incurred on an ordinary annuity.

Reducing-balance loans

In the previous section on compound interest loans, a principal amount of money was borrowed, interest was calculated and charged at regular compounding time periods and then the money was paid back, along with any interest charged, at the **end** of the loan.

In practice, it is very unusual for a borrower to wait until the end of the loan to repay the principal and interest to the bank. Instead, loans are usually repaid by making regular repayments that coincide with the compounding time periods. This has the effect of gradually reducing the **balance** of the loan over time, until it is fully repaid.

This kind of loan is called a **reducing-balance loan**. Home loans and other personal loans are examples of reducing-balance loans.

A recursive model for an ordinary annuity (e.g. reducing-balance loan)

Consider a reducing-balance loan for \$5000 that is charged interest at the rate of 8% per annum, compounding yearly. Repayments of \$1500 will be made each year.

Let A_n be the balance of the loan after n years.

The starting value of the recurrence relation is the principal value of the loan, $A_0 = 5000$.

Each year, the loan balance increases by the amount of interest charged, that is, 8% of the previous balance, and then reduces by the amount of the repayment. That is,

$$\begin{aligned}
 \text{balance next year} &= \text{balance this year} + \text{interest charged} - \text{repayment} \\
 &= \text{balance this year} + 8\% \text{ of the balance this year} - \text{repayment} \\
 &= 100\% \text{ of the balance this year} + 8\% \text{ of the balance this year} - \text{repayment} \\
 &= 108\% \text{ of the balance this year} - \text{repayment}
 \end{aligned}$$

In recurrence relation symbols: $A_{n+1} = 1.08 \times A_n - 1500$.

We now have a recurrence relation that can be used to model the balance of a reducing-balance loan for differing compounding periods.

A recurrence relation model for an ordinary annuity

Let A_n be the balance of an ordinary annuity after n compounding periods.

Let n be the total number of compounding periods.

Let i be the decimal interest rate per compounding period.

Let d be the periodic payment per compounding period.

A recurrence relation model for the balance of an ordinary annuity is:

$$A_0 = \text{principal of loan}, \quad A_{n+1} = r \times A_n - d$$

where $r = 1 + i$.

**Example 14** Modelling an ordinary annuity with a recurrence relation

Alyssa will borrow \$4800 and will be charged compound interest at the rate of 15% per annum, compounding monthly. She will make monthly repayments of \$300 to repay this loan.

- a** Construct a recurrence relation model for Alyssa's loan.
- b** Apply the recurrence relation to determine how much Alyssa will still owe on the loan after two repayments.

Solution

- a** $A_0 = 4800$ (principal of loan)
 $d = 300$ (monthly repayment)

$$\begin{aligned} &15\% \text{ per annum compounding monthly} \\ i &= \frac{15}{12 \times 100} \\ &= 0.0125 \end{aligned}$$

$$\begin{aligned} r &= 1 + i \\ &= 1 + 0.0125 \\ &= 1.0125 \end{aligned}$$

$$A_0 = 4800, A_{n+1} = 1.0125 \times A_n - 300$$

- b** $A_0 = 4800$
 $A_1 = 1.0125 \times 4800 - 300 = 4560$
 $A_2 = 1.0125 \times 4560 - 300 = 4317$

Explanation

Write down the values of A_0 and d .

Calculate the value of the decimal interest rate, i .

Calculate the value of r .

Write your answer.

Apply the recurrence relation two times to find A_2 .

1	4800	0	4800

Pressing '=' or enter once for A_1

Ans $\times$ 1.0125 – 300	4560

Pressing '=' or enter once again for A_2

Ans $\times$ 1.0125 – 300	4317

After two repayments, Alyssa will still owe \$4317.00.

Calculator recursion can also be used to find A_2 . Press **AC** (Casio) or **clear** (TI) to create a blank calculation screen.

Type **4800** and then press = (Casio) or enter (TI). Next, type $\times 1.0125 - 300$ and then press = (Casio), or enter (TI), twice to find A_2

Write your answer.

Calculating the total interest incurred on an ordinary annuity

When a repayment on an ordinary annuity is made, the first priority is to pay the interest that was charged after that compounding period. This interest amount is usually smaller than the periodic payment and so any remaining amount of the repayment will pay back some of the initial value of the annuity. In this way, the principal of the loan, that is, the amount owed after each compounding period, will gradually be reduced in value.

After n compounding periods, the reduction in value can be calculated as:

$$\text{Reduction in value} = A_0 - A_n.$$

The total repayment amount after n compounding periods can be calculated as:

$$\text{Total repayment} = n \times \text{repayment per compounding period}.$$

The total repayment amount pays both the reduction in value and the interest charged and so the total interest charged after n periodic payments can be calculated as:

$$\text{Total interest} = \text{total repayment} - \text{reduction in value}.$$

Total interest charged for ordinary annuities

Let A_0 be the initial value of an ordinary annuity.

Let A_n be the balance of the ordinary annuity after n compounding periods.

Let d be the periodic payment per compounding period.

Let I be the total interest charged after n compounding periods.

The reduction in value after n compounding periods $= A_0 - A_n$.

The total repayments made after n compounding periods $= n \times d$.

$$\begin{aligned} I &= \text{total repayments made} - \text{reduction in principal} \\ &= n \times d - (A_0 - A_n) \end{aligned}$$

Another way of writing this rule is

$$I = A_n + n \times d - A_0.$$

**Example 15** Analysing reducing-balance loans with a recurrence relation

Henry will borrow \$20 000 and will be charged compound interest at the rate of 8.4% per annum, compounding quarterly. He will make quarterly repayments of \$1500 to repay this loan.

- a** Construct a recurrence relation model for Henry's loan.
- b** Find how much interest Henry will pay in the first year of his loan.

Solution

$$\mathbf{a} \quad A_0 = 20\,000$$

$$i = \frac{8.4}{4 \times 100}$$

$$= 0.021$$

$$d = 1500$$

$$r = 1 + 0.021$$

$$= 1.021$$

$$A_0 = 20\,000, A_{n+1} = 1.021 \times A_n - 1500$$

Explanation

Write down the values of A_0 , i , and d .

Calculate the value of r .

Write your answer.

b

20 000	20 000

Pressing '=' or enter once for A_1

$\text{Ans} \times 1.021 - 1500$	18 920

Pressing '=' or enter three more times for A_4

$\text{Ans} \times 1.021 - 1500$	15542.00488

$$n = 4$$

$$d = 1500$$

$$A_0 = 20\,000$$

$$A_n = 15\,542$$

$$\begin{aligned}
 \text{Interest paid} &= n \times d - (A_0 - A_n) \\
 &= 4 \times 1500 \\
 &\quad - (20\,000 - 15\,542) \\
 &= 6000 - 4458 \\
 &= 1542
 \end{aligned}$$

In the first year of his loan, Henry will pay \$1542.00 in interest.

Since one year has four quarters, we apply the recurrence relation four times to find A_4 .

Using calculator recursion: Press **AC** (Casio) or **clear** (TI) to create a blank calculation screen. Type **20 000** and then press = (Casio) or **enter** (TI).

Next, type $\times 1.021 - 1500$ and then press = (Casio), or **enter** (TI), four times to find A_4 . Round your answer to the nearest cent if necessary.

Write down the value of n, d, A_0 and A_n .

Calculate the interest paid after one year (four repayments).

Write your answer.

Section Summary

- ▶ A **reducing balance loan** is a type of **ordinary annuity** that is repaid in regular repayments. The interest is calculated on the amount still owing after each periodic payment. The recurrence relation is:

$$A_0 = \text{initial value}, \quad A_{n+1} = r \times A_n - d, \quad \text{where } r = 1 + i$$

- ▶ The **total interest** incurred on an ordinary annuity after n periodic payments is the total repayments made - (initial value - balance after n repayments).

$$I = n \times d - (A_0 - A_n) \text{ or } I = A_n + n \times d - A_0$$

Exercise 7E

Modelling a recurrence relation model for an ordinary annuity

Example 14

- 1 Fergus borrows \$6500 and is charged compound interest at the rate of 14% per annum, compounding yearly. He will make a yearly repayment of \$1800. Construct a recurrence relation model for Fergus's loan.
- 2 Samay borrows \$14 000 and is charged compound interest at the rate of 11.2% per annum, compounding quarterly. Samay makes a quarterly repayment of \$2000.
 - a Construct a recurrence relation model for Samay's loan.
 - b Apply the recurrence relation model to find how much Samay owes after one, two and three compounding periods.
- 3 Edwina borrows \$22 000 and is charged compound interest of 7.2% per annum, compounding quarterly. Edwina makes a quarterly repayment of \$1000.
 - a Construct a recurrence relation model for Edwina's loan.
 - b Apply the recurrence relation model to determine the balance after three quarters.
- 4 Oli takes out a loan of \$85 000. The annual interest rate on the loan is 8.04% and interest compounds monthly. Oli makes monthly repayments of \$1800.
 - a Construct a recurrence relation model for Oli's loan.
 - b Apply the recurrence relation model to determine the balance after three months.
- 5 Max takes out a loan of \$150 000 to buy a new tractor. He is charged an annual interest rate of 6.48%, compounding monthly. Max makes a repayment of \$1700 each month. By first constructing a recurrence relation, determine how much Max owes after three compounding periods.
- 6 Jasmine borrows \$245 000 and is charged compound interest of 4.16% per annum. Interest compounds fortnightly and Jasmine makes repayments of \$1200 each fortnight. Determine how much Jasmine owes after three compounding periods.

Analysing a reducing-balance loan

Example 15

- 7 A reducing-balance loan can be modelled by the recurrence relation

$$A_0 = 2500, A_{n+1} = 1.08 \times A_n - 600$$
 where A_n is the balance of the loan after n repayments have been made.
 - a State the value of the repayment each compounding period.
 - b Use calculator recursion to determine how much is still owed on the loan after three repayments have been made.
 - c Find how much interest has been paid after three repayments.

- 8** A reducing-balance loan can be modelled by the recurrence relation

$$A_0 = 5000, A_{n+1} = 1.01 \times A_n - 860$$

where A_n is the balance of the loan after n repayments have been made.

- a** State the value of the repayment each compounding period.
- b** Use calculator recursion to determine how much is still owed on the loan after five repayments have been made.
- c** Find how much interest has been paid after five repayments.

- 9** A reducing-balance loan with interest compounding monthly and with monthly repayments can be modelled by the recurrence relation

$$A_0 = 14\,500, A_{n+1} = 1.0072 \times A_n - 1500$$

where A_n is the balance of the loan after n repayments have been made.

- a** Calculate the annual percentage rate of interest.
- b** Use your calculator to determine how much is owed on the loan after three months.
- c** Find how much interest has been paid after three months.

- 10** A reducing-balance loan with interest compounding monthly and with monthly repayments can be modelled by the recurrence relation

$$A_0 = 6300, A_{n+1} = 1.0095 \times A_n - 450$$

where A_n is the balance of the loan after n repayments have been made.

- a** Calculate the annual percentage rate of interest.
- b** Use your calculator to determine how much is owed on the loan after four months.
- c** Find how much interest has been paid after four months.

- 11** Andrea needs to borrow \$20 000. Her bank will charge interest at the annual percentage interest rate of 7.08%, compounding monthly. Andrea will be required to make monthly repayments of \$600, but Andrea thinks she can afford to pay \$800 instead.

- a** Construct a recurrence relation model for the loan with monthly repayments of \$600 and apply it to evaluate:
 - i** how much Andrea would owe after five months
 - ii** the total interest that Andrea would pay after five months.
- b** Construct a recurrence relation model for the loan with monthly repayments of \$800 and apply it to evaluate:
 - i** how much Andrea would owe after five months
 - ii** the total interest that Andrea would pay after five months.
- c** If Andrea makes monthly repayments of \$800 instead of \$600, determine how much interest she will save over the first five months of her loan.

- 12** Gus wants to borrow \$50 000. The bank offers him two options. Option A (\$1500) has an annual percentage interest rate of 6.4%, compounding quarterly. Option B (\$500) has an annual percentage interest rate of 6%, compounding monthly. Gus wishes to make repayments of \$6000 each year, either quarterly if he goes with Option A or monthly if he goes with Option B.
- By first constructing a recurrence relation model for the loan with an appropriate repayment amount, determine the balance of the loan under each option after one year, and the total amount of interest that has been repaid.
- 13** Helen plans to initially borrow \$30 000 at an interest rate of 4.8%, compounding monthly. Helen will make a repayment of either \$1200 or \$1500 each month. Calculate the amount of interest that she will save over six months if she makes the larger monthly repayment.

Paper 1-style multiple-choice questions

- 14** Wilma borrows \$15 000 and will be charged compound interest at the rate of 5.4% per annum, compounding quarterly. Wilma will repay the loan with quarterly repayments of \$2200.
- If W_n is the balance of the loan after n quarters, a recurrence relation model for this reducing balance loan is:
- A** $W_0 = 15\,000$, $W_{n+1} = 1.0054 \times A_n - 2200$
B $W_0 = 15\,000$, $W_{n+1} = 1.0135 \times A_n - 2200$
C $W_0 = 15\,000$, $W_{n+1} = 1.135 \times A_n - 2200$
D $W_0 = 15\,000$, $W_{n+1} = 1.054 \times A_n - 2200$
- 15** Tina has a reducing balance loan with an initial balance of \$28 000. Interest is calculated and compounds monthly and Wilma makes a repayment of \$800 each month. The balance after one month is \$27 424. The annual interest rate is closest to
- A** 0.90% **B** 0.96%
C 9.0% **D** 9.6%
- 16** Elias takes out a loan with an annual interest rate is 7.2% which compounds monthly. He makes a repayment of \$450 each month. The balance after one month is \$3574. The principal of the loan is closest to
- A** \$3145 **B** \$3381
C \$3838 **D** \$4000

7F Investigating reducing-balance loans

Learning intentions

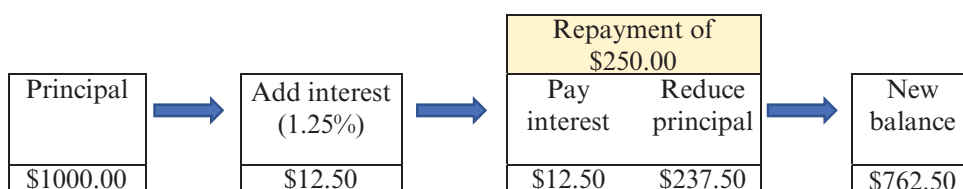
- ▶ To construct and analyse a repayment schedule for an ordinary annuity.
- ▶ To understand the effect of a change in the repayment amount or a lump sum payment on an ordinary annuity.

To understand how periodic payments impact an ordinary annuity, we can construct repayment schedules.

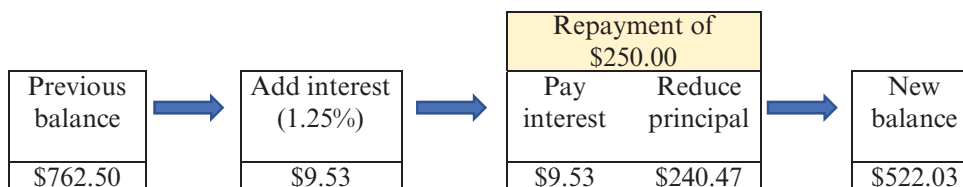
Repayment schedules for reducing-balance loans

Consider a loan with principal \$1000. Interest will be charged at the rate of 1.25% per month and a repayment of \$250 will be made every month.

The calculation of the new balance after the first repayment has been made is shown here.



The calculation of the new balance after the second repayment has been made is shown here.



The results of these calculations can be displayed in a table called a **repayment schedule**.

A repayment schedule for the first three months is shown below.

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	1000.00
1	250.00	12.50	237.50	762.50
2	250.00	9.53	240.47	522.03
3	250.00	6.53	243.47	278.56

Note: Some of the money values in the repayment schedule have been rounded to the nearest cent and may differ slightly to the values calculated using a recurrence relation model. The repayment schedule values are rounded after every calculation (if necessary) while the recurrence relation calculations are not.

Constructing a repayment schedule for a reducing-balance loan

At each step of the loan:

- 1 Interest paid = interest rate per compounding period $\times$ unpaid balance

For example, when repayment 2 is made:

$$\text{Interest paid} = 1.25\% \text{ of } \$762.50 = \$9.53$$

- 2 Principal reduction = repayment – interest

For example, when repayment 2 is made:

$$\text{Principal reduction} = \$250.00 - \$9.53 = \$240.47$$

- 3 Balance of loan = previous balance – principal reduction

For example, when repayment 2 is made:

$$\text{balance} = \$762.50 - \$240.47 = \$522.03$$

- 4 Total interest paid = total repayments made – (principal – balance)



Example 16 Constructing a repayment schedule for a reducing-balance loan

A repayment schedule for the first six repayments of a reducing-balance loan is shown in the table below. Interest is charged at the annual percentage interest rate of 7.68%, compounding monthly, with monthly repayments of \$600.

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	8500.00
1	600.00	54.40	545.60	7954.40
2	600.00	50.91	549.09	7405.31
3	600.00	47.39	A	6852.70
4	600.00	43.86	556.14	6296.56
5	600.00	B	559.70	5736.86
6	600.00	36.72	563.28	5173.58

- a State the principal value of this loan.
- b Calculate the value of A, the principal reduction from repayment number 3.
- c Calculate the value of B, the interest paid with repayment number 5.
- d Calculate the total interest paid after six repayments.

Solution

- a The principal value of the loan is \$8500.

Explanation

The principal of the loan is the balance after repayment number 0.

$$\begin{aligned}\text{b } A &= \$600.00 - \$47.39 \\ &= \$552.61\end{aligned}$$

$$\begin{aligned}\text{c } B &= \frac{7.68}{0.12} \times 6296.56 \\ &= 40.2979 \\ &= 40.30\end{aligned}$$

$$\begin{aligned}\text{d } \text{Total repayments} &= 6 \times 600 \\ &= 3600\end{aligned}$$

$$\begin{aligned}\text{Total interest} &= \text{total repayments} \\ &\quad - (\text{principal} - \text{balance}) \\ &= 3600 - (8500 - 5173.58) \\ &= 273.58\end{aligned}$$

The total interest paid on this loan after six repayments is \$273.58.

Principal reduction = repayment amount – interest charged

The interest paid with repayment number 5 is the interest rate percentage of the balance after repayment 4, rounded to the nearest cent.

Calculate the total of the repayments made.

Calculate the total interest paid.

Note: This answer can be verified by adding all of the values in the interest column of the repayment schedule.

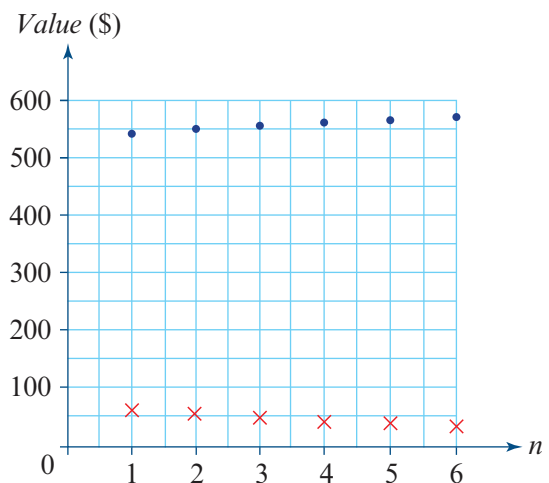
Write your answer.

Graphs of reducing-balance loans

The repayment schedule for a reducing-balance loan shows that after each successive repayment on the loan, the amount of interest that is charged decreases. It also shows that the principal reduction increases after each successive repayment.

The graph on the right shows the interest paid with each repayment as a red cross. This interest value decreases with each repayment.

The graph shows the reduction in the principal of the loan with each repayment as a blue dot. This value increases with each repayment.



The effect of the repayment amount on reducing-balance loans

A reducing-balance loan repayment schedule spreadsheet

The repayment schedule for a reducing-balance loan can be constructed using a spreadsheet. A spreadsheet allows further exploration and investigation of reducing-balance loans.



Spreadsheet activity 7F: Repayment schedule for a reducing-balance loan. Use the spreadsheet to analyse the effect the repayment amount has on a reducing-balance loan.

The effect of the repayment amount on reducing-balance loans

The larger a regular repayment amount for a particular reducing-balance loan, the quicker that loan will be repaid; that is, the shorter the term of the loan will be. The term of a reducing-balance loan can be reduced by increasing the regular repayment amount.

A lump sum repayment is a larger than usual repayment. A lump sum repayment can reduce the amount of interest that is paid in a reducing-balance loan, as well as the term of the loan, if it is large enough. The larger the lump sum repayment, the more interest is saved and the shorter the term of the loan will be.

The earlier that a lump sum repayment is paid, the more interest is saved, as the amount of interest charged towards the end of the loan is much smaller than the amount of interest charged towards the start of a loan; that is, the interest charged decreases after each repayment of a reducing-balance loan.

Home loans are an example of how this feature of reducing-balance loans can be used to great advantage by a borrower. Many people choose to repay more than is required by their home loan agreements because this will mean they will pay less interest overall than if they paid only the required amount. They may also choose to pay larger amounts, perhaps the funds from the sale of another property, into their home loan as lump sum repayments, thereby significantly reducing the balance and the interest charged overall.

The effect of the repayment amount on reducing-balance loans

In general, for reducing-balance loans:

- increasing the repayment amount
 - can mean the loan is repaid in a shorter time
 - will mean less interest is paid overall.
- lump sum repayments
 - can mean the loan is repaid in a shorter time
 - will mean less interest paid overall.
- the earlier a lump sum repayment is made, the less interest is paid overall (or the more interest is saved overall).

Section Summary

- A repayment schedule is a table that summarises the interest calculations at every compounding stage of a reducing-balance loan. It shows the repayment number, repayment amount, interest paid, principal reduction and loan balance after each repayment for some, or all, of the repayments of a loan.

Exercise 7F

Constructing a repayment schedule for a reducing-balance loan

Example 16

- 1** Misaki borrowed \$8400 and will be charged compounding interest at the rate of 11.4% per annum, compounding monthly. Misaki will make monthly repayments of \$500.
 - a** Construct a repayment schedule to determine the balance of Misaki's loan after five repayments.
 - b** Find how much interest in total Misaki paid after five repayments.
- 2** Vadik borrowed \$2000 and will be charged compounding interest at the rate of 18.2% per annum, compounding weekly. Vadik will make weekly repayments of \$100.
 - a** Construct a repayment schedule to determine the balance of Vadik's loan after five repayments.
 - b** Determine the total amount of interest Vadik paid after:
 - i** three weeks
 - ii** four weeks.
- 3** Sofia borrowed \$4600 and will be charged compounding interest at the rate of 4.8% per annum, compounding monthly. The first three repayments that Sofia made were for \$300. The next two repayments Sofia made were for \$450. Construct a repayment schedule to determine the balance of Sofia's loan after five repayments.
- 4** Jabulani borrowed \$7500 and will be charged compounding interest at the rate of 3.8% per annum, compounding quarterly. He made two repayments of \$300 and then doubled this amount for the next three repayments. Construct a repayment schedule to determine the balance of Jabulani's loan after five repayments.
- 5** Hyam borrowed \$6000 and will be charged compounding interest at the rate of 5.2% per annum, compounding quarterly. He made two repayments of \$500, then two repayments of \$700. Construct a repayment schedule to determine the total amount of interest that Hyam paid after four repayments.



Using a spreadsheet to analyse reducing-balance loans

- 6** Consider a reducing-balance loan of \$25 000. This loan is charged interest at the rate of 3.8% per annum, compounding monthly. Monthly repayments of \$2130 will be used to repay this loan, except for the final repayment.
- Enter these loan details into the repayment schedule for a reducing-balance loan spreadsheet.
 - Determine the balance of the loan after five repayments.
 - The balance of the loan is negative for the first time after twelve repayments. Explain what this means.
 - Determine the total amount of interest that has been charged on this loan.
 - The borrower made a lump sum repayment of \$3000 as repayment number four.
 - Determine the value of the final repayment required now.
 - Find how much interest in total has been charged on this loan with the lump sum repayment.
 - Find how much interest has been saved by making this lump sum repayment.
 - Determine the best time to make the lump sum repayment and explain why.

Paper 1-style multiple-choice questions

Consider the following repayment scheduled for Question 7 and 8. Interest is compounding monthly with repayments made each month.

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	12 000.00
1	900.00	72.00	828	11 172.00
2	900.00	67.03	832.97	10 399.03
3	900.00	62.03	A	9501.06
4	900.00	57.01	842.99	8658.07
5	900.00	51.95	848.05	7810.02
6	900.00	46.86	853.14	6956.88

- 7** The annual percentage interest rate is closest to
- A** 0.72% **B** 0.93% **C** 9.3% **D** 7.2%
- 8** The value of A is
- A** \$832.97 **B** \$837.97 **C** \$842.99 **D** \$962.03

7G Using the present value annuity formula to model the present value of an ordinary annuity

Learning intentions

- ▶ To use the present value annuity formula to find the present value of a reducing-balance loan.
- ▶ To use the present value annuity formula to find the present value of a retirement account.

The present value annuity formula

If the total number of repayments made on a reducing-balance loan is small, we can use a recurrence relation in reverse to find the present value, or principal value. But for many reducing-balance loans, such as home loans, the principal is very large and repayments are made over a long period of time, usually decades.

For example, a family may wish to take out a home loan over 25 years making monthly repayments of \$2500 and wish to know how much they can borrow. Analysing this reducing-balance loan using a recurrence relation would take $25 \times 12 = 300$ interest calculations or 300 rows in a repayment schedule table.

We can calculate the present value of an ordinary annuity, such as a reducing-balance loan or a retirement pension with periodic payments, where interest is calculated before the periodic payment is made, using the **present value annuity formula**.

The present value annuities formula

Let A_{PV} be the present value of an annuity.

Let n be the total number of compounding periods.

Let i be the decimal interest rate per compounding period.

Let d be the repayment made after each compounding period.

The present value of the annuity after n compounding periods is

$$A_{PV} = d \frac{(1 - (1 + i)^{-n})}{i}$$



Example 17 Using the present value annuities formula for a reducing-balance loan

Nelly takes out a loan for twelve months at an interest rate of 11.28% per annum, compounding monthly.

She makes monthly repayments of \$500. Determine the present value of the loan.

Round your answer to the nearest cent.

Solution	Explanation
$i = \frac{11.28}{12 \times 100} = 0.0094$ $d = 500$ $n = 12 \text{ (one year of monthly repayments)}$	Write down the values of i , d and n .
$A_{PV} = 500 \times \frac{(1 - (1 + 0.0094)^{-12})}{0.0094}$ $= 5648.931 \dots$	Apply the present value annuities formula to calculate A_{PV} .
The present value of the loan is \$5648.93.	Write your answer, rounding to the nearest cent.

Note that in the example, the amount that Nelly can borrow is less than $12 \times 500 = \$6000$. This is because each future payment is discounted back to its present value using the interest rate. That is, a payment far in the future is worth less today due to the time value of money.

To see this, we could check our answer by taking our principal of \$5648.931 and applying the appropriate recurrence relation to determine that the balance would be zero after 12 payments.



Example 18 Using the present value annuities formula for a retirement pension

Cooper plans to retire and hopes to withdraw \$1000 each month for 10 years from his retirement account. The account earns 6% interest per annum, compounding monthly. Determine the lump sum that Cooper requires in his retirement account today. Round your answer to the nearest cent.

Solution	Explanation
$i = \frac{6}{12 \times 100} = 0.005$ $d = 1000$ $n = 120 \text{ (ten years of monthly repayments)}$	Write down the values of i , d and n .
$A_{PV} = 1000 \times \frac{(1 - (1 + 0.005)^{-120})}{0.005}$ $= 90\,073.45 \dots$	Apply the present value annuities formula to calculate A_{PV} .
Cooper requires \$90 073.45 in his retirement account.	Write your answer, rounding to the nearest cent.

Section Summary

- The **present value annuity formula** allows us to determine the initial balance required for an annuity to provide a periodic payment of d over n compounding periods where i is the decimal interest rate per compounding period. The present value of an ordinary annuity is given by

$$A_{PV} = d \frac{(1 - (1 + i)^{-n})}{i}$$

It is commonly used for reducing-balance loans or retirement pensions.

Exercise 7G

Using the present value annuities formula for a reducing-balance loan

Example 16

- 1 A loan requires annual payments of \$5000 for 10 years. The annual interest rate is 6%, and interest compounds annually. Find the present value of the loan.
- 2 William wants to take out a loan with monthly payments of \$1200 for 15 years. The annual interest rate is 4.5%, and interest compounds monthly. Find the present value of the loan.
- 3 Mia agrees to pay \$2500 monthly for 20 years on a loan with an annual interest rate of 7% where interest compounds monthly. Find the present value of the loan.
- 4 Jack's loan requires quarterly payments of \$3000 for 8 years. The annual interest rate is 5% and interest compounds quarterly. Find the present value of the loan.
- 5 Isabella wishes to take out a 30-year loan for a property. She is able to make quarterly payments of \$1500. The annual interest rate is 3.8%, and interest compounds quarterly. Determine the amount that she can borrow for the loan.

Using the present value annuities formula for a retirement pension

Example 17

- 6 A retiree will receive annual pension payments of \$15 000 for 20 years. The annual interest rate is 5%, compounding annually. Find the present value of the pension.
- 7 Lachlan calculates that he would like to receive \$2000 monthly for 25 years once he retires. The annual interest rate for his retirement fund is 6%, compounding monthly. Determine how much he needs to have in his retirement fund.
- 8 Chloe hopes to receive monthly payments of \$1500 for 15 years once she retires. The annual interest rate for her fund is 7.5%, compounding monthly. Find the present value of her retirement pension.
- 9 Noah hopes to receive quarterly payments of \$5000 for 10 years. His fund pays interest of 4% annually, compounding quarterly. Find the initial value of his pension.

- 10** Charlotte determines that she will need to receive quarterly payments of \$3500 when she retires to cover her living costs. She estimates that she will need to receive this amount for 30 years. Her retirement pension pays 3.2% per annum, compounding quarterly. Determine the balance that Charlotte needs to have in her retirement fund when she retires.

Harder questions about present value annuities

- 11** Two banks offer reducing balance loans with the same term of 20 years.
Loan A: Monthly payment is \$1200, and the annual interest rate is 5.5%, compounding monthly.
Loan B: Monthly payment is \$1100, and the annual interest rate is 6.2%, compounding monthly.
- Calculate the present value of Loan A.
 - Calculate the present value of Loan B.
 - Determine which loan has the lower present value.
- 12** Two retirement pensions offer quarterly payments over 15 years.
Pension A: Quarterly payment is \$4500, and the annual interest rate is 4.8% (compounding quarterly).
Pension B: Quarterly payment is \$4200, and the annual interest rate is 5.5% (compounding quarterly).
- Calculate the present value of Pension A.
 - Calculate the present value of Pension B.
 - Determine which pension has the lower present value.
- 13** Two banks offer 20-year reducing-balance loans, each with monthly payments but with different terms.
Loan A: Payment of \$1500, with annual interest rate of 5.8%, compounding monthly.
Loan B: Payment of \$1400, with annual interest rate of 5.2%, compounding monthly for the first 10 years. After 10 years, the annual interest rate increases to 6.5% for the remaining 10 years.
Determine which loan has the lower present value.

Paper 1-style multiple-choice questions

- 14** Oliver has a loan with monthly payments of \$1000 over 10 years, and the annual interest rate is 6%, compounding monthly. The present value of the loan is
- A** \$83 946.89 **B** \$90 073.45 **C** \$93 654.22 **D** \$95 808.05
- 15** Emily receives a monthly pension of \$3500 for 18 years, and the annual interest rate is 4.2%, compounding monthly. The present value of the pension is
- A** \$529 838.99 **B** \$204 169.49 **C** \$83 321.81 **D** \$529 834

7H Solving practical problems involving the present value annuity formula

Learning intentions

- ▶ To find the periodic payment of an ordinary annuity.
- ▶ To find the total payments and total interest of an ordinary annuity.
- ▶ To use a formula to find the outstanding balance of an annuity.

The present value annuity formula can be rearranged to find the periodic payment of an ordinary annuity. This allows us to not only find the periodic payment, but also the total payment and the total interest of an ordinary annuity.

The periodic payment of an ordinary annuity

Let A_{PV} be the present value of an annuity.

Let n be the total number of compounding periods.

Let i be the decimal interest rate per compounding period.

Let d be the repayment made after each compounding period.

The periodic payment of the annuity after n compounding periods is

$$d = \frac{A_{PV} \times i}{1 - (1 + i)^{-n}}$$



Example 19 Using the present value formula to find the periodic payment in a reducing-balance loan

Riley has a reducing-balance loan of \$550 000 with an interest rate of 5.2%, compounding quarterly. The loan term is 15 years.

Find the periodic payment for this loan. Round your answer to the nearest cent.

Solution

$$A_{PV} = 550\,000$$

$$i = \frac{5.2}{4 \times 100} = 0.013$$

$$n = 4 \times 15 = 60$$

$$\begin{aligned} d &= \frac{550\,000 \times 0.013}{1 - (1 + 0.013)^{-60}} \\ &= 13\,258.32 \end{aligned}$$

Riley will pay \$13 258.32 each quarter.

Explanation

Write down the values of A_{PV} , i and n .

Apply the formula to calculate d .

Write your answer.


Example 20 Finding the total payment and the total interest in a reducing-balance loan

Ella has a reducing-balance loan with a present value of \$340 000. The loan has an annual interest rate of 6%, compounding monthly, and is to be repaid over 20 years.

- a** Find the monthly periodic payment.
- b** Find the total payment Ella makes over the life of the loan.
- c** Determine the total interest that Ella will pay.

Solution

$$\begin{aligned} \mathbf{a} \quad A_{PV} &= 340\,000 \\ i &= \frac{6}{12 \times 100} = 0.005 \\ n &= 12 \times 20 = 240 \end{aligned}$$

$$\begin{aligned} d &= \frac{340\,000 \times 0.005}{1 - (1 + 0.005)^{-240}} \\ &= 2435.87 \end{aligned}$$

Ella will pay \$2435.87 each quarter.

$$\begin{aligned} \mathbf{b} \quad \text{Total Payment} &= 2435.87 \times 240 \\ &= 584\,607.74 \end{aligned}$$

Ella will pay a total of \$584 607.74.

$$\begin{aligned} \mathbf{c} \quad \text{Total interest} &= A_{PV} - \text{total payment} \\ &= 584\,607.74 - 340\,000 \\ &= 244\,607.74 \end{aligned}$$

Ella will pay a total of \$244 607.74 interest over the life of the loan.

Explanation

Write down the values of A_{PV} , i and n .

Apply the formula to calculate d .

Write your answer, rounding to the nearest cent.

Multiply the monthly payment by the number of payments.

Write your answer.

Calculate the difference between the amount Ella pays and the present value of the loan.

Write your answer.

The present value formula for an annuity can be extended to allow us to find the outstanding balance of a present value annuity. That is, given the periodic repayment, the interest rate, compounding rate and number of periods, we can find the balance of the annuity at any point in the life of the annuity.

The outstanding balance of an ordinary annuity

Let OB_k be the outstanding balance of an annuity after k payments.

Let n be the total number of compounding periods.

Let i be the decimal interest rate per compounding period.

Let d be the repayment made after each compounding period.

The periodic payment of the annuity after n compounding periods is

$$OB_k = d \frac{(1 - (1 + i)^{-(n-k)})}{i}$$

**Example 21** Using a formula to find the outstanding balance of an annuity

Thomas has a retirement pension where he receives \$1500 each month from the fund. The interest rate is 6.6% annually, compounding monthly, and the fund is designed to last for 15 years.

Find the balance of the fund after 5 years.

Solution

$$i = \frac{6.6}{12 \times 100} = 0.0055$$

$$d = 1500$$

$$n = 180 \quad (15 \times 12)$$

$$k = 60 \quad (5 \times 12)$$

$$OB_k = 1500 \times \frac{(1 - (1 + 0.0055)^{-(180-60)})}{0.0055}$$

$$= 131\,512.68 \dots$$

Thomas will have \$131 512.68 in his retirement account after 5 years.

Explanation

Write down the values of i , d , n and k .

Apply the formula to calculate OB_k .

Write your answer, rounding to the nearest cent.

Section Summary

- ▶ The **periodic payment** of the annuity with n compounding periods, a present value of A_{PV} and a decimal interest rate per compounding period of i is

$$d = \frac{A_{PV} \times i}{1 - (1 + i)^{-n}}$$

- ▶ The **total payment** is found by calculating $n \times d$ and the **total interest** is the difference between the total payment and the present value of the annuity.
- ▶ The **outstanding balance of an annuity** gives the balance of the annuity after k periods, given d , n and i :

$$OB_k = d \frac{(1 - (1 + i)^{-(n-k)})}{i}$$

Exercise 7H

Using the present value formula to find a periodic payment

Example 19

- 1** John has a reducing-balance loan of \$450 000. The loan has an annual interest rate of 5.5%, compounding annually, and the term of the loan is 10 years. Find the annual payment required to pay off the loan.
- 2** Emma has a reducing-balance loan of \$750 000. The loan has an interest rate of 6% annually, compounding monthly, and the term of the loan is 20 years. Find the monthly payment required to pay off the loan.
- 3** A retirement pension fund has \$780 000 and pays an amount every quarter for 15 years. The interest rate is 4.5% per annum, compounding quarterly. Determine the quarterly payment.
- 4** Carlos has a reducing-balance loan of \$500 000. The loan has an annual interest rate of 7.5%, compounding monthly, and the term of the loan is 15 years. Find the monthly payment required to pay off the loan.
- 5** A retirement pension pays quarterly for 20 years. The balance of the pension is \$800 000, and the annual interest rate is 5%, compounding quarterly. Determine the quarterly payment required to ensure the pension lasts 20 years.
- 6** Liam has a reducing-balance loan of \$1 000 000. The loan has an annual interest rate of 8%, compounding monthly, and the term of the loan is 25 years. Find the monthly payment required to pay off the loan.
- 7** Sally has a retirement pension that pays weekly payments for 10 years. The fund has a balance of \$600 000 and the annual interest rate is 5.2%, compounding weekly. Find the weekly payment that Sally receives.

Finding the total payment and total interest paid

Example 20

- 8** Sienna takes out a reducing balance loan of \$300 000 with an interest rate of 6.5% annually, compounding monthly. The term of the loan is 15 years.
 - a** Find the monthly payment required to pay the loan.
 - b** Find the total payment made over the life of the loan.
 - c** Determine the total interest that Sienna pays on the loan.

- 9** Michael has a retirement pension fund that pays him \$3500 per month. The fund earns interest at a rate of 5% per annum, compounding monthly, and the pension will last for 20 years.
- a** Find the present value of the pension fund.
 - b** Find the total payment made over the life of the pension fund.
 - c** Determine the total interest that Michael earns from the pension fund.
- 10** Lucas has a reducing balance loan of \$450 000. The interest rate is 5% annually, compounding monthly, and the loan term is 10 years.
- a** Find the monthly payment required to pay the loan.
 - b** Find the total payment made over the life of the loan.
 - c** Determine the total interest that Lucas pays on the loan.
- 11** Lily's retirement pension pays her \$2200 per month for 12 years. The annual interest rate is 4.8%, compounding monthly.
- a** Find the present value of the pension fund.
 - b** Find the total payment made over the life of the pension fund.
 - c** Determine the total interest that Lily earns on the pension.

Finding the outstanding balance of an ordinary annuity

Example 21

- 12** Alex has a reducing-balance loan of \$400 000, with an interest rate of 7% per annum, compounding monthly. The loan is for 20 years, and Alex makes monthly payments of \$3101. Determine the outstanding balance after 5 years of payments.
- 13** Eva's retirement pension fund has an initial balance of \$600 000, with an interest rate of 5% per annum, compounding quarterly. The pension pays \$10 544.56 per quarter for 25 years. Determine the outstanding balance after 10 years.
- 14** James has a reducing-balance loan of \$500 000 with an interest rate of 6% per annum, compounding quarterly. The loan term is 15 years, and James makes quarterly payments of \$12 696. Determine the outstanding balance after 7 years.

Harder questions about present value annuities

- 15** Olivia has a reducing-balance loan with an initial balance of \$300 000. The loan has an annual interest rate of 3.6%, compounding monthly. Find the total interest that Olivia pays over the full 20 years.
- 16** Liam has a retirement pension fund with an initial balance of \$710 000. The fund has an annual interest rate of 6.4%, compounding quarterly over 20 years. Find the total interest that Liam earns from the fund over the 20 years.

- 17** Ethan wants to take out a loan to buy a house. The loan will have an annual interest rate of 7.2%, compounding monthly. He considers the two scenarios:
- Option A:** Monthly payment of \$4500, over 20 years
- Option B:** Monthly payment of \$3200, over 30 years.
- Determine which option results in paying less interest.
- 18** Samuel has a retirement pension fund with an initial balance of \$720 000. The fund has an interest rate of 4.2%, compounding monthly, and Samuel expects it to last for 15 years. Find the total interest earned by the fund over 15 years.

Paper 1-style multiple-choice questions

- 19** Samantha is taking out a loan of \$500 000 with an interest rate of 4.8% annually, compounding monthly. The loan is to be paid off over 25 years. Her monthly payment is
- A \$2640.23
B \$2864.98
C \$2775.01
D \$2885.36
- 20** Tom has a retirement pension fund with an initial balance of \$800 000. The interest rate is 5.2%, compounding monthly. The fund is designed to last for 20 years. The total amount he will receive over the life of the pension is closest to
- A \$1 040 000
B \$1 288 400
C \$1 104 000
D \$1 154 200
- 21** Amelia has a retirement pension with monthly payments of \$4000. The interest rate is 6%, compounding monthly, and the fund will last for 25 years. The total interest earned by the fund is
- A \$1 200 000
B \$620 827.45
C \$579 172.54
D \$580 000
- 22** John has a reducing-balance loan where he will make monthly payments of \$1 800 for 15 years. The interest rate is 5.4%, compounding monthly. The present value of this loan is closest to
- A \$315 000
B \$290 000
C \$285 000
D \$220 000

Key ideas and chapter summary

Recurrence relation

A **recurrence relation** is a relation that enables the value of the next term in a sequence to be obtained from one or more current terms. Examples include ‘to find the next term, add two to the current term’ and ‘to find the next term, multiply the current term by three and subtract five’.

Mathematical modelling

Mathematical modelling is the use of mathematical terms and symbols to describe or explain real-life situations. **Recurrence relations** can be used as models for many different real-life situations, including financial situations.

Interest

The fee that is added to a loan or the payment received for investing money is called **interest**.

Principal

The **principal** is the initial amount that has been invested or borrowed.

Simple interest

When a fixed amount of interest is added to a loan or investment at regular time intervals, the interest is called **simple interest**.

Compound interest

When interest is added to a loan or investment and then contributes to earning more interest, the interest is called **compound interest**.

Compounding period

The time period for the calculation of interest is called the compounding period. **Compounding periods** are usually daily, fortnightly, monthly, quarterly, six-monthly or annually.

Decimal rate of interest per compounding period

If x is the annual percentage rate of interest, and if k is the number of compounding periods per year, then the decimal rate of interest per compounding period is

$$i = \frac{x}{k \times 100}$$

Recursive model for compound interest

A recurrence relation can be used to determine the balance of a compound interest loan or investment after n compounding periods. If the number of compounding periods per year is k , and if the annual percentage interest rate is i , then the **recursive model** for the loan or investment is

$$A_0 = \text{principal of loan or investment}, A_{n+1} = r \times A_n$$

where $r = 1 + i$

Nominal interest rate The annual percentage interest rate for a loan or investment is called the **nominal interest rate**.

Effective annual rate of interest The **effective annual rate of interest** $i_{\text{effective}}$ is used to compare the interest paid on loans or investments with different nominal rates of interest and compounding periods.

If the number of compounding periods per year is k , then the annual decimal rate of interest is

$$i_{\text{effective}} = (1 + i)^k - 1$$

The annual percentage rate of interest is

$$i_{\text{effective}} = ((1 + i)^k - 1) \times 100\%$$

Future value The **future value** (A) of a compound interest loan or investment is the balance of that loan or investment after some number of compounding periods.

Rule for the future value of a compound interest loan or investment If P is the principal of the loan or investment and i is the decimal rate of interest per compounding period, then the future value of the loan after n compounding periods is:

$$A = P \times (1 + i)^n$$

Rule for the principal of a compound interest loan or investment If A is the future value of a loan or investment and i is the decimal rate of interest per compounding period, then the principal of the loan or investment after n compounding periods is

$$P = \frac{A}{(1 + i)^n}$$

Rule for the decimal interest rate per compounding period of a compound interest loan or investment If P is the principal of the loan or investment and A is the future value after n compounding periods, then the decimal interest rate per compounding period is

$$i = \left(\frac{A}{P}\right)^{\frac{1}{n}} - 1$$

Reducing-balance loan A **reducing-balance loan** is a type of compound interest loan that is repaid in regular repayments. The interest on a reducing-balance loan is calculated on the amount still owing after each repayment is made.

Recursive model for a reducing-balance loan

A recurrence relation can be used to determine the balance of a reducing-balance loan after n compounding periods.

Given the decimal rate of interest per compounding period, i , and the regular repayment amount, R , then the recursive model for a reducing-balance loan is

$$A_0 = \text{principal of loan}, \quad A_{n+1} = r \times A_n - R$$

where $r = 1 + i$

Repayment schedule for a reducing-balance loan

A table that summarises the interest calculations at every compounding stage is called a **repayment schedule**. It shows the repayment number, repayment amount, interest paid, principal reduction and loan balance after each repayment for some, or all, of the repayments of a loan.

Annuities formula

The **annuities formula** gives the present value of an annuity (such as a reducing-balance loan or a retirement pension fund) given the periodic payment (d), decimal rate of interest per compounding period (i) and the number of compounding periods (n). The annuities formula is

$$A_{PV} = d \frac{(1 - (1 + i)^{-n})}{i}$$

Total interest

The **total interest** paid on a reducing-balance loan after n repayments = total repayments made – principal

Skills checklist

Download this checklist from the Interactive Textbook, then print it and fill it out to check your skills.

**7A**

- 1** I can use a recurrence relation to model a compound interest loan or investment.



See Example 1 and 2 and Exercise 7A Question 1 and 3

7A

- 2** I can convert annual interest rates for different compounding periods.



See Example 3 and 4 and Exercise 7A Question 5 and 6

7A

- 3** I can construct recurrence relation models for compound interest loans and investments.



See Example 5 and 6 and Exercise 7A Question 7 and 9

7B

- 4** I can use the compound interest formula to model a compound interest loan or investment.



See Example 7 and 8 and Exercise 7B Question 1 and 3

- 7C** **5** I can compare loans and investments with effective annual rates of interest. ☐
See Example 9 and Exercise 7C Question 1
- 7C** **6** I can use spreadsheet calculators to compare loans or investments using effective interest rates. ☐
See Example 10 and Exercise 7C Question 2
- 7D** **7** I can use the rule for compounding interest loans and investments to find the principal or interest rate. ☐
See Example 11 and 12 and Exercise 7D Question 1 and 6
- 7D** **8** I can use technology to find an unknown value in compound interest problems. ☐
See Example 13 and Exercise 7D Question 11
- 7E** **9** I can model a reducing-balance loan with a recurrence relation. ☐
See Example 14 and Exercise 7E Question 1
- 7E** **10** I can calculate the total interest charged for a reducing-balance loan. ☐
See Example 15 and Exercise 7E Question 7
- 7F** **11** I can construct and analyse a repayment schedule for an ordinary annuity. ☐
See Example 16 and Exercise 7F Question 1
- 7F** **12** I can understand the effect of a change in the repayment amount of a lump sum payment for an ordinary annuity. ☐
See Spreadsheet Activity and Exercise 7F Question 6
- 7G** **13** I can use the present value annuities formula for a reducing-balance loan. ☐
See Example 17 and Exercise 7G Question 1
- 7G** **14** I can use the present value annuities formula for a retirement pension. ☐
See Example 18 and Exercise 7G Question 6
- 7H** **15** I can use the present value formula to find the periodic payment in an ordinary annuity. ☐
See Example 19 and Exercise 7H Question 1
- 7H** **16** I can find the total payment and the total interest in an ordinary annuity. ☐
See Example 20 and Exercise 7H Question 8
- 7H** **17** I can use a formula to find the outstanding balance of an annuity. ☐
See Example 21 and Exercise 7H Question 12

Multiple-choice questions

- 1 A recurrence relation defines a sequence with starting value 4 and the rule ‘multiply by 3 and subtract 1’. This can be represented as:

A $A_1 = 1, A_{n+1} = 3A_n - 4$	B $A_1 = 4, A_{n+1} = 3A_n - 1$
C $A_1 = 4, A_{n+1} = A_n - 3$	D $A_1 = 1, A_{n+1} = 4A_n - 3$
- 2 A recurrence relation is given by $A_1 = 25, A_{n+1} = 2A_n - 30$. The number of positive terms is

A 1	B 2	C 3	D 4
- 3 An investment that compounds monthly is modelled by a recurrence relation $A_0 = 12\,000, A_{n+1} = 1.0055 \times A_n$. The principal and the annual interest rate are

A \$1000, 0.0055%	B \$1000, 6.6%
C \$12\,000, 1.0055%	D \$12\,000, 6.6%
- 4 The annual percentage rate of interest for a compound interest loan is 12.6% per annum, compounding monthly.
 The balance of this loan after n months, A_n , can be modelled by the recurrence relation:
 $A_0 = 4000, A_{n+1} = 1.0105 \times A_n$.
 The loan and interest is fully repaid after 5 months. The total that will be paid is

A \$4127.33	B \$4170.66	C \$4214.46	D \$4258.71
- 5 A compound interest investment of principal \$12\,000 will earn interest at the rate of 10.8% per annum, compounding every six months. The recurrence relation model for A_n , the balance of the investment, after n six-month periods is given by

A $A_0 = 12\,000, A_{n+1} = 1.054 \times A_n$	B $A_0 = 12\,000, A_{n+1} = 1.108 \times A_n$
C $A_0 = 12\,000, A_{n+1} = 1.15 \times A_n$	D $A_0 = 12\,000, A_{n+1} = 10.8 \times A_n$
- 6 Kaelan invested \$6800 for 9 years at an interest rate of 9.6% per annum, compounding quarterly. The amount of interest earned in the 9th year was closest to

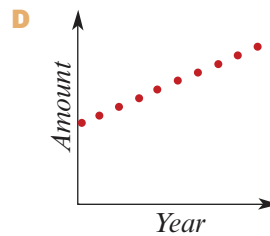
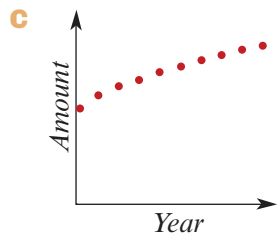
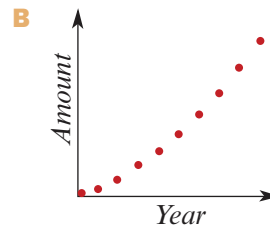
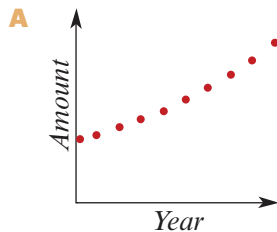
A \$197	B \$1618
C \$1445	D \$1618
- 7 Jackson has a two year loan of \$7200 that compounds monthly and has an interest rate of 4.5% per annum. The effective percentage interest rate is closest to

A 4.59%	B 0.046%
C 1.046%	D 0.375%

- 8 Issa has some money to invest and would like to earn as much interest as possible in the first year of the investment. The interest rate that gives him the best return is

A 3.1% per annum, compounding weekly
B 3.1% per annum, compounding monthly
C 3.2% per annum, compounding quarterly
D 3.2% per annum, compounding monthly

- 9 An investment of \$50 000 is made at a fixed rate of interest compounding annually over a number of years. The graph that best represents the value of the investment at the end of each year is given by



- 10 A compound interest investment earns interest that compounds monthly. The balance of this investment after n months, A_n , can be found using the recurrence relation:

$$A_0 = 15\,000, A_{n+1} = 1.0024 \times A_n.$$

The balance of the investment can also be found using the rule:

A $A_n = 1.0024 \times (15\,000)^n$ **B** $A_n = 15\,000 \times (2.88)^n$
C $A_n = 15\,000 \times (1.24)^n$ **D** $A_n = 15\,000 \times (1.0024)^n$

- 11 A principal of \$2000 is invested and will earn compound interest at the rate of 5.4% per annum, compounding quarterly. The effective annual rate of interest for this investment is closest to:

A 5.3% **B** 5.4% **C** 5.5% **D** 5.6%

- 12 A compound interest investment earns interest at the rate of 3.6% per annum, compounding quarterly. If the balance after 4 years is \$10 964.33, the principal investment amount is closest to:

A \$6200 **B** \$9500 **C** \$10 400 **D** \$10 600

- 13** A compound interest investment of \$10 000 earns \$590.69 interest over a period of 15 months. If interest compounds monthly, the annual percentage rate of interest is closest to:
- A** 1.5% **B** 3.8% **C** 4.6% **D** 7.3%
- 14** A reducing-balance loan is modelled by the recurrence relation shown below.
 $A_0 = 25\,000, A_{n+1} = 1.007 \times A_n - 400$
 where A_n is the balance of the loan after n months.
 The balance of the loan after five months is:
- A** \$23 626.15 **B** \$23 859.14
C \$24 090.51 **D** \$25 707.38
- 15** Hermione borrows \$18 000 and will be charged compound interest at the rate of 6.96% per annum compounding monthly. Hermione will repay the loan with monthly repayments of \$850.
 If A_n is the balance of the loan after n months, a recurrence relation model for this reducing balance loan is:
- A** $A_0 = 18\,000, A_{n+1} = 1.0058 \times A_n - 850$
B $A_0 = 18\,000, A_{n+1} = 1.0174 \times A_n - 850$
C $A_0 = 18\,000, A_{n+1} = 1.0696 \times A_n - 850$
D $A_0 = 18\,000, A_{n+1} = 1.0696 \times A_n - 1242$

Use the following information to answer Questions 16, 17 and 18.

A repayment schedule for the first five repayments of a reducing-balance loan is shown below.

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	15 000.00
1	500.00	97.50	402.50	14 597.50
2	500.00	94.88	405.12	14 192.38
3	500.00	92.25	A	13 784.63
4	500.00	89.60	410.40	13 374.23
5	500.00	86.93	413.07	12 961.16

- 16** The principal of the loan is
- A** \$97.50 **B** \$402.50
C \$500.00 **D** \$15 000

- SE

SE

- SE

- 3** The following recurrence relation can be used to model a compound interest loan. The interest is calculated and added to the loan weekly.

$$A_0 = 1600, A_{n+1} = 1.0035 \times A_n.$$

In this recurrence relation, A_n is the balance of the investment after n weeks.

- a** Determine the principal of this loan.
 - b** Calculate the annual percentage rate of interest.
 - c** Apply the recurrence relation to find the balance of the loan after one, two and three weeks.
 - d** Determine the number of weeks until the balance of the loan will first exceed \$1650.
- 4** Hansie has borrowed \$2200 and will be charged compound interest at the rate of 15.2% per annum, compounding quarterly. Let the balance of Hansie's loan be A_n after n quarters.
- a** Determine the quarterly interest rate for Hansie's loan.
 - b** Construct a recurrence relation that models Hansie's loan.
 - c** Hansie fully repays her loan with one payment after the first year. Determine the amount of money will she need to repay.
- 5** Rodney will borrow \$1500. He will be charged compound interest at the rate of 11.28% per annum compounding monthly.
- a** Use the compound interest rule to find the balance of Rodney's loan after two months.
 - b** Find the total interest that has been charged after two months.
- 6** A reducing-balance loan is modelled using the recurrence relation shown below.
- $$A_0 = 9500, A_{n+1} = 1.0035 \times A_n - 250$$
- In the recurrence relation, A_n is the balance of the loan after n fortnightly repayments.
- a** State the principal of this loan.
 - b** Find the value of the fortnightly repayments.
 - c** Determine the balance of this loan after six repayments.
- 7** Barry is considering borrowing \$250 000 to buy a house. His bank will charge interest at the rate of 5.88% per annum, compounding monthly. Barry can afford to make repayments of \$2400 per month.
- Let A_n be the balance of Barry's loan after n months.
- a** State the monthly percentage rate of interest for Barry's loan.
 - b** Construct a recurrence relation model for Barry's loan.
 - c** Find the balance of Barry's loan after 6 months.
 - d** Determine the number of months it will take for Barry's loan to have a balance below \$240 000 for the first time.

- 8** Eva borrows \$15 000 and will be charged compound interest at the rate of 4.6% per annum.
- a** Calculate the balance of the loan after two years.
 - b** Find the total interest that will be charged after two years.
- 9** A repayment schedule for a reducing-balance loan repaid with monthly repayments is shown below.

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	3500.00
1	600.00	32.20	567.80	2932.20
2	600.00	26.98	573.02	2359.18
3	600.00	21.70	578.30	1780.88
4	600.00	16.38	583.62	1197.26
5	600.00	11.01	588.99	608.27
6	600.00	5.60	594.40	13.87

- a** State the principal of this loan.
 - b** Find the value of the monthly repayment.
 - c** Calculate the:
 - i** monthly interest rate, rounded to two decimal places
 - ii** annual interest rate, rounded to two decimal places.
 - d** If the loan was fully repaid with repayment number six, determine the size of that repayment.
- 10** A reducing-balance loan of \$125 000 is to be repaid with monthly repayments of \$1000.
- The annual percentage rate of interest for this loan is 3.72%.
- a** Construct a repayment schedule that shows the first four repayments of this loan.
 - b** Use the annuities formula to determine the balance of the loan after two years.
- 11** Chelsea has borrowed \$75 000 and will be charged interest at the rate of 7.212% per annum, compounding monthly. Chelsea will repay this loan with monthly repayments over 3 years.
- a** Find Chelsea's monthly repayment.
 - b** Find the total amount that Chelsea pays.
 - c** Determine the total interest that Chelsea pays.

- 12** Ilana uses a credit card to buy a dress that costs \$300. Interest on this loan will be compounded monthly. If Ilana does not make any repayments on her credit card, she will need to repay \$323.98 after five months.
Use the compound interest rule to determine the annual percentage interest rate for Ilana's credit card. Round your answer to one decimal place.
- 13** Anton borrowed \$149 000 to buy an apartment and has been paying \$1000 per fortnight to repay this loan. The balance of Anton's loan is \$84 987.19 after 3 years of payments. Find the annual percentage rate of interest for this loan. Round your answer to two decimal places.
- 14** Meghan has \$5000 to invest.
- Bank A offers investment accounts that pay interest at the rate of 6.3% per annum, compounding quarterly.
 - Bank B offers investment accounts that pay interest at the rate of 6.1% per annum, compounding monthly.
- a** Calculate the effective annual rate of interest for each of these investment accounts. Round your answers to two decimal places.
- b** Determine which investment account Meghan should choose. Justify your answer by explaining how you compared the two investment options.
- c** Calculate the extra interest that Meghan will earn in one year by choosing the investment account in part **b**, compared to the other account. Write your answer to the nearest dollar.
- 15** Darius has \$25 000 to invest. He has two investment options:
- Bank A offers to pay 8.2% per annum, compounding six-monthly.
 - Bank B offers to pay 8.1% per annum, compounding quarterly.
- Darius would like his money to remain invested for a period of 18 months.
- a** Determine which of the two investment options would earn Darius the most interest. Justify your answer by explaining how you compared the two investment options.
- b** Calculate the difference between the total interest earned by both investment options. Round your answer to the nearest dollar.
- 16** Eva invests \$15 000 and will earn compound interest at the annual percentage interest rate of 4.6% per annum.
- a** Calculate the amount of money Eva has invested after one year if the interest earned compounds:
- i quarterly ii monthly iii weekly
- b** Determine the financial principle that is used to compare the investment conditions in part **a** above.
- c** Write a paragraph to explain to Eva why weekly compounding interest on her investment will be of most benefit to her.

- 17** Chi borrowed some money and will be charged interest that compounds weekly. The balance of Chi's loan after n weeks can be found from the rule $A_n = 2300 \times 1.0034^n$. In the following questions, round your answers to the nearest cent.
- a** Determine the principal of Chi's loan.
 - b** Calculate the annual interest rate for Chi's loan.
 - c** Find the total amount of interest that has been charged after 1 year.
- 18** A 'payday loan' company offers short term loans. Interest on these loans is charged at the annual percentage interest rate of 12.48%.
- Lucille needs \$2500 to pay for urgent repairs for her car.
- a** Complete the following table that shows how much Lucille will have to pay back after different periods of time and for quarterly and monthly compounding periods.

Month	Quarterly compounds	Monthly compounds
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		

- b i** Construct a graph that shows the values in the table from part **a** above.
- ii** Quarterly compounding interest is better for Lucille than monthly compounding interest. Describe the features of the table and graph you have drawn in question **b i** above that support this statement.
- c** Lucille will repay the principal and all interest on her loan after one year. Find how much extra interest Lucille will pay if she is charged interest that compounds monthly instead of quarterly.

Chapter 8

Loans, investments and annuities 2

Chapter questions

UNIT 4 INVESTING AND NETWORKING

Topic 2: Loans, investments and annuities 2

- ▶ How do we use a recursive model for an annuity?
- ▶ How do we use the future value annuity formula to model the future value of an ordinary annuity?
- ▶ How do we solve practical problems involving the future value of an ordinary annuity?
- ▶ How do we use the perpetuity formula to solve problems?
- ▶ How can we solve practical problems involving perpetuities, including finding the total amount, the periodic payment and the interest rate?

In this chapter we explore annuities and perpetuities. We model them using recursion and use formulas to solve problems that allow us to find the balance, interest rate and the periodic payment.

8A A recursive model for annuities

Learning intentions

- ▶ To model an annuity during the deposit phase.
- ▶ To model an annuity in the withdrawal phase with a recurrence relation.
- ▶ To analyse annuities with a recurrence relation.

Annuities

An **annuity** is a type of investment that can be used to provide a regular income to the investor. The **principal** of the annuity is the amount of money initially invested. **Interest** is calculated and added to the investment at regular time periods. An investor may also deposit a payment at regular time periods, which adds to the principal and will continue to earn interest over the life, or term, of the annuity. The balance of the annuity will increase during the **deposit phase** of the investment.

After some time, an investor may stop making deposits and instead choose to withdraw a payment at regular time periods. If the payment withdrawn is larger than the interest earned, then the balance of the annuity will decrease during this **withdrawal phase** of the investment.

Superannuation is an example of an annuity investment. During a person's working life, an employer makes regular payments into a fund, causing the account to grow with the deposits and interest. After retirement, the individual begins withdrawing from the account to provide retirement income.

A recursive model for an annuity

The deposit and withdrawal phases of an annuity have different recursive models.

A recursive model for the deposit phase of an annuity

Consider an annuity with a principal of \$100 000 with a deposit of \$15 000 made each year. Interest is paid at the annual percentage interest rate of 6%, compounding annually.

Let A_n be the balance of the annuity after n years.

The starting value of the recurrence relation is the principal value, $A_0 = 100\,000$.

Each year, the annuity balance increases by the amount of interest that is earned, that is, 6% of the previous balance, and then increases by the amount of the deposit.

$$\begin{aligned}
 \text{So : balance next year} &= \text{balance this year} + \text{interest earned} + \text{deposit} \\
 &= \text{balance this year} + 6\% \text{ of the balance this year} + \text{deposit} \\
 &= 100\% \text{ of the balance this year} + 6\% \text{ of the balance this year} + \text{deposit} \\
 &= 106\% \text{ of the balance this year} + \text{deposit} \\
 &= 1.06 \times \text{balance this year} + \text{deposit}
 \end{aligned}$$

In recurrence relation symbols:

$$A_{n+1} = 1.06A_n + 15\,000$$

A recursive model for the withdrawal phase of an annuity

Consider an annuity with a current balance of \$200 000. Interest will be earned at the annual percentage interest rate of 6%, compounding yearly. Each year, \$24 000 will be withdrawn from this annuity.

Let A_n be the balance of the annuity after n years.

The starting value of the recurrence relation is the principal value, $A_0 = 200\,000$.

Each year, the annuity balance increases by the amount of interest that is earned, that is, 6% of the previous balance, and then reduces by the amount withdrawn.

So : balance next year = balance this year + interest earned – withdrawal
 $=$ balance this year + 6% of the balance this year – withdrawal
 $=$ 100% of the balance this year + 6% of the balance this year – withdrawal
 $=$ 106% of the balance this year – withdrawal
 $= 1.06 \times$ balance this year – withdrawal

In recurrence relation symbols:

$$A_{n+1} = 1.06A_n - 24\,000$$

We now have a recurrence relation that can be used to model the balance of an annuity during the deposit and withdrawal phases of that investment.

A recurrence relation model for an annuity

Let A_n be the balance of an annuity after n compounding periods.

Let n be the total number of compounding periods.

Let i be the decimal interest rate per compounding period for the annuity.

Let d be the payment deposited or withdrawn each compounding period.

A recurrence relation model for the balance of an annuity investment is:

Deposit phase: $A_0 =$ balance of annuity investment, $A_{n+1} = rA_n + d$

Withdrawal phase: $A_0 =$ balance of annuity investment, $A_{n+1} = rA_n - d$

where $r = 1 + i$


Example 1 Modelling an annuity during the deposit phase

Julian's employer will deposit \$1500 each month into a superannuation account. Interest is paid at the rate of 4.5% per annum, compounding monthly.

- Construct a recurrence relation model for this annuity.
- Apply the recurrence relation to determine the balance of Julian's superannuation account after 5 months.

Solution

a $A_0 = 0$

$$i = \frac{4.5}{12 \times 100}$$

$$i = 0.00375$$

$$r = 1 + 0.00375$$

$$= 1.00375$$

$$A_0 = 0, A_{n+1} = 1.00375A_n + 1500$$

b

0	0
ans $\times$ 1.00375 + 1500	
1500	

Pressing '=' or enter four more times for A_5

ans $\times$ 1.00375 + 1500	7556.461333

After five months, Julian's superannuation account will have a balance of \$7556.46.

Explanation

Write down the initial balance of the annuity, A_0 .

Calculate the value of the decimal interest rate, i , given the interest rate.

Calculate the value of r using the formula $r = 1 + i$.

Write your answer.

Use calculator recursion to determine the value of A_5 (balance after five months), by first entering in the principal, 0.

Write your answer (rounding to the nearest cent if necessary).


Example 2 Modelling an annuity investment with a recurrence relation

Reza invests \$12 000 in an annuity investment, earning interest at the rate of 6% per annum, compounding monthly. He withdraws a payment of \$1500 per month.

- Construct a recurrence relation model for this annuity investment.
- Apply the recurrence relation to find the balance of the investment after three months.
- Find how many payments of \$1500 can Reza receive from this investment.
- After all payments of \$1500 have been received, Reza can receive one final payment to close the annuity. Determine the value of the final payment.

Solution

a $A_0 = 12\,000$

$$i = \frac{6}{12 \times 100}$$

$$i = 0.005$$

$$r = 1 + 0.005$$

$$= 1.005$$

$$A_0 = 12\,000, A_{n+1} = 1.005A_n - 1500$$

b

12 000	12 000
Ans $\times 1.005 - 1500$	10 560
	...
	7658.364

Pressing '=' or enter three times for A_3

After three months, Reza's annuity will have a balance of \$7658.36.

c Pressing '=' 8 times

Ans $\times 1.005 - 1500$	276.3713495

Reza can withdraw 8 payments of \$1500.

d The balance of the annuity after 8 payments of \$1500 is \$276.37.

$$\text{Final payment} = 276.3713495 \times 1.005$$

$$= \$277.75$$

Reza's final payment will be \$277.75, after which there will be nothing left in his annuity.

Explanation

Write down the principal of the investment.

Calculate the value of the decimal interest rate, i , given the interest rate.

Calculate the value of r using the formula $r = 1 + i$.

Write your answer.

Use calculator recursion, by first typing in the principal, **12 000**, and pressing press = (Casio) or **enter** (TI).

Type **$\times 1.005 - 1500$** then press = (Casio), or **enter** (TI), three times to find A_3 .

Write your answer (rounding to the nearest cent if necessary).

Use calculator recursion to count the number of payments before the balance of the annuity is less than \$1500.

Write your answer.

Use calculator recursion to apply the recurrence relation one last time.

The final balance after 8 full payments is found by multiplying by 1.005, to include the interest earned in the final period.

Write your answer.

Calculating the total interest earned for annuities

Before a payment is withdrawn from an annuity, the annuity earns interest. This amount is usually smaller than the payment amount so that the balance of the annuity is declining after each compounding period. The total interest accrued to an annuity can be calculated in a similar way to how we calculated the interest on an ordinary annuity.

Total interest earned by annuities

Let A_0 be the principal amount of an annuity.

Let A_n be the balance of the annuity after n compounding periods, where n is the total number of compounding periods.

Let d be the payment deposited or withdrawn each compounding period.

Let I be the total interest earned after n compounding periods.

Deposit phase:

The increase in principal after n compounding periods $= A_n - A_0$.

The total amount deposited after n compounding periods $= n \times d$.

I = increase in principal – total amount deposited

$$= (A_n - A_0) - n \times d, \text{ or alternatively,}$$

$$= A_n - n \times d - A_0$$

Withdrawal phase:

The reduction in principal after n compounding periods $= A_0 - A_n$.

The total payments withdrawn after n compounding periods $= n \times d$.

I = total payments withdrawn – reduction in principal

$$= n \times d - (A_0 - A_n), \text{ or alternatively,}$$

$$= A_n + n \times d - A_0$$



Example 3 Analysing annuities with a recurrence relation

Diego has an annuity with a current balance of \$120 000, with the investment earning interest at the rate of 7.68% per annum, compounding monthly.

For the next three months, Diego will deposit \$500 per month into his annuity, after which he will withdraw \$2500 per month for a further five months.

- a** Construct a recurrence relation model for Diego's annuity during the deposit phase.
- b** Use the recurrence relation to find the balance of Diego's account after three months.
- c** Construct a recurrence relation model for Diego's annuity during the withdrawal phase.
- d** Find how much interest Diego's investment earned during these eight months.

Solution

a $A_0 = 120\,000$

7.68% per annum compounding
monthly

$$i = \frac{7.68}{12 \times 100}$$

$$= 0.0064$$

$$r = 1 + i$$

$$= 1 + 0.0064$$

$$= 1.0064$$

$$d = 500$$

$$A_0 = 120\,000, A_{n+1} = 1.0064A_n + 500$$

b

120 000	120 000
ans $\times$ 1.0064 + 500	121 268
ans $\times$ 1.0064 + 500	122 544.1152
ans $\times$ 1.0064 + 500	123 828.3975

Pressing '=' or enter three times for A_3

The balance of Diego's investment is
\$123 828.40 after three months.

c $A_0 = 123\,828.40$

$$i = 0.0064$$

$$r = 1.0064$$

$$d = 2500$$

$$A_0 = 123\,828.40,$$

$$A_{n+1} = 1.0064 \times A_n - 2500$$

Explanation

Write down the values of A_0 .

Calculate the values of i , r and d .

Write your answer.

Use calculator recursion to apply the recurrence relation three times to find A_3 , by typing in the initial value, pressing enter, then typing in the formula and pressing enter three times.

Write your answer, rounding to the nearest cent if necessary.

The value of A_0 for this recurrence relation is the balance of the investment after the deposit phase.

Write down the values of i , r and d .

Write your answer.

$$\begin{aligned}
 \text{d } I &= A_n - n \times d - A_0 \\
 &= 123\,828.40 - 3 \times 500 - 120\,000 \\
 &= \$2328.40
 \end{aligned}$$

123 828.40	123 828.40
ans $\times 1.0064 - 2500$	
	122 120.90176
	...
	115 180.92728

$$\begin{aligned}
 I &= A_n + n \times d - A_0 \\
 &= 115\,180.93 + 5 \times 2500 - 123\,828.40 \\
 &= \$3852.53
 \end{aligned}$$

$$\begin{aligned}
 \text{Total interest} &= \$2328.40 + \$3852.53 \\
 &= \$6180.93
 \end{aligned}$$

Calculate the amount of interest earned during the deposit phase.

Calculate the balance of the annuity after five months of withdrawals, A_5 , by entering in the balance at the start of the phase and then applying the recurrence relation five times.

Calculate the amount of interest earned during the withdrawal phase.

Calculate the total interest earned.

Section Summary

- An **annuity** is a type of investment where interest is earned on the balance and there can either be a deposit made or a withdrawal in each period. A recurrence relation model for the balance of an annuity investment is:

Deposit phase: A_0 = balance of annuity investment, $A_{n+1} = rA_n + d$

Withdrawal phase: A_0 = balance of annuity investment, $A_{n+1} = rA_n - d$

where A_n is the balance of the annuity after n compounding periods, n is the compounding periods, i is the decimal interest rate per compounding period, d is the payment deposited or withdrawn and $r = 1 + i$.

- The total interest earned by an annuity can be calculated by finding the amount earned in the deposit and withdrawal phase separately.

Deposit phase:

$$I = (A_n - A_0) - n \times d \text{ or alternatively, } I = A_n - n \times d - A_0$$

Withdrawal phase:

$$I = n \times d - (A_0 - A_n) \text{ or alternatively, } I = A_n + n \times d - A_0$$

Exercise 8A

Annuities during the deposit phase

Example 1

- 1 The table below shows the principal amount, annual percentage interest rate, compounding period and deposit amount per compounding period for six annuities.

	Principal	Annual percentage interest rate	Compounding period	Deposit per compounding period
a	\$0	2.5%	Yearly	\$5000
b	\$0	6.4%	Quarterly	\$6500
c	\$320 000	3.6%	Quarterly	\$8000
d	\$460 000	6.96%	Monthly	\$4200
e	\$845 000	4.92%	Monthly	\$7500
f	\$1 250 000	6.24%	Weekly	\$2700

For each of these investments:

- construct a recurrence relation model
- apply the recurrence relation model to find how much is left in the annuity after three compounding periods.

Annuities during withdrawal phase

- 2 The table below shows the principal amount, annual percentage interest rate, compounding period and withdrawal amount per compounding period for six annuities.

	Principal	Annual percentage interest rate	Compounding period	Withdrawal per compounding period
a	\$120 500	2.8%	Yearly	\$8000
b	\$276 000	5.04%	Quarterly	\$4600
c	\$358 000	5.72%	Quarterly	\$25 000
d	\$440 000	4.32%	Monthly	\$5000
e	\$845 000	8.04%	Monthly	\$9600
f	\$1 360 000	7.8%	Weekly	\$2900

For each of these investments:

- construct a recurrence relation model
- apply the recurrence relation model to find the balance of the annuity after three compounding periods.

Applying a recurrence relation model to analyse an annuity

Example 2

- 3** An annuity can be modelled by the recurrence relation

$$A_0 = 5000 \quad A_{n+1} = 1.01 \times A_n - 1030$$

where A_n is the balance of the annuity after n payments have been received.

- a** Explain how we can tell this annuity is in the withdrawal phase.
- b** State the payment that is withdrawn from this annuity each compounding period.
- c** Use calculator recursion to apply the recurrence relation and determine the amount left in the annuity after three payments have been received.
- d** Find how much interest has been earned after three payments have been received.

- 4** An annuity can be modelled by the recurrence relation

$$A_0 = 6000 \quad A_{n+1} = 1.005 \times A_n + 300$$

where A_n is the balance of the investment after n payments have been withdrawn.

- a** Explain how we can tell this annuity is in the deposit phase.
- b** State the payment that is deposited into this annuity each compounding period.
- c** Use calculator recursion to apply the recurrence relation and determine the balance of the annuity after five deposits.
- d** Find how much interest has been earned after five deposits have been made.

Example 3

- 5** An annuity can be modelled by the recurrence relations below.

Deposit phase : $A_0 = 40\,000, \quad A_{n+1} = 1.0018 \times A_n + 4500$

Withdrawal phase : $A_0 = P, \quad A_{n+1} = 1.0018 \times A_n - 7500$

where A_n is the balance of the investment after n monthly payments have been withdrawn or deposited.

- a** For the deposit phase, calculate:
 - i** the annual percentage rate of interest for this investment.
 - ii** the balance of the annuity after five months.
- b** After five months, the annuity will enter the withdrawal phase.
 - i** State the monthly withdrawal amount.
 - ii** State the value of P .
 - iii** Find the balance of the annuity after five withdrawals.
- c** Find how much interest has been earned:
 - i** during the deposit phase.
 - ii** during the withdrawal phase for five withdrawals.
 - iii** in total over this period of ten months.

- 6** An annuity can be modelled by the recurrence relations below.

Deposit phase : $A_0 = 265\,000$, $A_{n+1} = 1.0031 \times A_n + 750$

Withdrawal phase : $A_0 = P$, $A_{n+1} = 1.0031 \times A_n - 1800$

where A_n is the balance of the investment after n monthly payments have been withdrawn or deposited.

- a** For the deposit phase, calculate:

- i** the annual percentage rate of interest for this investment.
- ii** the balance of the annuity after three months.

- b** After three months, the annuity will enter the withdrawal phase.

- i** State the monthly withdrawal amount.
- ii** State the value of P .
- iii** Find the balance of the annuity after three withdrawals.

- c** Find how much interest has been earned:

- i** during the deposit phase.
- ii** during the withdrawal phase for three withdrawals.
- iii** in total over this period of six months.

- 7** A superannuation annuity has a current balance of \$125 000 and will earn interest at the annual percentage interest rate of 3.38% per annum, compounding fortnightly. Each fortnight, an employer deposits \$695 into this account. Let A_n be the balance of the annuity after n deposits.

- a** Construct a recurrence relation model for this annuity.

- b** Use calculator recursion to determine the number of deposits that are required to raise the balance of the annuity above \$130 000.

- c** Find how much interest is earned by this annuity after nine fortnights.

- 8** A sum of \$32 000 is invested in an annuity and will earn interest at the annual percentage interest rate of 4.32% per annum, compounding monthly. Monthly payments of \$3500 will be withdrawn.

Let A_n be the balance of the investment after n payments.

- a** Construct a recurrence relation model for this annuity.

- b** Use calculator recursion to determine the number of payments until the balance of the annuity first falls below \$20 000.

- c** Find how many payments of \$3500 can be withdrawn before the annuity is exhausted.

- d** After all payments of \$3500 have been received, state how much money is left in the annuity.

- e** Determine the final payment that can be received to fully exhaust the annuity.

- 9** The amount of \$54 000 is invested in an annuity. This investment will increase with monthly deposits of \$1500 for a period of six months. After these six months, monthly payments of \$1800 will be withdrawn, for six months.
- When deposits are being made, interest is earned at the annual percentage interest rate of 7.68%. When withdrawals are being made, interest is earned at the annual percentage interest rate of 7.56%.
- Let A_n be the balance of the investment after n payments.
- Construct a recurrence relation model for the deposit phase of this annuity.
 - Find the balance of the annuity at the end of the deposit phase.
 - Construct a recurrence relation model for the withdrawal phase of this annuity.
 - Determine how much interest is earned in total after twelve months of investment, (six months deposit and six months withdrawal).

Paper 1-style multiple-choice questions

- 10** An annuity can be modelled by the recurrence relation

$$\text{Deposit phase : } A_0 = 200\,000, \quad A_{n+1} = 1.0035 \times A_n + 600$$

$$\text{Withdrawal phase : } A_0 = P, \quad A_{n+1} = 1.0035 \times A_n - 2000$$

where A_n is the balance of the investment after n monthly payments have been withdrawn or deposited. If the annuity enters the withdrawal phase after three months in the deposit phase, the balance of the annuity after three months in the withdrawal phase is

- A** \$200 000 **B** \$200 041 **C** \$201 300 **D** \$203 914

- 11** An annuity with an initial balance of \$10 000 receives monthly deposits of \$200 for a period of 12 months. After this, \$100 is withdrawn each month. The interest rate during the deposit phase is 7.2% per annum and 6% per annum during the withdrawal phase. The annuity can be modelled by the following recurrence relation, where A_n is the balance of the investment after n payments.

A Deposit phase : $A_0 = 10\,000, \quad A_{n+1} = 1.072 \times A_n + 200$

Withdrawal phase : $A_0 = P, \quad A_{n+1} = 1.06 \times A_n - 100$

B Deposit phase : $A_0 = 10\,000, \quad A_{n+1} = 1.006 \times A_n + 200$

Withdrawal phase : $A_0 = P, \quad A_{n+1} = 1.0072 \times A_n - 100$

C Deposit phase : $A_0 = 10\,000, \quad A_{n+1} = 1.006 \times A_n + 200$

Withdrawal phase : $A_0 = P, \quad A_{n+1} = 1.005 \times A_n - 100$

D Deposit phase : $A_0 = 10\,000, \quad A_{n+1} = 1.06 \times A_n + 200$

Withdrawal phase : $A_0 = P, \quad A_{n+1} = 1.05 \times A_n - 100$

8B Investigating annuities

Learning intentions

- To interpret a payment schedule for an annuity in the withdrawal phase.

Like with ordinary annuities, we can construct an amortisation table so we can investigate the payment schedule and the impact on the balance after each compounding period.¹

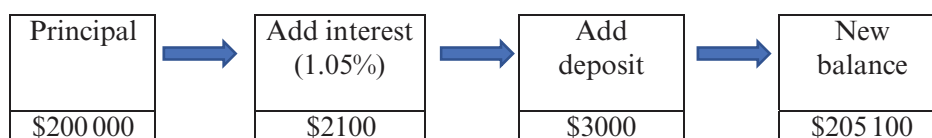
Payment schedules for annuities

As outlined in the previous section, annuities can have a deposit and/or a withdrawal phase. We consider each one separately.

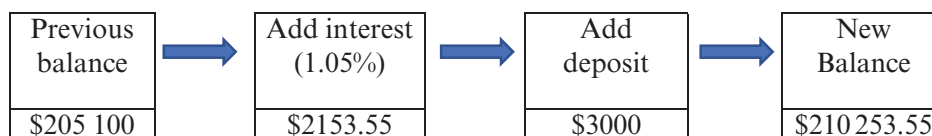
Deposit phase

Consider an annuity with a principal of \$200 000. Interest will be charged at the rate of 1.05% per month and a deposit of \$3000 will be made every month.

Using a recurrence relation or technology, the calculation of the balance of the annuity after the first deposit has been made is shown here.



Similarly, the calculation of the balance of the annuity after the second deposit has been made is shown here.



It is convenient to record all of the results of these calculations in a payment schedule.

A payment schedule for the first three months of this annuity is shown on the following page.

Note: Some of the money values in the payment schedule have been rounded to the nearest cent and may differ slightly to the values calculated using a recurrence relation model. The payment schedule values are rounded after every calculation (if necessary), while the recurrence relation calculations are not.

¹ While this section is not explicitly listed in the study design, a understanding of amortisation tables helps to build a greater understanding of both recurrence models and annuities more generally.

Payment number	Deposit amount	Interest	Principal increase	Balance of annuity
0	0	0	0	200 000.00
1	3000.00	2100.00	5100.00	205 100.00
2	3000.00	2153.55	5153.55	210 253.55
3	3000.00	2207.66	5207.66	21 561.21

Constructing a payment schedule for an annuity (deposit phase)

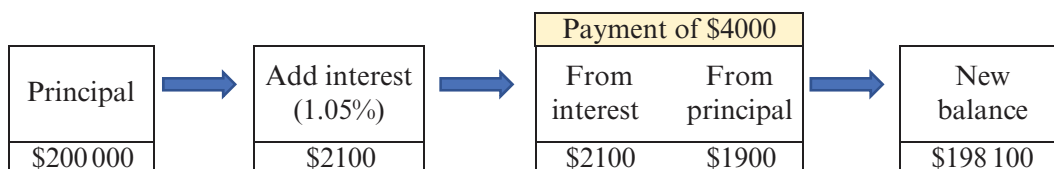
At each step of the investment:

- 1 interest earned = interest rate per compounding period $\times$ current balance
For example: interest earned = 1.05% of \$205 100.00 = \$2153.55
- 2 principal increase = deposit + interest
For example: principal increase = \$3000 + \$2153.55 = \$5153.553
- 3 balance of annuity = previous balance + principal increase
For example: balance = \$205 100.00 + \$5153.55 = \$210 253.55
- 4 total interest earned = (balance – principal) – total deposits made

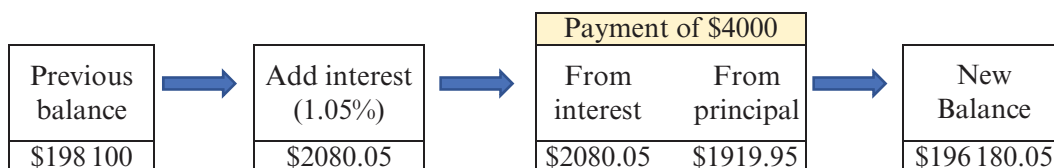
Withdrawal phase

Consider an annuity with a principal of \$200 000. Interest will be charged at the rate of 1.05% per month and a payment of \$4000 will be withdrawn every month.

To calculate the amount still invested in the annuity after the first payment:



To calculate the amount still invested in the annuity after the second payment:



A payment schedule for the first three months of this annuity is shown on the following page. Note that the payment schedule and repayment schedules differ only in the headings and interpretations of the calculation results. The calculations are all performed in exactly the same way.

Payment number	Payment withdrawn	Interest	Principal reduction	Balance of annuity
0	0	0	0	200 000.00
1	4000.00	2100.00	1900.00	198 100.00
2	4000.00	2080.05	1919.95	196 180.05
3	4000.00	2059.89	1940.11	194 239.94

Constructing a payment schedule for an annuity (withdrawal phase)

At each step of the investment:

- 1 interest earned = interest rate per compounding period $\times$ current balance
For example: interest earned = 1.05% of \$198 100 = \$2080.05
- 2 principal reduction = payment – interest
For example: principal reduction = \$4000 – \$2080.05 = \$1919.95
- 3 balance of annuity = previous balance – principal reduction
For example: balance = \$198 100 – \$1919.95 = \$196 180.05
- 4 total interest earned = total payments received – (principal – balance)



Example 4 Interpreting a payment schedule for an annuity in withdrawal phase

The payment schedule for the first six payments from an annuity is shown in the table below. The interest compounds monthly and payments are also withdrawn monthly.

Payment number	Payment withdrawn	Interest	Principal reduction	Balance of annuity
0	0	0	0	65 000.00
1	1500.00	455.00	1045.00	63 955.00
2	1500.00	447.69	A	62 902.69
3	1500.00	440.32	1059.68	61 843.01
4	1500.00	432.90	1067.10	60 775.91
5	1500.00	B	1074.57	59 701.34
6	1500.00	417.91	1082.09	58 619.25

- a State the principal value of this annuity.
- b Calculate the annual percentage rate of interest for this annuity.
- c Calculate the value of A, the principal reduction from payment number 2.
- d Calculate the value of B, the interest earned before payment number 5.
- e Determine the total interest earned after 6 payments have been made.

Solution

a The principal value of the annuity is \$65 000.

b Choose month 1.

$$\begin{aligned}\text{Interest rate} &= \frac{\text{interest amount}}{\text{previous balance}} \times 100\% \\ &= \frac{455.00}{65\,000} \times 100\% \\ &= 0.007(0.7\%)\end{aligned}$$

$$\begin{aligned}\text{Interest rate} &= 0.007 \times 12 \\ &= 0.084 \\ &= 8.4\% \text{ per annum}\end{aligned}$$

The annual percentage rate of interest for this annuity is 8.4%.

c $A = \$1500 - \$447.69 = \$1052.31$

$$\begin{aligned}\text{d } B &= \frac{8.4}{12 \times 100} \times \$60\,775.91 \\ &= \$425.43\end{aligned}$$

$$\begin{aligned}B &= \$1500 - \$1074.57 \\ &= \$425.43\end{aligned}$$

e Total payments received = $6 \times \$1500$
= \$9000

$$\begin{aligned}\text{Total interest} &= \text{total payments} - (\text{principal} - \text{balance}) \\ &= \$9000 - (\$65\,000 - \$58\,619.25) \\ &= \$2619.25\end{aligned}$$

Explanation

The principal of the annuity is the balance after payment number 0.

Choose any month.

Write the interest amount as a percentage of the previous balance.

Note: Because the values in the table have been rounded to the nearest cent, the interest rates calculated using each of the years may be slightly different.

Convert this monthly interest rate to an annual interest rate by multiplying by 12 (12 months per year).

Write your answer.

Principal reduction = payment withdrawn – interest charged

The interest earned before payment number 5 is the interest rate percentage of the balance after payment number 4, rounded to the nearest cent.

Alternatively, Interest = payment – principal reduction

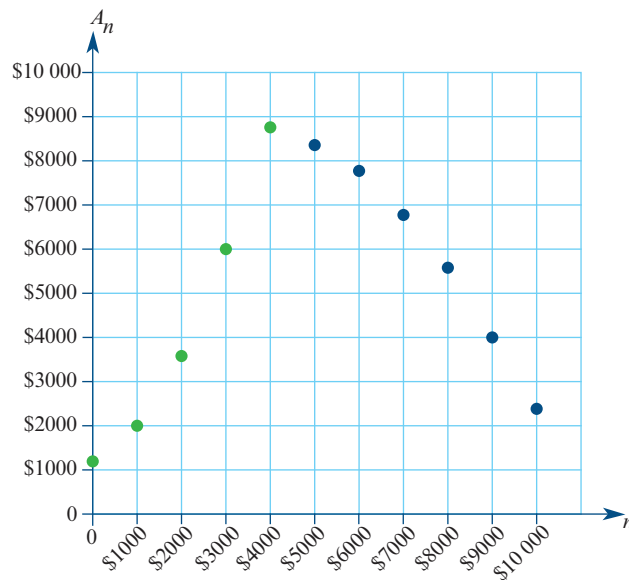
Calculate the total payments received.

Calculate the total interest earned.

Note: this answer can be verified by adding the interest amounts from the interest column of the payment schedule.

The payment schedule for an annuity in deposit phase shows that after each successive deposit, the amount of interest that is earned increases. This is because the balance of the investment is increasing over time, meaning more and more interest is earned. The green dots represent the balance of the investment during the deposit phase. After each additional deposit, the amount of interest earned increases.

The payment schedule for an annuity in withdrawal phase shows that after each successive withdrawal, the amount of interest that is earned decreases. This is because the balance is decreasing over time, meaning less and less interest is earned. The blue dots represent the balance of the investment during the withdrawal phase. After each withdrawal is made, the amount of interest earned decreases.



If the annuity continues in deposit phase, it will continue to grow in value indefinitely. At some point in the investment, however, it is usual for it to enter a withdrawal phase, typically to provide an income for the investor. The balance of the annuity decreases more rapidly later in the life of the annuity than it did in the early stages, as shown in the example graph above.

Section Summary

- ▶ A **payment schedule** for an annuity in either deposit or withdrawal phase shows the calculations for each period. In particular, it shows the amount deposited or withdrawn, along with the interest added in each period, the amount that the principal increases/decreases and the balance of the annuity.
- ▶ Calculations can differ from a recurrence relation because all values are rounded to two decimal places at each step when forming a payment schedule.

Exercise 8B

Interpreting a payment schedule for an annuity in withdrawal phase

Example 4

- 1** A payment schedule for the first five withdrawals from an annuity is shown below. The interest compounds monthly and payments will be withdrawn after each month.

Payment number	Payment withdrawn	Interest	Principal reduction	Balance of loan
0	0	0	0	164 000.00
1	3500.00	787.20	2712.80	161 287.20
2	3500.00	774.18	2725.82	158 561.38
3	3500.00	761.09	2738.91	155 822.47
4	3500.00	747.95	2752.05	153 070.42
5	3500.00	734.74	2765.26	150 305.16

- State the principal of the investment.
- Find the annual percentage rate of interest for the investment. Round your answer to one decimal place.
- State the balance of the investment after three payments have been withdrawn.
- Determine the amount of interest that was received before the fourth payment.
- Determine how much the fifth payment reduced the balance of the loan by.
- Complete the following.
 - Calculate the total interest earned after five payments.
 - Verify your answer to part **i** above by adding values from the interest column of the table.
- Construct the next two rows of this payment schedule.

Constructing a payment schedule for an annuity in deposit phase

- Create a payment schedule for an annuity in the deposit phase, with five deposits when the initial balance is \$50 000 earning interest at the rate of 3.2% per annum, compounding quarterly, with deposits of \$4000 per quarter.
- Create a payment schedule for an annuity in the deposit phase with five deposits when the initial balance is \$135 000 earning interest at the rate of 4.32% per annum, compounding monthly, with deposits of \$1200 per month.

Constructing a payment schedule for an annuity in withdrawal phase

- 4** Create a payment schedule for an annuity in the withdrawal phase, with five withdrawals when the initial balance is \$25 000 earning interest at the rate of 6.9% per annum, compounding monthly, with payments of \$1000 per month.
- 5** Create a payment schedule for an annuity in the withdrawal phase, with five withdrawals when the initial balance is \$380 000 earning interest at the rate of 4.8% per annum, compounding quarterly, with payments of \$12 000 per quarter.

Constructing a payment schedule for an annuity with both phases

- 6** Create a payment schedule for an annuity with an initial balance of \$10 000 and an interest rate of 5.1% per annum, compounding annually. There are two years where \$1000 is added each year and then two years where \$1000 is withdrawn in each year.
- 7** Create a payment schedule for an annuity with an initial balance of \$30 000 and an interest rate of 4.8% per annum, compounding quarterly. There are three quarters where \$1500 is added each quarter and then three quarters where \$1500 is withdrawn in each year.

Paper 1-style multiple-choice questions

The following repayment schedule is to be used for the next two questions.

Payment number	Payment withdrawn	Interest	Principal reduction	Balance of loan
0	0	0	0	100 000.00
1	6000.00	5000	1000	99 000.00
2	6000.00	B	1050	97 950.00
3	6000.00	A	1102.50	96 847.50

- 8** From the repayment schedule above, the value of A is
A \$6000 **B** \$5000 **C** \$4950 **D** \$4897.50
- 9** The total interest earned after three payments is
A \$5000.00 **B** \$15 000.00 **C** \$14 847.50 **D** \$3152.50

8C Using the future value annuity formula

Learning intentions

- ▶ To use the future value annuity formula to find the future value of an annuity.
- ▶ To use the future value annuity formula to find the future value of an annuity in the deposit phase when the balance isn't zero.
- ▶ To use the future value annuity formula to find the future value of an annuity in the withdrawal phase.

If the total number of deposits for an annuity is small, it is convenient to use a recurrence relation to model the annuity. But for many annuities, such as a superannuation account, the principal becomes very large over a worker's lifetime, and the deposits are made over a long period of time, usually decades.

For example, a superannuation account may receive a payment each month over 45 years. Analysing this annuity using a recurrence relation would take $45 \times 12 = 450$ calculations. Similarly, a repayment schedule for this annuity would require 450 rows.

There is a formula that can be used to calculate the future balance of any annuity, after any number of deposits, and it is called the **future value annuity formula**.

The future value annuity formula

Let A_{FV} be the future value of the annuity.

Let n be the total number of compounding periods.

Let i be the decimal interest rate per compounding period.

Let d be the deposit made after each compounding period.

The balance of the annuity after n compounding periods is

$$A_{FV} = d \frac{((1 + i)^n - 1)}{i}$$



Example 5 Using the future value annuity formula

Betsy's employer contributes \$500 into her superannuation fund each month. The fund has an interest rate of 7.2% per annum, compounding monthly.

Determine the balance of the fund after one year.

Round your answer to the nearest cent.

Solution	Explanation
$d = 500$ $n = 12$ $i = \frac{7.2}{12 \times 100} = 0.006$ $A_{FV} = 500 \times \frac{((1 + 0.006)^{12} - 1)}{0.006}$ $= 6202.0139...$ After 1 year (12 months), the balance of the annuity is \$6202.01.	Write down the values of d , n and i . Apply the future value annuity formula to calculate A_{FV} . Write your answer, rounding to the nearest cent.

The general future value annuities formula

When the annuity has a balance greater than zero, we must also consider the principal of the annuity and any interest that it earns, in addition to the interest earned off the regular deposits. Additionally, we can also consider what happens to the future value of an annuity during a withdrawal phase.

We can extend our earlier definition of the future value annuity formula to account for these.

The general future value annuities formula

Let P be the principal amount of the annuity, n be the total number of compounding periods, A be the balance of the annuity after n deposits, i be the decimal interest rate per compounding period and d be the amount of each deposit or withdrawal.

The future value of the annuity after n compounding periods is

Deposit phase:

$$A_{FV} = P(1 + i)^n + d \frac{((1 + i)^n - 1)}{i}$$

Withdrawal phase:

$$A_{FV} = P(1 + i)^n - d \frac{((1 + i)^n - 1)}{i}$$

In the deposit phase, we can see that the future value of an annuity is found through two steps. First, we calculate the value of the principal after it earns interest in each compounding period, $P(1 + i)^n$. Second, we use our original future value annuity formula, $d \frac{((1 + i)^n - 1)}{i}$, to find the amount that is added to the value from the regular deposits that earn compound interest.

The withdrawal phase is similar with two components to be considered. The only difference here is that we **subtract** $d \frac{((1 + i)^n - 1)}{i}$ as we are withdrawing instead of depositing.


Example 6 Using the future value annuities formula when the balance isn't zero

Edward is starting a new job. His salary each week will be \$495 and his employer will pay 9% of this into a superannuation account for him. This superannuation fund will earn interest at the rate of 4.16% per annum, compounding weekly. He already has \$23 500 in his superannuation account.

- a** Find how much money Edward's employer will deposit into the account each week.
- b** Assume that the interest rate for this account does not change. Calculate the balance of Edward's superannuation account after 15 weeks of work.

Solution

$$\begin{aligned}\mathbf{a} \quad d &= 9\% \text{ of } \$495.00 \\ &= \frac{9}{100} \times \$495.00 \\ &= \$44.55\end{aligned}$$

Each week, Edward's employer will deposit \$44.55 into Edward's superannuation account.

$$\begin{aligned}\mathbf{b} \quad &4.16\% \text{ per annum compounding weekly} \\ i &= \frac{4.16}{52 \times 100} \\ &= 0.0008 \\ n &= 15\end{aligned}$$

$$\begin{aligned}A_{FV} &= P(1+i)^n + d \frac{((1+i)^n - 1)}{i} \\ &= 23\,500 \times (1 + 0.0008)^{15} \\ &\quad + 44.55 \times \frac{(1 + 0.0008)^{15} - 1}{0.0008} \\ &= 23\,783.5846 \dots + 672.0052 \dots \\ &= 24\,455.59\end{aligned}$$

After 15 weeks, the balance of Edward's superannuation account is \$24 455.59.

Explanation

Each deposit, d , is 9% of Edward's salary.

Write down your answer.

Write down the value of i and n .

Apply the annuities formula to calculate A_{FV} .

Write your answer, rounding to the nearest cent.


Example 7 Using the future value annuity formula in the withdrawal phase

Wendy invested her superannuation funds of \$375 000 in an annuity that will pay her interest at the rate of 6.12% per annum, compounding monthly. Wendy will withdraw a monthly payment of \$4000 from this investment. Use the annuities formula to calculate the balance of Wendy's investment after 10 years.

Solution

$$P = 375\,000.00$$

$$i = \frac{6.12}{12 \times 100} = 0.0051$$

$$d = 4000.00$$

$$\begin{aligned} n &= 10 \times 12 \text{ (10 years of monthly repayments)} \\ &= 120 \end{aligned}$$

$$\begin{aligned} A &= P(1+i)^n - d \frac{((1+i)^n - 1)}{i} \\ &= 37\,5000 \times (1 + 0.0051)^{120} \\ &\quad - 4000 \frac{((1 + 0.0051)^{120} - 1)}{0.0051} \\ &= 30\,664.83067 \end{aligned}$$

After 10 years, the balance of Wendy's investment is \$30 664.83.

Explanation

Write down the values of P , i , d and n .

Apply the annuities formula to calculate A_{FV} .

Write your answer, rounding to the nearest cent.

Section Summary

- The **future value annuity formula** can be used to find the value of an annuity after a long period of time. It is given by

$$A_{FV} = d \frac{((1+i)^n - 1)}{i}$$

where A_{FV} is the future value of the annuity, n is the total number of compounding periods, i is the decimal interest rate per compounding period and d is the deposit made after each compounding period.

- The formula can be extended for the **deposit phase** when the initial balance (principal, P) is not zero:

$$A_{FV} = P(1+i)^n + d \frac{((1+i)^n - 1)}{i}$$

Or for the **withdrawal phase**:

$$A_{FV} = P(1+i)^n - d \frac{((1+i)^n - 1)}{i}$$

Exercise 8C

Using the future value annuity formula when the principal is zero

Example 5

- 1** Consider an annuity with an initial balance of zero, over a period of five years. The annuity has an annual percentage rate of interest is 5% per annum, compounding each year where \$1000 is deposited each year.
 - a** State the value of d , n and i .
 - b** Apply the future value annuity formula to calculate the value of the annuity after five years. Round your answer to the nearest cent.
- 2** Consider an annuity with an initial balance of zero that earns an annual rate of interest of 4.14%, compounding monthly. Find the future value of the annuity after 8 months if \$800 is added to the annuity each month. Round your answer to the nearest cent.
- 3** Consider an annuity with an initial balance of zero that earns an annual rate of interest of 6.03%, compounding quarterly. Find the future value of the annuity after 8 years if \$2500 is added to the annuity each quarter. Round your answer to the nearest cent.
- 4** Consider an annuity with an initial balance of zero that earns an annual rate of interest of 6.14%, compounding weekly. Find the future value of the annuity after 2 years if \$100 is added to the annuity each week. Round your answer to the nearest cent.

Using the future value annuity formula when the principal is zero

Example 6

- 5** Consider an annuity with an initial balance of \$5000 over a period of five years. The annuity has an annual percentage rate of interest is 5% per annum, compounding each year where \$1000 is deposited each year.
 - a** State the value of P , d , n and i .
 - b** Apply the future value annuity formula to calculate the value of the annuity after five years. Round your answer to the nearest cent.
- 6** Consider an annuity with an initial balance of \$22 000 that earns an annual rate of interest of 3.9%, compounding monthly. Find the future value of the annuity after 6 months if \$1200 is added to the annuity each month. Round your answer to the nearest cent.
- 7** Consider an annuity with an initial balance of \$81 500 that earns an annual rate of interest of 5.42%, compounding quarterly. Find the future value of the annuity after 3 years if \$2650 is added to the annuity each quarter. Round your answer to the nearest cent.

Using the future value annuity formula in the withdrawal phase

Example 7

- 8** Consider an annuity with an initial balance of \$10 000 over a period of five years. The annuity has an annual percentage rate of interest is 5% per annum, compounding each year where \$1000 is withdrawn each year.
- State the value of P , d , n and i .
 - Apply the future value annuity formula to calculate the value of the annuity after five years. Round your answer to the nearest cent.
- 9** Consider an annuity with an initial balance of \$160 000 that earns an annual rate of interest of 4.14%, compounding monthly. Find the future value of the annuity after 8 months if \$2500 is withdrawn from the annuity each month. Round your answer to the nearest cent.
- 10** Consider an annuity with an initial balance of \$415 000 that earns an annual rate of interest of 6.03%, compounding quarterly. Find the future value of the annuity after 8 years if \$8400 is withdrawn from the annuity each quarter. Round your answer to the nearest cent.

Using the future value annuity formula for harder problems

- 11** Consider an annuity with an initial balance of \$100 000 that earns an annual rate of interest of 6.2%, compounding quarterly. Find the future value of the annuity after 10 years if \$2000 is deposited each quarter for 5 years and then \$2000 is withdrawn from the annuity each quarter for 5 years. Round your answer to the nearest cent.

Paper 1-style multiple-choice questions

- 12** An annuity can be modelled by the recurrence relation
 Deposit phase : $A_0 = 150\,000$, $A_{n+1} = 1.04 \times A_n + 500$
 where A_n is the balance of the investment after n yearly payments have been withdrawn or deposited. The future value of the annuity after 24 years is closest to
- A** \$364 954 **B** \$384 496 **C** \$404 036 **D** \$404 037
- 13** An annuity with an initial balance of \$180 000 receives monthly deposits of \$1200 for a period of 4 years. After this, \$1500 is withdrawn each month for 4 years. The interest rate during the deposit and withdrawal phase is 7.2% per annum, compounding monthly. The future value of the annuity after 8 years is closest to
- A** \$180 000 **B** \$325 148 **C** \$306 392 **D** \$474 823

8D Solving practical problems involving the future value annuity formula

Learning intentions

- ▶ To use technology to solve problems involving the future value annuity formula.
- ▶ To find the total interest in an annuity.
- ▶ To find the interest rate in an annuity.

The annuity formula used in the previous section is very easy to use to calculate the balance of an annuity investment given all of the other values for the investment. Unfortunately, it isn't easy to use to calculate other values, such as the number of compounding periods, withdrawal amount or interest rate. Instead, we must use technology to solve them.

Solving annuity problems using technology

Annuity calculations involve six different values:

- Principal, P
- Annual interest rate, i
- Number of compounds per year
- Future balance after n compounds
- n , the number of compounds that will be considered
- Withdrawal or deposit amount, d

If the number of compounds per year is known, along with any four other values, the sixth can be calculated using technology such as a CAS calculator, online calculation tool or a spreadsheet.



Spreadsheet activity 8D: Annuity Calculator

Using the Annuity Calculator spreadsheet

The Annuity Calculator is shown here. It works in the same way as the ordinary annuity calculator from Chapter 7.

Click the 'Clear' button before every calculation.

Enter the five known quantities into the spreadsheet and click the 'Calculate' button next to the quantity that you want to find. It will appear in the box.

	A	B	C	D	E
1	Annuity Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				
3					
4		number of compounds			Calculate
5		annual percentage rate of interest			Calculate
6		compounds per year			
7		payment withdrawn			Calculate
8		principal			Calculate
9		future balance			Calculate



Example 8 Solving annuity problems using a spreadsheet

Nino has \$475 000 to invest in an annuity. His investment will earn interest at the rate of 5.2% per annum, compounding quarterly and Nino plans to withdraw \$12 000 from his investment every quarter.

- Verify that Nino's investment will last for a period of at least ten years.
- Determine the number of payments of \$12 000 that Nino will be able to withdraw.
- Determine the final withdrawal that Nino can make.
- If Nino withdraws \$10 000 each month instead of \$12 000, determine how much longer his investment will last.

Solution

- Principal = \$475 000
Compounds per year = 4
Percentage annual interest rate = 5.2%
Number of compounding periods = 40
Quarterly withdrawal = \$12 000
Calculate the balance after ten years by first clicking 'Clear' on the annuity spreadsheet and enter the known quantities without the dollar sign or the thousands comma.

	A	B	C	D	E
1	Annuity Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				
3					Clear
4		number of compounds	40	Calculate	
5		percentage rate of interest	5.2000000%	Calculate	
6		compounds per year	4	Calculate	
7		payment withdrawn	\$12,000.00	Calculate	
8		principal	\$475,000.00	Calculate	
9		future balance		Calculate	

Click 'Calculate' next to future balance. The value of the balance after ten years will be automatically entered.

Note: this balance amount is a rounded value.

	A	B	C	D	E
1	Annuity Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				
3					Clear
4		number of compounds	40	Calculate	
5		annual percentage rate of interest	5.2000000%	Calculate	
6		compounds per year	4	Calculate	
7		payment withdrawn	\$12,000.00	Calculate	
8		principal	\$475,000.00	Calculate	
9		future balance	\$171,920.49	Calculate	

Explanation

Identify the known quantities, noting that there are 4 compounds in a year when interest compounds quarterly, meaning that there are $10 \times 4 = 40$ compounding periods over 4 years.



The balance after ten years is positive.

The balance of Nino's investment after ten years is positive and so his investment will last for at least ten years.

- b** Click 'Clear' and add all known values to calculate the number of compounding periods (number of withdrawals). Make the future balance equal to zero to calculate the number of possible withdrawals.

	A	B	C	D	E
1	Annuity Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4		number of compounds			Calculate
5		annual percentage rate of interest	5.2000000%		Calculate
6		compounds per year	4		
7		payment withdrawn	\$12,000.00		Calculate
8		principal	\$475,000.00		Calculate
9		future balance	\$0.00		Calculate

Click 'Calculate' next to number of compounds.

	A	B	C	D	E
1	Annuity Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4		number of compounds	55.956569		Calculate
5		annual percentage rate of interest	5.2000000%		Calculate
6		compounds per year	4		
7		payment withdrawn	\$12,000.00		Calculate
8		principal	\$475,000.00		Calculate
9		future balance	\$0.00		Calculate

Nino's investment will allow 55 withdrawals of \$12 000.

If the balance of the annuity is positive, the annuity will last for at least ten years.

Write your answer.

Write your answer.

Note: the number of compounds must be rounded down to 55. The 56th withdrawal will need to be smaller than usual.

- c** Use the spreadsheet to calculate the future balance after 56 withdrawals.

	A	B	C	D	E
1	Annuity Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4		number of compounds	56	Calculate	
5		annual percentage rate of interest	5.2000000%	Calculate	
6		compounds per year	4		
7		payment withdrawn	\$12,000.00	Calculate	
8		principal	\$475,000.00	Calculate	
9		future balance	-\$517.96	Calculate	

$$\begin{aligned}\text{Final withdrawal} &= \$12\,000 - \$517.96 \\ &= \$11\,482.04\end{aligned}$$

- d** Use the spreadsheet to calculate the number of \$10 000 withdrawals that are possible.

	A	B	C	D	E
1	Annuity Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4		number of compounds	74.40460397	Calculate	
5		annual percentage rate of interest	5.2000000%	Calculate	
6		compounds per year	4		
7		payment withdrawn	\$10,000.00	Calculate	
8		principal	\$475,000.00	Calculate	
9		future balance	\$0.00	Calculate	

$$\begin{aligned}\$12\,000 \text{ withdrawals will last for } & 55 \text{ quarters} \\ & +1 \text{ smaller withdrawal. } \$10\,000 \text{ withdrawals} \\ & \text{will last for } 74 \text{ quarters } +1 \text{ smaller withdrawal.} \\ \text{Extra withdrawals} &= 74 - 55 \\ &= 19\end{aligned}$$

Nino's investment will last an extra 19 quarters if he withdraws \$10 000 per quarter instead of \$12 000.

Calculate the 56th and final withdrawal amount.

Compare the two options.
Note: The same result is obtained by subtracting the total number of repayments. The number of extra withdrawals
 $= 75 - 56 = 19$.

Write your answer.


Example 9 Finding the total interest in an annuity

Following on from the previous example, Nino has \$475 000 to invest in an annuity. His investment will earn interest at the rate of 5.2% per annum, compounding quarterly and Nino plans to withdraw \$12 000 from his investment every quarter. Determine the total amount of interest earned over the 55 periods plus the final payment period.

Solution

$$\begin{aligned}\text{Total withdrawn} &= 55 \times 12\,000 + 11\,482.04 \\ &= 671\,482.04\end{aligned}$$

$$\begin{aligned}\text{Total interest} &= 671\,482.04 - 475\,000 \\ &= 196\,482.04\end{aligned}$$

The total amount of interest earned over the 56 payments is \$196 482.04.

Explanation

Nino made 55 withdrawals of \$12 000 and one withdrawal of \$11 482.04.

Recall that the original balance of the annuity was \$475 000.

Write your answer.


Example 10 Finding the interest rate in an annuity

Mingjia has \$20 000 in an annuity and withdraws \$1200 per year. She wants her annuity to last 20 years. Find the interest rate required for this to occur if interest compounds annually. Round your answer to three decimal places.

Solution

Enter the known quantities into the spreadsheet. Click 'Calculate' next to percentage rate of interest.

	A	B	C	D	E
1	Annuity Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4		number of compounds	20		Calculate
5		annual percentage rate of interest	1.8029584%		Calculate
6		compounds per year	1		
7		payment withdrawn	\$1,200.00		Calculate
8		principal	\$20,000.00		Calculate
9		future balance	\$0.00		Calculate

The interest rate is 1.803%.

Explanation

Identify and enter the known quantities.

Write your answer.

Section Summary

- Annuity calculations involve six different variables: the principal, annual interest rate, number of compounds per year, future balance, the number of compounds and the amount being withdrawn. Technology can solve annuity problems provided the number of compounds per year is known along with any four other values.

Exercise 8D**Example 8**

- 1** Andrea has \$200 000 invested in an annuity that earns interest of 4.15% per annum, compounding annually. Andrea receives a withdrawal of \$20 000 each year. Verify that Andrea's investment will last for a period of at least 13 years.
- 2** Taylor has \$320 000 to invest in an annuity. Her investment will earn interest at the rate of 4.8% per annum, compounding quarterly and Taylor plans on withdrawing \$10 000 from her investment each quarter.
 - a** Verify that Taylor's investment will last for a period of at least ten years.
 - b** Determine the number of payments of \$10 000 that Taylor will be able to withdraw.
 - c** Determine the final withdrawal that Taylor can make.
- 3** Raj invests \$85 500 in an annuity that will pay interest at the rate of 7.25% per annum, compounding quarterly.
If Raj receives a regular quarterly payment of \$5000, find how many payments, in total, he will receive, including the final payment.

Example 9

- 4** Leigh invests \$64 000 in an annuity and will be charged interest at the rate of 6.25% per annum, compounding monthly. Leigh will withdraw \$1275 per month.
 - a** Find how many withdrawals of \$1275 Leigh will be able to make.
 - b** His final withdrawal will be smaller than \$1275. Find the value of this final payment.
 - c** What is the total interest that Leigh has earned from this investment?
- 5** Stephanie invests \$40 000 in an annuity and would like to receive a monthly payment for exactly 10 years. Interest on her investment is earned at the rate of 7.5% per annum, compounding monthly.
 - a** Find how many payments Stephanie will receive.
 - b** The last of these payments will be smaller than all of the others.
 - i** Determine the value of the usual payment. Round your answer to the nearest cent.
 - ii** Determine the value of the final payment. Round your answer to the nearest cent.
 - c** Find the total interest that Stephanie will earn from her investment.
- 6** Kaspar has an annuity investment that earns interest at the rate of 8.16% per annum, compounding monthly, from which he receives a monthly payment of \$3600. The balance of Kaspar's investment was \$391 262.50 after the first two years.
 - a** Find the principal of Kaspar's investment.
 - b** Determine how much interest Kaspar earned after two years.

Example 10

- 7** Simon has an annuity investment with a principal of \$480 000, from which he receives a payment of \$1250 per week. The balance of Simon's investment is \$439 252.37 after the first year of payments.
- a** Find the annual interest rate for Simon's investment.
 - b** Calculate the balance of Simon's investment after two years.
 - c** Determine how many payments, in total, Simon can expect before his investment is fully exhausted.
 - d** If the interest rate on Simon's investment decreases by 0.2%, and if Simon continues to withdraw \$1250 per week, find how many fewer payments he will receive compared to the number of payments calculated in part **c** above.
- 8** Olek invests \$100 000 into an annuity with interest compounding monthly.
- a** Find the interest rate that would allow Olek to withdraw \$2500 each month for 4 years. Round your answer to two decimal places.
 - b** Assume the interest rate is 6% and Olek receives a regular payment of \$2000 followed by a smaller final payment. Find how many months the annuity will last.
- 9** Sophia invests \$300 000 into an annuity, paying 4.3% compound interest per annum, compounding quarterly. She wishes to receive a regular quarterly payment of \$5000.
- a** Determine how many quarters, including the final payment, that Sophia's annuity will last.
 - b** Find the final payment of the annuity. Round your answer to the nearest cent.
- 10** Kai invests \$500 000 in an annuity. The annuity earns interest at the rate of 4.7% per annum, compounding monthly and Kai receives a payment each month. The balance of Kai's annuity at the end of the first year of investment is \$474 965.28. Find how much interest Kai's annuity earned in the first year. Round your answer to the nearest cent.

Paper 1-style multiple-choice questions

- 11** Monthly withdrawals of \$220 are made from an account that has an opening balance of \$35 300, invested at 7% per annum, compounding monthly. The balance of the account after 1 year is closest to:
- A** \$32 660 **B** \$33 500 **C** \$35 125 **D** \$35 211
- 12** Audrey invests \$85 000 in an annuity, paying 6.3% per annum, compounding monthly. Audrey receives a regular monthly payment from the annuity. If the value of the annuity after one year is \$71 983.41, the amount of interest earned in the first year is closest to
- A** \$1500 **B** \$4983 **C** \$13 017 **D** \$31 017

8E Perpetuities

Learning intentions

- ▶ To use the perpetuity formula to calculate the payment withdrawn from a perpetuity.
- ▶ To calculate the payment withdrawn from a perpetuity using a recurrence relation.

Introduction to perpetuities

An annuity where the amount withdrawn has a value equal to the interest earned after one compounding period is called a **perpetuity**. The balance of an annuity will remain constant forever, or *in perpetuity*, if the payment withdrawn is the same as the interest earned. This means that, while the investment remains in place, the regular payment can be withdrawn for as long as required.

Perpetuities

Let A be the future value of the perpetuity after n payments have been withdrawn.

Let n be the total number of compounds.

Let i be the decimal rate of interest per compounding period.

Let d be the payment withdrawn after each compounding period.

The balance of a perpetuity investment is constant and equal to the principal, $A = P$

$$d = i \times A \qquad A = \frac{d}{i} \qquad i = \frac{d}{A}$$



Example 11 Calculating the payment withdrawn from a perpetuity

A university has invested \$80 000 into a perpetuity, the interest from which will provide an annual prize for one of their students. The investment earns interest at the rate of 8.4% per annum, compounding yearly.

Find the value of the student prize.

Solution

$$i = \frac{8.4}{100} = 0.084$$

$$A = 80\,000$$

$$\begin{aligned} d &= 0.084 \times 80\,000 \\ &= \$6720 \end{aligned}$$

The annual student prize has value \$6720.

Explanation

State the value of i and A .

Use the formula $d = i \times A$.

Write your answer.

A formula can be used to find d , A or i . An alternative is to use the recurrence relation.



Example 12 Calculating the payment withdrawn from a perpetuity using a recurrence relation

The balance of an annuity investment after n months, A_n , is modelled by the recurrence relation below.

$$A_0 = 500\,000, A_{n+1} = 1.005 \times A_n - d$$

Determine the payment that should be withdrawn from this investment every month if the payments are to be withdrawn in perpetuity.

Solution

$$A_{n+1} = 1.005 \times A_n - d$$

$$A = 1.005 \times A - d$$

$$500\,000 = 1.005 \times 500\,000 - d$$

$$d = 1.005 \times 500\,000 - 500\,000$$

$$d = 2500$$

The payment withdrawn from this perpetuity investment should be \$2500. Alternatively, the formula can be used, noting that $r = 1.005 = 1 + i$, so $i = 0.005$. Thus, $d = i \times A$

$$\begin{aligned} &= 0.005 \times 500\,000 \\ &= 2500 \end{aligned}$$

Explanation

The balance of the perpetuity is always equal to A , so write the recurrence relation with A_n and A_{n+1} both equal to A .

Use the value of A (A_0) from the recurrence relation.

Solve for d

Write your answer.

Note that perpetuities are just special cases of annuities in withdrawal phase. The payment that is withdrawn just happens to be the same as the interest that is earned after each compounding period.

Section Summary

- ▶ A **perpetuity** is a special case of an annuity in the withdrawal phase where the amount withdrawn each period is equal to the interest earned. This keeps the balance constant. The formula is given by $A = \frac{d}{i}$ where A is the balance, d is the payment that is withdrawn each period and i is the decimal interest per compounding period.
- ▶ A recurrence relation can also be used to model and solve a perpetuity.

Exercise 8E

Calculating the payment withdrawn from a perpetuity using a formula

Example 11

- 1 Dom has a trust fund with a balance of \$75 000 in perpetuity. The trust fund earns interest at the rate of 4.1% per annum, compounding yearly. Find the value of the yearly payment that Dom receives.
- 2 Brian retires and notes that he has \$520 000 in his fund. He sets up a perpetuity for the full amount that earns interest at the rate of 4.9% per annum, compounding yearly. Find the amount that Brian receives each year.
- 3 Craig has won \$1 000 000 in a lottery and has decided to invest this money in a perpetuity that pays interest at the annual percentage rate of interest of 5.75%, compounding monthly. Find the monthly payment that Craig can withdraw from his investment.
- 4 Mia has \$670 000 in her retirement fund that pays out a monthly payment in perpetuity. The fund earns interest at the annual percentage rate of 6%, compounding monthly. Find the monthly amount that Mia receives from her fund.
- 5 Suzie has invested her inheritance of \$642 000 in a perpetuity that pays interest at the rate of 6.1% per annum, compounding quarterly.
 - a Find the quarterly payment that Suzie receives.
 - b State the balance of the perpetuity after Suzie has received five payments.

Calculating the payment withdrawn from a perpetuity using a recurrence relation

Example 12

- 6 The balance of an annuity investment after n years, A_n , is modelled by the recurrence relation below.

$$A_0 = 200\,000, A_{n+1} = 1.035 A_n - d$$
 Determine the payment that should be withdrawn from this investment each year if the payments are to be withdrawn in perpetuity.
- 7 The balance of an annuity investment after n years, A_n , is modelled by the recurrence relation below.

$$A_0 = 240\,000, A_{n+1} = 1.042 A_n - d$$
 Determine the payment that should be withdrawn from this investment each year if the payments are to be withdrawn in perpetuity.

- 8** The balance of Letty's annuity investment after n months, A_n , is modelled by the recurrence relation below.

$$A_0 = 180\,000, A_{n+1} = 1.0055A_n - d$$

Determine the total amount that Letty receives over one year if the annuity investment is a perpetuity.

- 9** The balance of Han's annuity investment after n months, A_n , is modelled by the recurrence relation below.

$$A_0 = 980\,000, A_{n+1} = 1.004A_n - d$$

He withdraws $\$d$ each month, giving him a total of $\$47\,040$ for the year. Determine if Han's annuity investment is a perpetuity, providing an explanation for your answer.

Paper 1-style multiple-choice questions

- 10** Aaliyah invests $\$120\,000$ in a perpetuity from which she'll receive a regular monthly payment. The perpetuity has a compound interest rate of 5.4% per annum and compounds monthly. The amount that Aaliyah will receive from the perpetuity in the first two years is closest to

A $\$520$ **B** $\$540$ **C** $\$6280$ **D** $\$12\,960$

- 11** Determine which of the following recurrence relations could be used to model the value of a perpetuity investment, A_n , after n months.

A $A_0 = 100\,000, \quad A_{n+1} = 1.005A_n + 500$

B $A_0 = 100\,000, \quad A_{n+1} = 1.005A_n - 500$

C $A_0 = 200\,000, \quad A_{n+1} = 1.003A_n + 6000$

D $A_0 = 200\,000, \quad A_{n+1} = 1.003A_n - 6000$

- 12** Roman invests $\$535\,400$ in a perpetuity from which he will receive a regular monthly payment. The perpetuity has a compound interest rate of 4.2% per annum and compounds monthly. The recurrence relation that models Roman's perpetuity is

A $A_0 = 535\,400, \quad A_{n+1} = 1.0035A_n - 1873.90$

B $A_0 = 535\,400, \quad A_{n+1} = 1.0035A_n - 22\,468.80$

C $A_0 = 535\,400, \quad A_{n+1} = 1.0042A_n + 1873.90$

D $A_0 = 535\,400, \quad A_{n+1} = 1.0042A_n - 22\,468.80$

8F Solving practical problems involving perpetuities

Learning intentions

- ▶ To calculate the total amount required to establish a perpetuity.
- ▶ To calculate the interest rate required for a perpetuity.
- ▶ To use technology to solve a problem involving a perpetuity investment.

The formula for finding the regular payment from a perpetuity, $d = i \times A$, can be rearranged to allow us to find the total amount required to establish a perpetuity or the interest rate per compounding period.



Example 13 Calculating the investment required to establish a perpetuity

Find how much money will need to be invested in a perpetuity account, earning interest at the rate of 4.8% per annum compounding monthly, if \$225 will be withdrawn every month.

Solution

$$i = \frac{4.8}{12 \times 100} = 0.004$$

$$d = 225$$

$$A = \frac{d}{i}$$

$$= \frac{225}{0.004}$$

$$= 56\,250$$

The principal invested in this perpetuity should be \$56 250.

Explanation

Write down the values of i and d .

Use the rule to calculate the principal.

Write your answer as a sentence.


Example 14 Calculating the interest rate required for a perpetuity

A university mathematics faculty has \$30 000 to invest. It intends to award an annual mathematics prize of \$1500 to a student using the interest from this investment. If the award is to be made in perpetuity, determine the annual interest rate required for this investment.

Solution

$$A = 30\,000$$

$$d = 1500$$

$$\begin{aligned} i &= \frac{d}{A} \\ &= \frac{1500}{30\,000} \\ &= 0.05 \end{aligned}$$

Annual percentage interest rate
 $= 0.05 \times 1 \times 100\%$ (one compound per year) $= 5\%$

The annual interest rate required for this perpetuity investment is 5% per annum.

Explanation

Write down the values of A and d .

Use the rule to calculate the decimal rate of interest per compounding period.

Convert the value of i to an annual percentage rate.

Write your answer.

While the formulas can be used to find d , A or i , technology can also be used to solve perpetuity problems. The spreadsheet ‘Annuity Calculator’ used in the earlier sections can be used to solve problems involving perpetuities.

When using the spreadsheet calculator, note that the number of compounds is always entered as ‘1’, as the balance of the perpetuity will be the same no matter how many compounding periods are considered. Additionally, the future balance is the same as the principal for a perpetuity, as the balance remains unchanged each period. The compounds per year is 1 because the payment is withdrawn after one compounding period.


Spreadsheet activity 8F: Annuity Calculator



Example 15 Using technology to solve a problem involving a perpetuity investment

A university mathematics faculty has \$30 000 to invest. It intends to award an annual mathematics prize of \$1500 to a student using the interest from this investment. If the award is to be made in perpetuity, use technology to find the annual interest rate required for this investment.

Solution

$$P = 30\,000$$

$$d = 1500$$

Enter values into the spreadsheet.

	A	B	C	D	E
1	Annuity Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4		number of compounds	1	Calculate	
5		annual percentage rate of interest		Calculate	
6		compounds per year	1		
7		payment withdrawn	\$1,500.00	Calculate	
8		principal	\$30,000.00	Calculate	
9		future balance	\$30,000.00	Calculate	

Click 'Calculate' to find the annual interest rate required.

	A	B	C	D	E
1	Annuity Calculator				
2	Click the clear button to begin. Enter the five compound interest values that are known into the boxes and then click the 'Calculate' button next to the required value.				Clear
3					
4		number of compounds	1	Calculate	
5		annual percentage rate of interest	5.0000000%	Calculate	
6		compounds per year	1		
7		payment withdrawn	\$1,500.00	Calculate	
8		principal	\$30,000.00	Calculate	
9		future balance	\$30,000.00	Calculate	

An interest rate of 5% per annum is required for this perpetuity investment.

Explanation

Write down the values of P and d .

Write your answer.

Section Summary

- ▶ The formula for perpetuities can be rearranged to solve for d , A or i , where A is the balance, d is the payment that is withdrawn each period and i is the decimal interest per compounding period.

$$d = i \times A \qquad A = \frac{d}{i} \qquad i = \frac{d}{A}$$



Exercise 8F

Calculating the investment required to establish a perpetuity

Example 13

- 1** Geoff would like to establish a perpetuity, the interest from which will be donated to the RSPCA. He would like the annual payment to be \$2500. The perpetuity account will pay interest at the rate of 2.5% compounding annually.
 - a** Use a rule to verify that Geoff will need \$100 000 for this investment.
 - b** Geoff only has \$80 000 to invest. Use a rule to determine the interest rate that Geoff would need to provide the annual payment to the RSPCA.
 - c** Verify your answer to part **b** above using technology.
- 2** Donna receives a lump sum and decides to invest it into a perpetuity, which earns interest at the rate of 3.6% per annum, compounding monthly. If Donna receives a payment of \$2160 each month, determine the initial amount that Donna invested.
- 3** Michael is thinking about retirement and decides that he would like to receive a monthly payment of \$3850 from his superannuation fund. He predicts that the annual interest rate on the account will be 4.2%, compounding monthly. Determine the amount that Michael should aim to have invested in his superannuation fund if he wishes for it to be a perpetuity.

Calculating the interest rate required for a perpetuity

- 4** Cathy wishes to maintain an ongoing donation of \$5500 per year to the Brisbane Broncos. If the interest on the initial investment averages 2.75% per annum, compounding annually, find how much she should she invest.

Example 14

- 5** Barbara would like to establish a scholarship that will reward the hardest working mathematics student in Year 12 each year with a \$500 prize.
 - a** The interest on her investment is 4% per annum. Find how much Barbara should invest into the scholarship fund.
 - b** Barbara has \$8000 to invest in the perpetuity. Find the annual interest rate that Barbara requires in order to pay the prize in perpetuity.
- 6** On retiring from work, Tyson received a superannuation payout of \$694 400. If Tyson invests the money in a perpetuity, he would then receive \$3645.60 each month for the rest of his life. Find the annual percentage rate of interest earned by this perpetuity.
- 7** Benjamin has \$12 000 to invest in a perpetuity to provide a prize of \$750 each year. Determine the interest rate that he requires in order to pay the prize in perpetuity if interest compounds annually.

Using technology to solve perpetuity investment problems

Example 15

- 8** Sandra has \$80 000 to invest. Using technology, find the minimum interest rate she requires to provide an annual donation of \$2400 in perpetuity if interest is compounding annually.
- 9** Marco invests \$350 000 in a perpetuity which earns interest at the rate of 6% per annum, compounding monthly. After one year, the annual interest rate declines to 4.8% per annum, compounding monthly. Determine how much less Marco receives each year after the decline in interest rates.
- 10** Omar inherits \$920 000 and splits the money between a perpetuity and an annuity. The perpetuity pays \$2340 each month based on an interest rate of 5.2% per annum that compounds monthly. The annuity investment has an interest rate of 4.8% per annum that compounds monthly. He realises that he only needs \$2000 from his perpetuity each month and so puts the remaining money as an additional payment into the compounding annuity investment. Calculate the total balance of his perpetuity and his compound investment after three years.

Paper 1-style multiple-choice questions

- 11** Drew invests a lumpsum into a perpetuity from which he'll receive a regular quarterly payment. The perpetuity has a compound interest rate of 5.2% per annum and compounds quarterly. Over two years, Drew receives a total of \$1747.20. The original amount invested was
- A** \$16 800
B \$33 600
C \$8400
D \$4200
- 12** Zihan invests \$200 000 in a perpetuity and receives \$560 each month. The recurrence relation that models the value of Zihan's perpetuity A_n , after n months is
- A** $A_0 = 200\,000$, $A_{n+1} = 1.0028A_n + 560$
B $A_0 = 200\,000$, $A_{n+1} = 1.0028A_n - 560$
C $A_0 = 200\,000$, $A_{n+1} = 1.0336A_n + 560$
D $A_0 = 200\,000$, $A_{n+1} = 1.0336A_n - 560$

Key ideas and chapter summary

**Annuity**

An **annuity** is a type of compound interest investment from which either a regular payment is withdrawn, or into which a regular payment is deposited. Each period, interest on an annuity is calculated on the balance of the investment before the regular payment is withdrawn or deposited.

Recursive model for an annuity

A **recurrence relation** can be used to determine the balance of an annuity after n compounding periods.

If A_n is the balance of the annuity after n compounding periods, i is the decimal rate of interest per compounding period, $r = 1 + i$ and the regular amount that is deposited or withdrawn after each compounding period is d , then the recursive model for an annuity can have one of two forms:

Deposit phase:

$$A_0 = \text{principal of investment}, \quad A_{n+1} = r \times A_n + d$$

Withdrawal phase:

$$A_0 = \text{principal of investment}, \quad A_{n+1} = r \times A_n - d$$

Payment schedule for an annuity

A table that summarises the interest calculations for an annuity is called a **payment schedule**. A payment schedule shows the payment number, withdrawal amount, interest paid, principal reduction and investment balance after each payment for some, or all, of the payments of an annuity.

Annuities formula

The **annuities formula** can be used to calculate the balance of an annuity after n compounding periods, given the principal (P), decimal rate of interest per compounding period (i), total number of compounding periods (n), and payment amount (d). The annuities formula is:

$$\text{Deposit : } A = d \frac{((1 + i)^n - 1)}{i} \quad (\text{for principle} = 0)$$

$$A = P(1 + i)^n + d \frac{((1 + i)^n - 1)}{i}$$

$$\text{Withdrawal : } A = P(1 + i)^n - d \frac{((1 + i)^n - 1)}{i}$$

Total interest

An annuity will earn interest in both deposit and withdrawal phases. If I is the **total interest** earned by the annuity, A_n is the balance of the annuity after n compounding periods and d is the amount of each regular deposit or withdrawal, then:

$$I = A_n - n \times d - A_0 \text{ during the deposit phase}$$

$$I = A_n + n \times d - A_0 \text{ during the withdrawal phase}$$

Perpetuity

A **perpetuity** is a special case of an annuity. The regular payment withdrawn from a perpetuity is equal to the interest earned by the principal. The value of a perpetuity will remain constant.

The future value of a perpetuity is the same as the principal, $A = P$.

Given the decimal rate of interest per compounding period (i), the regular payment amount (d) or the principal (P), the rules that allow the calculation of d , P and i are:

$$d = i \times P \qquad P = \frac{d}{i} \qquad i = \frac{d}{P}$$

Skills checklist



Download this checklist from the Interactive Textbook, then print it and fill it out to check your skills.

**8A****1** I can model an annuity during the deposit phase.

See Example 1 and Exercise 8A Question 1

8A**2** I can model an annuity with a recurrence relation.

See Example 2 and Exercise 8A Question 3

8A**3** I can analyse annuities with a recurrence relation.

See Example 3 and Exercise 8A Question 5

8B**4** I can interpret a payment schedule for an annuity in withdrawal phase.

See Example 4 and Exercise 8B Question 1

8C**5** I can use the future value annuity formula.

See Example 5 and Exercise 8C Question 1

8C**6** I can use the future value annuities formula when the balance isn't zero.

See Example 6 and Exercise 8C Question 5

- 8C** **7** I can use the future value annuities formula in the withdrawal phase. ☐
See Example 7 and Exercise 8C Question 8
- 8D** **8** I can solve annuity problems using a spreadsheet. ☐
See Example 8 and Exercise 8D Question 1
- 8D** **9** I can find the total interest in an annuity. ☐
See Example 9 and Exercise 8D Question 4
- 8D** **10** I can find the interest rate in an annuity. ☐
See Example 10 and Exercise 8D Question 7
- 8E** **11** I can calculate the payment withdrawn from a perpetuity. ☐
See Example 11 and Exercise 8E Question 1
- 8E** **12** I can calculate the payment withdrawn from a perpetuity using a recurrence relation. ☐
See Example 12 and Exercise 8E Question 6
- 8F** **13** I can calculate the investment required to establish a perpetuity. ☐
See Example 13 and Exercise 8F Question 1
- 8F** **14** I can calculate the interest rate required for a perpetuity. ☐
See Example 14 and Exercise 8F Question 4
- 8F** **15** I can use technology to solve a problem involving a perpetuity investment. ☐
See Example 15 and Exercise 8F Question 8

Multiple-choice questions

- 1** Julie has started a new job. She has a new superannuation account and her employer will deposit \$850 each month into this account. Assume that the money in this account will earn interest at the rate of 8.28% per annum, compounding monthly. A recurrence relation model for the balance of this investment after n months, A_n is
- A** $A_0 = 0, A_{n+1} = 1.0069 \times A_n + 850$
B $A_0 = 0, A_{n+1} = 1.069 \times A_n + 850$
C $A_0 = 0, A_{n+1} = 1.0828 \times A_n + 850$
D $A_0 = 0, A_{n+1} = 1.69 \times A_n + 850$

- 2** The value of an annuity investment, in dollars, after n years, V_n , can be modelled by the recurrence relation shown below

$$V_0 = 54\,000, \quad V_{n+1} = 1.0055V_n + 1500$$

The value of the regular payment added to the principal of this annuity investment is

- A** \$55 **B** \$297 **C** \$1500 **D** \$5400

- 3** The value of an annuity investment, in dollars, after n quarters, V_n , can be modelled by the recurrence relation shown below

$$V_0 = 36\,000, \quad V_{n+1} = 1.008V_n + 200$$

The increase in the value of this investment in the third quarter is closest to

- A** \$200 **B** \$288 **C** \$488 **D** \$496

Use the following information to answer Questions 4, 5 and 6.

An annuity is modelled by the recurrence relation

$$A_0 = 386\,000, \quad A_{n+1} = 1.0065 \times A_n - 5600$$

where A_n is the balance of the investment after n months.

- 4** The balance of the investment after four months is:
- A** \$370 342.77 **B** \$373 514.93 **C** \$376 666.59 **D** \$379 797.91
- 5** The annual percentage rate of interest for this investment is:
- A** 5.4% **B** 6.5% **C** 7.8% **D** 17.4%
- 6** The total interest earned after four months is closest to
- A** \$9915 **B** \$12 485 **C** \$22 400 **D** \$25 090
- 7** An investment of \$18 000 earns compound interest at the rate of 6.8% per annum, compounding yearly. Regular additions of \$2500 are made each year. The annuity can be modelled with a recurrence relation where V_n is the value of the investment after n years.
- The recurrence relation is
- A** $V_0 = 18\,000, V_{n+1} = 1.006V_n - 2500$
- B** $V_0 = 2500, V_{n+1} = 1.068V_n - 18\,000$
- C** $V_0 = 18\,000, V_{n+1} = 1.068V_n + 2500$
- D** $V_0 = 18\,000, V_{n+1} = 1.068V_n - 2500$

- 8** James invests \$50 000 in an annuity from which he receives a regular monthly payment of \$925.30.

The balance of the annuity, in dollars, after n months, J_n , can be modelled by a recurrence relation of the form

$$J_0 = 50\,000, \quad J_{n+1} = 1.0035J_n - 925.30$$

The balance of the annuity after six months is closest to

- A** \$45 289 **B** \$45 458 **C** \$45 459 **D** \$56 659

Use the following information to answer Questions 9, 10 and 11.

A payment schedule for the first five payments from an annuity is shown below.

Payment number	Payment withdrawn	Interest earned	Principal reduction	Balance of investment
0	0	0	0	285 000.00
1	5500.00	1653.00	3847.00	281 153.00
2	5500.00	1630.69	3869.31	277 283.69
3	5500.00	1608.25	3891.75	273 391.94
4	5500.00	1585.67	3914.33	A
5	5500.00	1562.97	3937.03	265 540.58

- 9** The principal amount of this annuity is
A \$3847 **B** \$5500 **C** \$281 153 **D** \$285 000
- 10** The value of A, the balance of the annuity after payment 4 is
A \$267 891.94 **B** \$269 477.61 **C** \$269 500.19 **D** \$271 806.27
- 11** The interest charged on this annuity compounds monthly and monthly payments are received.
 The annual percentage rate of interest for this annuity is closest to
A 1.35% **B** 5.80% **C** 6.96% **D** 16.12%
- 12** Michelle will spend one year travelling the world. She invests \$25 000 into an annuity that earns interest at the rate of 7.08% per annum, compounding monthly. Michelle expects to receive exactly 12 payments from this investment, the first 11 of which are equal in value, before it is fully exhausted.
 The monthly payment that Michelle receives is closest to
A \$1160 **B** \$2160 **C** \$3160 **D** \$4160

- 13** An annuity with a principal of \$285 000 earns interest at the rate of 6.36% per annum, compounding monthly. Monthly payments of \$3200 are received from this investment. The number of payments until the balance of this investment is first below \$270 000 is
A 7 **B** 8 **C** 9 **D** 10
- 14** Monthly withdrawals of \$220 are made from an account that has an opening balance of \$35 300, invested at 7.2% per annum, compounding monthly. The balance of the account after one year is closest to
A \$37 927 **B** \$35 291 **C** \$32 660 **D** \$35 198
- 15** Tilly invests \$5000 in an account that pays interest at the rate of 3.9% per annum, compounding annually. She makes an additional payment of \$1200 each year. The number of years that it will take the investment to first reach a balance of \$20 000 is
A 1 **B** 2 **C** 9 **D** 10
- 16** Pham invests \$74 000 from which he'll receive a regular monthly payment. The perpetuity has a compound interest rate of 4.8% per annum and compounds monthly. The amount that Pham will receive from the perpetuity in the first two years is closest to
A \$296 **B** \$3233 **C** \$3552 **D** \$7104
- 17** Francesca invests \$6000 into an annuity with an interest rate of 5.8% per annum, compounding quarterly. She receives a quarterly payment of \$1554.76 for the first three quarters and then a final payment in the fourth quarter. After three quarters, the value of the annuity is \$1532.56. The payment that Francesca receives in the fourth quarter is
A \$0.02 **B** \$0.03 **C** \$1510.34 **D** \$1554.79
- 18** Benjamin invests \$75 000 in an annuity, paying 7.8% per annum, compounding monthly. Benjamin receives a payment of \$2326 each month from the annuity. The value of the annuity after two years is closest to
A \$27 415 **B** \$130 824 **C** \$399 071 **D** \$19 176
- 19** A perpetuity will be set up to provide an annual prize of \$400 to the best mathematics student in a school. Interest will be earned on the principal of the investment at a rate of 3.4% per annum and will be used to pay for the prize every year. The amount that must be invested is closest to
A \$400 **B** \$800 **C** \$1176 **D** \$11 765

- 20** A perpetuity has a balance of \$120 000 after six years. Interest is earned at the percentage annual interest rate of 5% per annum. The perpetuity is used to provide an annual prize of value \$6000.
- After a further six years, the balance of the perpetuity is
- A** \$84 000.00 **B** \$86 654.33 **C** \$120 000.00 **D** \$123 563.76

Short-response questions

- 1** An annuity in withdrawal phase is modelled using the recurrence relation shown below. sf
- $$A_0 = 624\,000, A_{n+1} = 1.0013 \times A_n - 2500$$
- In the recurrence relation, A_n is the balance of the investment after n weekly repayments.
- State the principal of this investment.
 - State the value of the weekly payments withdrawn.
 - Find the balance of this loan after six payments have been withdrawn.
- 2** Carys has \$345 000 to invest in an annuity. Interest will be paid at the annual percentage interest rate of 4.6%, compounding quarterly. Carys will withdraw a payment of \$12 000 per quarter from the investment.
- Let A_n be the balance of Carys' annuity after n quarters.
- Construct a recurrence relation model for the balance of Carys' investment after n quarters.
 - Apply the recurrence relation to calculate the balance of Carys' investment after six payments have been withdrawn.
 - Find the total interest earned after six payments have been received.
- 3** Leigh invests \$64 000 in an annuity, with interest of 6.3% per annum, compounding monthly. He receives payments of \$1275 per month.
- Determine how long his annuity will last. Give your answer correct to the nearest month.
 - Find the balance of Leigh's investment after six months.
 - Find how much interest Leigh will earn over the first six months.
- 4** Raj invests \$85 500 in an annuity, with interest of 7.25% per annum, compounding quarterly. Raj receives a regular quarterly payment of \$5000.
- State the balance of Raj's annuity after one year.
 - Find the total amount of interest that Raj has earned after one year.
 - Find how long Raj's annuity will pay the full amount each quarter.

- 5** Stephanie invests \$40 000 in an annuity, with interest paid at 7.5% per annum compounded monthly. She wishes to receive a monthly payment for 10 years. Determine the amount that Stephanie can receive each month, to the nearest cent.
- 6** The payment schedule for an annuity with monthly payments is shown below.

Payment number	Payment withdrawn	Interest earned	Principal reduction	Balance of investment
0	0	0	0	84 000.00
1	14 500.00	394.80	14 105.20	69 894.80
2	14 500.00	328.51	14 171.49	55 723.31
3	14 500.00	261.90	14 238.10	41 485.21
4	14 500.00	194.98	14 305.02	27 180.19
5	14 500.00	127.75	14 352.25	12 827.94
6		60.29		0.00

- a** State the principal of this investment.
- b** State the value of the monthly payment.
- c** Calculate the:
- i** monthly interest rate.
 - ii** annual interest rate.
- d** The annuity was fully exhausted by the sixth payment. Determine the value of the payment.
- 7** Michael's employer has started a superannuation fund for him. Each month, \$250 will be placed in this account. The superannuation fund earns interest at the annual percentage interest rate of 3.89%, compounding monthly. Determine the balance of the account after five years. Round your answer to the nearest cent.
- 8** A university would like to establish a perpetuity, the interest from which will fund a scholarship for a talented student each year. The value of this prize should be \$5400 per year.
- a** If the interest on this investment is 3.6% per annum, determine how much the university will need to invest.
- The university only has \$60 000 to invest.
- b** Find the annual interest rate that is required to pay the scholarship in perpetuity. Round your answer to two decimal places.
- c** Use technology to verify your answer to part **b** above.

- 9** Samantha inherited \$150 000 from her aunt. She decides to invest this money into an account paying 6.6% per annum interest, compounding monthly.
- a** If Samantha deposited her money into a perpetuity, find the monthly payment that she would receive.
 - b** If Samantha deposited her money into an annuity and withdrew \$1000 per month, find the amount she would have in the account after one year.
 - c** If Samantha deposited her money into an annuity and withdrew \$2000 per month, find how long would it take for the value of her investment to first drop below \$100 000.
 - d** If Samantha deposited her money into an annuity and withdrew \$4000 per month:
 - i** Determine how many payments of \$4000 Samantha could receive.
 - ii** Including interest, determine the value of Samantha's last withdrawal.
- 10** Helene has won \$750 000 in a lottery. She decides to place the money in an investment account that pays 4.5% per annum interest, compounding monthly.
- a** Find the balance that Helene has in the investment account after 10 years.
 - b** After the 10 years are up, Helene decides to use her money to invest in an annuity, which pays 3.6% per annum, compounding monthly. Determine how long the annuity will last if Helene withdraws \$6000 per month for her living expenses.
 - c** Helene's accountant suggests that rather than purchase an annuity she place \$1 175 000 in a perpetuity so that she will be able to leave some money to her grandchildren. If the perpetuity pays 3.6% per annum compounding monthly, find the monthly payment that Helene will receive.
- 11** Mary has just started a new job. She will be paid a salary of \$63 000 per year. Mary receives her salary as 12 equal payments on the first of every month. Mary's employer will pay 8.4% of her monthly salary into a superannuation account each month. The money in this account earns interest at the rate of 1.8% per annum, compounding monthly.
- a** Find Mary's monthly salary.
 - b** Find how much Mary's employer will pay into the superannuation account each month.
 - c** Use the annuity formula (principal = \$0) to determine the balance of Mary's superannuation account after:
 - i** one full year of work
 - ii** ten years of work.
 - d** Determine the total amount of interest that Mary's superannuation fund earns after ten years of work.

- 12** Henry has just started a new job. He will be paid a salary of \$63 960 per year. Henry receives his salary as 26 equal fortnightly payments. Every fortnight, Henry's employer will pay 7.8% of his fortnightly salary into a superannuation account. The money in this account earns interest at the rate of 3.9% per annum, compounding fortnightly.
- State how much Henry will receive each fortnight.
 - Find how much Henry's employer will pay into the superannuation account each fortnight.
 - Use the annuity formula (principal = \$0) to determine the balance of Henry's superannuation account after:
 - one year of work
 - five years of work.

Henry will add \$100 of his own money into the superannuation account each fortnight.
 - Use the annuity formula (principal = \$0) to determine:
 - the balance of the superannuation account after one year with these extra payments
 - how much higher the balance of the superannuation account will be after one year with extra payments compared to the balance with no extra payments.
- 13** Ludwig inherited \$150 000 from his aunt. He decided to invest this money into an account that pays interest at the rate of 5.76% per annum, compounding monthly.
- If Ludwig's account was a perpetuity, find the monthly payment that he would receive.
 - If Ludwig's account was an annuity and he withdrew \$2000 per month:
 - write a recurrence relation to model this annuity.
 - find how much money would be left in the account after 6 months.
 - determine how many months it would take for the balance of the investment to first fall below \$130 000.
 - If Ludwig's account was an annuity and he withdrew \$4500 per month:
 - find how many payments of \$4500 he could receive.
 - determine the value of his final repayment if it was smaller than all the others.
- 14** Jagathi receives monthly payments of \$5250 from an annuity that is earning interest at the rate of 5.28% per annum, compounding monthly. The balance of Jagathi's investment is \$376 623.14 after three years of investment.
- State the principal of Jagathi's investment.
 - Find how much interest Jagathi earned after three years of investment.
 - Find how many more payments of \$5250 Jagathi can withdraw.
 - Determine the final amount Jagathi can withdraw to fully exhaust his annuity.

- 15** Ebrahim has recently retired and will invest his superannuation payment into an annuity that will earn interest at the rate of 5.04% per annum, compounding monthly. The principal amount of this investment will be \$534 000. Ebrahim decided to decrease his monthly payment after four years of investment. He will withdraw monthly payments of \$4500 until his investment is fully exhausted.
- a** Find how many payments of \$4500 Ebrahim can expect.
 - b** Find the final amount that Ebrahim can withdraw to fully exhaust his annuity.
- 16** Byron has begun a new job and a superannuation account was opened by his employer. Byron will be paid \$1400 per fortnight. His employer will deposit 8% of his salary into the superannuation account every fortnight. Assume that the superannuation account earns interest at the annual percentage interest rate of 2.1%, compounding fortnightly. After one year of work, Byron's employer offers him a change of salary conditions. Byron has a choice:
- Option 1** – a salary of \$1800 per fortnight with no change to superannuation conditions.
- Option 2** – a salary of \$1600 per fortnight with 8.2% of salary deposited into the superannuation account each fortnight.
- a** Determine which option would result in the greater amount of money deposited in the superannuation account.
- Assume that Byron chooses Option 1 and that the annual percentage interest rate of the superannuation account remains constant at 2.1% per annum, compounding fortnightly. After ten years, Byron changes jobs. He transfers his superannuation balance to a new account that earns interest at the rate of 2.4% per annum, compounding monthly. Byron's new job pays a monthly salary of \$3500 and his new employer will contribute 9.4% of this salary to the superannuation account each month.
- b i** Calculate the balance of Byron's superannuation account after ten more years.
 - ii** Determine the total interest that Byron will earn over the twenty-one years he will hold the superannuation account.
- 17** Jarrod's superannuation account has a balance of \$250 000. His employer adds \$1000 to this account every month. Jarrod's superannuation fund pays him interest at the annual percentage interest rate of 3.36% with monthly compounds, and he intends to work for a further 3 years before retiring. Jarrod then plans to invest the money from his superannuation fund in an annuity that will pay interest at the annual percentage interest rate of 4.08% with quarterly compounds. He would like this money to last for at least 15 years. Determine the quarterly payment that he can withdraw and the total amount of interest that his annuity fund will earn during the 15 years.

Chapter 9

Graphs and networks

Chapter questions

UNIT 4 INVESTING AND NETWORKING

Topic 3: Graphs and networks

- ▶ How do we identify the features of a graph?
- ▶ How do we draw a graph?
- ▶ How do we apply graphs in practical situations?
- ▶ How do we construct an adjacency matrix from a graph?
- ▶ How do we define and draw a planar graph?
- ▶ How do we identify the type of path on a graph?
- ▶ How do we draw and use Eulerian semi-Eulerian graphs?
- ▶ How do we draw and use Hamiltonian and semi-Hamiltonian graphs?
- ▶ How do we find the shortest path between two vertices of a graph?

In this chapter graphs and their use as networks representing connections between objects will be introduced, in addition to exploring their properties and applications.

9A Graphs and associated terminology

Learning intentions

- ▶ To be able to define and identify a graph, vertex, edge and loop.
- ▶ To be able to find the degree of a vertex and the sum of degrees.
- ▶ To be able to describe the features of a graph.

Representing connections with graphs

There are many situations in everyday life that involve connections between people or objects. Towns are connected by roads, computers are connected to the internet and people connect to each other through being friends on social media. A diagram that shows these connections is called a **graph**.



Networks: Basic concepts Watch the video in the Interactive Textbook for an illustration of the terms and concepts in action.

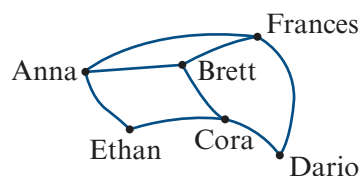
Vertices and edges

The graph opposite represents the connections between friends on a social media website.

There are six people in this graph and each person is represented by a dot called a **vertex**.

Each vertex in the graph is joined to some of the other *vertices* (plural of vertex) by a line called an **edge**.

These lines represent the connection between the people represented by the vertices.

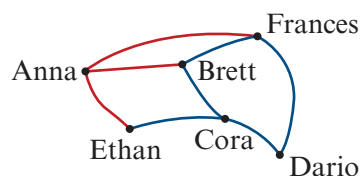


For example, the vertex for Anna is connected by an edge to the vertex for Brett, which means Anna and Brett are connected as friends on the website.

There is no edge between the vertices for Frances and Cora, which means they are not connected as friends on the website.

The degree of a vertex

The graph of the social media connections opposite shows that Anna has three friends, Frances, Brett and Ethan. There are three edges that connect Anna to other people. This number is called the **degree of the vertex** representing Anna. It is the number of times an edge connects to that vertex.



In symbolic form, the degree of the vertex representing Anna can be written as $\deg(\text{Anna}) = 3$.

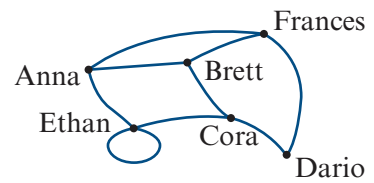
We say a vertex is **even** if the degree of the vertex is even, and we say a vertex is **odd** if the degree of the vertex is odd.

Loops

Imagine that Ethan is able to add himself as a friend on the social media website.

The edge representing this connection would connect the vertex representing Ethan back to itself. This type of edge is called a **loop**.

A loop is attached twice to a vertex and so it will add two to the degree for that vertex. In this graph, $\deg(\text{Ethan}) = 4$.

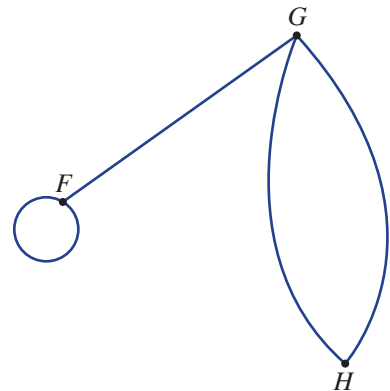


Representing connections with graphs

- A graph consists of a number of vertices and a number of edges. The edges join some pairs of vertices.
- The number of edges attached to a vertex is called the **degree** of that vertex.
- The degree of vertex V is written as $\deg(V)$.
- A vertex is **even** if the degree of the vertex is even, and a vertex is **odd** if the degree of the vertex is odd.
- A **loop** is an edge that connects a vertex to itself.
- A loop connects twice to a vertex and so it adds two to the degree of that vertex.

For example, consider the graph opposite.

- There are 3 vertices.
- There are 4 edges.
- The vertex F has a loop.
- The degree of vertex F is 3, so we write $\deg(F) = 3$.
- The degree of vertex G is 3, so we write $\deg(G) = 3$.
- The degree of vertex H is 2, so we write $\deg(H) = 2$.
- The sum of degrees for this graph is $2 \times 4 = 8$, twice the number of edges. (In this case there are 4 edges.)

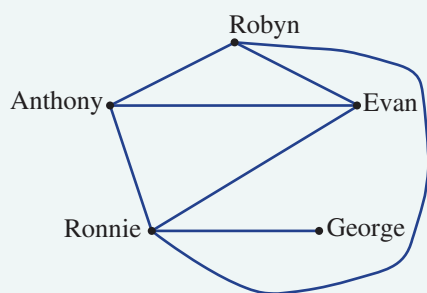



Example 1 Drawing a graph to represent connections

Five people, Anthony, Ronnie, Robyn, George and Evan have accounts on a social media website.

- Ronnie is a friend of everyone.
- Anthony is a friend of Robyn, Ronnie and Evan.
- George is a friend of Ronnie only.
- Robyn is a friend of Anthony, Ronnie and Evan.
- Evan is a friend of Anthony, Ronnie and Robyn.

- a Draw a graph to represent the connections between the five people above.
- b Write down the value of $\deg(\text{Anthony})$.
- c Which person has the vertex with the:
 - i smallest degree?
 - ii largest degree?
- d Which vertices have even degree?

Solution
a


The graph must have an edge between:

- Anthony and Robyn
- Anthony and Ronnie
- Anthony and Evan
- Ronnie and Robyn
- Ronnie and George
- Ronnie and Evan
- Robyn and Evan.

There will be one edge for every pair of people that are friends on the social media website.

Note: The position of the vertices representing the people do not have to be in the same position as they are shown in the diagram. As long as the edges connecting the people are the same, the graph can be drawn with vertices in many different positions.

- b $\deg(\text{Anthony}) = 3$

The vertex representing Anthony has three edge connections to it.

- c i George has the vertex with the smallest degree. $\deg(\text{George}) = 1$

The vertex representing George has only one edge connection to it, less than all the other vertices, so it has the smallest degree.

- ii Ronnie has the vertex with the largest degree. $\deg(\text{Ronnie}) = 4$

The vertex representing Ronnie has four edge connections to it, more than all the other vertices, so it has the largest degree.

- d The vertex Ronnie has degree 4.

Describing graphs

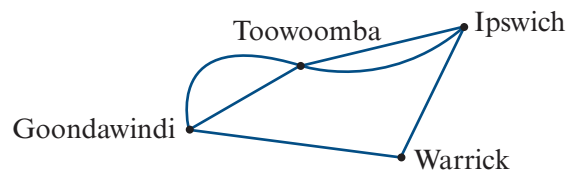
Graphs that represent connections between objects can take different forms and have different features. This means that there is a variety of ways to describe these graphs.

Multiple edges

The graph below shows six cities in Queensland represented as vertices, with the major highways connecting these towns represented as edges.

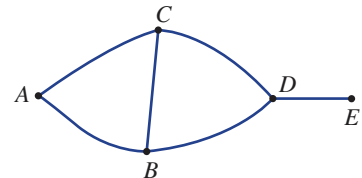
There are two different major highways that can be travelled to drive from Ipswich to Toowoomba and the graph shows this using **multiple edges**.

Multiple edges connect the same two vertices in a graph.



Simple graphs

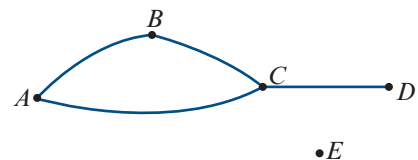
Simple graphs do not have any loops. There are no duplicate or **multiple edges** either.



Isolated vertex

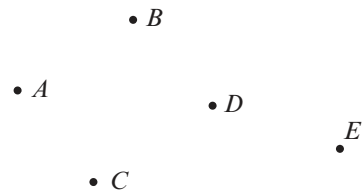
A graph has an **isolated vertex** if there is a vertex that is not connected to another vertex by an edge.

The isolated vertex in this graph is *E*, because it is not connected to any other vertex by an edge. The degree of an isolated vertex is 0.



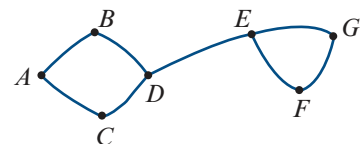
Degenerate graphs

Degenerate graphs have all vertices isolated. This means that there are no edges in the graph at all.

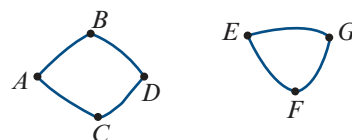


Connected graphs and bridges

A **connected graph** has every vertex connected to every other vertex, either directly or indirectly via other vertices. The graph on the right is connected. A **bridge** is an edge in a connected graph that, if removed, will cause the graph to be disconnected. The graph on the right has a bridge connecting vertex *D* to vertex *E*.

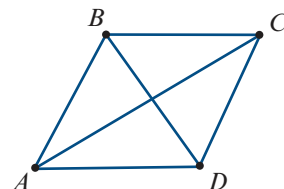


The graph on the right shows the bridge from vertex D to vertex E removed. There are now two separate sections of the graph that are not connected to each other.

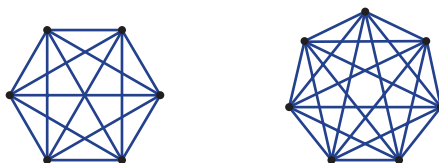


Complete graphs

If there is an edge between every pair of vertices, the graph is called a **complete graph**. Every vertex in the graph is connected directly by an edge to every other vertex in the graph.



Shown below are the complete graphs with 6 and 7 vertices.

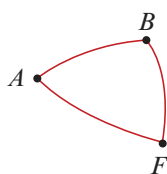
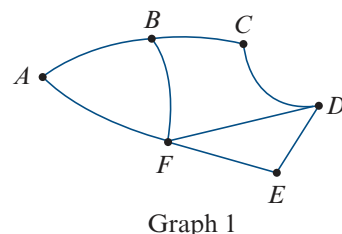


Note that the degree of each vertex is one less than the number of vertices. From this observation we see that the number of edges for a complete graph with n vertices is $\frac{1}{2} \times n \times (n - 1)$.

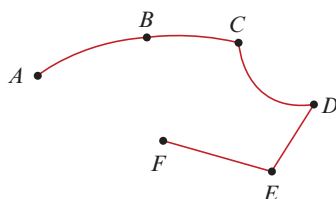
Subgraphs

A **subgraph** is a part of a larger graph. All of the edges and vertices in the subgraph must exist in the original graph.

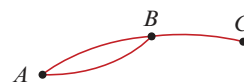
If there are extra edges or vertices, the graph will not be a subgraph of the larger graph.



Graph 2



Graph 3



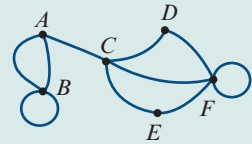
Graph 4

Graphs 2 and 3 above are subgraphs of graph 1. All of the vertices and edges in graphs 2 and 3 exist in graph 1.

Graph 4 above is not a subgraph of graph 1. There are two edges connecting vertex A to vertex B , but in graph 1 there is only one.

**Example 2** Describing graphs

A connected graph is shown on the right.



- What is the degree of vertex C ?
- Which vertices have a loop?
- What is the degree of vertex F ?
- A bridge exists between two vertices. Which vertices are they?
- Draw a subgraph of this graph that involves only vertices A , B and C .

Solution

- The degree of vertex C is 4.
 $\deg(C) = 4$
- Vertex B and vertex F have loops.
- The degree of vertex F is 5.
 $\deg(F) = 5$
- A bridge exists between vertex A and vertex C .
-

Explanation

Count the number of times an edge connects to vertex C . There are four connections.

A vertex has a loop if an edge connects it to itself.

Count the number of times an edge connects to vertex F . Remember that a loop counts as two degrees.

Look for an edge that, if removed, would disconnect the graph.

There are a few possible answers for this question. Some are shown on the left.

Investigation 9A: The shaking hands problem

The investigation above showed that each edge of a graph contributes two to the sum of the degrees of the vertices. This can easily be verified by the fact that in order for an edge to exist, it must involve two vertices.

In any graph, the sum of the degrees of the vertices is always twice the number of edges in that graph.

Section Summary

- ▶ A graph consists of a number of vertices and a number of edges. The edges join some pairs of vertices.
- ▶ The number of edges attached to a vertex is called the **degree** of that vertex.
- ▶ The **degree** of vertex V is written as $\deg(V)$.
- ▶ A vertex is **even** if the degree of the vertex is even, and a vertex is **odd** if the degree of the vertex is odd.

- ▶ A **loop** is an edge that connects a vertex to itself.
- ▶ A loop connects twice to a vertex and so it adds two to the degree of that vertex.
- ▶ A **connected graph** is a graph that has no isolated vertices and no separate parts.
- ▶ A **bridge** is an edge that, if removed, would cause the graph to no longer be connected.
- ▶ An **isolated vertex** is a vertex in a graph that is not connected to any other vertex by an edge.
- ▶ The degree of an isolated vertex is zero.
- ▶ **Multiple edges** connect the same two vertices of a graph.
- ▶ **Simple graphs** do not have loops and do not have multiple edges.
- ▶ **Complete graphs** have an edge between every pair of vertices. If the graph has n vertices, each vertex has degree $n - 1$ and there are $\frac{1}{2} \times n \times (n - 1)$ edges.
- ▶ **Subgraphs** are a small section of an existing graph, with no extra vertices and no extra edges.
- ▶ In any graph, the sum of the degrees of the vertices is always twice the number of edges in that graph.



Exercise 9A

Example 1

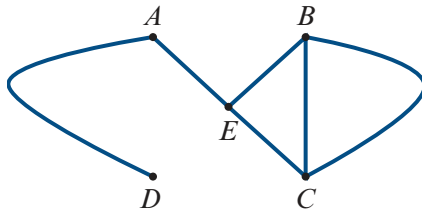
- 1 Six people – Albert, Bryn, Charles, David, Elizabeth and Francis – are connected on a social media website. In the graph,
 - Albert is a friend of Bryn, Charles and Francis
 - Charles is a friend of Albert, Bryn and David
 - David is a friend of Bryn and Charles
 - Elizabeth is a friend of Bryn
 - Bryn is a friend of Albert, Charles, David and Elizabeth
 - Francis is a friend of Albert and Bryn.
 - a Draw a graph to represent the connections between the six people above.
 - b Write down the value of $\deg(\text{Albert})$.
 - c Which person has the vertex with the:
 - i smallest degree?
 - ii largest degree?
 - d Which people have vertices with odd degree?
 - e What is the degree of the vertex for Elizabeth?
- 2 Six people are seated at a round table. For each of the following, draw a graph that models the situation:
 - a Each person shakes hands with the two people they are sitting next to.
 - b Each person shakes hands with the person they are sitting opposite.
 - c Each person shakes hands with every other person.

Describing graphs

Example 2

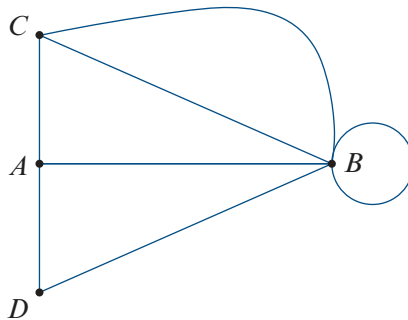
3 For each graph shown, complete the associated statements by filling in the boxes.

a



- i The graph has vertices.
- ii The graph has edges.
- iii The graph has loops.
- iv $\deg(A) =$
- v $\deg(E) =$
- vi The graph has odd vertices.
- vii The graph has even vertices.

b



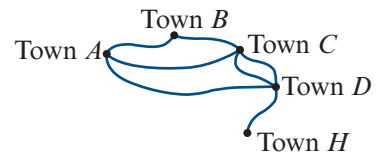
- i The graph has vertices.
- ii The graph has edges.
- iii The graph has loops.
- iv $\deg(B) =$
- v $\deg(D) =$
- vi The graph has odd vertices.
- vii The graph has even vertices.

Example 2

4 The graph below shows five towns, A, B, C, D and H, represented as vertices and the roads between the towns are represented by edges.

a Write down the degree of the vertex representing:

- i Town A
- ii Town B
- iii Town H



- b i How many edges are in this graph?
- ii What is the sum of the degrees of the vertices in this graph?
- iii Verify your answer to b ii by adding all the degrees of the vertices.
- c Draw a subgraph of this graph that contains only towns H, D and C.
- d Which two towns are connected by the bridge in this graph?

Drawing graphs

5 Draw a graph that has:

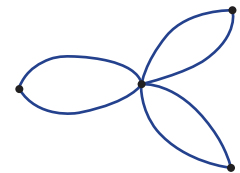
- a three vertices, two of which have an odd degree
- b four vertices and five edges, one of which is a loop
- c six vertices, one of which is isolated, and eight edges
- d six vertices, two of which have an odd degree, and which contain at least one subgraph that is a triangle.

Properties of graphs

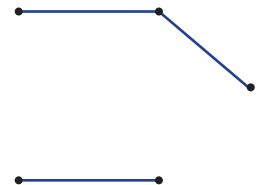
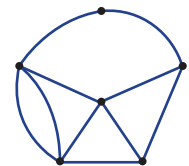
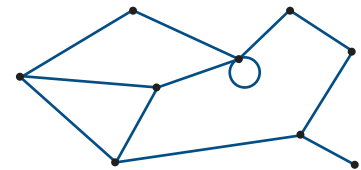
- 6** Complete the following for a graph.
- a** Draw a complete graph that has 5 vertices.
 - b** How many edges does this graph have?
- 7** A simple connected graph has five vertices and seven edges. Find the sum of the degrees of the vertices.
- 8** Find the number of edges needed to make a complete graph with six vertices.

Paper 1-style multiple-choice questions

- 9** Consider the following graph.
Which one of the following statements is **not** true for this graph?

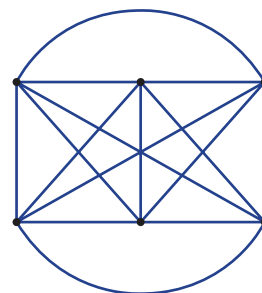


- A** There are three loops.
 - B** All vertices have an even degree.
 - C** Three of the vertices have the same degree.
 - D** The sum of the degrees of the vertices is twelve.
- 10** Consider the graph opposite.
The number of vertices the graph has is
- A** 5 **B** 7 **C** 9 **D** 11
- 11** Consider the graph opposite.
The number of vertices with a degree of 4 is
- A** 1 **B** 2 **C** 3 **D** 4
- 12** Consider the graph shown opposite.
The minimum number of edges that must be added to make this a complete graph is
- A** 12 **B** 2 **C** 6 **D** 8
- 13** What is the maximum number of edges in a simple graph with 6 vertices?
- A** 6 **B** 12
 - C** 15 **D** 18



- 14** The following graph with six vertices is a complete graph. Edges are removed so that the graph will have the minimum number of edges to remain connected. The number of edges that are removed is

A 6 **B** 8
C 10 **D** 12



9B The adjacency matrix

Learning intentions

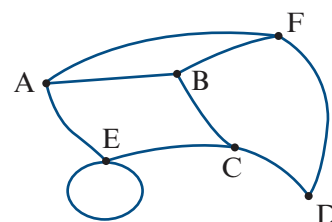
- ▶ To be able to use matrices to describe graph.
- ▶ To be able to draw a graph from an adjacency matrix.

Vertices in a graph that are connected by an edge are said to be *adjacent* to, or next to, each other. A **matrix** can be used to record which vertices are adjacent to each other and also the number of connections between adjacent vertices. This matrix is called an **adjacency matrix**.

Consider the simple graph that shows the social media connections between people from earlier in the chapter. For the purposes of creating an adjacency matrix for this graph, we will represent the vertices with the first letter of the names of the people. For example, Anna will be represented by *A*, Brett will be represented by *B* and so on.

The adjacency matrix and graph for that graph are shown below.

$$\begin{array}{c}
 A \quad B \quad C \quad D \quad E \quad F \\
 \begin{array}{l}
 A \left[\begin{array}{cccccc} 0 & 1 & 0 & 0 & 1 & 1 \end{array} \right] \\
 B \left[\begin{array}{cccccc} 1 & 0 & 1 & 0 & 0 & 1 \end{array} \right] \\
 C \left[\begin{array}{cccccc} 0 & 1 & 0 & 1 & 1 & 0 \end{array} \right] \\
 D \left[\begin{array}{cccccc} 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right] \\
 E \left[\begin{array}{cccccc} 1 & 0 & 1 & 0 & 1 & 0 \end{array} \right] \\
 F \left[\begin{array}{cccccc} 1 & 1 & 0 & 1 & 0 & 0 \end{array} \right]
 \end{array}
 \end{array}$$



The adjacency matrix has:

- six rows and six columns, one for each vertex of the graph
- row and column labels that match the vertices of the graph
- a '0' in the intersection of row *A* and column *C* because there is no edge connecting *A* to *C*

- a '1' in the intersection of row B and column F because there is one edge connecting B to F
- a '0' in the intersection of row D and column D because there is no edge connecting D to D (that is, there is no loop at vertex D)
- a '1' in the intersection of row E and column E because there is one edge connecting E to E (that is, there is a loop at vertex E).

The number of edges between every other pair of vertices in the graph is recorded in the adjacency matrix in the same way. Notice that the adjacency matrix is symmetric about the main diagonal. This means that the number in row m , column n is the same as the number in row n , column m .



Example 3 Drawing a graph from an adjacency matrix

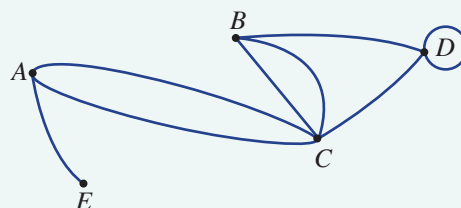
Draw the graph that has this adjacency matrix.

	A	B	C	D	E
A	0	0	2	0	1
B	0	0	2	1	0
C	2	2	0	1	0
D	0	1	1	1	0
E	1	0	0	0	0

Solution

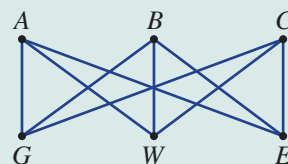
- Draw a dot for each vertex and label them A to E .
- There is a '2' in the intersection of row A and column C . This means that there are two edges connecting vertex A and vertex C . These will be multiple edges.
- There is a '1' in the intersection of row D and column D . This means that there is a loop at vertex D (vertex D is connected to itself by a loop).
- Look at every intersection of row and column in the matrix and add edges to the graph, if they do not already exist.

Note: The graph has been drawn so that the edges do not cross over each other. This is not strictly necessary.



**Example 4**

Construct an adjacency matrix that can be used to represent the graph opposite. This graph represents the ways that three houses A, B and C are connected to three utility outlets, gas (G), water (W) and electricity (E).

**Solution**

	A	B	C	G	W	E
A	0	0	0	1	1	1
B	0	0	0	1	1	1
C	0	0	0	1	1	1
G	1	1	1	0	0	0
W	1	1	1	0	0	0
E	1	1	1	0	0	0

Explanation

The convention used to enter the values is the same as discussed above.

The graph in Example 4 is called a bipartite graph as the set of vertices is separated into two sets of objects Houses (A, B, C) and Utility outlets (G, W, E) with each edge connecting a vertex in each set. You will meet bipartite graphs again in Chapter 11, when we study allocation problems.

Adjacency matrices

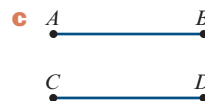
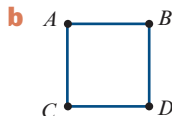
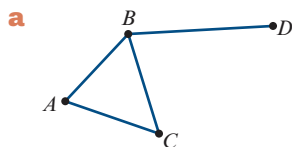
An **adjacency matrix** is a square matrix that summarises the connections between vertices of a graph.

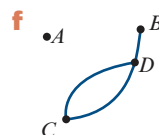
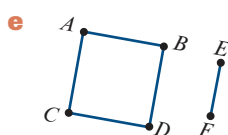
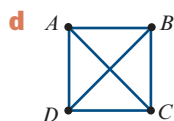
- The entry in row m and column n shows the number of edges that join the vertices from this row and column.
- The adjacency matrix is symmetric about the main diagonal.

Loops are counted as one edge.

Exercise 9B**Writing adjacency matrices for graphs**

- 1 For each of the following graphs, write down the adjacency matrix.





Drawing graphs from adjacency matrices

Example 3

2 Draw a graph for the following adjacency matrices.

a

	A	B	C
A	0	1	1
B	1	0	1
C	1	1	0

b

	A	B	C	D
A	1	0	0	0
B	0	0	1	1
C	0	1	1	0
D	0	1	0	0

c

	A	B	C	D
A	0	1	2	1
B	1	0	1	1
C	2	1	0	0
D	1	1	0	0

Properties of graphs

3 The adjacency matrix on the right has a row and column for vertex C that contains all zeros. What does this tell you about vertex C ?

	A	B	C
A	0	1	0
B	1	0	0
C	0	0	0

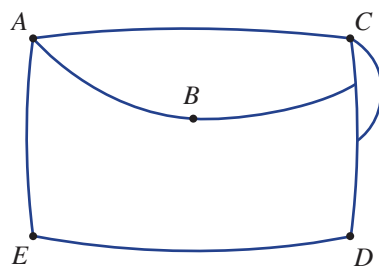
4 Every vertex in a graph has one loop. What feature of the adjacency matrix would tell you this information?

5 A graph has five vertices: A, B, C, D and E . It has no duplicate edges and no loops. If this graph is complete, write down the adjacency matrix for the graph.

Paper 1-style multiple-choice questions

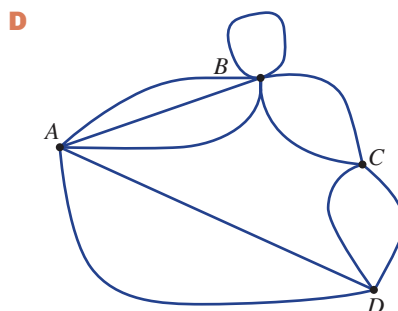
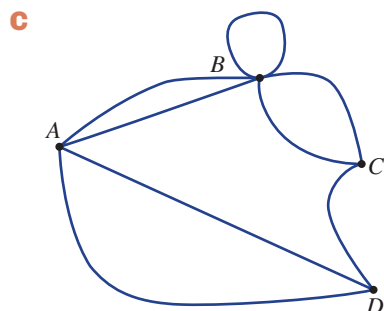
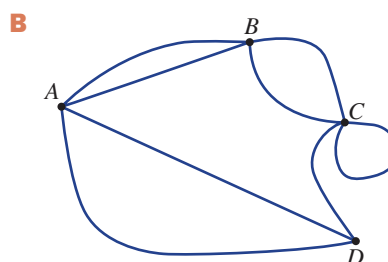
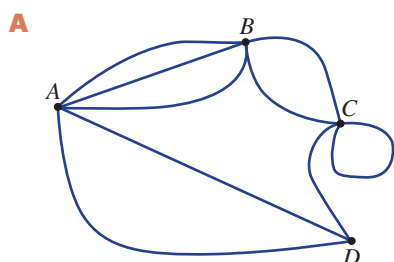
6 The network opposite shows the pathways between five buildings: A, B, C, D and E . An adjacency matrix for this network is formed. The number of zeros in this matrix is

- A** 9 **B** 10 **C** 11 **D** 12



7 The adjacency matrix opposite shows the number of pathways between four points: A, B, C and D . A network of pathways that could be represented by the adjacency matrix is

	A	B	C	D
A	0	3	0	2
B	3	0	2	0
C	0	2	1	1
D	2	0	1	0



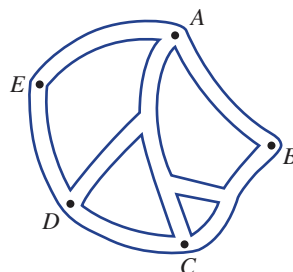
Use the following information to answer Questions 8 and 9.

- 8** The network shows the pathways between five schools: A, B, C, D and E.

An adjacency matrix for this network is formed.

Of the 25 elements in the adjacency matrix, the number '1' appears

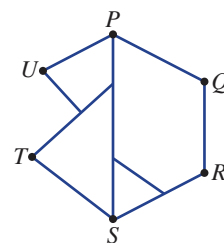
- A** 7 times **B** 8 times
C 9 times **D** 10 times



- 9** Of the 25 elements in the adjacency matrix, the numbers '2' or '3' appear

- A** 6 times **B** 7 times **C** 8 times **D** 9 times

- 10** The map opposite shows all the road connections between six towns, P, Q, R, S, T and U. The road connections could be represented by the adjacency matrix



A

	P	Q	R	S	T	U
P	0	1	1	1	1	1
Q	1	0	1	0	0	0
R	1	1	0	1	1	1
S	1	0	1	1	1	1
T	1	0	1	1	0	1
U	1	0	1	1	1	0

B

	P	Q	R	S	T	U
P	0	1	1	1	1	1
Q	0	1	1	1	1	1
R	0	1	1	1	1	1
S	1	0	1	1	2	2
T	1	0	1	2	0	1
U	1	0	1	2	1	0

C

	P	Q	R	S	T	U
P	0	2	1	2	1	0
Q	1	0	1	0	0	0
R	1	1	0	2	1	1
S	2	0	2	1	3	2
T	1	0	1	3	0	1
U	2	0	1	2	1	0

D

	P	Q	R	S	T	U
P	0	1	1	2	1	2
Q	1	0	1	0	0	0
R	1	1	0	2	1	1
S	2	0	2	1	3	2
T	1	0	1	3	0	1
U	2	0	1	2	1	0

9C Directed graphs and their adjacency matrices

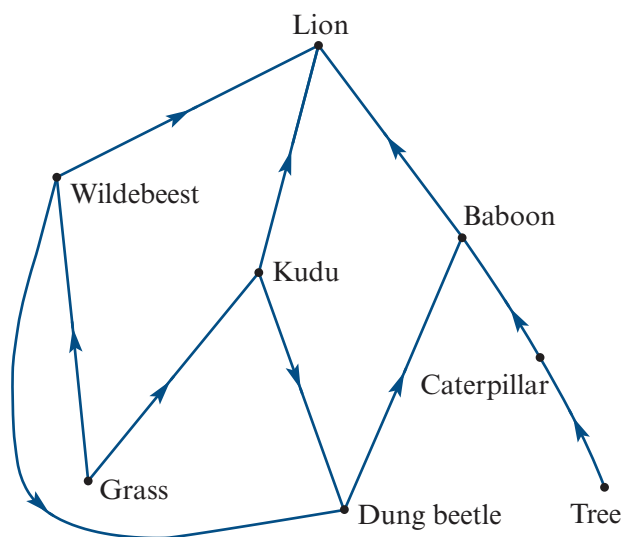
Learning intentions

- ▶ To be able to draw and interpret directed graphs.
- ▶ To be able to use matrices to describe directed graphs.
- ▶ To be able to draw a directed graph from an adjacency matrix.

Directed graphs (digraphs)

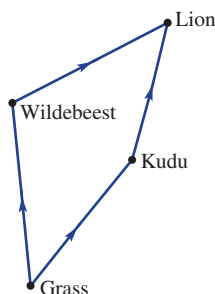
A **directed graph**, or **digraph**, is a graph where there is a direction associated with the edge. An edge of a directed graph is sometimes called an arc or a directed edge.

The directed graph below shows the food connections between some African animals and plants. The arrow on the directed edge (arc) between the vertex representing wildebeest and the vertex representing lion points towards the lion. This means that the lion eats the wildebeest.

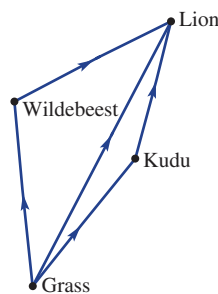


Subgraphs of directed graphs

A **subgraph** is part of a larger graph. All of the edges and vertices in the subgraph must exist in the original graph.



The directed graph above on the left shows only the Wildebeest, Lion, Kudu and Grass. It is a small part of the graph from above. All of the vertices and directed edges (arcs) exist in the original graph and so this is a subgraph of the original food connection graph.



The directed graph above on the right contains the same vertices, Wildebeest, Lion, Kudu and Grass, but there is an extra directed edge (arc) that shows the lion also eats grass. This edge (arc) was not in the original graph and so this is not a subgraph of the original food connection graph.

Adjacency matrices for directed graphs

An adjacency matrix can also be drawn to show the directed connections between vertices in a directed graph. The adjacency matrix for a directed graph may be, but is not necessarily, symmetric about the main diagonal.

The matrix labels for the rows of the matrix are the *origin* vertices for the arcs. The column labels in the matrix are the destination vertices. The arrow on the arc will point from the origin vertex (row) to the destination vertex (column).

In the graph for this adjacency matrix:

- there is one arc from vertex B to vertex D , shown by a '1' in row B , column D
- there is no arc from vertex D back to vertex B , shown by a '0' in row D , column B .

$$\begin{array}{c} \text{origin} \end{array} \begin{array}{ccccc} & A & B & C & D & E \\ \begin{array}{c} A \\ B \\ C \\ D \\ E \end{array} & \left[\begin{array}{ccccc} - & - & - & - & - \\ - & - & - & 1 & - \\ - & - & - & - & - \\ - & 0 & - & - & - \\ - & - & - & - & - \end{array} \right] \end{array}$$



This means the arc between vertex B and vertex D has a directional arrow from B to D .

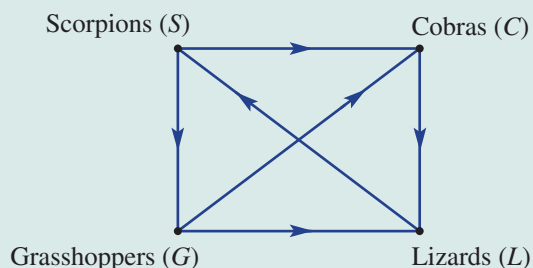


Example 5 Writing an adjacency matrix from a directed graph

A directed graph is used to represent the winners in a round robin sporting competition. Each of the teams in the competition are represented by a vertex and the arc represents the game between the two teams that it connects. The arrow on the arc points to the winner of the game.

Use the graph on the right to answer the following questions.

- Which team was the winner in the game between the Grasshoppers and the Cobras?
- Which is the only team to beat the Scorpions?
- Write down the adjacency matrix for this graph.



Solution

- The Cobras won the game against the Grasshoppers.
- The Lizards are the only team to beat the Scorpions.

$$\begin{array}{c} \mathbf{c} \end{array} \quad \begin{array}{c} S \quad G \quad C \quad L \\ \begin{array}{l} S \left[\begin{array}{cccc} 0 & 1 & 1 & 0 \end{array} \right] \\ G \left[\begin{array}{cccc} 0 & 0 & 1 & 1 \end{array} \right] \\ C \left[\begin{array}{cccc} 0 & 0 & 0 & 1 \end{array} \right] \\ L \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \end{array} \right] \end{array} \end{array}$$

Explanation

Look at the arc that joins the vertices for Grasshoppers and Cobras. It is pointing to the Cobras.

Look at all the arcs connected to the vertex for Scorpions. The only one pointing to Scorpions is from Lizards.

In the matrix, use a '1' for an arc that starts at the row team and points towards the column team. Otherwise, use a '0'.

Note: the vertices can be written in any order within the row headings and column headings of the matrix, but it is usual to keep the order consistent between the rows and columns.

Section Summary

- ▶ **Directed graphs** (digraphs) have a directional meaning associated with the edges.
- ▶ The edges of a directed graph are called **arcs** or **directed edges**.

An **adjacency matrix** is a square matrix that summarises the connections between vertices of a graph.

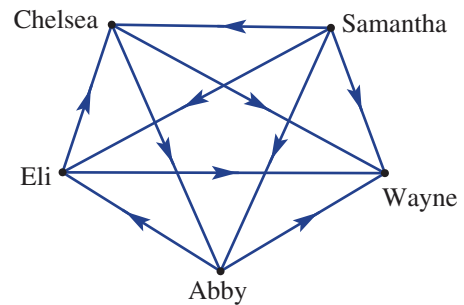
For directed graphs:

- ▶ the entry in row m and column n shows the number of arcs from the vertex in this row to the vertex in this column
- ▶ the adjacency matrix may be, but is not necessarily, symmetric about the main diagonal.

Loops are counted as one edge.

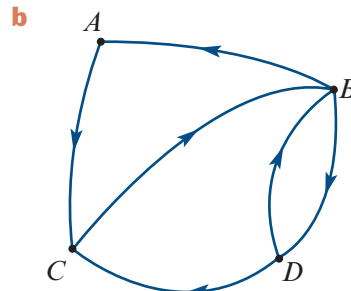
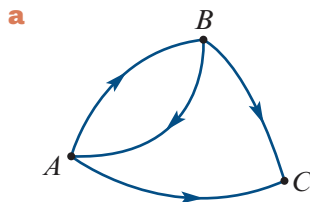
Exercise 9C

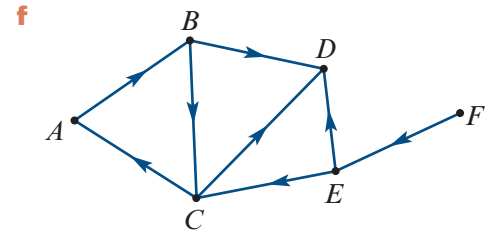
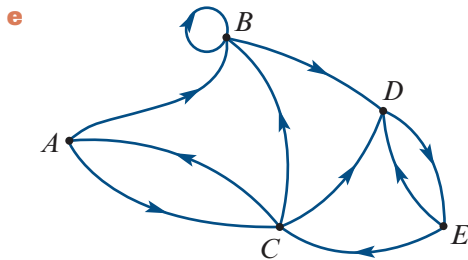
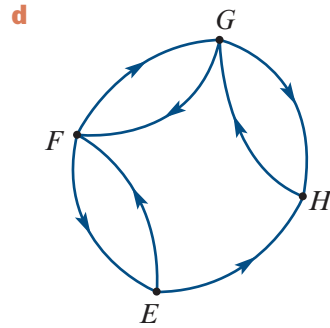
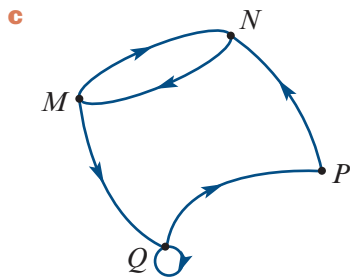
- 1** The directed graph on the right shows the results of a squash tournament. The vertices represent the people involved in the tournament and the arcs represent the games between the two people whose vertices are joined by it. The arrow on the arc points to the winner of the game.



- a** Who was the winner in the game between Chelsea and Samantha?
- b** Which player won all of their games?
- c** Which player lost all their games?
- d** Write down:
- $\deg(\text{Samantha})$
 - $\deg(\text{Eli})$
- 2** A national park in Africa contains a number of animal species. In the park:
- lions eat impala
 - leopards eat impala and warthogs
 - warthogs eat lizards
 - lizards eat flies
 - eagles eat lizards and small birds
 - small birds eat lizards and flies.
- a** Draw a directed graph to represent the information above. Use a vertex to represent each animal and an arc to represent the connection between them. The arrow of the arc should point to the animal that eats the other.
- b** Write down $\deg(\text{Warthogs})$.
- c** Draw a subgraph that contains only lizards, flies, small birds and eagles.
- 3** For each of the following directed graphs, write down the adjacency matrix.

Example 5





4 Draw a directed graph for the following adjacency matrices.

a

	A	B	C
A	0	1	1
B	0	1	0
C	1	1	1

b

	S	T	U
S	1	0	1
T	0	0	2
U	0	1	0

c

	P	Q	R	S
P	0	1	1	1
Q	1	0	2	0
R	1	0	0	1
S	0	0	1	0

d

	P	Q	R	S
P	1	1	0	1
Q	0	0	1	1
R	1	1	0	1
S	1	0	0	0

e

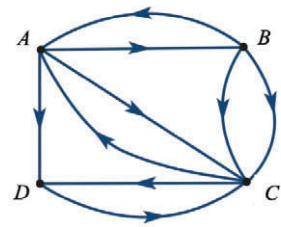
	P	Q	R	S	T
P	0	1	1	1	0
Q	0	0	1	0	0
R	0	0	0	1	1
S	0	0	0	0	0
T	0	0	1	1	0

f

	A	B	C	D	E
A	0	1	1	1	1
B	0	1	0	1	1
C	1	0	1	0	1
D	0	0	0	0	0
E	0	0	1	1	0

Paper 1-style multiple-choice questions

- 5 An adjacency matrix for the directed graph shown opposite is



A

	A	B	C	D
A	0	1	1	1
B	1	0	2	0
C	1	0	0	1
D	0	0	1	0

B

	A	B	C	D
A	0	0	1	1
B	1	0	2	0
C	1	0	0	1
D	0	0	1	1

C

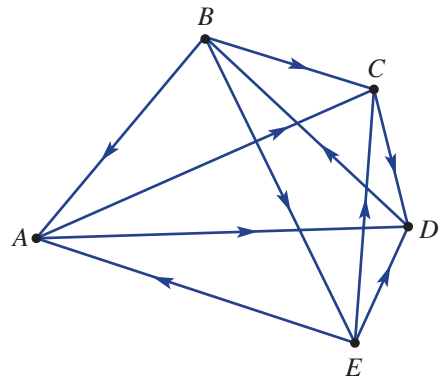
	A	B	C	D
A	0	0	1	1
B	0	1	1	0
C	1	0	0	1
D	0	0	1	0

D

	A	B	C	D
A	1	0	0	1
B	1	0	2	0
C	1	0	0	1
D	0	0	1	1

- 6 The directed graph shown opposite shows the results of a competition between five players A, B, C, D and E. For example it shows B defeats A and A defeats D. Two of the players have three victories; these are

- A** A and B
B D and E
C C and D
D B and E



- 7 A connected directed graph has 20 vertices. The number of entries in the corresponding adjacency matrix is

- A** 20
B 40
C 200
D 400

9D Planar graphs and Euler's formula

Learning intentions

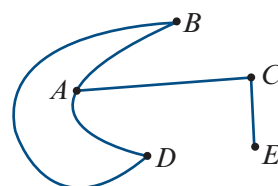
- ▶ To be able to define and identify a planar graph and its faces.
- ▶ To be able to apply Euler's formula and use it to verify if a graph is planar.

Planar graphs

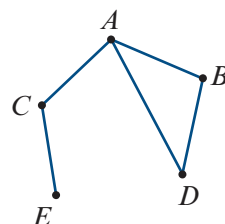
Equivalent graphs

All of the graphs shown in the diagram below contain the same information. For example, the edge between vertex E and C exists in all three of the graphs.

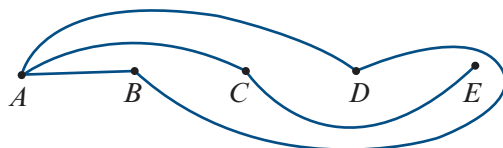
The physical location of the vertices and edges in the diagram is unimportant. As long as the connection information is represented accurately, the graph can be drawn with the vertices in any location.



The first of the graphs has some curved edges and the second has all straight edges. The third has the vertices arranged in a straight line.



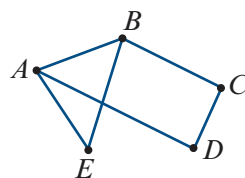
All of them, regardless of how they are drawn, contain exactly the same connections between vertices and so these graphs are considered to be *equivalent* to each other.



Equivalent graphs contain identical information and are sometimes called **isomorphic graphs**.

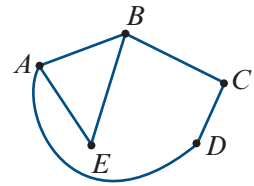
Planar graphs

The graph opposite has two edges that cross over each other (EB and AD). It is helpful to think of edges that cross like this as insulated electrical wires. It is quite safe to cross two insulated wires because the wires themselves never touch and never interfere with each other. We can think of crossing edges in a graph in a similar way.



The edges that cross over in this diagram are similar in that they do not intersect. It is important to note that there is no vertex at the point where these edges cross over.

Graphs with edges that cross in this way *may* be able to be redrawn so that the edges no longer cross. In this diagram the edge between vertices A and D has been moved, but none of the information in the graph has changed.



Planar Graph

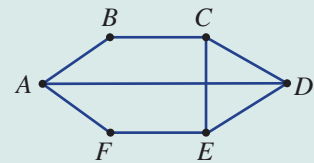
A graph is **planar** if it can be drawn in the plane in such a way that no two edges 'cross over'. If a graph is drawn so that no edges cross over, then it is said to be drawn in **planar form**.

If it is impossible to draw an equivalent graph without crossing edges, then that graph is called a non-planar graph. It is impossible to draw a non-planar graph in planar form.

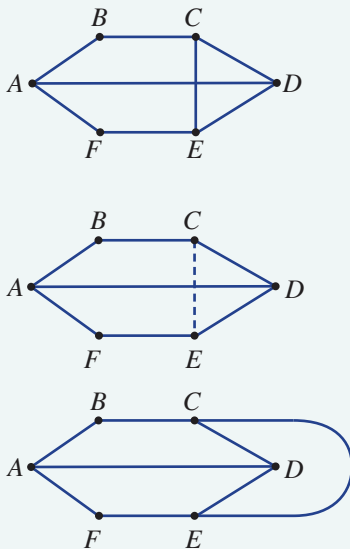


Example 6 Redrawing a graph in planar form

Show that this graph is planar by redrawing it so that no edges cross.



Solution



Explanation

Choose one of the edges that crosses over another edge. For example, choose the edge between vertex C and vertex E . Alternatively, the edge between vertex A and vertex D could be chosen.

Remove this edge temporarily from the graph.

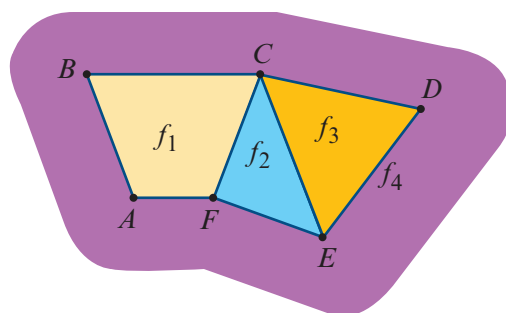
Redraw the edge between the same two vertices (C and E), but without the edge crossing any other edge.

Euler's formula

Leonard Euler (pronounced 'oiler') was one of the most prolific mathematicians of all time. He contributed to many areas of mathematics and his proof of the rule named after him is considered to be the beginning of the branch of mathematics called topology.

Faces

A planar graph defines separate regions of the paper it is drawn on. These regions, called **faces**, could be coloured in as you can see in the diagram shown here. There are three faces inside the graph, one coloured cream, one blue and one gold, but there is also a fourth face coloured purple that totally surrounds the graph.



In the diagram, the faces are labelled f_1 , f_2 , f_3 and f_4 .

A face is an area in a graph that can only be reached by crossing an edge.

Euler's formula

For any connected planar graph, we can count the number of vertices (v), the number of faces (f) and the number of edges (e). There is a relationship between these numbers, called **Euler's formula**.

In words: 'the number of vertices + the number of faces – the number of edges = 2'

In symbols: $v + f - e = 2$

Euler's formula

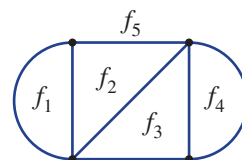
For any connected planar graph:

$$v + f - e = 2$$

where v is the number of vertices, f is the number of faces and e is the number of edges in the graph.

The connected planar graph on the right has four vertices, seven edges and five faces. Therefore $v = 4$, $e = 7$ and $f = 5$. We can verify that

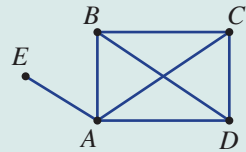
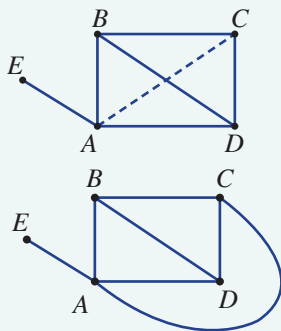
$$\begin{aligned} v - e + f &= 4 - 7 + 5 \\ &= 2 \end{aligned}$$



**Example 7** Verifying Euler's formula

For the graph shown on the right:

- redraw the graph in planar form
- verify Euler's formula.

**Solution****a**

- In this graph, there are: five vertices, four faces and seven edges.
 $v + f - e = 5 + 4 - 7 = 2$
 Euler's formula is verified.

Explanation

Temporarily remove an edge that crosses another edge and redraw it so that it does not cross another edge.

Count the number of vertices, faces and edges in the graph.
 Substitute into Euler's formula to verify.

**Example 8** Using Euler's formula

A connected planar graph has six vertices and nine edges. How many faces does this graph have? Draw a connected planar graph with six vertices and nine edges.

Solution

$$v = 6 \text{ and } e = 9$$

$$v + f - e = 2$$

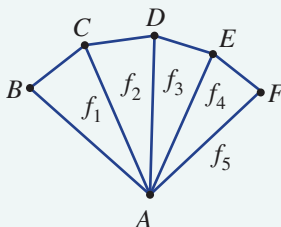
$$6 + f - 9 = 2$$

$$f - 3 = 2$$

$$f = 2 + 3$$

$$f = 5$$

This graph has five faces.

**Explanation**

Write down the known values.

Substitute into Euler's formula and solve for the unknown value.

Note: There are other possible graphs.

Section Summary

► A graph is **planar** if it can be drawn in the plane in such a way that no two edges 'cross over'. If a graph is drawn so that no edges cross over, then it is said to be drawn in **planar form**.

► For any connected planar graph:

$$v + f - e = 2$$

where v is the number of vertices, f is the number of faces and e is the number of edges in the graph.

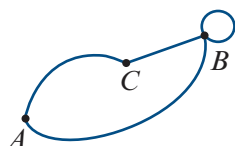


Exercise 9D

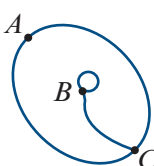
Equivalent (isomorphic) graphs

1 In each question below, three of the graphs are equivalent (isomorphic) and the fourth is not. Identify the graph which is not equivalent (isomorphic) to the others.

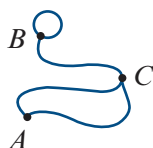
a i



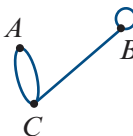
ii



iii

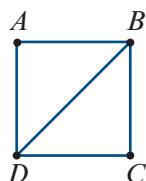


iv

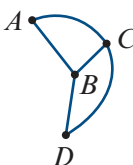


b

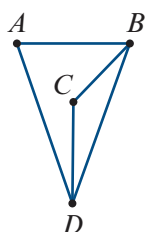
i



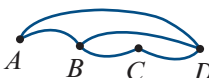
ii



iii

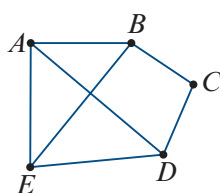


iv

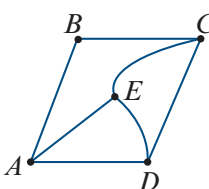


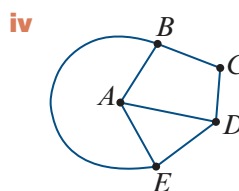
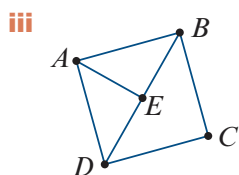
c

i



ii

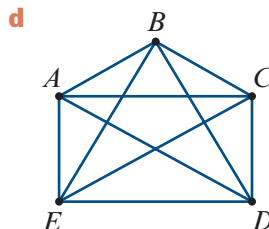
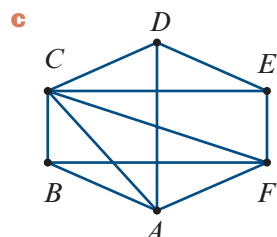
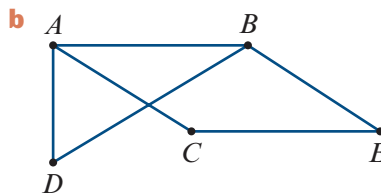
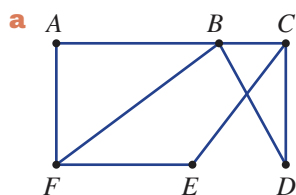




Drawing graphs in planar form

Example 6

- 2 Where possible, show that the following graphs are planar by redrawing them in a suitable planar form.

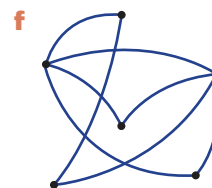
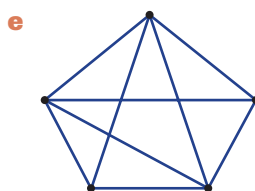
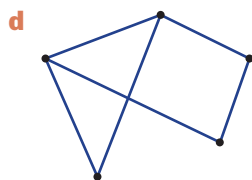
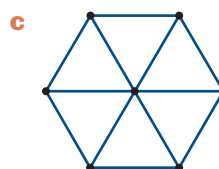
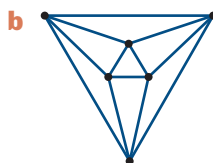
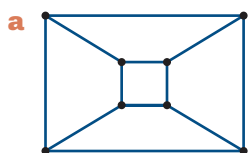


Euler's formula

Example 7

- 3 For each of the following graphs:

- state the values of v , e and f
- verify Euler's formula.



Example 8

- 4 For a planar connected graph, determine:

- f , if $v = 8$ and $e = 10$
- v , if $e = 14$ and $f = 4$
- f , if $v = 5$ and $e = 14$
- e , if $v = 10$ and $f = 11$

Properties of graphs

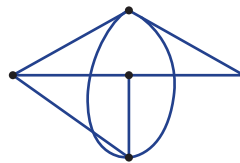
- 5** A connected planar graph has eight vertices and thirteen edges. Determine the number of faces of this graph.

CE

Paper 1-style multiple-choice questions

- 6** A planar graph has four faces. The graph could have
- | | |
|---|--|
| A Seven vertices and seven edges | B Seven vertices and four edges |
| C Seven vertices and five edges | D Five vertices and seven edges |

Use the following information for Questions 7 and 8.



- 7** The number of faces in the graph above is
- | | | | |
|------------|------------|------------|------------|
| A 3 | B 4 | C 5 | D 6 |
|------------|------------|------------|------------|
- 8** Consider the following five statements about the graph above:
- | | |
|---------------------------|---|
| ■ The graph is planar. | ■ The graph contains a bridge. |
| ■ It is a simple graph. | ■ The sum of degrees of the vertices is 16. |
| ■ It is a complete graph. | |
- How many of these statements are true?
- | | | | |
|------------|------------|------------|------------|
| A 1 | B 2 | C 3 | D 4 |
|------------|------------|------------|------------|
- 9** A connected planar graph has 6 vertices and 8 edges. The number of faces is
- | | | | |
|------------|------------|------------|------------|
| A 2 | B 3 | C 4 | D 5 |
|------------|------------|------------|------------|
- 10** The complete graph with 4 vertices is a connected planar graph. The number of faces this graph has is
- | | | | |
|------------|------------|------------|------------|
| A 5 | B 6 | C 7 | D 4 |
|------------|------------|------------|------------|

9E Exploring a graph

Learning intentions

- ▶ To be able to define, identify and work with walks, trails and paths.
- ▶ To be able to define, identify and work with closed and open walks, trails and paths.

Travelling

Graphs can be used to model and analyse problems involving exploring and **travelling**. These problems include minimising the distance travelled or time taken between different locations using different routes. For example, a courier driver would like to know the shortest route to use for deliveries, and a tour guide would like to know the quickest route that allows tourists to see a number of sights without retracing their steps.

To solve these types of problems, you will need to learn the language we use to describe the different ways of navigating through a graph, from one vertex to another.

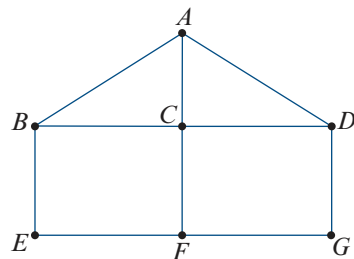


Travelling through a network: Watch the video in the Interactive Textbook to see the five types of routes that can be travelled through networks.

Walks, trails, paths, circuits and cycles

The different ways of navigating through graphs, from one vertex to another, are described as **walks, trails, paths, circuits** and **cycles**.

The graph opposite will be used to explain and define each of these terms.



Walks

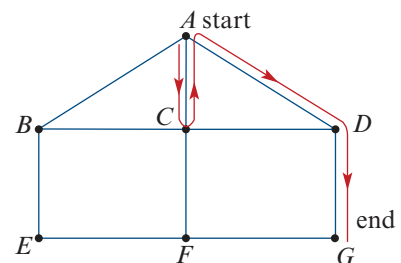
In a graph, a **walk** is sequence of vertices such that from each vertex there is an edge to the next vertex in the sequence.

A walk that starts and finishes at different vertices is said to be an **open walk**.

The red line in the graph opposite traces out an open walk. This walk can be written down by listing the vertices in the order they are visited: $A - C - A - D - G$.

A walk that starts and finishes at the same vertex is said to be **closed walk**

The **length** of a walk is the number of edges in the walk. The walk $A - C - A - D - G$ has length 4.

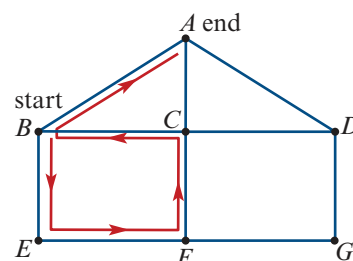


Trails and circuits

A **trail** is a walk with no repeated edges.

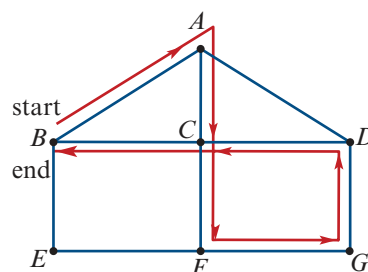
An **open trail** is a trail that starts and finishes at different vertices.

The red line in the graph opposite traces out an open trail. This trail can be written down by listing the vertices in the order they are visited: $B - E - F - C - B - A$.



Note: There are *no repeated edges* in this trail, but one vertex (B) is repeated.

A **closed trail or circuit** is a trail that starts and finishes at the same vertex. The red line in the graph opposite traces out a closed trail (circuit). This closed trail can be written down by listing the vertices in the order they are visited: $A - C - F - G - D - C - B - A$.



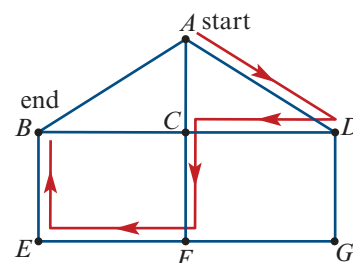
Note: There are *no repeated edges* in this circuit, but one vertex, C , is repeated. The start and end vertices are repeated. This is a circuit.

Paths and cycles

A **path** is a walk with no repeated edges and no repeated vertices.

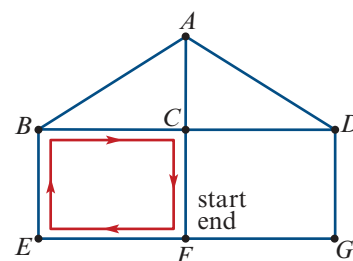
An **open path** is a path that starts and finishes at different vertices.

The red line in the graph opposite traces out an open path. This path can be written down by listing the vertices in the order they are visited: $A - D - C - F - E - B$.



A **closed path or cycle** is a path that starts and finishes at the same vertex. The start and end vertex is an exception to the 'no repeated vertices' condition.

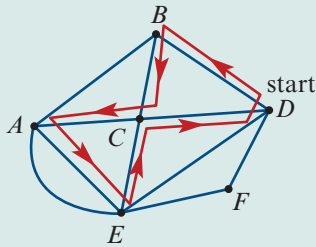
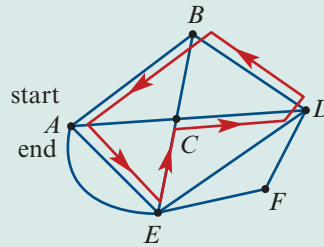
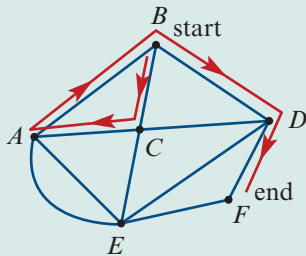
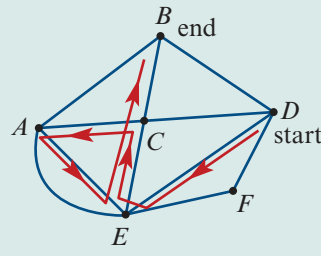
The red line in the graph opposite traces out a closed path (cycle). This closed path can be written down by listing the vertices in the order they are visited: $F - E - B - C - F$.



Note: There are *no repeated edges* and *no repeated vertices* in this closed graph, except for the start and end vertices.


Example 9 Identifying types of walks

Identify the walk in each of graphs below as a trail, path, circuit, cycle or walk only.

a

b

c

d

Solution

- a** This walk starts and ends at the same vertex so it is either a circuit or a cycle. The walk passes through vertex C twice without repeated edges, so it must be a circuit.
- b** This walk starts and ends at the same vertex so it is either a circuit or a cycle. The walk has no repeated vertex or edge so it is a cycle.
- c** This walk starts at one vertex and ends at a different vertex, so it is not a circuit or a cycle. There is one repeated vertex (B) and no repeated edge, so it must be a trail.
- d** This walk starts at one vertex and ends at a different vertex, so it is not a circuit or a cycle. There are repeated vertices (C and E) and repeated edges (the edge between C and E), so it must be a walk only.

Section Summary

Walks, trails, paths, circuits and cycles

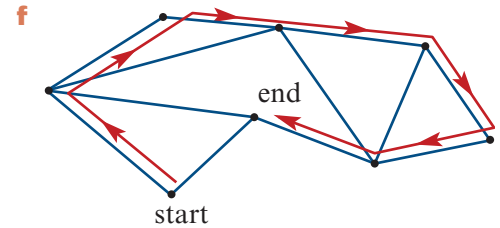
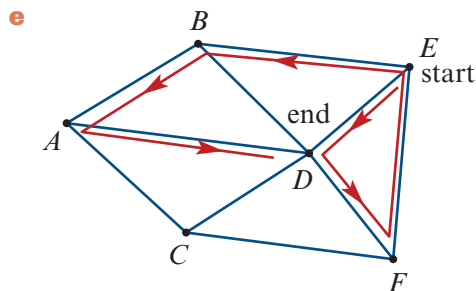
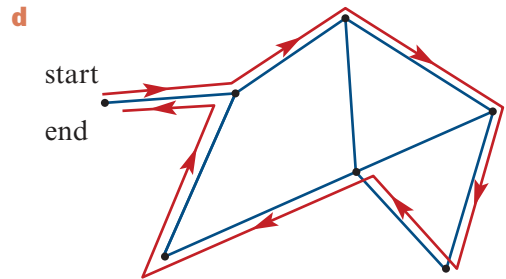
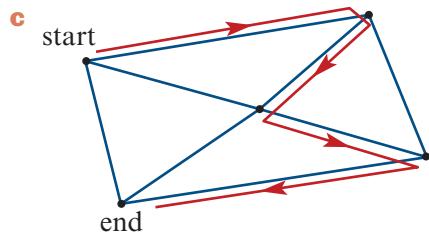
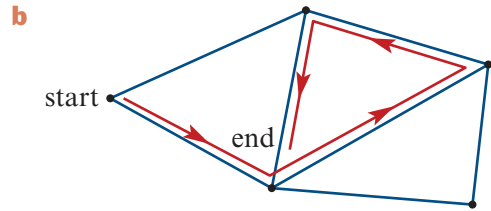
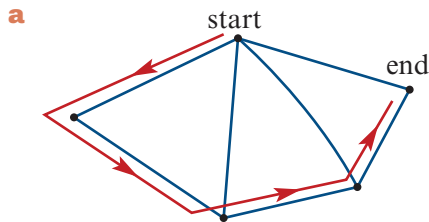
- ▶ A **walk** is a sequence of edges, linking successive vertices in a graph.
- ▶ A **trail** is a walk with no repeated edges.
 - ▷ An **open trail** is a trail that starts and finishes at different vertices.
 - ▷ A **closed trail** or **circuit** is a trail that starts and finishes at the same vertex.
- ▶ A **path** is a walk with no repeated edges and no repeated vertices.
 - ▷ An **open path** is a path that starts and finishes at different vertices.
 - ▷ A **closed path** or **cycle** is a path that starts and finishes at the same vertex.

Exercise 9E

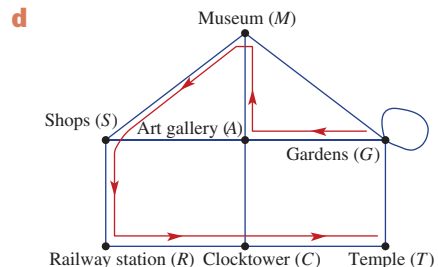
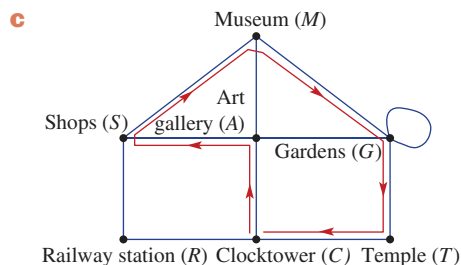
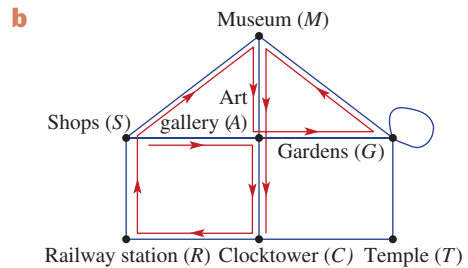
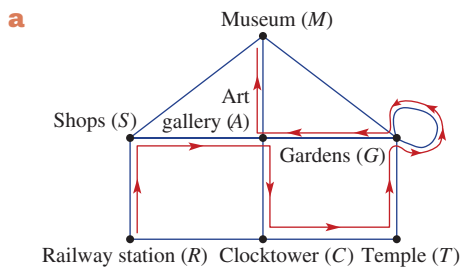
Describing walks through a graph

Example 9

- 1 Describe the walk shown in each of the following graphs as a trail, path, closed trail, cycle, open walk only or closed walk only.

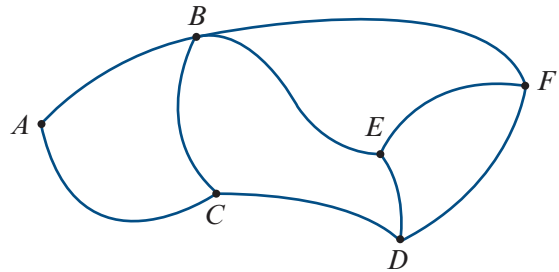


- 2 Identify the walk shown in each of the graphs below as a trail, a path or a walk only.



- 3** Use the graph below to describe the walks below as a trail, path, closed path, closed trail, cycle, open walk only or closed walk only.

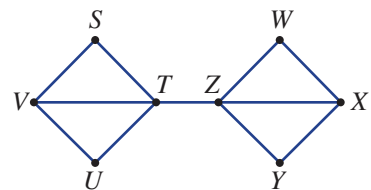
- a** $B - E - D - C - A - B$
b $A - B - E - B - C - D - E$
c $F - E - B - F - E - D$
d $A - B - C - D - E - B - F$
e $C - A - B - C - D - E - B - C$
f $D - E - B - A - C - D$



Paper 1-style multiple-choice questions

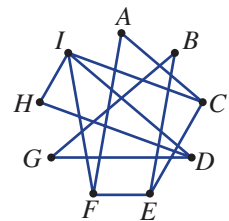
- 4** There is one path of length three from vertex S to vertex Y . This is $S - T - Z - Y$. There are three edges in the path. The number of paths of length five is

- A** 1
B 2
C 3
D 4



- 5** Which one of the following walks is an open path of the graph.

- A** $A - C - E - B - G - D - I - C - A$
B $A - C - E - B - G$
C $A - C - I - F - A$
D $C - E - B - G - D - I - C - A$



9F Eulerian graphs and applications

Learning intentions

- ▶ To be able to identify a walk as an Eulerian trail or Eulerian circuit.
- ▶ To be able to use the degrees of the vertices to identify if an Eulerian trail or circuit is possible.

The explorer problem

Consider a national park that contains a number of campsites and some roads that lead between them. A park ranger might need to travel along each of these roads in order to check their condition. It is in the best interests of the park ranger to minimise the number of times each road is travelled and to avoid backtracking wherever possible. Ideally, each road should be travelled only once.

This situation is an example of the *explorer problem*. The explorer problem asks the question ‘Is it possible to travel every edge in a network only once?’

Investigation 9F: The explorer problem

Whether or not the explorer problem has a solution for a particular graph depends upon the degrees of the vertices in that graph.

We note that in the section we are interested in trails.

Eulerian graphs and Eulerian trails

In Investigation 9F above you may have noticed that if the degree of every vertex in the graph was even, then the ranger could begin at any campsite, travel along every road only once, and would end up back at the starting campsite. This is a closed trail (circuit) that involves every edge of the graph. Such closed trails have a special name.

Eulerian graphs and Eulerian trails

A connected graph is **Eulerian** if it has a closed trail (starts and ends at the same vertex) and includes every edge. Such a trail is called an **Eulerian trail** or **Eulerian circuit**.

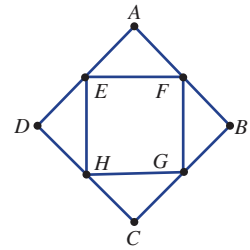
We note:

- an Eulerian trail may include repeated vertices
- an Eulerian graph has every vertex of even degree, and conversely
- a connected graph for which every vertex is of even degree is an Eulerian graph.

The graph shown here is an Eulerian Graph. Vertices A, B, C and D have degree 2 and vertices E, F, G and H have degree 4.

One Eulerian trail (circuit) is:

$A - F - B - G - C - H - D - E - F - G - H - E - A$



Eulerian trails and circuits: Watch the video in the Interactive Textbook to see them in action.

Semi-Eulerian graphs and semi-Eulerian trails

If there were two odd vertices in the graphs from Investigation 9F, the ranger could start at either one of these campsites, travel along every road only once, and would end up at the other odd-degree campsite. Such trails have a special name.

Semi-Eulerian graph and semi-Eulerian trail

A connected graph is **semi-Eulerian** if it has an open trail that includes every edge. Such a trail is called a **semi-Eulerian trail**.

We note:

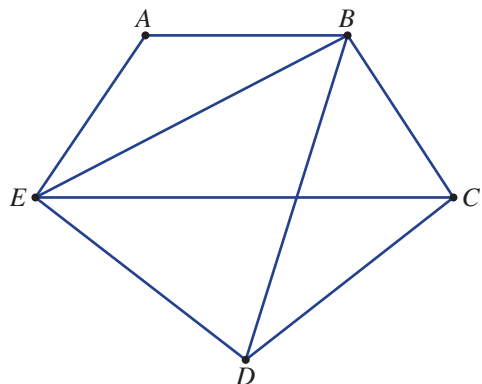
- a semi-Eulerian trail may include repeated vertices
- a semi-Eulerian graph has exactly two vertices with odd degree, and conversely
- a connected graph with exactly two vertices of odd degree is a semi-Eulerian graph
- in a semi-Eulerian graph the semi-Eulerian trail must start at one of the vertices of odd degree and finish at the other vertex of odd degree.

The graph shown here is a semi-Eulerian Graph. Vertex A has degree 2 and vertices B and E have degree 4.

Vertices C and D have degree 3. One semi-Eulerian trail (circuit) is:

$D - E - A - B - E - C - D - B - C$

Note: If another edge $C - D$ is added to the graph, then all vertices are even and the graph is Eulerian.



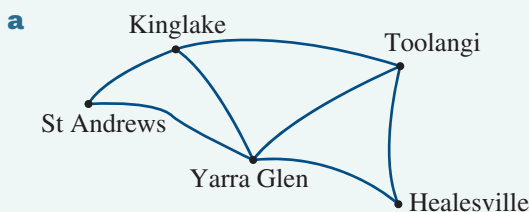

Example 10 The explorer problem – Eulerian graphs

A map showing the towns of St Andrews, Kinglake, Yarra Glen, Toolangi and Healesville is shown below.



Map data © 2025 Google

- a** Draw a graph with a vertex representing each of these towns and with edges representing the direct road connections between towns. Ignore any towns on the map not listed in the question.
- b** Explain why this graph is semi-Eulerian and not Eulerian.
- c i** Write down an open trail that begins at Toolangi and follows every edge only once.
- ii** Explain how you could tell this trail would end at Kinglake.

Solution


- b** Graphs that have exactly two vertices with odd degrees are semi-Eulerian.
- c i** Toolangi – Healesville – Yarra Glen
– Toolangi – Kinglake – Yarra Glen
– St Andrews – Kinglake

Explanation

A road connection exists between:

- St Andrews and Kinglake
- St Andrews and Yarra Glen
- Kinglake and Yarra Glen
- Kinglake and Toolangi
- Yarra Glen and Toolangi
- Yarra Glen and Healesville
- Healesville and Toolangi.

The graph has exactly two vertices with odd degrees (Kinglake and Toolangi).

This question has many answers, one of which is shown.

- ii An open trail through a semi-Eulerian graph will start at one of the vertices with odd degree and end at the other. Toolangi was one of the vertices with odd degree and Kinglake was the other, so the open trail must end at Kinglake.

Section Summary

- ▶ A connected graph is **Eulerian** if it has a closed trail (starts and ends at the same vertex) and includes every edge. Such a trail is called an **Eulerian trail** or **Eulerian circuit**.
- ▶ A connected graph is **semi-Eulerian** if it has an open trail that includes every edge. Such a trail is called a **semi-Eulerian trail**.
- ▶ An Eulerian trail may include repeated vertices.
- ▶ An Eulerian graph has every vertex of even degree, and conversely
- ▶ A connected graph for which every vertex is of even degree is an Eulerian graph.
- ▶ A semi-Eulerian graph has exactly two vertices with odd degree, and conversely
- ▶ A connected graph with exactly two vertices of odd degree is a semi-Eulerian graph.

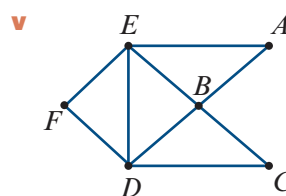
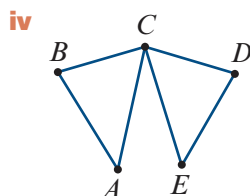
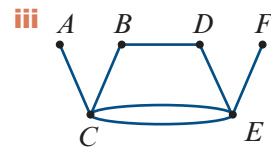
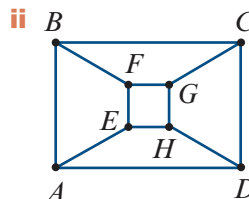
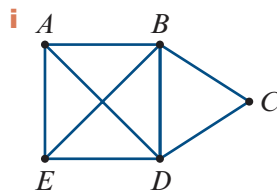


Exercise 9F

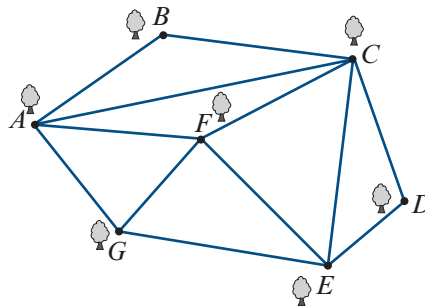
Eulerian and semi-Eulerian graphs

Example 10

- 1 For each of the graphs shown below:
- a identify whether the graph is Eulerian, semi-Eulerian or neither
 - b name any open or closed Eulerian trails found.

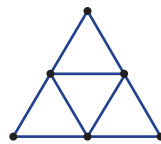


- 2 A housing estate has large open parklands that contain seven large trees. The trees are denoted as vertices A to G on the graph below.

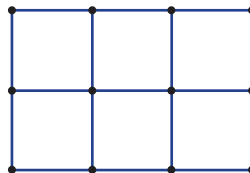


Walking tracks link the trees and are shown as edges on the graph.

- a Determine the degree of each of the vertices in the graph.
 - b One day, Jamie decided to go for a walk that will take him along each path only once at most.
 - i At which vertices could Jamie start?
 - ii At which vertex will Jamie end?
 - c A new track is to be made between two trees. This track will mean that Jamie could start at any vertex, walk along each of the tracks only once and return to his starting point.
 - i Between which two vertices should the new track be made?
 - ii What is the name of the walk that Jamie will follow once the new track is made?
- 3 The graph below consists of $1 + 2 + 3 = 6$ vertices arranged in a grid to give the **triangular grid graph**, T_3 .



- a Draw the triangular grid graph T_4 .
 - b Briefly explain why T_n has an Euler circuit for all $n \geq 1$.
- 4 The 3×4 vertices are arranged in an array below to create the 3×4 **grid graph**.

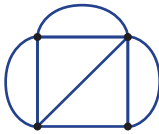
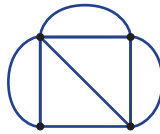
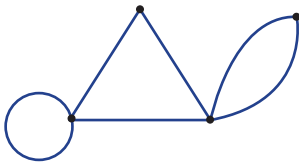
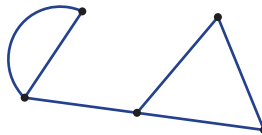


For what values of m and n will an $m \times n$ grid graph have:

- a a semi-Eulerian trail?
- b an Euler circuit?

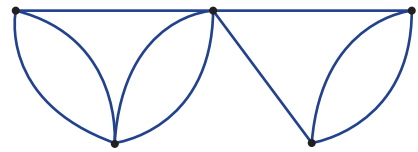
Paper 1-style multiple-choice questions

- 5 Which one of the following graphs has an Eulerian circuit?

A**B****C****D**

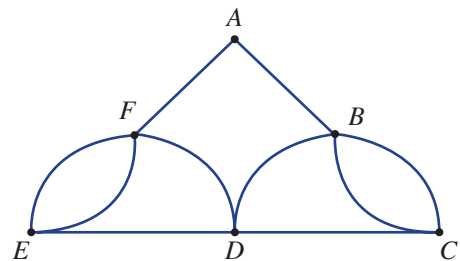
- 6 A semi-Eulerian trail for the graph opposite will be possible if only one edge is removed. In how many different ways could this be done?

A 2 **B** 3 **C** 4 **D** 5



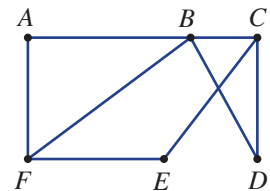
- 7 The graph will have an Eulerian circuit if an edge could be added between the vertices

A E and C
B A and B
C A and F
D A and D



- 8 The graph opposite has a trail. An edge can be added between two of its vertices so that it is an Eulerian circuit. These two vertices are

A A and D **B** B and D
C E and A **D** C and F



- 9 Shown below are adjacency matrices for five different graphs. Which of these graphs will *not* have an Eulerian circuit?

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Graph A

$$\begin{bmatrix} 0 & 3 & 3 \\ 3 & 0 & 1 \\ 3 & 1 & 0 \end{bmatrix}$$

Graph B

$$\begin{bmatrix} 0 & 3 & 1 \\ 3 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Graph C

$$\begin{bmatrix} 0 & 2 & 1 \\ 2 & 0 & 3 \\ 1 & 3 & 0 \end{bmatrix}$$

Graph D

A Graph A **B** Graph B **C** Graph C **D** Graph D

9G Hamiltonian graphs and applications

Learning intentions

- To be able to identify a walk as a Hamiltonian path or Hamiltonian cycle.

The traveller problem

You have seen an application of graphs that involved tourists visiting a number of sites of interest. A traveller in such a situation would be much more interested in seeing the sites at the vertices of the graph than the roads that connect them. *The traveller problem* involves a situation where every vertex in a graph is visited once.

Unlike the explorer problem, the traveller problem does not have a set of rules to define whether it has a solution. Rather, finding a route that allows a traveller to visit all vertices in a graph relies purely on inspection.

Investigation 9G: The traveller problem

Hamiltonian graphs

In 1857, Hamilton's puzzle in Investigation 9G above was commercially produced as a board game called 'The Icosian Game'. The object of the game was to find a route through the graph that:

- starts and ends at the same vertex
- visits all vertices exactly once (except for the start and end vertex).



© The Puzzle Museum – J. Dalgety 2019

The game consisted of a wooden board with holes at vertices. Numbered ivory plugs were put into the board to mark out the route. There are only four of these games known to still exist.

The solution to Hamilton's puzzle is a path because it has no multiple edges and no multiple vertices. The solution is also a cycle, because it starts and ends at the same vertex.

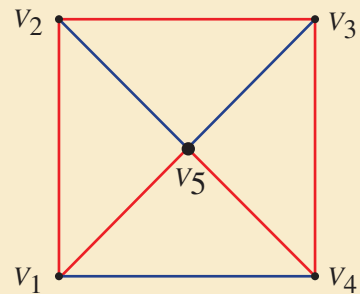
A **Hamiltonian graph** is a connected graph that has a cycle (closed path) that includes every vertex only once. This cycle is called a **Hamiltonian cycle**.

A **semi-Hamiltonian graph** is a connected graph that has an open path that includes every vertex only once. This path is called a **Hamiltonian path**.

Hamiltonian and semi-Hamiltonian graphs

Hamiltonian graphs:

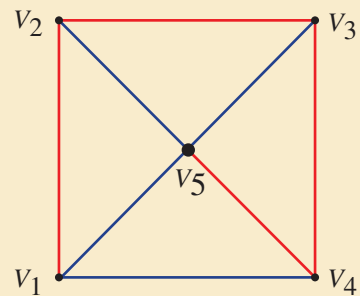
- contain a cycle (closed path) called a **Hamiltonian cycle** that starts and finishes at the same vertex and visits every other vertex exactly once
- allow travellers to begin at any vertex, visit every vertex only once and return to the starting vertex.



The path $(v_1 - v_2 - v_3 - v_4 - v_5 - v_1)$ is a Hamiltonian cycle.

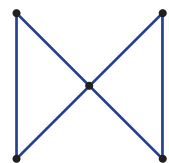
Semi-Hamiltonian graphs:

- contain an open path called a **Hamiltonian path** that involves every vertex of the graph
- allow travellers to begin at one vertex and to visit every other vertex only once but not return to the starting vertex.

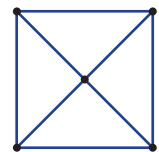


The path $(v_1 - v_2 - v_3 - v_4 - v_5)$ is a Hamiltonian path.

Notes: 1 A graph can have an Euler circuit, but no Hamiltonian cycle. This can be seen in the graph to the right. It has all vertices with even degree.



2 A graph can have a Hamiltonian cycle, but no Euler circuit. This can be seen in the graph to the right. It has vertices with odd degree.



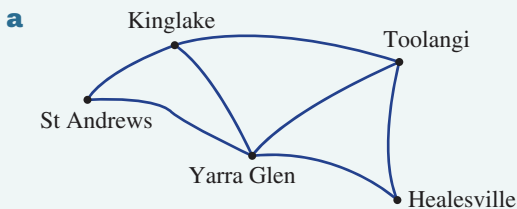

Example 11 Solving a traveller problem

A map showing the towns of St Andrews, Kinglake, Yarra Glen, Toolangi and Healesville is shown below.



Map data © 2025 Google

- a** Draw a graph with a vertex representing each of the towns and edges representing the direct road connections between the towns. Ignore any towns on the map not listed in the question.
- b** Explain why this graph is Hamiltonian.
- c** Write down one:
 - i** Hamiltonian path
 - ii** Hamiltonian cycle that begins at Healesville.

Solution


- b** It is possible to visit every town exactly once and return to the starting town, so the graph is Hamiltonian.

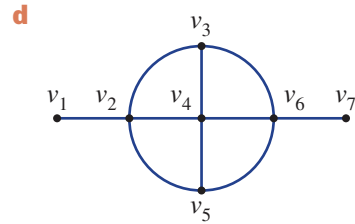
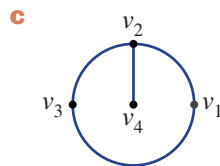
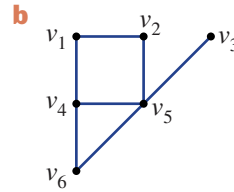
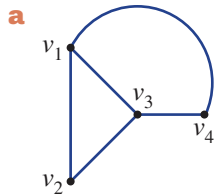
Explanation

A road connection exists between:

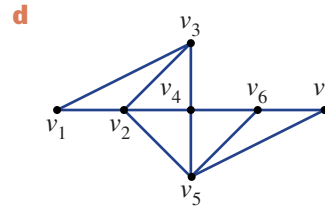
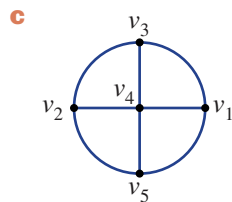
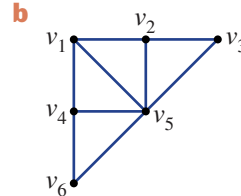
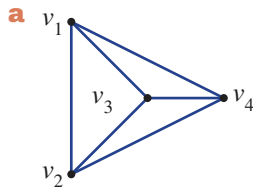
- St Andrews and Kinglake
- St Andrews and Yarra Glen
- Kinglake and Yarra Glen
- Kinglake and Toolangi
- Yarra Glen and Toolangi
- Yarra Glen and Healesville
- Healesville and Toolangi.

Can a closed walk be found that visits every town?

2 List a Hamiltonian path for each of the following graphs:

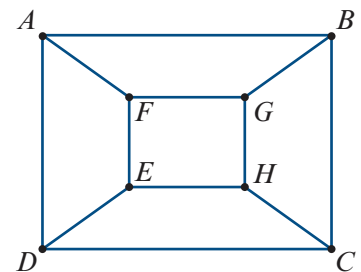


3 List a Hamiltonian cycle starting from v_1 for each of the following graphs:

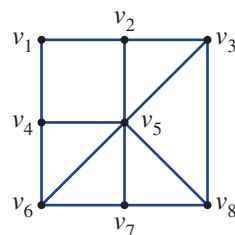


4 List a Hamiltonian path for this graph:

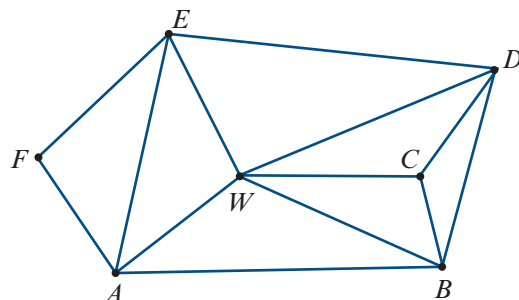
- a** starting at vertex A and finishing at vertex D
b starting at vertex F and finishing at vertex G .



- 5** Eight towns are represented using vertices $v_1, v_2, \dots, v_8$. The roads that connect these towns are represented as edges. Starting and returning at v_1 , how can a salesperson visit every town exactly once? Illustrate a path on the graph.



- 6** The network diagram opposite shows the location of a warehouse at vertex W . This warehouse supplies equipment to six factories: A, B, C, D, E and F .

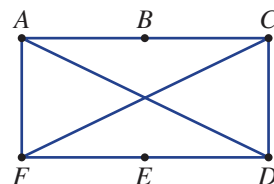


- a** What is the degree of vertex W ?
- b** A salesperson plans to leave factory E , first visit the warehouse, W , and then visit every other factory. They will visit each location only once and will not return to factory E .
- i** Write down the mathematical term used to describe the planned route.
- ii** Write down an order in which the salesperson can visit the factories.

Paper 1-style multiple-choice questions

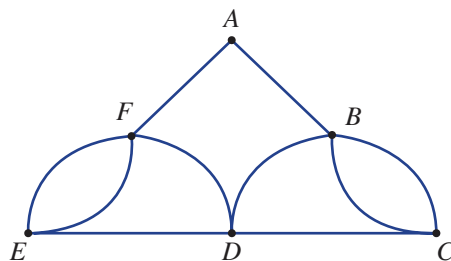
- 7** Consider the graph opposite. Which one of the following is **not** a Hamiltonian cycle for this graph?

- A** $ABCFEDA$ **B** $BADEFCEB$
C $CDEFABC$ **D** $DEFACBD$



- 8** Which one of the following is a Hamiltonian path for the graph shown opposite?

- A** $AFEDBA$ **B** $ABCDEF$
C $AFEDBCBA$ **D** $ABDFEDCBA$



9H Weighted graphs, networks and shortest path problems

Learning intentions

- To be able to work with weighted graphs to help solve weighted graph problems.

Weighted graphs

The edges of a graph represent the connections between the vertices of that graph. Sometimes there is more information known about that connection. For example, if the edge of a graph represents a road between two towns, the length of the road, or perhaps the time it takes to travel that road, might be known.

Extra numerical information about the edge can be written next to the edge in a graph. Graphs that have numerical information on each edge are called **weighted graphs** and the numbers themselves are called the **weights** of the edges. Weighted graphs in which the weights are physical quantities, such as distance, time or cost are called **networks**.

The **total weight** of a walk through a network is the sum of all the weights for the edges that are travelled in that walk.

The travelling salesperson problem

The travelling salesperson problem is solved using a weighted graph.

The solution to the travelling salesperson problem for a particular network is the Hamiltonian cycle in the network that has the smallest total weight. While there are some algorithms (mathematical procedures) that can be used to find this cycle, you will solve travelling salesperson problems, and shortest path problems, using observation and trial-and-error only.

Shortest path problems

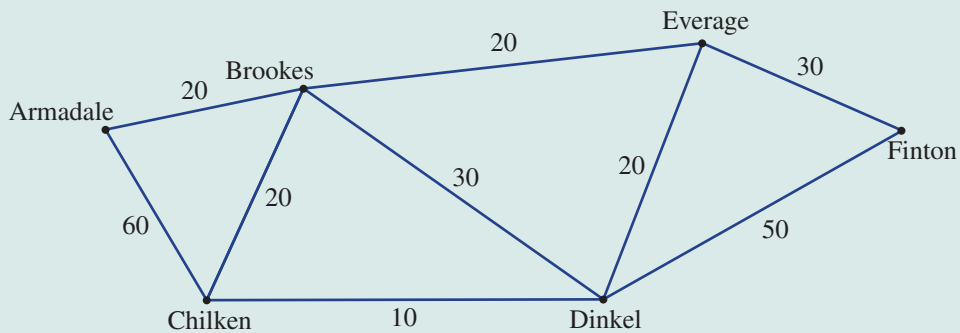
Shortest path problems involve finding the shortest path from one vertex to another. This path does not have to be Hamiltonian; that is, it does not need to visit every vertex in the network. The length of the shortest path will be the sum of the weights for every edge that the path covers.

Shortest path problems are easy to solve by inspection if the network has a small number of vertices. Some edges may have very large weights compared to others and are probably best avoided, while other edges may have very small weights compared to others. It may be useful to try and incorporate these small weights into any shortest path.

Careful inspection of the network can help rule out any routes that are unlikely to contribute to the shortest path overall.


Example 12 Finding a shortest path

The graph below shows six towns represented by vertices and the roads between those towns represented by edges.



The weights on each of the edges show the travel times in minutes between each town.

Determine the total time it takes to travel the shortest path from Armadale to Finton.

Solution

The shortest path from A to B is 20 (direct).
 The shortest path from A to C is 40 (via B).
 The shortest path from A to D is 50 (via B).
 The shortest path from A to E is 40 (via B).
 The shortest path from A to F is 70 (via B and E).

The shortest path from A to F is
 $A - B - E - F$.

$$\begin{aligned}\text{Total time} &= 20 + 20 + 30 \\ &= 70 \text{ minutes}\end{aligned}$$

The shortest time to travel from Armadale to Finton is 70 minutes, via Brookess and Everage.

Explanation

Consider smaller sections of the network, gradually moving through the network until the ending vertex is reached.

Write down the shortest path.

Add the weights to find the total time.

Write your answer.

Section Summary

Extra numerical information about the edge can be written next to the edge in a graph.

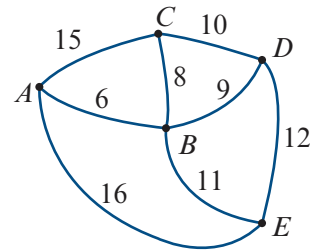
- ▶ Graphs that have numerical information on each edge are called **weighted graphs** and the numbers themselves are called the **weights** of the edges.
- ▶ Weighted graphs in which the weights are physical quantities, such as distance, time or cost are called **networks**.
- ▶ The **total weight** of a walk through a network is the sum of all the weights for the edges that are travelled in that walk.

Exercise 9H

Weighted graphs and networks

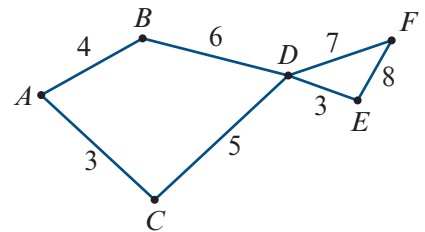
Example 12

- 1** The graph on the right shows towns A, B, C, D and E represented by vertices. The edges represent road connections between the towns and the weights on these edges are the average time, in minutes, it takes to travel along each road.
- Which two towns are 12 minutes apart by road?
 - How long will it take to drive from C to D via B ?
 - A motorist intends to drive from D to E via B . How much time will they save if they travel directly from D to E ?
 - Determine the shortest time it would take to start at A , finish at E and visit every town exactly once.

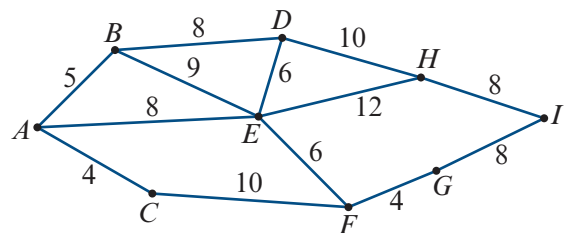


Shortest path problems

- 2** Determine the length of the shortest path from A to E in the graph on the right.

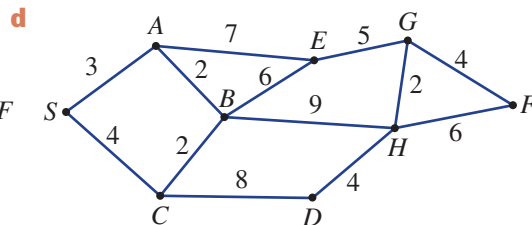
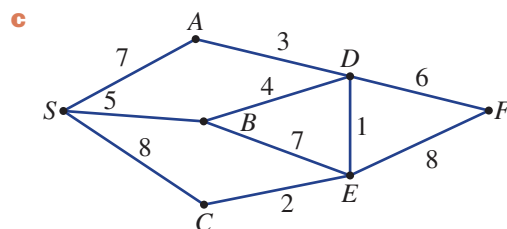
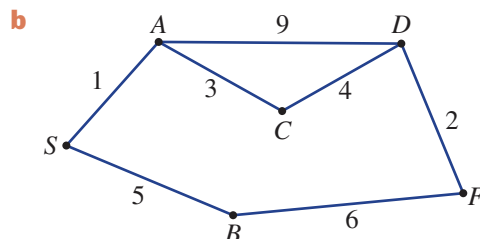
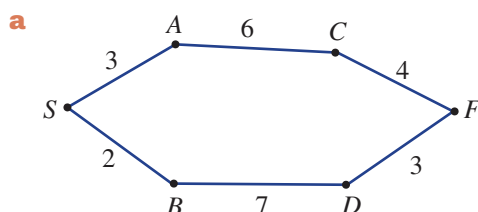


- 3** The network on the right shows the distance, in kilometres, along walkways that connect the landmarks A, B, C, D, E, F, G, H and I in a national park.

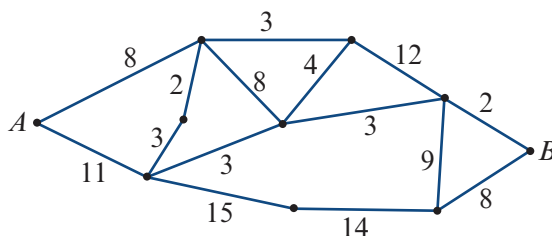


- What distance is travelled on the path $A - B - E - H - I$?
- What distance is travelled on the closed trail $F - E - D - H - E - A - C - F$?
- Determine the shortest path from A to I .

- 4 Determine the shortest path from S to F in the following networks. Write down the length of the shortest path.



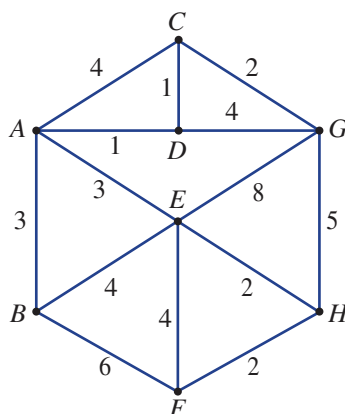
- 5 In the network on the right, the vertices represent small towns and the edges represent roads. The weights on the edges indicate the distance, in kilometres, between towns.



Determine the length of the shortest path between towns A and B.

Paper 1-style multiple-choice questions

Use the following information to answer Questions 6 and 7.



- 6 What is the shortest path between C and E?

A C – A – E

B C – D – A – E

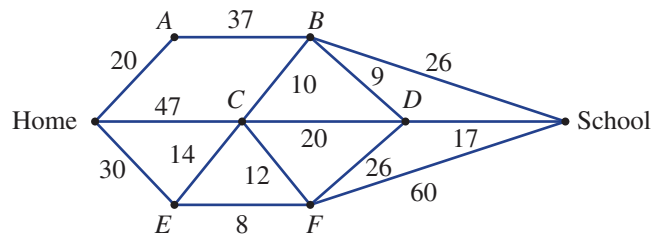
C C – D – G – E

D C – G – E

7 What is the shortest path between B and G ?

- A** $B - E - G$
- B** $B - E - H - G$
- C** $B - A - D - G$
- D** $B - A - D - C - G$

Use the following information to answer Questions 8 and 9.



8 Victoria rides her bike to school each day. The edges of the network represent the roads that Victoria can use to ride to school from her home. The numbers on the edges give the time taken, in minutes, to travel along each road.

The fastest Victoria can ride from home to school is

- A** 80 minutes
 - B** 81 minutes
 - C** 83 minutes
 - D** 84 minutes
- 9 Which of the following represent the fastest route for Victoria's journey from home to school.
- A** $Home - A - B - School$
 - B** $Home - A - B - D - School$
 - C** $Home - C - D - School$
 - D** $Home - E - C - B - D - School$

Key ideas and chapter summary

**Graph**

A **graph** is a diagram that shows the connections between objects, represented by vertices, as edges between those vertices.

Vertex

A **vertex** is a point in a graph that represents an object.

Edge

An **edge** is a line in a graph that connects two vertices and represents the connection between them.

Degree of a vertex

The **degree of a vertex** is the number of edges that are attached to that vertex. The degree of vertex A is written $\deg(A)$.

Loop

A **loop** is an edge that connects a vertex of a graph to itself. A loop contributes two to the degree of that vertex.

Multiple edge

Sometimes a graph has two or more edges that connect the same vertices. These are called **multiple edges**.

Isolated vertex

An **isolated vertex** is one that is not connected to any other vertex. Isolated vertices have degree zero.

Simple graph

A **simple graph** does not have any loops and does not have multiple edges.

Connected graph

A **connected graph** is a graph that has no isolated vertices and no separate parts.

Bridge

A **bridge** is an edge in a connected graph that, if removed, would leave the graph no longer connected.

Directed graph (digraph)

A **directed graph** (digraph) is a graph with a direction associated with the edges.

Subgraph

A **subgraph** is a graph that is part of a larger graph and has some of the same vertices and edges as that larger graph. A subgraph does not have any extra vertices or edges that do not appear in the larger graph.

Adjacency matrix

An **adjacency matrix** is a square matrix that uses a zero or an integer to record the number of edges connecting each pair of vertices in the graph.

Equivalent (isomorphic) graphs

Graphs that contain identical information (connections between vertices) to each other are called **equivalent graphs** or **isomorphic graphs**.

Planar graph	A planar graph is one that can be drawn so that no two edges cross over.
Face	A face is an area in a graph enclosed between edges. It can only be reached by crossing an edge. One face on a graph is always the area surrounding the graph.
Euler's formula	Euler's formula applies to planar graphs. It states that: 'the number of vertices plus the number of faces minus the number of edges always equals 2'. If v = the number of vertices, f = the number of faces and e = the number of edges in a planar graph, then $v + f - e = 2$.
Exploring a graph	Movement through a graph from one vertex to another along the edges of the graph is called exploring the graph.
Walk	A walk is a route through a graph, from one vertex to another, along the edges of the graph. An open walk will start and end at different vertices. A closed walk will start and end at the same vertex.
Trail	A trail is a walk that has no repeated edges. It may contain repeated vertices.
Path	A path is a walk that has no repeated edges and no repeated vertices.
Closed trail	A closed trail is a trail (no repeated edges) that starts and ends at the same vertex.
Cycle	A cycle is also known as a closed path. It is a path (no repeated edges and no repeated vertices) that starts and ends at the same vertex.
Eulerian trail	A closed trail (no repeated edges) that involves every edge of the graph is called an Eulerian trail .
Eulerian graph	An Eulerian graph is a graph that contains an Eulerian trail.
Semi-Eulerian graph	A semi-Eulerian graph is a graph that contains an open trail that involves every edge of the graph.
Hamiltonian cycle	A Hamiltonian cycle is a closed path (no repeat edges, no repeat vertices) that visits every vertex of the graph. It will start and end at the same vertex.
Hamiltonian graph	A Hamiltonian graph is a graph that contains a Hamiltonian cycle.

Hamiltonian path	A Hamiltonian path is an open path (no repeat edges, no repeat vertices) that visits every vertex of the graph.
Semi-Hamiltonian graph	A semi-Hamiltonian graph is a graph that contains a Hamiltonian path.
Weighted graph	A weighted graph has numbers, called weights, associated with the edges of a graph. The weights often represent physical quantities as additional information to the edge, such as time, distance or cost.
Network	A network is a weighted graph where the weights represent physical quantities such as time, distance or cost.
The travelling salesperson problem	The travelling salesperson problem involves finding the Hamiltonian cycle in a graph (visits all vertices in a graph once) that has the smallest total weight. It often refers to minimising the total distance travelled through a graph to return to the starting vertex.
The shortest path problem	The shortest path problem involves finding the shortest path from one vertex to another. The shortest path does not have to visit every vertex.

Skills checklist



Download this checklist from the Interactive Textbook, then print it and fill it out to check your skills.



9A

1 I can identify edges, vertices and loops in a graph.



See Example 1, Example 2 and Exercise 9A Questions 1, 2 and 3

9A

2 I can determine the degree of a vertex in a graph.



See Example 1, Example 2 and Exercise 9A Questions 1, 2 and 3

9A

3 I can define and identify simple graphs, isolated vertices, degenerate graphs, connected graphs, bridges and subgraphs.



See Example 1, Example 2 and Exercise 9A Questions 1, 2 and 3

9B

4 I can construct a graph from an adjacency matrix.



See Example 3 and Exercise 9B Question 2

9B

5 I can write an adjacency matrix from a graph.



See Example 4 and Exercise 9B Question 1

9C

6 I can write an adjacency matrix from a directed graph.



See Example 5 and Exercise 9C Question 3

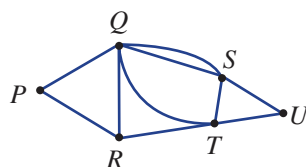
- 9C** **7** I can construct a graph from an adjacency matrix. ☐
See Exercise 9C Question 4
- 9D** **8** I can recognise equivalent (isomorphic) graphs. ☐
See Exercise 9D Question 1
- 9D** **9** I can redraw graphs in planar form. ☐
See Example 6 and Exercise 9D Question 2
- 9D** **10** I can use Euler's formula. ☐
See Example 7, Example 8 and Exercise 9D Questions 3 and 4
- 9E** **11** I can define walks, trails, paths, circuits and cycles through a graph. ☐
See Example 9 and Exercise 9E Question 1
- 9F** **12** I can identify Eulerian trails and circuits through graphs. ☐
See Example 10 and Exercise 9F Question 1
- 9F** **13** I can determine whether an Eulerian trail or circuit exists in a graph. ☐
See Example 10 and Exercise 9F Question 1
- 9G** **14** I can identify Hamiltonian paths and cycles through graphs. ☐
See Example 11 and Exercise 9G Question 1
- 9H** **15** I can define a weighted graph. ☐
See Example 12
- 9H** **16** I can calculate the shortest path from one vertex to another by inspection. ☐
See Example 12 and Exercise 9H Question 1

Multiple-choice questions

- 1 The minimum number of edges for a graph with seven vertices to be connected is:

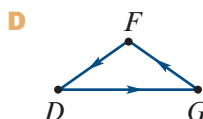
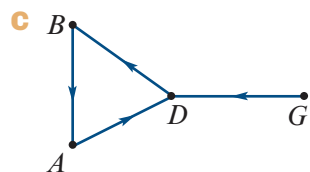
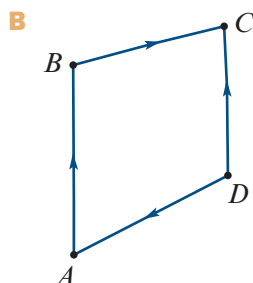
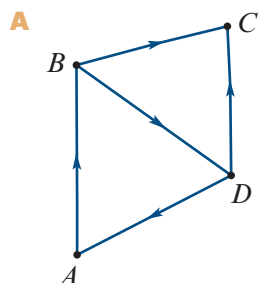
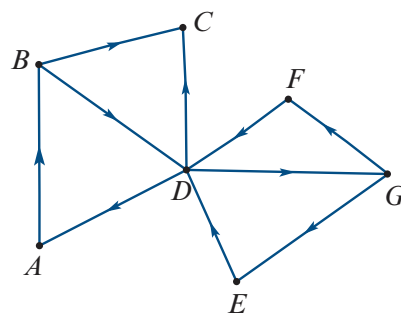
A 4 **B** 5 **C** 6 **D** 7

- 2 For the graph shown below, which vertex has degree 5?



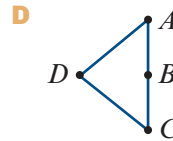
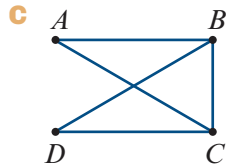
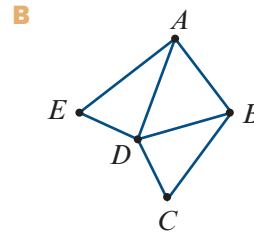
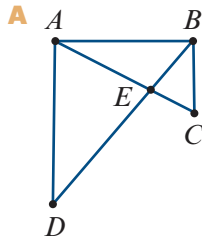
A Q **B** T **C** S **D** R

- 3 Which one of the following is **not** a subgraph of the graph shown on the right?

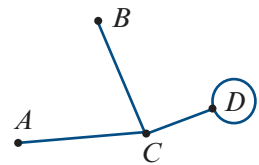


- 4 The graph that has been drawn from the adjacency matrix shown on the right is:

	A	B	C	D
A	0	1	0	1
B	1	0	1	0
C	0	1	0	1
D	1	0	1	0



- 5** The adjacency matrix that corresponds to the graph on the right is:



A

	A	B	C	D
A	0	0	1	0
B	0	0	1	0
C	1	1	0	1
D	0	0	1	1

B

	A	B	C	D
A	0	1	0	1
B	1	0	1	0
C	0	1	0	1
D	1	0	1	0

C

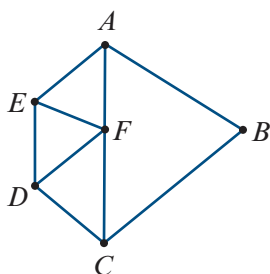
	A	B	C	D
A	0	0	1	1
B	0	0	1	0
C	1	1	0	1
D	0	0	1	3

D

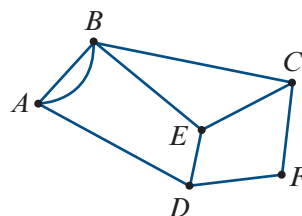
	A	B	C	D
A	1	0	0	1
B	0	0	1	0
C	0	1	0	1
D	1	0	1	1

- 6** A connected graph with 15 vertices divides the plane into 12 faces. How many edges does this graph have?
- A** 15
B 23
C 24
D 25
- 7** A connected planar graph divides a plane into a number of faces. If the graph has seven vertices and these are linked by twelve edges, then how many faces does it have?
- A** 5
B 6
C 7
D 8

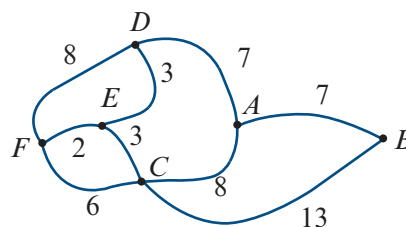
Use the graph below to answer Questions 8 and 9.



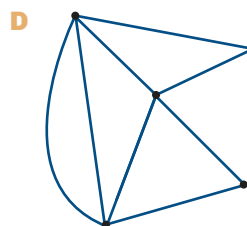
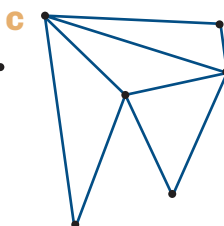
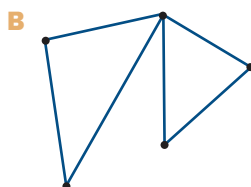
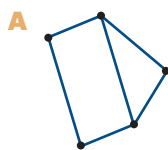
- 8 The walk $A - E - D - C - B - A$ is best described as a:
- A** cycle **B** Hamiltonian path
C Eulerian trail **D** closed trail
- 9 The walk $E - A - F - C - D - E$ is best described as a:
- A** Eulerian trail **B** closed path
C Hamiltonian cycle **D** closed trail
- 10 For the graph shown on the right, which edge could be removed to result in a semi-Eulerian graph?



- 11 The length of the shortest path from F to B in the network shown on the right is:
- A** 17
B 18
C 19
D 20



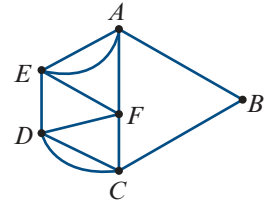
- 12 Which of the following graphs does *not* have an Eulerian circuit?



- 13 A connected planar graph divides the plane into a number of faces. If the graph has 8 vertices and these are linked by 13 edges, then the number of faces is:
- A** 5 **B** 6 **C** 7 **D** 8

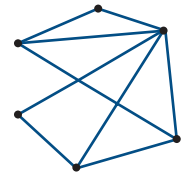
- 14** For the graph shown, which of the following paths is a Hamiltonian cycle?

- A** $A - B - C - D - C - F - D - E - F - A - E - A$
B $A - E - F - D - C - B - A$
C $A - F - C - D - E - A - B - A$
D $A - B - C - D - E - A$



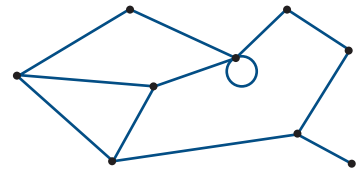
- 15** The graph opposite has:

- A** four faces
B five faces
C six faces
D seven faces

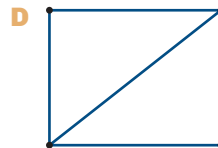
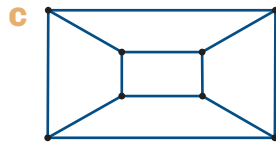
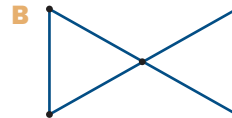
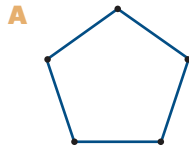


- 16** The sum of the degrees of the vertices on the graph shown here is:

- A** 20
B 21
C 23
D 24



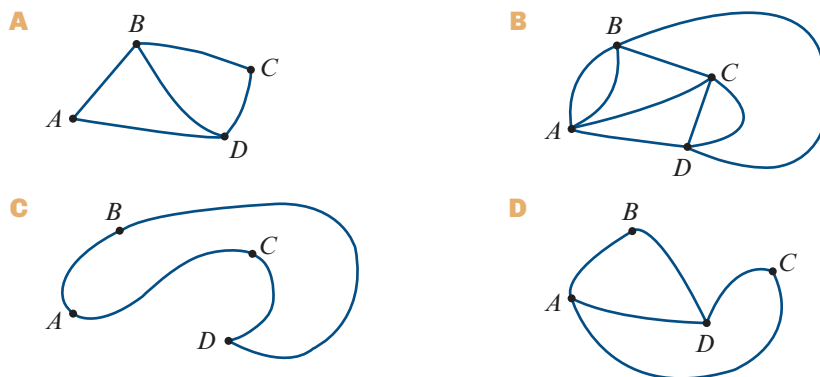
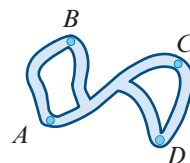
- 17** Of the following graphs, which one has both an Eulerian circuit and Hamiltonian cycles?



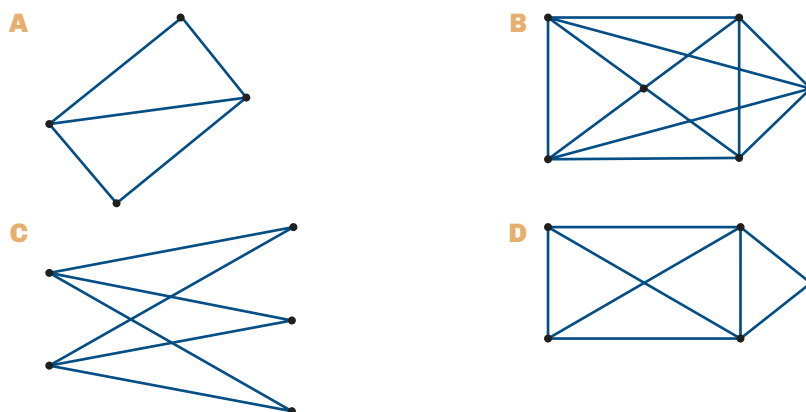
- 18** A network graph with six vertices is connected with the minimum number of edges. The minimum number of extra edges needed to make this a complete graph is

- A** 5
B 6
C 10
D 14

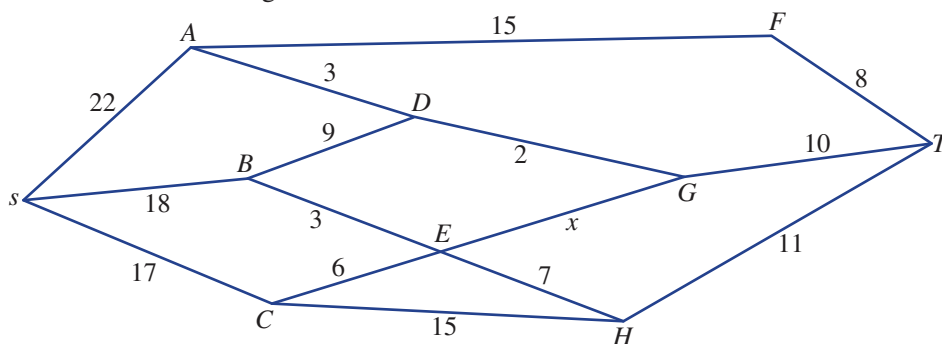
- 19** Four towns, A , B , C and D , are linked by roads as shown. Which of the following graphs could be used to represent the network of roads? Each edge represents a route between two towns.



- 20** Which *one* of the following graphs has an Eulerian circuit?



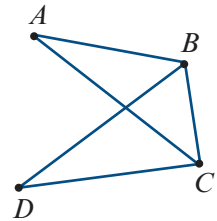
- 21** The network below shows the distance, in metres, between points. The shortest path between S and T has length 36 m. The value of x is.



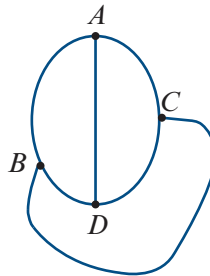
- A** 4 **B** 5 **C** 6 **D** 7

Short-response questions

- 1 Consider the graph shown on the right.
 - a What is the degree of vertex C ?
 - b How many odd-degree vertices does this graph have?
 - c Write down the vertices that have an even degree.
 - d Redraw the graph in planar form.



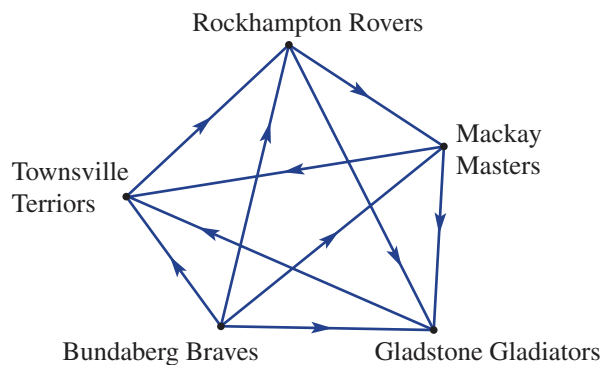
- 2 Construct an adjacency matrix for the graph below.



- 3 Draw a directed graph using the adjacency matrix below.

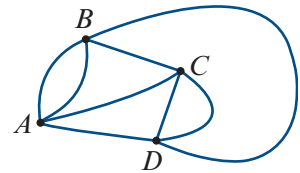
$$\begin{array}{c}
 A \quad B \quad C \quad D \\
 \begin{array}{l}
 A \left[\begin{array}{cccc} 0 & 1 & 0 & 1 \end{array} \right] \\
 B \left[\begin{array}{cccc} 1 & 0 & 1 & 0 \end{array} \right] \\
 C \left[\begin{array}{cccc} 1 & 0 & 0 & 0 \end{array} \right] \\
 D \left[\begin{array}{cccc} 0 & 0 & 2 & 0 \end{array} \right]
 \end{array}
 \end{array}$$

- 4 The directed graph on the right shows the results of a round-robin darts tournament. Each team is represented by a vertex and the arrows point to the winner of that match.



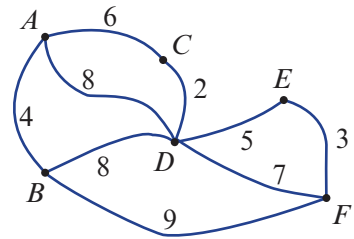
- a Which team(s) beat the Rockhampton Rovers?
- b How many teams did the Mackay Masters beat?
- c Which team lost all of its games?

- 5 Consider the graph on the right.
- How many vertices does this graph have?
 - How many faces does this graph have?
 - How many edges does this graph have?
 - Verify Euler's formula for this graph.



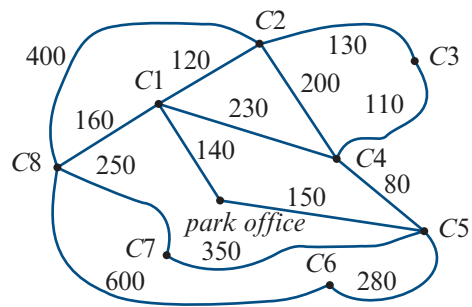
- 6 Consider the graph below. The vertices represent cities in a particular state. The numbers on the arcs shows the time take, in hours, to drive between each city.

- In hours, which two cities are two hours driving time apart?
- In hours, what is the shortest driving time between *E* and *B*?
- In hours, what is the shortest driving time between *F* and *A*?

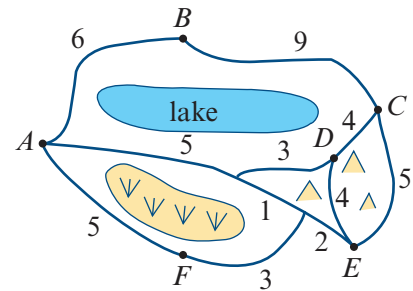


- 7 The network opposite shows the walking tracks in a small national park. The tracks, represented by arcs, connect campsites to each other and to the park office. The weights on the arcs show the distance, in metres, between each location.

- The graph shown is planar. Explain what this means.
- Verify Euler's formula for this network.
- A ranger at campsite *C8* plans to visit campsites *C1*, *C2*, *C3*, *C4* and *C5* on her way back to the park office. What is the shortest distance she will have to travel?
- Each day, the ranger on duty has to inspect each of the tracks to make sure that they are all safe to walk on.
 - Is it possible for her to do this starting and finishing at the park office, while only walking along a path once? Explain your answer.
 - Identify one route that she could take.
- Another ranger wants to inspect each of the campsites but not pass through any campsite more than once on the inspection route. He wants to start and finish his inspection route at the park office.
 - What is the mathematical name of the route he wants to take?
 - With the present layout of tracks, he cannot inspect all the tracks without passing through at least one campsite twice. Suggest where an additional track could be added to solve this problem.
 - With this new track, write down a route that he could follow.

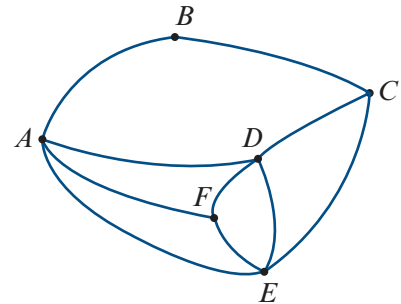


- 8 The map on the right shows six campsites, A, B, C, D, E and F , which are joined by tracks. The numbers on the edges give the lengths, in kilometres, of those sections of track.



- a Complete the following questions related to the map of the campsites.

- i Complete the graph opposite, which shows the shortest direct distances between campsites. The campsites are represented by vertices and the tracks are represented by edges.

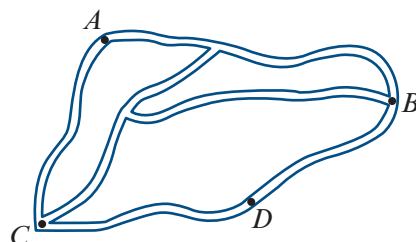


- ii Fill in the missing entries for the adjacency matrix of the graph in part a i.

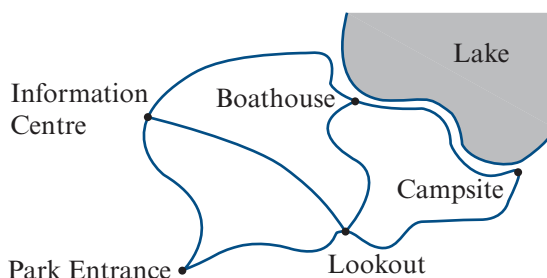
	A	B	C	D	E	F
A	0	1	0	1	1	1
B	1	0	1	0	0	0
C	0	1	0	1	1	0
D	1	0	—	—	—	—
E	1	0	—	—	—	—
F	1	0	—	—	—	—

- b A walker follows the route $A - B - A - F - E - D - C - E - F - A$.
- How far does this person walk if they take the shortest track between each point?
 - Why is the route not a Hamiltonian cycle?
 - Write down a route that a walker could follow that is a Hamiltonian cycle.
 - Determine the distance walked by following this Hamiltonian cycle.
- c It is impossible to start at A and return to A by going along each track exactly once. An extra track joining two campsites can be constructed so that this is possible. Which two campsites need to be joined by a track to make this possible?

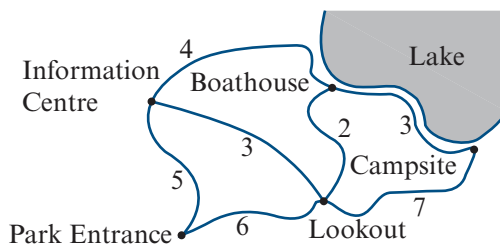
- 9 Four children each live in a different town. The diagram opposite is a map of the roads that link the four towns, A , B , C and D .



- How many different trails are there from town A to town D ?
 - How many different ways can a vehicle travel between town A and town B without visiting any other town?
 - Draw this map as a graph by representing towns as vertices and each different route between two towns as an edge.
 - Explain why a vehicle at A could not follow an Eulerian trail through this graph.
- 10 A national park contains five locations connected by bushwalking tracks. The graph below shows the park entrance, information centre, lookout, boathouse and camping site represented by vertices and the bushwalking tracks represented by edges.



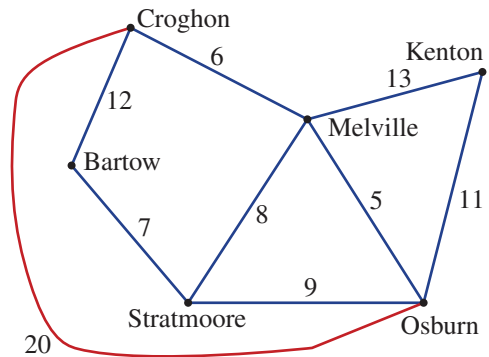
- If a bushwalker starts at the park entrance, write down two Hamiltonian cycles that he can follow.
- Is a semi-Eulerian trail possible for this graph? Explain your answer with mathematical reasoning.
- The network below has the distances between locations (in kilometres) added to each edge.



The bushwalker can walk at a speed of 4 km/hour.

- How long would it take the bushwalker to walk directly from the lookout to the boathouse?
- Determine the shortest time it would take the bushwalker to walk between the park entrance and the camping site.

- 11** The network on the right shows six villages represented as vertices of the graph. The edges represent the roads connecting the villages. The weights on the edges are the distances, in kilometres, along each of the roads.



- What is the degree of the vertex representing Melville?
- Determine the sum of the degrees of the vertices of this graph.
- Verify Euler's formula for this graph.

A salesperson might need to travel to every village in this network to conduct business.

- If the salesperson follows the path Stratmoore – Melville – Kanton – Osborn – Melville – Croghon – Bartow – Stratmoore, has the salesperson followed a Hamiltonian cycle? Give a reason to justify your answer.
- If the salesperson follows the path Croghon – Bartow – Stratmoore – Melville – Kanton – Osborn, what is the mathematical term for this path?

It would make sense for the salesperson to avoid visiting a certain village more than once, and it would also make sense for them to return 'home' after travelling the shortest distance possible.

- If the salesperson can start and end at any village in the network, what is the shortest route possible?

A road inspector must travel along every road connecting the six villages.

- Explain why the inspector could not follow an Eulerian circuit through this road network.
- The inspector may start and end their route at different villages, but would like to travel along each road once only. Which villages can the inspector start their route from? Write down a path the inspector could take to complete their work.
- The speed limit for each of these roads is 60 km/hr. If the inspector must complete their work by 5 p.m, what is the latest time that the inspector can begin their work?

Chapter 10

Networks and decision mathematics 1

Chapter questions

UNIT 4 INVESTING AND NETWORKING

Topic 4: Networks and decision mathematics 1

- ▶ How do we define a tree and a spanning tree?
- ▶ How do we draw a tree from any graph?
- ▶ How do we determine a minimum spanning tree in a weighted connected graph?
- ▶ How do we solve practical problems involving minimum spanning trees?
- ▶ How do we construct a project network diagram?
- ▶ How do we use forward and backward scanning to determine the earliest starting time (EST) and latest starting time (LST) for each activity in project?
- ▶ How do we use ESTs and LSTs to locate the critical path/s for a project and the minimum time to complete this project?
- ▶ How do we calculate float times in a project?

In this chapter graphs are used to consider minimum connector problems and scheduling problems.

10A Trees and connector problems

Learning intentions

- ▶ To be able to identify a tree.
- ▶ To be able to find a spanning tree for a graph.
- ▶ To be able to find the minimum spanning tree for a network using Prim's algorithm.

Trees

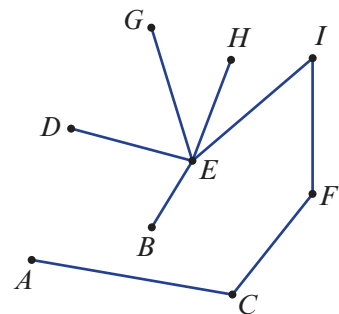
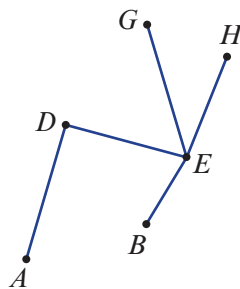
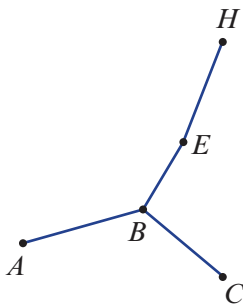
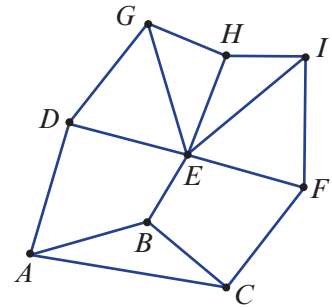
A **tree** is a connected graph that has no loops, no multiple edges, and no cycles. Recall that a cycle starts and ends at the same vertex and has no repeated vertices, nor repeated edges. Because a tree has no cycle, it will be impossible to find a walk in a tree that starts and ends at the same vertex without repeating an edge.

The number of edges in a tree will always be one less than the number of vertices.

Every connected graph will contain at least one subgraph that is a tree.

Some of the trees in the graph on the right are shown below.

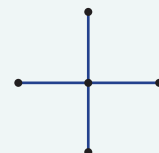
In all three of the graphs below, a cycle does not exist. The trees are all subgraphs of the original graph. The first two trees involve some of the vertices of the original graph, but the third tree involves every vertex from the original graph.



Example 1

Draw the three trees with five vertices.

Solution



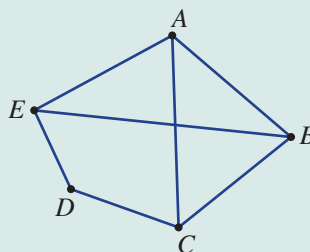
Spanning trees

A **spanning tree** of a graph is a tree that includes every vertex of that graph. Every connected graph has a spanning tree and it may have more than one spanning tree.



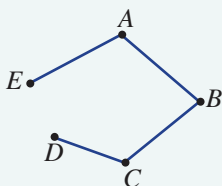
Example 2

Find two spanning trees for the graph shown here.

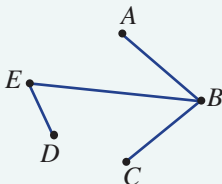


Solution

Spanning tree 1



Spanning tree 2



Explanation

The graph has five vertices and seven edges. A spanning tree will have five vertices and four edges.

- To form a spanning tree, remove any three edges, provided that all the vertices remain connected and there are no cycles, multiple edges or loops.
- Spanning tree 1 is formed by removing edges EB , ED and CA .
- Spanning tree 2 is formed by removing edges EA , AC and CD .

Spanning trees of weighted graphs

A spanning tree can also be found for a weighted graph. The **weight** of the spanning tree is the sum of all the weights in that spanning tree.

A spanning tree can be found by removing edges of the graph until there is one less edge than the number of vertices in the graph, making sure that all vertices remain connected.

Minimum spanning trees

For any graph, there may be more than one spanning tree possible. If the weight of every spanning tree is found, there will always be a tree, or trees, that has a smaller total weight than all the others. This tree is called the **minimum spanning tree** for that graph.

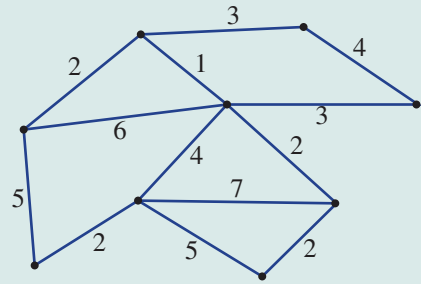


A guide to trees: Watch the video in the Interactive Textbook to see trees, spanning trees and minimum spanning trees in action.



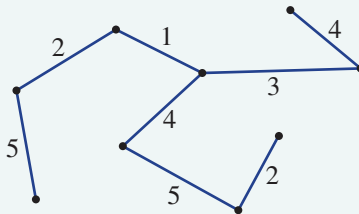
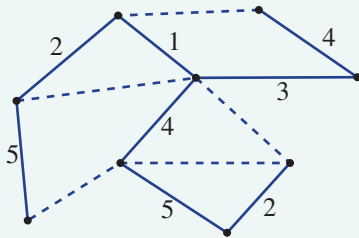
Example 3 Finding the weight of a spanning tree

- a** Draw one spanning tree of weight 26 for the graph shown on the right.
- b** Draw the minimum spanning tree for this graph and calculate its weight.



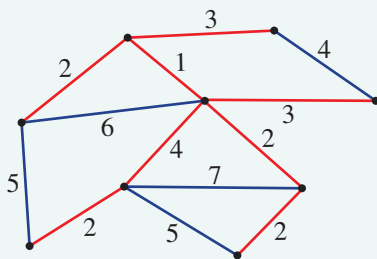
Solution

- a** There are 9 vertices and 13 edges.
The spanning tree will have $9 - 1 = 8$ edges.
Remove $13 - 8 = 5$ edges.



$$\begin{aligned} \text{weight} &= 5 + 2 + 1 + 4 + 5 + 2 + 3 + 4 \\ &= 26 \end{aligned}$$

b



The weight of the minimum spanning tree: $2 + 1 + 3 + 3 + 4 + 2 + 2 + 2 = 19$

Explanation

Count the number of vertices and edges in the graph. Calculate the number of edges in the spanning tree. Calculate how many edges must be removed. Choose edges to remove. Make sure that no vertex is left isolated.

Add the weights of the remaining edges.

The tree must pass through every vertex and have 8 edges. We start at the vertex to the far left and proceed by always choosing the edge with the lowest weight. An algorithm for this procedure is shown on the following page.

For small graphs, the minimum spanning tree can be found by inspection, but for larger trees, an **algorithm** must be used. An algorithm is a series of instructions that can be used to solve a particular problem. The algorithm used to find the minimum spanning tree for a network is called **Prim's algorithm**.

Prim's algorithm

Prim's algorithm is used to find the minimum spanning tree for a network.

- Choose a starting vertex:
 - the spanning tree will contain every vertex and so any vertex can be chosen as the starting vertex.
- Inspect the edges that are connected to the starting vertex:
 - choose the edge that has the lowest weight
 - if there is more than one edge that has the lowest weight, it does not matter which one you choose
 - the starting vertex, the edge chosen, and the vertex connected by this edge form the beginning of the minimum spanning tree.
- Inspect the edges that are connected to the vertices chosen so far:
 - choose the one that has the lowest weight, but ignore any that would connect the tree back to itself
 - add the chosen edge and the vertex connected by this edge to the minimum spanning tree.

Repeat the process until all vertices have been added to the minimum spanning tree.

This algorithm is also very useful with smaller graphs and it can be described as a common sense approach.

The algorithm always produces a minimum spanning tree but there may be other minimum spanning trees for a given graph. For example, if all the edge weights of a given graph are the same, then every spanning tree of that graph is minimum.

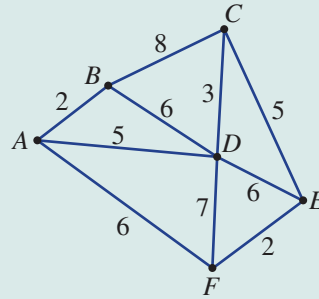
If all the edge weights in a connected graph are distinct then there is only one minimum spanning tree.

Of course it is not true that if there is only one minimum spanning tree then all edge weights are equal. For example if the graph is a tree which does not have distinct edge weights then the graph itself is a minimum spanning tree.



Example 4 Finding the minimum spanning tree

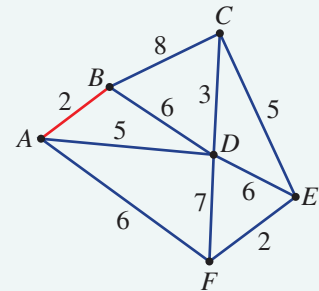
Apply Prim's algorithm to find the minimum spanning tree for the graph shown on the right. Write down the total weight of the minimum spanning tree.



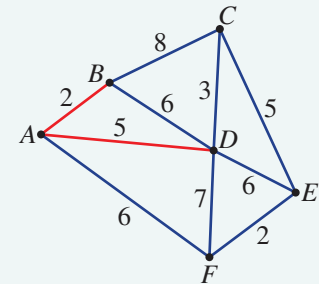
Solution

Start with vertex A .

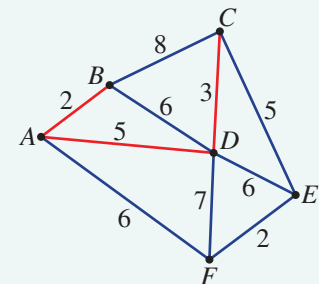
The smallest weighted edge from vertex A is to B with weight 2.



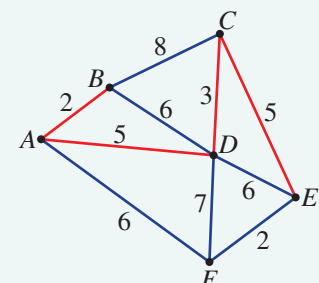
Look at vertices A and B . The smallest weighted edge from either vertex A or vertex B is from A to D with weight 5.



Look at vertices A , B and D . The smallest weighted edge from vertex A , B or D is from D to C with weight 3.

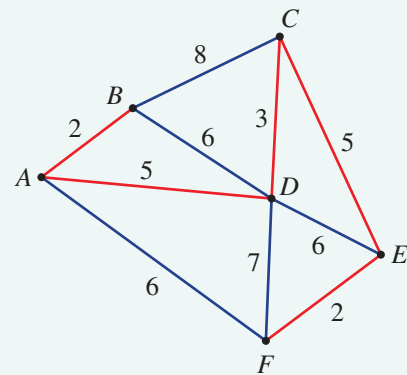


Look at vertices A , B , D and C . The smallest weighted edge from vertex A , B , D or C is from C to E with weight 5.



Look at vertices A, B, D, C and E . The smallest weighted edge from vertex A, B, D, C or E is from E to F with weight 2.

All vertices have been included in the graph.
This is the minimum spanning tree.



Add the weights to find the total weight of the minimum spanning tree.

The total weight of the minimum spanning tree is $2 + 5 + 3 + 5 + 2 = 17$.

Minimum connector problems

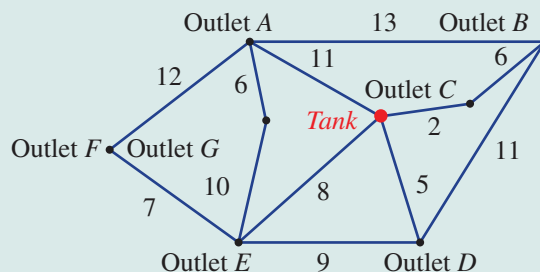
Minimum spanning trees represent the least weight required to keep all of the vertices connected in the graph. If the edges of a graph represent the cost of connecting towns to a gas pipeline, then the total weight of the minimum spanning tree would represent the minimum cost of connecting the towns to the gas. This is an example of a **connector problem**, where it is important to make the cost of keeping towns or other objects connected together as low as possible.



Example 5 Solving a connector problem

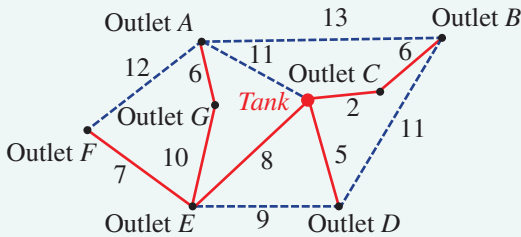
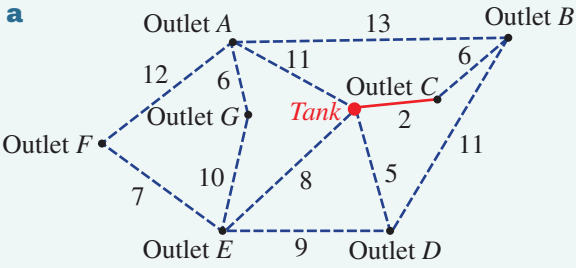
Water is to be piped from a water tank to seven outlets on a property. The distances (in metres) of the outlets from the tank and from each other are shown in the network below.

Starting at the tank, the aim is to find the minimum length of pipe, in metres, required in order to have water piped to all outlets in the property.



- On the diagram, show where the water pipes will be placed in order to minimise the length required.
- Calculate the total length, in metres, of the water pipe that is required to obtain this minimum length.

Solution



- b** The length of water pipe required is
 $2 + 6 + 5 + 8 + 7 + 10 + 6 = 44$ metres.

Explanation

The water pipes will be a minimum length if they are placed on the edges of the minimum spanning tree for the network.

A starting point for Prim's algorithm is the vertex that is connected to the tank by the edge with the smallest weight. The starting vertex (Tank), the edge and the vertex it connects to form the beginning of the minimum spanning tree.

Follow Prim's algorithm to find the minimum spanning tree.

Add the weights of the minimum spanning tree. Write your answer.

Section Summary

- ▶ A tree has no loops, multiple edges or cycles.
- ▶ If a tree has n vertices, it will have $n - 1$ edges.
- ▶ A **spanning tree** is a tree that connects all of the vertices of a graph.
- ▶ The **weight** of a spanning tree for a weighted graph is the sum of the weights of the edges in that tree.
- ▶ For a given connected weighted graph, a tree with the minimum weight is called a **minimum spanning tree**.
- ▶ **Prim's algorithm** can be used to find a minimum spanning tree.

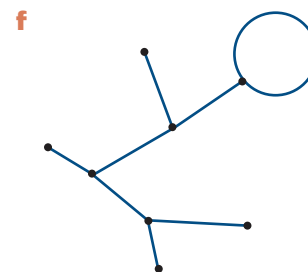
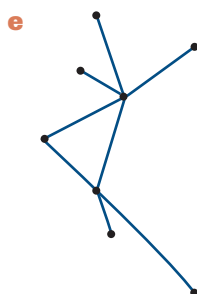
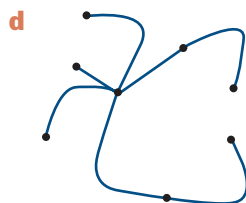
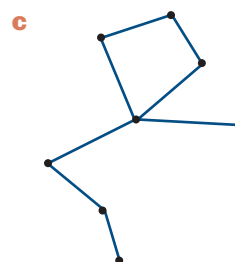
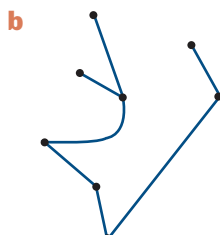
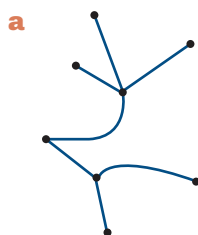
Exercise 10A

Trees and spanning trees

Example 1

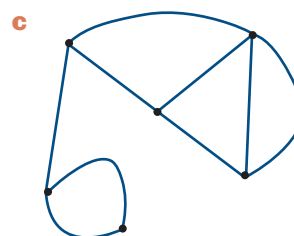
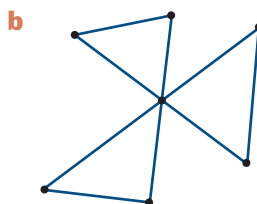
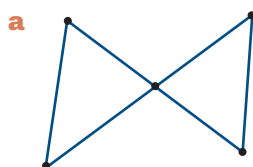
- 1** Complete the following for the different trees.
 - a** How many edges are there in a tree with 12 vertices?
 - b** How many vertices are there in a tree with 8 edges?
 - c** Draw two different trees that have 5 vertices.

- 2** Which of the following graphs are trees?

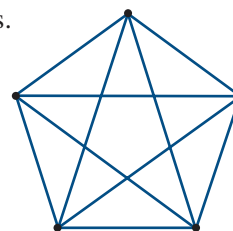


Example 2

- 3** For each of the following graphs, draw two different spanning trees.



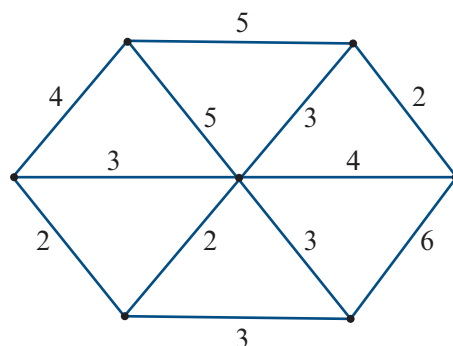
- 4** Draw a spanning tree for the complete graph on 5 vertices.



Weighted spanning trees

Example 3

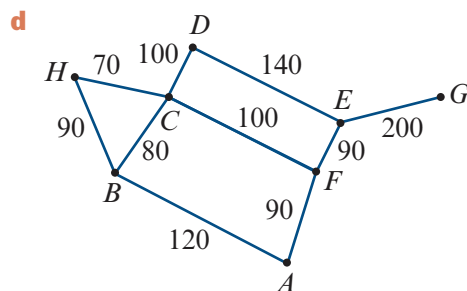
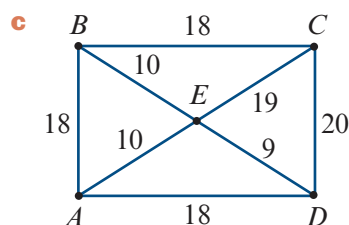
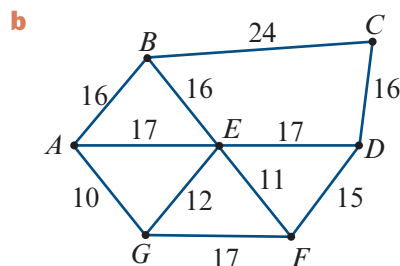
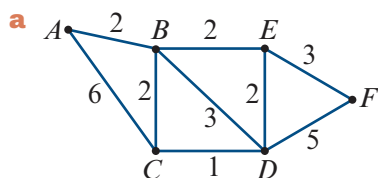
- 5** A network is shown on the right.
- How many edges must be removed in order to leave a spanning tree?
 - Remove some edges to form two different spanning trees.
 - For each tree in part **b**, find the total weight.



Minimum spanning tree

Example 4

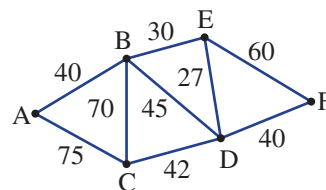
- 6** Determine a minimum spanning tree for each of the following graphs and then write down their weight. (Use Prim's algorithm if needed.)



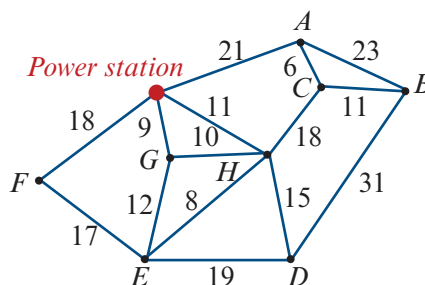
Connector problems

Example 5

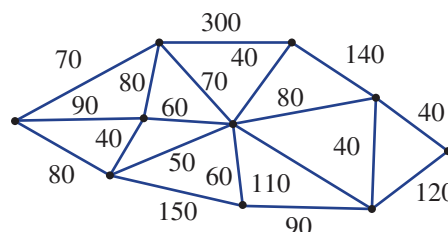
- 7** Six towns are to be connected with a new fibre-optic cable system. The hub of the system is to be town E. Distances are in kilometres. What is the minimum length of cable needed to connect the towns?



- 8** Power is to be connected by cable from a power station to eight substations (A to H). The distances (in kilometres) of the substations from the power station and from each other are shown in the network to the right. Determine the minimum length of cable needed to provide all substations with power.

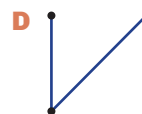
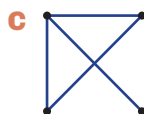
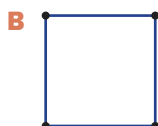
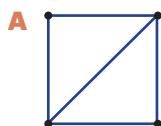


- 9** In the network opposite, the vertices represent water tanks on a large property and the edges represent pipes used to move water between these tanks. The numbers on each edge indicate the lengths of pipes (in m) connecting different tanks. Determine the shortest length of pipe needed to connect all water storages.



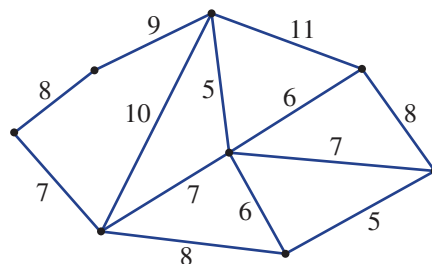
Paper 1-style multiple-choice questions

- 10** Which of the following graphs is a tree?



- 11** For the network opposite the length of the minimum spanning tree is

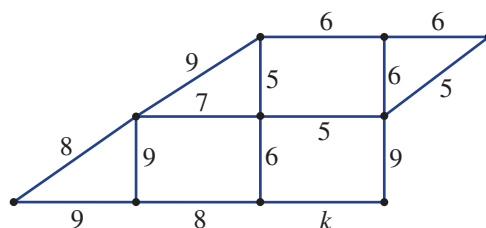
A 44 **B** 45 **C** 46 **D** 47



- 12** The minimum spanning tree for the network below includes the edge with weight labelled k .

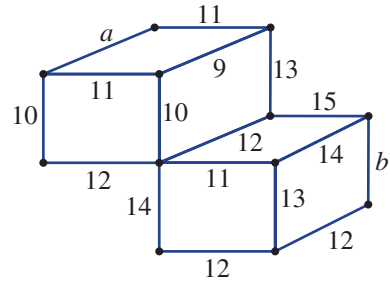
The total weight of all edges for the minimum spanning tree is 58. The value of k is

A 6 **B** 7 **C** 8 **D** 9



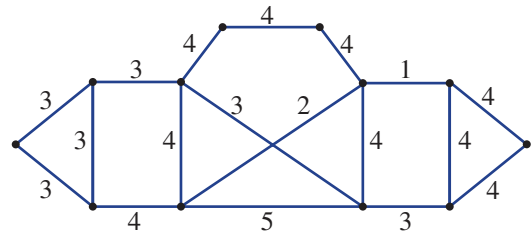
- 13** The minimum spanning tree for the network opposite includes two edges with weights a and b . The length of the minimum spanning tree is 124. The values of a and b could be

A $a = 6$ and $b = 17$
B $a = 12$ and $b = 12$
C $a = 10$ and $b = 14$
D $a = 10$ and $b = 15$



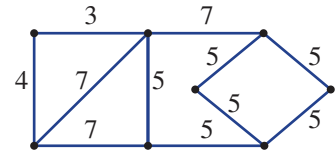
- 14** A minimum spanning tree is to be drawn for the weighted graph opposite. How many edges with weight 4 will **not** be included in any particular minimum spanning tree?

A 2 **B** 3 **C** 4 **D** 5



- 15** Consider the weighted graph opposite. How many different minimum spanning trees are possible?

A 1 **B** 2 **C** 3 **D** 4



10B Project planning – precedence tables and activity networks

Learning intentions

- ▶ To identify activities in a project.
- ▶ To understand the precedence that some activities have over others in a project.
- ▶ To identify immediate predecessors of activities from an activity network.
- ▶ To draw an activity network from a precedence table.

Project planning

Building a house, manufacturing a product, and organising a wedding are all examples of a *project*, that is, a task that involves a number of individual steps, or *activities*, that must be completed. The individual activities often rely upon each other and some cannot be performed until others are completed.

For example, in the organisation of a wedding, invitations would be sent out to guests, but a plan for seating people at the tables during the reception cannot be completed until the

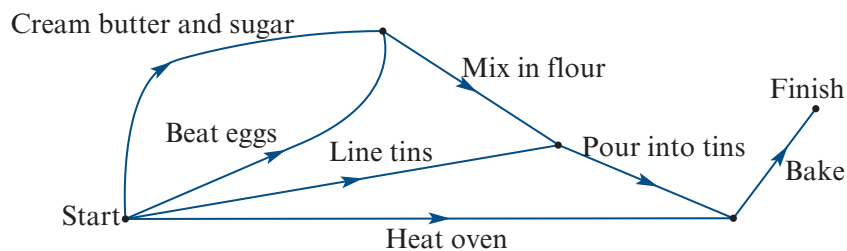
invitations have been accepted. When building a house, the plastering of the walls cannot begin until the house has been sealed from the weather.

Project planning involves the analysis of the requirements of each of the activities of a project to determine the order in which they must be completed.

Activity networks

Projects are represented using a directed graph called an **activity network**. Activity networks have a vertex labelled *start* and another labelled *finish*. Each **activity** within the project is represented by an edge, and so in activity networks it is the edges that must be labelled, not the vertices. The edges are arranged to display the order in which activities must be completed. Activity networks do not have multiple edges.

The activity network shown below represents the project of making a sponge cake. There is an edge labelled for each of the steps of the recipe. The vertices have not been labelled, except for *start* and *finish*.

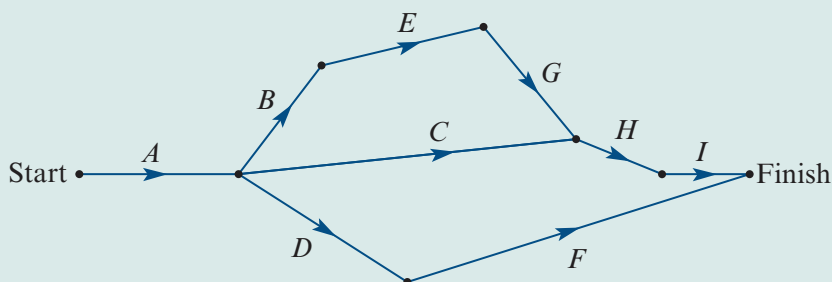


- The edges for activities ‘Cream butter and sugar’ and ‘Beat eggs’ both end at the vertex where the activity ‘Mix in flour’ begins. This shows that the butter and sugar must be creamed and the eggs beaten before the flour can be mixed in.
- The activity ‘Mix in flour’ cannot begin until the other two activities are completed.
- Similarly, the activity ‘Pour into tins’ cannot begin until the activity ‘Line tins’ is also completed.
- Finally, the activity ‘Bake’ cannot begin until all the other activities, including ‘Heat oven’, are completed.

Activity networks show the *precedence* that activities have over each other. The activity ‘Mix in flour’ must be completed before the activity ‘Pour into tins’ can begin and so ‘Mix in flour’ is called an **immediate predecessor** of activity ‘Pour into tins’.


Example 6 Interpreting activity networks

The activity network for a project is shown below.



- How many activities are involved in this project?
- Which activity is an immediate predecessor of activity F ?
- Activity B is an immediate predecessor of which activity?
- How many immediate predecessors does activity H have?

Solution

- This project has nine activities.
- Activity D is an immediate predecessor of activity F .
- Activity B is an immediate predecessor of activity E .
- Activity H has two immediate predecessors, G and C .

Explanation

Count the number of edges in the network.

An immediate predecessor of activity F ends at the vertex at which activity F begins.

Any activity that begins at the same vertex that activity B ends on has activity B as an immediate predecessor.

Count the number of activities that end at the vertex at which activity H begins.

Precedence tables

The activities within a project can have multiple immediate predecessors and these can be recorded in a table called a **precedence table**.

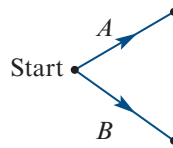
This precedence table shows some of the activities involved in a project and their immediate predecessors.

The information in the precedence table can be used to draw an **activity network**.

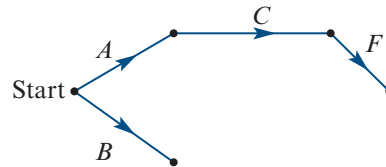
Activity networks do not have labelled vertices, other than the *start* and *finish* of the project. The activities in the project are represented by the edges of the diagram and so it is the edges that must be labelled, not the vertices.

Activity	Immediate predecessors
A	–
B	–
C	A
D	B
E	B
F	C, D
G	E, F

Activities A and B have no immediate predecessors. These activities can start immediately and can be completed at the same time.

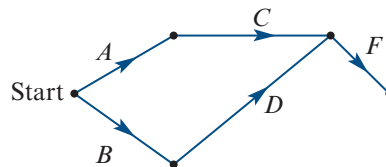


Activity A is an immediate predecessor of activity C , so activity C must follow immediately after activity A .



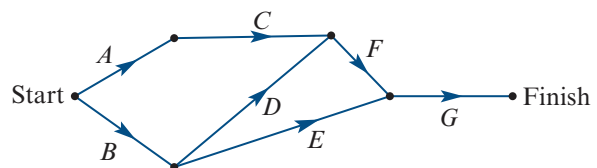
Activity C is an immediate predecessor of activity F , so activity F must follow immediately after activity C .

Activity D has immediate predecessor activity B , so it follows immediately after activity B .



Activity D is also an immediate predecessor of activity F , so activity F must follow immediately after activity D .

Activity E has immediate predecessor activity B , so it will follow immediately after activity B .



Activity G has immediate predecessor activity F and activity E , so it must follow immediately after both of these activities.

Activity G is not an immediate predecessor for any activity, so the project is finished after this activity is complete.

Activity networks

When activity A must be completed before activity B can begin, activity A is called an immediate **predecessor** of activity B .

A table containing the activities of a project, and their immediate predecessors, is called a **precedence table**.

An **activity network** can be drawn from a precedence table. Activity networks have edges representing activities. The vertices are not labelled, other than the start and finish vertices.


Example 7 Drawing an activity network from a precedence table

Draw an activity network from the precedence table shown below.

Activity	Immediate predecessors
<i>A</i>	–
<i>B</i>	<i>A</i>
<i>C</i>	<i>A</i>
<i>D</i>	<i>A</i>
<i>E</i>	<i>B</i>
<i>F</i>	<i>C</i>
<i>G</i>	<i>D</i>
<i>H</i>	<i>E, F, G</i>

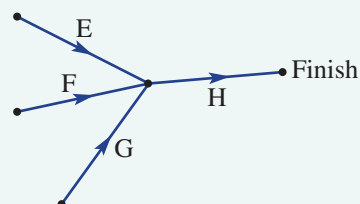
Solution

The activity network can be drawn in any order, beginning with any activity. In this solution, the activity network will be drawn from the finishing vertex back to the starting vertex.

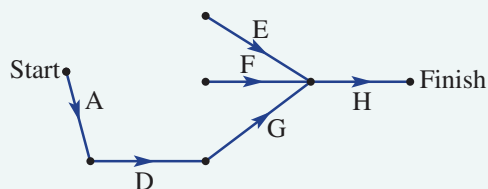
H is not an immediate predecessor for any other activity, so it will lead to the finish of the project.



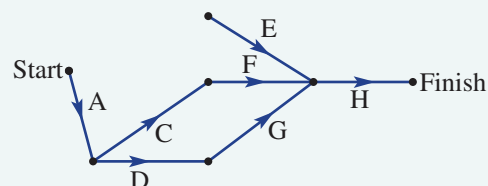
H has immediate predecessors *E, F* and *G*, so these three activities will lead into activity *H*.



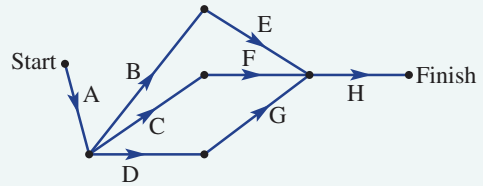
Activity *D* is an immediate predecessor of activity *G* and has immediate predecessor activity *A*. There will be a path through activity *A*, activity *D* and then activity *G*.



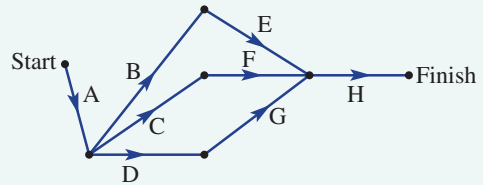
Activity *C* is an immediate predecessor of activity *F* and has immediate predecessor activity *A*. There will be a path through activity *A*, activity *C* and then activity *F*.



Activity *B* is an immediate predecessor of activity *E* and has immediate predecessor activity *A*. There will be a path through activity *A*, activity *B* and then activity *E*.



Activity *A* has no immediate predecessors, so it is the start of the project.



Sketching activity networks

Activities that have no immediate predecessors follow from the start vertex.

Activities that are not immediate predecessors for other activities lead to the finish vertex.

For every other activity, look for:

- activities for which it is an immediate predecessor
- which activities it has as immediate predecessors.

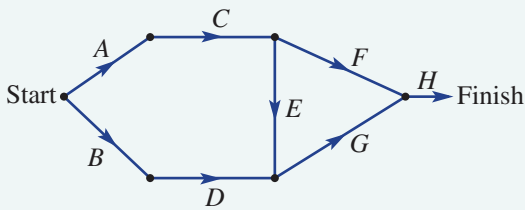
Construct the activity network from this information.



Example 8 Constructing an activity network

Draw an activity network from the precedence table shown below.

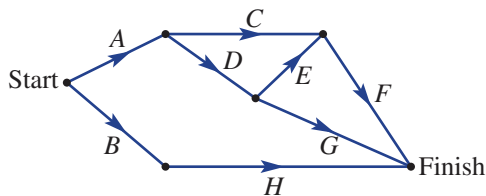
Activity	Immediate predecessors
<i>A</i>	—
<i>B</i>	—
<i>C</i>	<i>A</i>
<i>D</i>	<i>B</i>
<i>E</i>	<i>C</i>
<i>F</i>	<i>C</i>
<i>G</i>	<i>E, D</i>
<i>H</i>	<i>F, G</i>

Solution**Explanation**

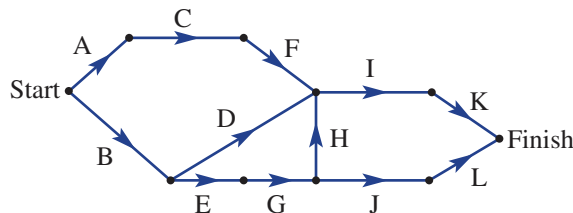
- A and B will lead from the start vertex.
- H will lead to the end vertex.

Section Summary

- ▶ A project is made up of individual activities.
- ▶ An **activity network** has edges that represent the activities of a project. The vertices of an activity network are not labelled, except for the *start* and *finish* vertices.
- ▶ Activity networks do not have multiple edges.
- ▶ When activity *M* must be completed before activity *N* begins, activity *M* is called an **immediate predecessor** of activity *N*.
- ▶ A **precedence table** records all of the immediate predecessors for each activity of a project.

Exercise 10B**Analysis of activity networks and precedence tables****Example 6****1**

- a How many activities are involved in this project?
- b Which activities are immediate predecessors of activity *F*?
- c Activity *B* is an immediate predecessor of which activity?
- d How many immediate predecessors does activity *C* have?

2 Consider the following activity network.

- a Which activities are immediate predecessors of activity *I*?
- b Which activities must be completed before activity *I* can commence?

Drawing activity networks from precedence tables

Example 7

Example 8

3 Draw an activity network from the precedence tables below.

a

Activity	Immediate predecessors
A	–
B	A
C	A
D	B
E	C

b

Activity	Immediate predecessors
P	–
Q	–
R	P
S	Q
T	R, S

c

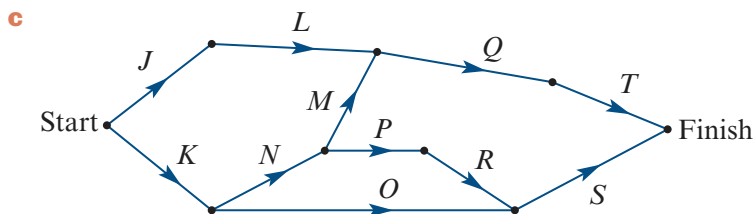
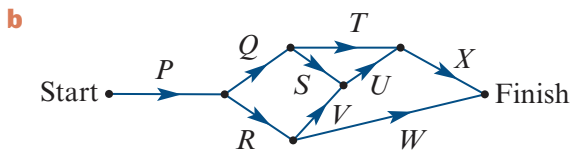
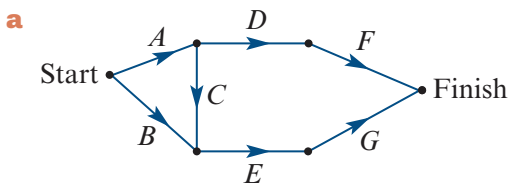
Activity	Immediate predecessors
T	–
U	–
V	T
W	U
X	V, W
Y	X
Z	Y

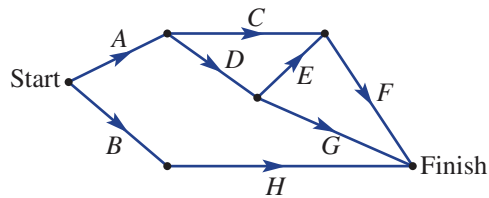
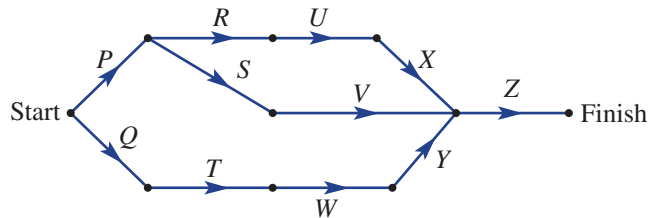
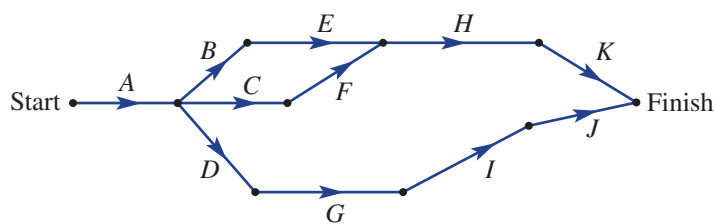
d

Activity	Immediate predecessors
F	–
G	–
H	–
I	F
J	G, I
K	H, J
L	K

Writing precedence tables from activity networks

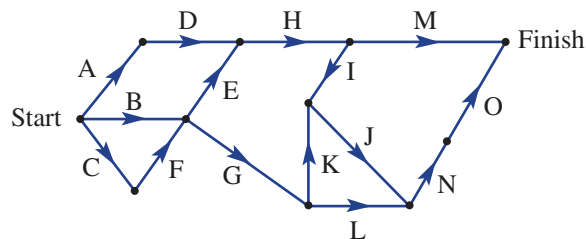
4 Write a precedence table for the activity networks shown below.



d**e****f**

Analysis of activity networks and precedence tables

5 Consider the following activity network for a project.

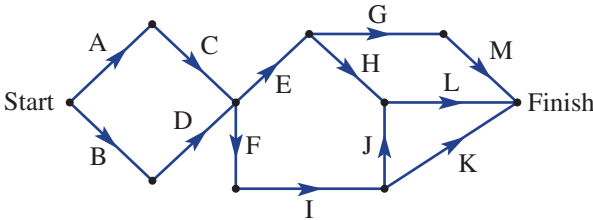


- a** Write down a precedence table for the network above.
- b** Write down the two paths from *start* to *finish* that begin with activity A.
- c** Write down the four paths from *start* to *finish* that begin with activity B.
- d** Write down the four paths from *start* to *finish* that begin with activity C.

Paper 1-style multiple-choice questions

Use the following information for Questions 6 and 7.

The activity network below shows the sequence of activities required to complete a project.



- 6 Beginning with activity *B*, the number of paths from start to finish is
- A** 1 **B** 2
C 3 **D** 4
- 7 The immediate predecessors of activity *L* are
- A** *M* and *K* **B** *J* and *H*
C *I* and *K* **D** *G* and *M*

Use the following information to answer Questions 8 and 9

A project involves nine activities, *A* to *I*.

The immediate predecessor(s) of each activity is shown in the table opposite.

- 8 To complete activity *H* you first have to complete
- A** activities *A* and *D* and no others.
B activities *C* and *F* and no others.
C activities *A*, *C* and *F* and no others.
D activities *A*, *D*, *C* and *F* and no others.
- 9 The number of paths from start to finish is
- A** 4 **B** 5
C 6 **D** 7

Activity	Immediate predecessors
<i>A</i>	–
<i>B</i>	–
<i>C</i>	–
<i>D</i>	<i>A</i>
<i>E</i>	<i>C</i>
<i>F</i>	<i>C</i>
<i>G</i>	<i>B</i> , <i>E</i>
<i>H</i>	<i>D</i> , <i>F</i>
<i>I</i>	<i>D</i> , <i>F</i>
<i>J</i>	<i>G</i> , <i>H</i>

10C Scheduling problems

Learning intentions

- ▶ To understand activity networks that include weights (durations) of each activity.
- ▶ To determine the EST for activities using forward scanning.
- ▶ To determine the LST for activities using backward scanning.
- ▶ To calculate and understand the existence of float times for some activities.
- ▶ To identify the critical path and completion time of a project.

Scheduling

Projects that involve multiple activities are usually completed against a time schedule. Knowing how long individual activities within a project are likely to take allows managers of such projects to hire staff, book equipment and also to estimate overall costs of the project. Allocating time to the completion of activities in a project is called scheduling. *Scheduling problems* involve analysis to determine the minimum overall time it would take to complete a project.

Weighted precedence tables

Weighted precedence tables show the completion times, or durations, of each of the activities that make up a particular project.

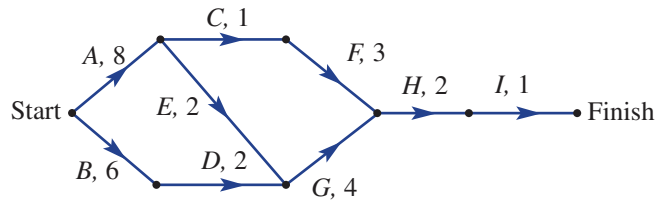
The precedence table on the right shows the activities of a project, the duration of each activity and the immediate predecessors of each of the activities.

Activity	Duration (days)	Immediate predecessors
<i>A</i>	8	–
<i>B</i>	6	–
<i>C</i>	1	<i>A</i>
<i>D</i>	2	<i>B</i>
<i>E</i>	2	<i>A</i>
<i>F</i>	1	<i>C</i>
<i>G</i>	4	<i>D, E</i>
<i>H</i>	1	<i>F, G</i>
<i>I</i>	2	<i>H</i>

The durations are recorded on the activity network that is drawn from the precedence table. It is usual to record the name of the activity followed by a comma and then the duration of that activity on the edge that represents it.

For example, Activity *D* would be labelled *D, 2*.

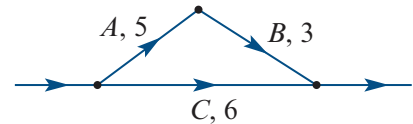
The activity network for this project is shown on the following page.



Float times

The diagram below shows a small section of an activity network. Three activities are shown, A , B and C , along with their individual durations, in hours.

Activity A and B form a small sequence of activities. Activity B cannot begin until activity A has finished. The minimum time it would take to complete activity A and B would be $5 + 3 = 8$ hours.



Activity C can begin at the same time as activity A , but must be completed no later than activity B . The activity network shows that activity C can be completed at the same time as the sequence of activities $A - B$. Activity C has a duration of 6 hours, which is 2 hours less than the time for the sequence $A - B$ and so there is some flexibility around when activity C could start. This value is called the **float time** for activity C . The float time is sometimes called the *slack time*.

The flexibility around the timing of activity C is shown in the diagram below.

	A	A	A	A	A	B	B	B
Start at same time	C	C	C	C	C	C	Slack	Slack
Delay C by 1 hour	Slack	C	C	C	C	C	C	Slack
Delay C by 2 hour	Slack	Slack	C	C	C	C	C	C

The five red squares represent the 5 hours it takes to complete activity A . The three green squares represent the 3 hours it takes to complete activity B .

This six yellow squares represent the 6 hours it takes to complete activity C . Activity C does not have to start at same time as activity A because it has some slack time available (2 hours).

Activity C should not be delayed by more than 2 hours because this would cause delays to the project. The next activity requires B and C to be complete before it can begin.

Critical path analysis

Scheduling problems are concerned with minimising the total time it takes to complete a project and so it is essential that all the activities in a project begin at the earliest possible time.

Critical path analysis is the process of analysing the timing of activities of a project to find the overall minimum time to complete the project. It also involves identifying those activities that have float time and those that do not. Activities that have no float time are said to be **critical** activities because if they are delayed, the whole project will not finish in the minimum possible time.

Critical path analysis begins with determining the **earliest starting time (EST)** and the latest finishing time (LFT) for each activity.

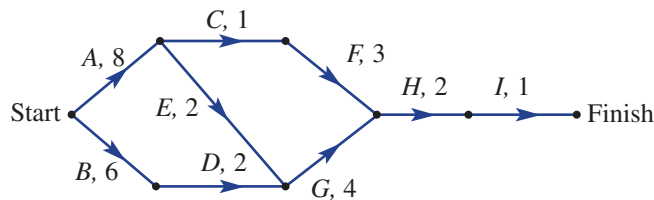
EST values indicate the earliest possible time after the start of the project that a particular activity can begin and still allow the project to be completed in minimum time. For example, an EST of 8 hours means that the activity can begin, at the earliest, 8 hours after the start of the project.

LFT values indicate the latest possible time after the start of the project that a particular activity can finish and still allow the project to be completed in minimum time. For example, an LFT of 14 hours means that at the very latest, the activity can finish 14 hours after the start of the project.

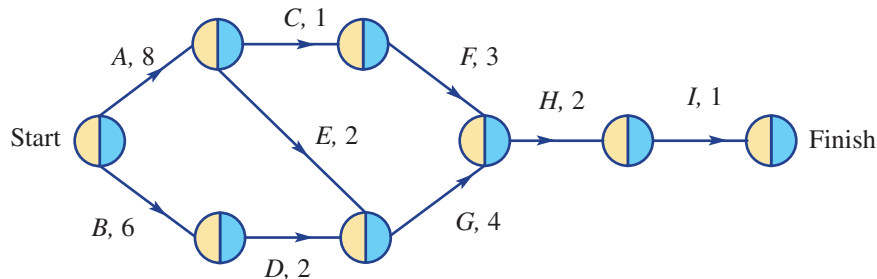
EST values are determined using a process called strongforward scanning.

Forward scanning

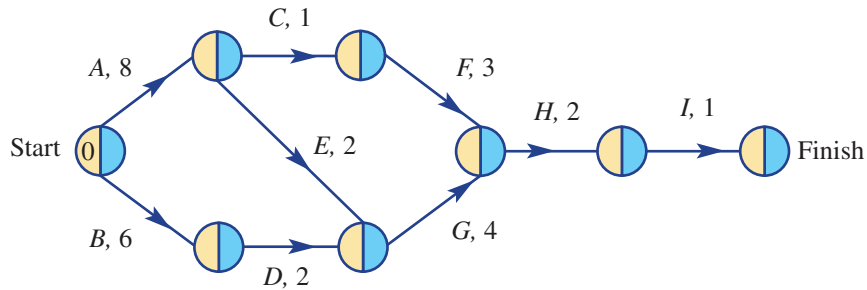
Forward scanning will be demonstrated using the activity network shown below.



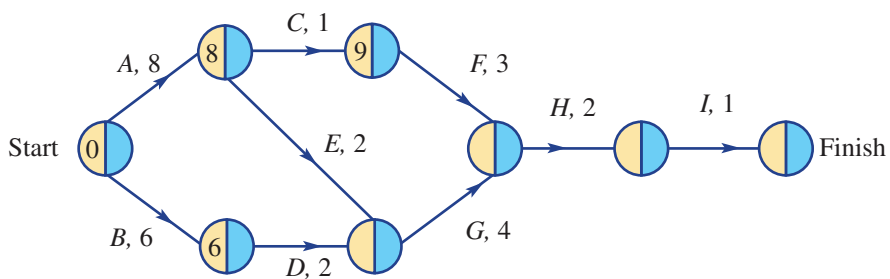
- 1 Draw a circle, split into two cells, next to each vertex of the activity network, as shown. The cells in the circles in the diagram are coloured yellow and blue to help identify which cell we are using. The left cell (yellow) at any vertex will contain the EST for any activity that begins at that vertex.



- 2** Put a zero (0) in the left cell (yellow) of the circle at the start vertex. This represents the start of the entire project. It also represents the EST for activities *A* and *B* because they start at this vertex.

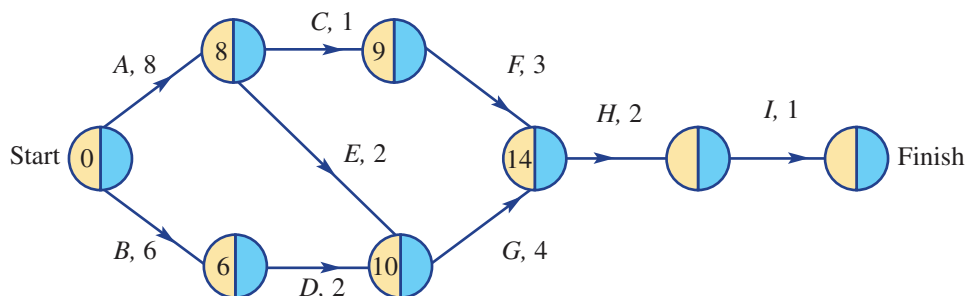


- 3** Each activity will have a box at the vertices at either end of the edge that represents that activity. Take the left cell (yellow) value of the box at the start of the activity, add it to the duration of the activity and write the answer in the left cell (yellow) of the box at the end of the activity. This is the EST for the activity or activities that follow.



Notes:

- a** The cell at the end of activity *A* has value $0 + 8 = 8$. This is the EST for activity *C* and *E*.
 - b** The cell at the end of activity *B* has value $0 + 6 = 6$. This is the EST for activity *D*.
 - c** The cell at the end of activity *C* has value $8 + 1 = 9$. This is the EST for activity *F*.
- 4** If the edges representing more than one activity end at the same vertex, the left cell (yellow) of the box at this vertex must contain the *largest* of the possible values because this activity must wait for all predecessor activities to be completed before it can begin.



Notes:

- 1** The cell at the end of activity *D* and *E* could be:

- from activity *D*: $6 + 2 = 8$
- from activity *E*: $8 + 2 = 10$

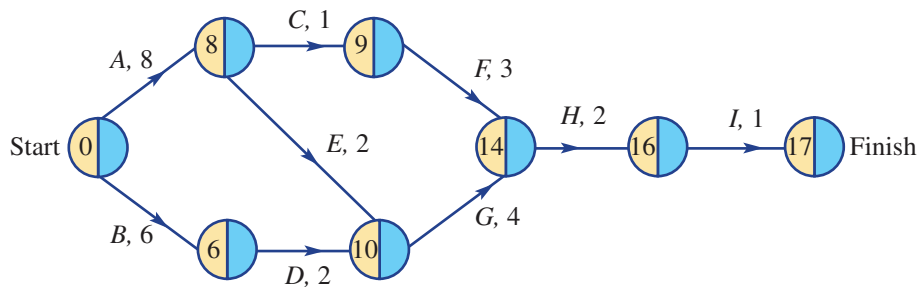
The largest of these options is 10. This is the EST for activity *G*.

- 2** The cell at the end of activity *F* and *G* could be:

- from activity *E*: $9 + 3 = 12$
- from activity *F*: $10 + 4 = 14$

The largest of these options is 14. This is the EST for activity *H*.

- 5** Continue adding the previous EST value to the duration to calculate the following EST values until the final cell is reached.

**Identifying minimum project completion time**

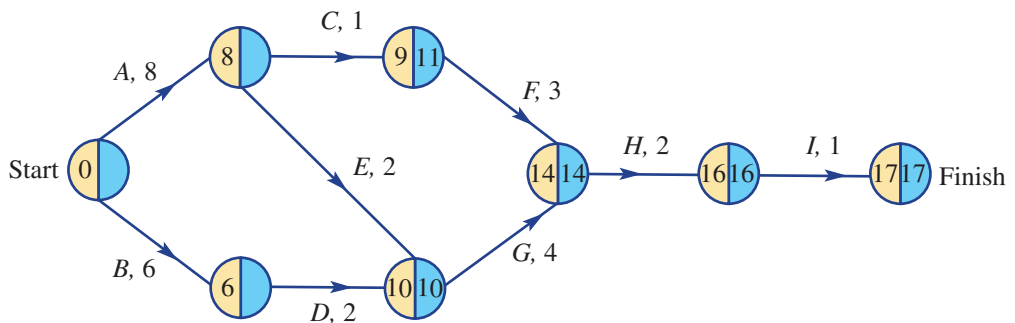
The box for the finish vertex above contains the EST for the next activity, but there is no activity left to begin. So, this value represents the overall minimum completion time for the entire project. This project can be completed in a minimum of 17 days.

To complete the analysis, latest finishing time (LFT) values need to be calculated for each activity.

LFT values are determined using a process called **backward scanning**.

Backward scanning

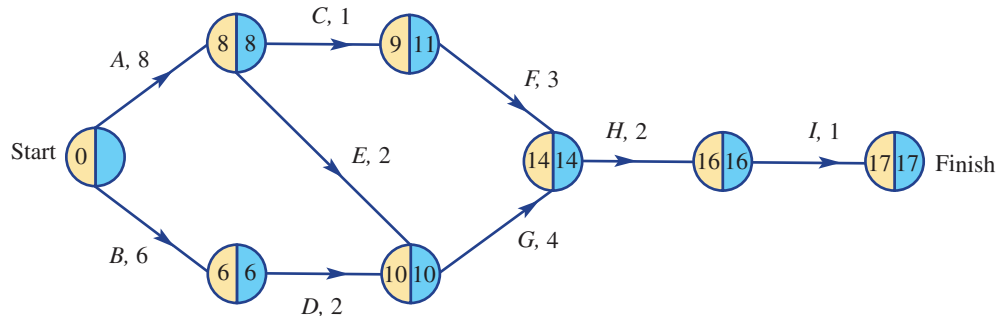
- 1** Copy the minimum time to complete the project into the final right cell shown shaded blue in the diagram.



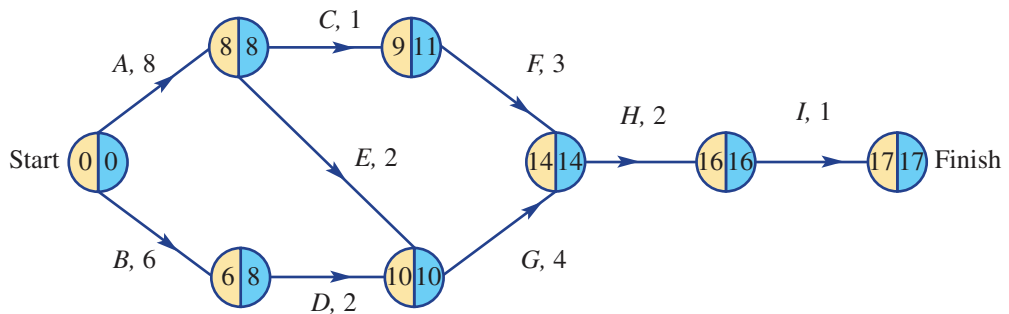
For each activity, take the LFT from the right cell (blue) value of the box at the end of the activity and subtract the duration of the activity. Write the answer in the right cell (blue) of the box at the start of the activity. This will give the LFT for any activity that ends at this vertex.

Notes:

- 1** The LFT for activity *I* is 17.
 - 2** The cell at the start of activity *I* has value $17 - 1 = 16$. This is the LFT for activity *H*.
 - 3** The cell at the start of activity *H* has value $16 - 2 = 14$. This is the LFT for activity *F* and *G*.
 - 4** The cell at the start of activity *F* has value $14 - 3 = 11$. This is the LFT for activity *C*.
 - 5** The cell at the start of activity *G* has value $14 - 4 = 10$. This is the LFT for activity *D*.
- 2** If the edges representing more than one activity start at the same vertex, the right cell (blue) of the box at this vertex must contain the smallest of the possible values. This ensures that the longest of the activities that follow this vertex will have time to be completed.

**Notes:**

- 1** The values in the cell at the start of activity *C* and *E* to be considered are:
 - from activity *C*: $11 - 1 = 10$
 - from activity *E*: $10 - 2 = 8$
 The smallest of these options is 8. This is the LFT for activity *A*.
- 3** Complete the backward scanning for all remaining activities.

**Notes:**

- 1** The cell at the start of activity *D* has value $10 - 2 = 8$. This is the LFT for activity *B*.
- 2** The cell at the start vertex has value, either:
 - from activity *A*: $8 - 8 = 0$
 - from activity *B*: $8 - 6 = 2$
 The smallest of these options is 0.
- 3** The box at the start vertex will always contain a zero in both the left (yellow) and right (blue) cells.

Determining latest starting time (LST)

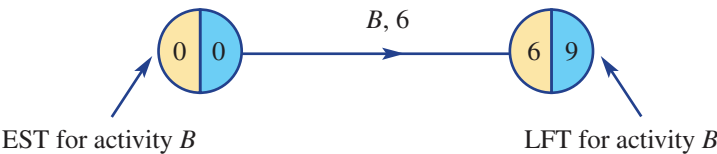
The forward scanning process identified the EST, earliest starting time, for each activity.

The backward scanning process identified the LFT, latest finishing time, for each activity.

The **latest starting time (LST)** for an activity is the latest possible time it could start after the beginning of the project and still allow the project to be completed in minimum time.

The LST for an activity is determined by a simple calculation:

$LST = LFT - \text{duration}$. Activity *B* from the completed forward and backward scanning process above is shown in the diagram below.



- The EST for activity *B* is in the left cell (yellow) in the box at the vertex where activity *B* begins.
- The LFT for activity *B* is in the right cell (blue) in the box at the vertex where activity *B* ends.
- The LST for activity *B* = $LFT - \text{duration}$
 $= 9 - 6$
 $= 3 \text{ days}$

Activity *B* must be completed, at the latest, after 9 days. Since it has a duration of 6 days, it can begin 3 days after the start of the project and still finish in time. The latest time it can start (LST) is 3 days.

The LST for all activities in the project can be found using similar calculations, the results of which are shown in the table below.

$LST\ A = 8 - 8 = 0$

$LST\ B = 8 - 6 = 2$

$LST\ C = 11 - 1 = 10$

$LST\ D = 10 - 2 = 8$

$LST\ E = 10 - 2 = 8$

$LST\ F = 14 - 3 = 11$

$LST\ G = 14 - 4 = 10$

$LST\ H = 16 - 2 = 14$

$LST\ I = 17 - 1 = 16$

Activity	Duration	EST	LFT	LST
A	8	0	8	0
B	6	0	8	2
C	1	8	11	10
D	2	6	10	8
E	2	8	10	8
F	3	9	14	11
G	4	10	14	10
H	2	14	16	14
I	1	16	17	16

Determining float time

Some of the activities from the project on the previous page have the same EST and LST values in the table. The earliest start time and the latest start time are exactly the same, which means that there really is only one start time that is possible in order to make sure the project is completed in minimum time.

These activities are the critical activities described earlier. They have no flexibility in their starting time and so have a float time of zero. The critical activities in this project are *A, E, G, H* and *I*.

The non-critical activities *B, C, D, F* have float time that is not zero. Float time can easily be calculated using this rule.

$$\text{Float time} = LST - EST$$

$$\text{Float time for activity } B = 2 - 0 = 2$$

$$\text{Float time for activity } C = 10 - 8 = 2$$

$$\text{Float time for activity } D = 8 - 6 = 2$$

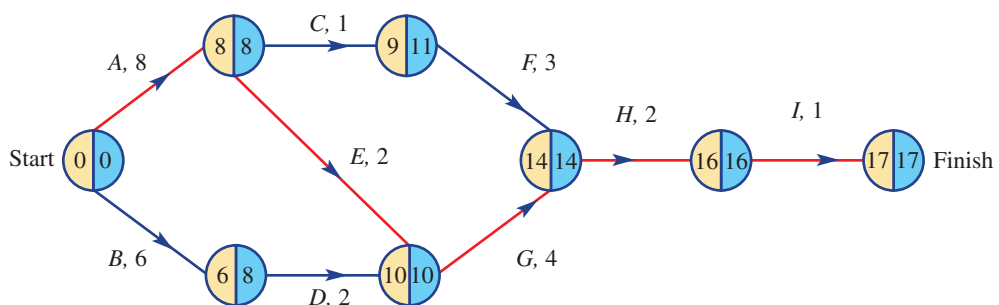
$$\text{Float time for activity } F = 11 - 9 = 2$$

Identifying the critical path

The critical activities have already been identified as those activities that have zero float time; that is, those activities that have equal EST and LST.

The critical path through the network is the sequence of these critical activities, from the start of the project through to the finish.

The critical path has been highlighted in red on the diagram below.



In most projects, there will be a single critical path from start to finish, but it is possible for a project to have a critical path that branches.


Example 9 Critical path analysis

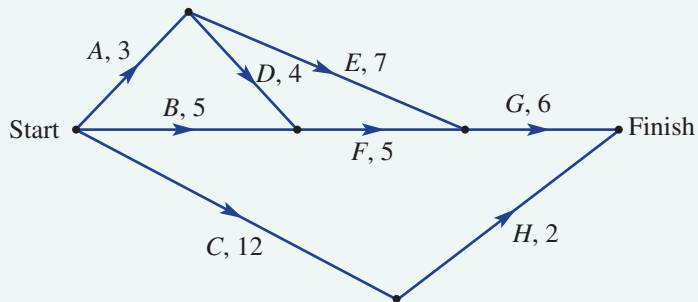
A project has eight activities as shown in the precedence table below.

Activity	Duration (weeks)	Immediate predecessors
A	3	—
B	5	—
C	12	—
D	4	A
E	7	A
F	5	B, D
G	6	E, F
H	2	C

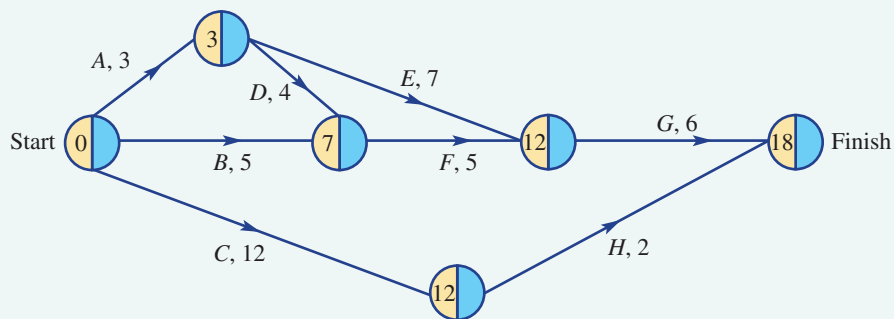
- a** Draw an activity network.
- b** Complete the forward scanning process to identify the minimum time it will take to complete this project.
- c** Complete the backward scanning process.
- d** What is the earliest starting time for activity *E*?
- e** What is the latest starting time for activity *E*?
- f** Identify the critical path for this project.
- g** The person responsible for completing activity *E* falls sick three weeks into the project. If he will be away from work for two weeks, will this cause the entire project to be delayed?

Solution

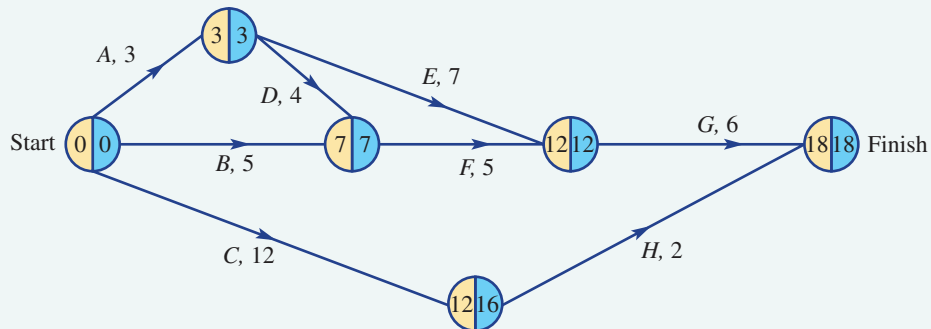
a



b The forward scanning process results are shown in the diagram.

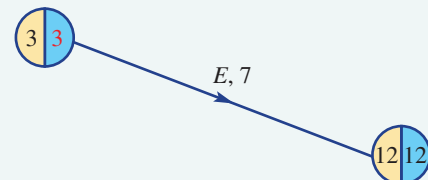


c The backward scanning process results are shown in the diagram.



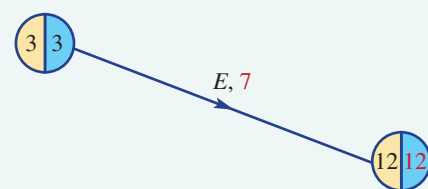
d The EST for activity $E = 3$.

Earliest starting time is the left cell value of the box at the beginning of the activity.



e The LST for activity $E = 12 - 7 = 5$.

Latest starting time is found by subtracting the duration of the activity from the right cell box at the end of the activity.



f The critical path for this project is $A - D - F - G$.

The critical path joins all of the activities that have the same EST and LST, and therefore which have zero float time.

$$\begin{aligned}
 \text{g Float } E &= \text{LST} - \text{EST} \\
 &= 5 - 3 \\
 &= 2 \text{ weeks}
 \end{aligned}$$

The person will be away for two weeks, starting three weeks into the project. This is equal to the float time for activity *E*, and so delaying the start of activity *E* until the person comes back to work will not affect the overall completion time of the project.

Calculate the float for activity *E*. This tells us how long the start of activity *E* can be delayed, without delaying the entire project.

If the float time is more or equal to the delay in the start of activity *E*, the project will not be affected.

Another approach to Scheduling diagrams

Another method for representing a scheduling problem with a network diagram is to record the EST and LST on the diagram rather than EST and LFT. The calculation for calculating LFT is done while completing the diagram. It is the same process that was done with a table previously. This requires a change in notation. The end result is the same. We give an example here illustrating this approach.

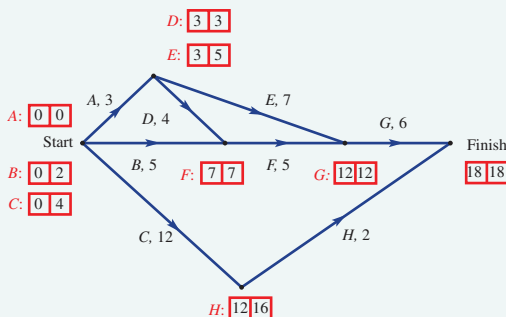
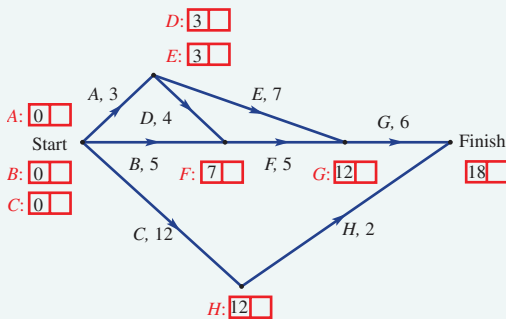
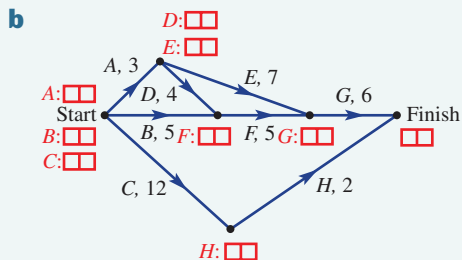
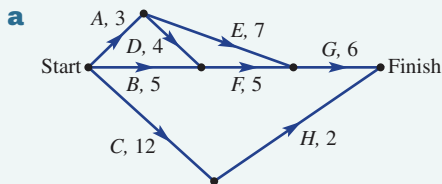


Example 10 Finding the critical path using a diagram with EST and LST shown

A project has eight activities as shown in the precedence table opposite.

- a** Draw an activity network for this project.
- b** Complete the critical path analysis to calculate the EST and LST for each activity.
- c** What is the earliest starting time for activity *H*?
- d** What is the latest starting time for activity *H*?
- e** What is the float time of activity *H*?
- f** Write down the critical path of this project.
- g** What is the minimum time required to complete the project?
- h** The person responsible for completing activity *E* falls sick three weeks into the project. If he will be away from work for two weeks, will this cause the entire project to be delayed?

Activity	Duration (weeks)	Immediate predecessors
<i>A</i>	3	–
<i>B</i>	5	–
<i>C</i>	12	–
<i>D</i>	4	<i>A</i>
<i>E</i>	7	<i>A</i>
<i>F</i>	5	<i>B</i> , <i>D</i>
<i>G</i>	6	<i>E</i> , <i>F</i>
<i>H</i>	2	<i>C</i>

Solution**Explanation**

A, B and C have no predecessors and so can begin at the same time. Continue drawing the network as outlined by the table on the previous page, including arrowheads, activity labels, duration labels and correct immediate predecessors.

Begin by drawing boxes, split into two cells, at the beginning of each activity. Label them with the name of each activity. You must also include a box, split into two cells, at the final vertex where the project finishes.

Use forward scanning to identify the EST for each activity. Activities with no immediate predecessors always have an EST of zero. Add the left cell value at the start of the activity to the duration and write the result in the left cell at the end of the activity. Use the largest of the possibilities if there is more than one activity ending at the same vertex. Identify the minimum project completion time as the left cell value at the finish vertex

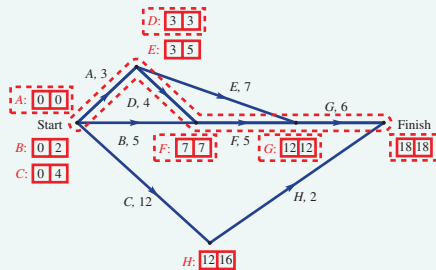
Use backward scanning to identify the LST for each activity. Subtract the duration from the right cell value at the end of the activity and write the result in the right cell at the start of the activity. Use the smallest of the possibilities if there is more than one activity beginning at the same vertex.

c The EST for activity H is 12 weeks.

d The LST for activity H is 16 weeks.

e $\text{Float } H = LST - EST$
 $= 16 - 12$
 $= 4 \text{ weeks}$

f



The critical path for this project is
 $A - D - F - G$.

g 18 weeks

h The person will be away for two weeks, starting three weeks into the project. This is equal to the float time for activity E , and so delaying the start of activity E until the person comes back to work will not affect the overall completion time of the project.

EST values are in the left cell at the start of each activity.

LST values are in the right cell at the end of each activity.

$\text{Float} = LST - EST$

The critical path joins all of the activities that have the same EST and LST, and therefore which have zero float time.

The minimum time required to complete the project is the EST (also, always equal to the LST) at the finish vertex.

If the float time is more or equal to the delay in the start of activity E , the project will not be affected.

Section Summary

Critical path

- ▶ A **critical path** is the longest or equal longest path in an activity network.
- ▶ There can be more than one critical path in an activity network.
- ▶ The critical path is the sequence of activities that cannot be delayed without affecting the overall completion time of the project.

Critical path analysis

Critical path analysis is the process of analysing the timing of activities in a project to determine the critical path of a project.

Perform critical path analysis by: analysis by drawing a box with two cells next to each vertex of the activity network.

Earliest Starting Time, EST

- ▶ Calculate the EST for each activity by forward scanning:

$$\text{EST} = \text{EST of predecessor} + \text{duration of predecessor}$$
- ▶ If an activity has more than one predecessor, the EST is the *largest* of the alternatives.
- ▶ The minimum overall completion time of the project is the EST value at the end vertex.

Latest Finishing Time, LFT

- ▶ Calculate the LFT for each activity by backward scanning
- ▶ Subtract the duration from the right cell value at the end of the activity.
- ▶ Write the result in the right cell at the start of the activity.
- ▶ Use the smallest of the possibilities if there is more than one activity beginning at the same vertex.

EST values are in the left cell at the start of each activity.

LFT values are in the right cell at the end of each activity.

Latest Starting Time, LST

$\text{LST} = \text{LFT} - \text{duration}$. Critical activities will have zero float; that is, $\text{LST} = \text{EST}$.

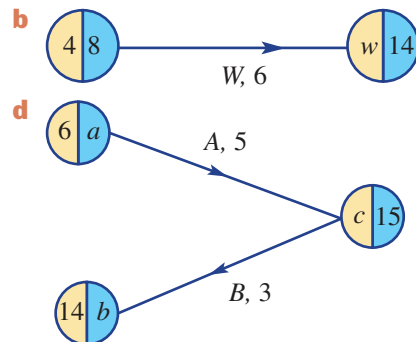
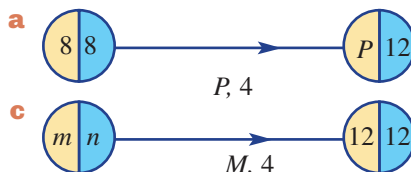
The critical path is the sequence of critical activities through the activity network.

Exercise 10C



Calculations from elements of an activity network

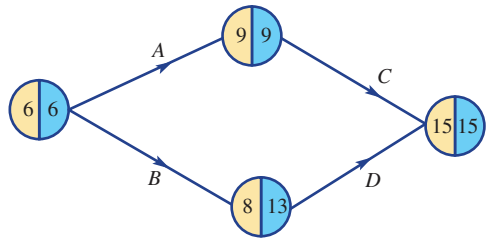
- 1 Write down the value of each pronumeral in the sections of activity networks below.



Interpreting completed forward and backward scanning

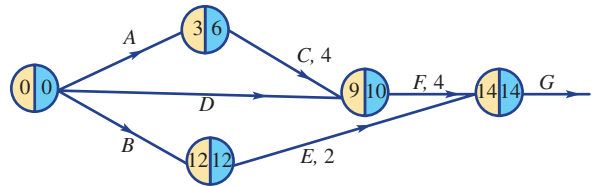
2 Consider the section of an activity network shown in the diagram below.

- What is the duration of activity *A*?
- What is the float time of activity *B*?
- What is the duration of activity *D*?
- What is the latest time that activity *D* can start?
- Write down the critical path through this section of the activity network.



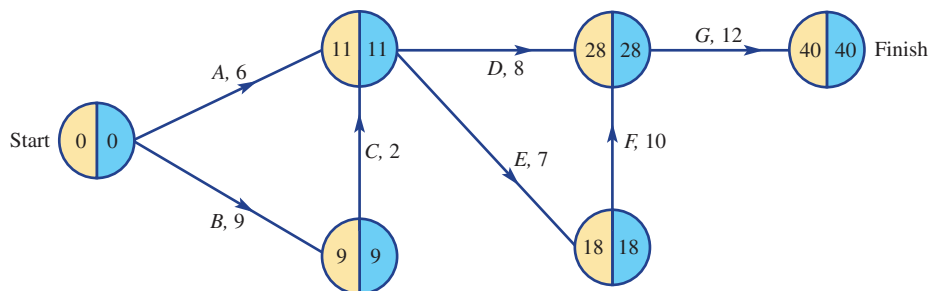
3 Consider the section of an activity network shown in the diagram below.

- What is the duration of activity *B*?
- What is the latest start time for activity *F*?
- What is the earliest time that activity *F* can start?
- What is the float time for activity *F*?
- What is the duration of activity *A*?
- What is the duration of activity *D*?



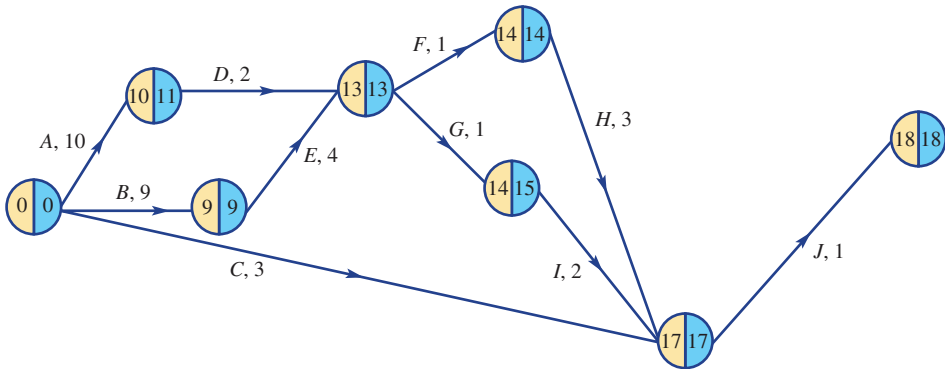
Example 9

4 The activity network below shows the results of forward and backward scanning.



- What is the minimum completion time for this project?
- What is the duration of activity *D*?
- What is the EST of activity *D*?
- What is the LFT of activity *D*?
- Calculate the LST of activity *D*.
 - Explain how you know that activity *D* is not a critical activity.
- There is one other non-critical activity in the project. Which one is it?

- 5 Consider the activity network shown in the diagram below. Earliest starting times and latest finishing times are shown.



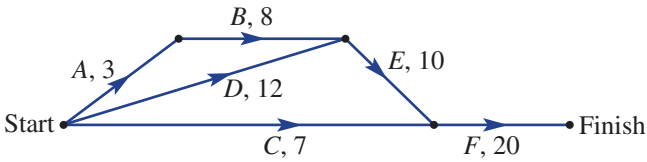
- a Complete the table of durations, EST, LFT, LST and float below.

Activity	Duration	EST	LFT	LST	Float
A	10		11	1	1
B		0	9		
C	3	0		14	14
D	2		13	11	
E	4	9		10	0
F		13	14		0
G	1		15	14	1
H		14		14	
I	2	14	17		1
J	1		18	17	0

- b Write down the critical path for the project.

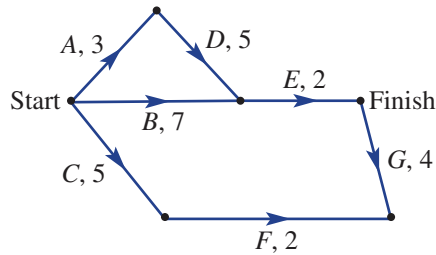
Critical path analysis from a given activity network

- 6 Consider the activity network in the diagram shown opposite.

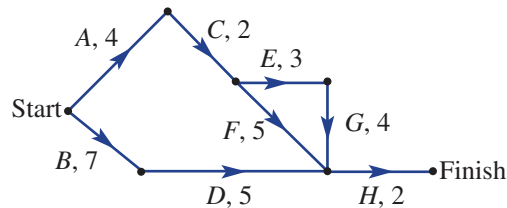


- a Complete forward and backward scanning for this activity network.
- b What is the minimum completion time for this project?
- c Write down the critical path for this project.
- d For each non-critical activity, calculate the:
- i LST
 - ii float time

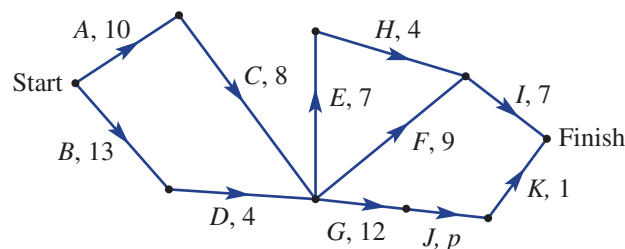
- 7 Consider the following activity network for a project. The duration of each activity is given in the network, in days.



- Determine the earliest start time for activity E .
 - Find the minimum completion time for this project.
 - Write down the critical path for this project.
 - Which activity has a float time of two days?
- 8 Consider the following activity network for a project.



- Write down the three activities that are immediate predecessors of activity H .
 - Determine the earliest start time of activity H .
 - For activity H , the earliest start time and the latest start time are the same. What does this tell us about activity H ?
 - Determine the minimum completion time, in hours, for this project.
 - Which activity could be delayed for the longest time without affecting the minimum completion time of the project?
- 9 Consider the following activity network for a project.



- Determine the earliest start time for the following activities:
 - H
 - I
 - J
- Determine the value of p , in weeks, that would create more than one critical path.
- If the value of p is 3 weeks, what will be the float time, in weeks, of activity G .

- 10** Consider the precedence table for the activities in a project shown on the right.
- a** Draw an activity network for this project.
 - b** Complete the forward and backward scanning for this project.
 - c** What is the shortest time, in weeks, in which this project could be completed?
 - d** Use the activity network and results of scanning to write down the critical path for this project.
 - e** Complete the table below to find the float times for every activity in the project.

Activity	Duration (weeks)	Immediate predecessors
<i>I</i>	2	–
<i>J</i>	3	–
<i>K</i>	5	–
<i>L</i>	4	<i>I</i>
<i>M</i>	8	<i>J, N</i>
<i>N</i>	1	<i>K</i>
<i>O</i>	6	<i>L, M</i>
<i>P</i>	6	<i>J, N</i>
<i>Q</i>	7	<i>J, N</i>
<i>R</i>	5	<i>K</i>
<i>S</i>	1	<i>O</i>
<i>T</i>	9	<i>Q, R</i>

Activity	Duration	EST	LFT	LST	Float
<i>I</i>	2				
<i>J</i>	3				
<i>K</i>	5				
<i>L</i>	4				
<i>M</i>	8				
<i>N</i>	1				
<i>O</i>	6				
<i>P</i>	6				
<i>Q</i>	7				
<i>R</i>	5				
<i>S</i>	1				
<i>T</i>	9				

- f** Use the table to verify that you identified the critical path correctly.

Paper 1-style multiple-choice questions

- 11** Activities *I* and *J* are the critical activities for a project.

Activity	Duration (weeks)	Immediate predecessors
<i>I</i>	4	–
<i>J</i>	8	<i>I</i>

What are the earliest starting time (EST) and latest starting time (LST) for Activity *J*?

- A** EST: 4; LST 4 **B** EST: 4; LST 8
C EST: 8; LST 8 **D** EST: 8; LST 12

- 12** The duration, in minutes, of all activities in a project are shown.

Activity	I	J	K	L	M	N	O
Duration	24	32	54	40	18	11	22

The critical path for the project is I – K – L – O.

What is the earliest completion time for the project if it starts at 9:00 a.m.?

- A** 12:30 p.m. **B** 11:20 a.m. **C** 3:30 p.m. **D** 3:45 p.m.

- 13** The table shows information for a project with four activities.

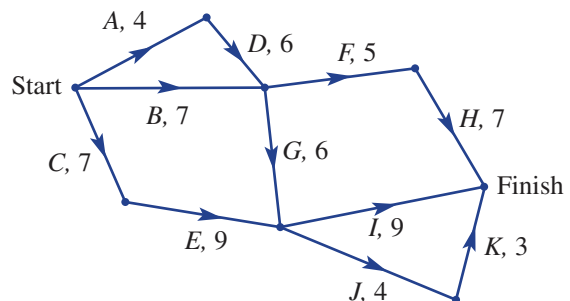
Activity	Duration (minutes)	Prerequisite	EST	LFT	LST
<i>A</i>	8	–	0	8	0
<i>B</i>	6	–	0	9	3
<i>C</i>	1	<i>A</i>	8	9	8
<i>D</i>	2	<i>B, C</i>	9	11	9

What is the float time for activity *B*, in minutes?

- A** 0 **B** 1 **C** 3 **D** 7

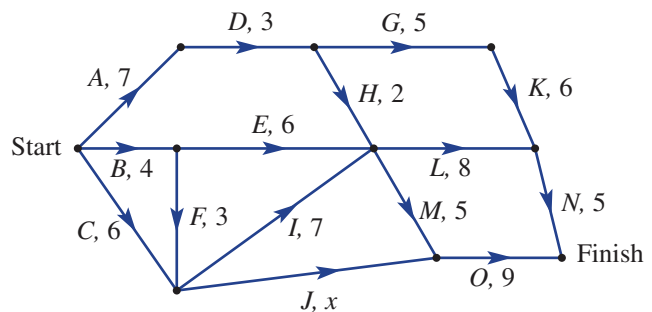
Use the following information to answer Questions 14, 15 and 16

The directed network opposite shows the sequence of eleven activities that are needed to complete a project. The time, in days, that it takes to complete each activity is also shown.



- 14** The earliest starting time, in days, for activity J is
A 13 **B** 14 **C** 15 **D** 16
- 15** The number of activities that have exactly two immediate predecessors is
A 1 **B** 2 **C** 3 **D** 4
- 16** How many of these activities could be delayed without affecting the minimum completion time of the project?
A 3 **B** 4 **C** 5 **D** 6

- 17** The directed graph opposite shows the sequence of activities required to complete a project. The time taken to complete each activity, in weeks, is also shown.



The minimum completion time for this project is 28 weeks. The time taken to complete activity J is labelled x . The maximum value of x is

- A** 12 **B** 10 **C** 8 **D** 4
- 18** A project consists of ten activities, A to J . The table below shows the immediate predecessor(s) and earliest start time, in days, of each activity.

Activity	Immediate predecessors	Earliest starting time
A	—	0
B	—	0
C	—	0
D	A	6
E	B	5
F	B	5
G	C	4
H	D, E	13
I	F, G	14
J	H, I	25

It is known that activity H has a completion time of ten days. The project can still be completed in minimum time if activity D is delayed. The maximum length of the delay for activity D is

- A** one day **B** two days **C** six days **D** eight days

10D Applications of critical path analysis

To complete your study of problem planning and scheduling problems, you will see some realistic applications of critical path analysis.



Example 11 Application of critical path analysis

Linda is building a new house and has decided to manage the project herself.

She has decided on the major activities involved in this build and has been given some advice regarding the length of time each activity is expected to take.

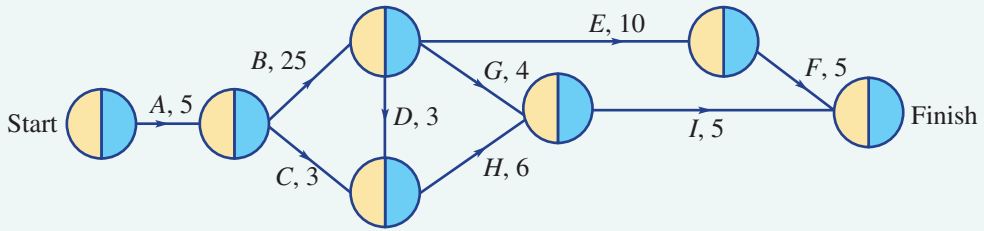
These activities and their durations are shown in the table below.

Activity	Description	Duration (days)	Immediate predecessors
<i>A</i>	Preparing site and laying slab	5	-
<i>B</i>	Constructing frame and roof	25	<i>A</i>
<i>C</i>	Preparing floor	3	<i>A</i>
<i>D</i>	Finalise utility installation locations	3	<i>B</i>
<i>E</i>	Landscaping gardens	10	<i>B</i>
<i>F</i>	Installing plants and lawn	5	<i>E</i>
<i>G</i>	Installing electrical	4	<i>B</i>
<i>H</i>	Installing plumbing	6	<i>D, C</i>
<i>I</i>	Finishing off ready to move in	5	<i>G, H</i>

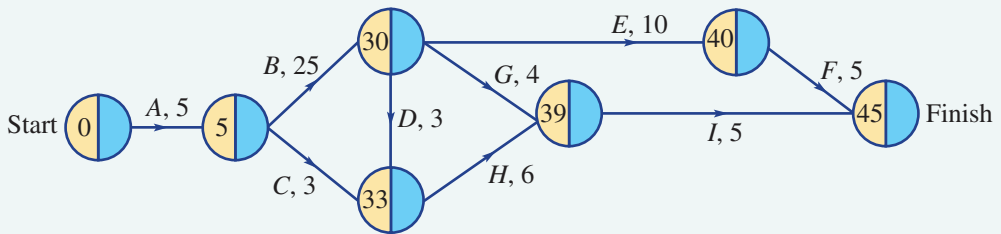
- a** Construct an activity network for this project.
- b** Apply the forward scanning technique to determine the shortest time in which Linda can expect completion of her house.
- c** Apply the backward scanning technique and then complete the following.
 - i** Construct a table that shows the EST, LFT, LST and float for each activity.
 - ii** Which of the activities will cause a delay in the entire project if they take longer than expected?
 - iii** Write down the critical path of this project.
- d**
 - i** If the electrical work takes two days longer than expected, what effect will this have on the overall project?
 - ii** Assuming that all other activities have no delay, what is the maximum delay possible for the floor preparation so that the project as a whole is not delayed?

Solution

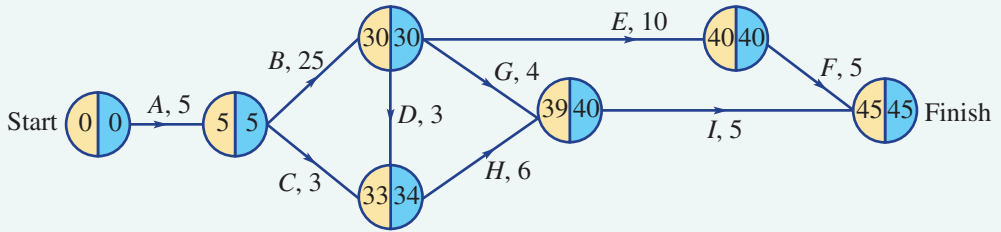
a



b



c



i

Activity	Duration (days)	EST	LFT	LST	Float
A	5	0	5	0	0
B	25	5	30	5	0
C	3	5	34	31	26
D	3	30	34	31	1
E	10	30	40	30	0
F	5	40	45	40	0
G	4	30	40	36	6
H	6	33	40	34	1
I	5	39	45	40	1

- ii** The activities with zero float time will cause a delay to the entire project if they take longer than expected. These activities are preparing site and laying slab (A), constructing frame and roof (B), landscaping gardens (E) and installing the plants (F).

- iii** The critical path of this project is A – B – E – F.

- d i** The activity installing electrical is activity G. Since activity G is not on the critical path, a delay will not necessarily affect the completion time of the project. Two days is less than the float time of activity G and so there will be no effect.
- ii** The activity preparing floor is activity C. Since C is not on the critical path, it may be delayed without affecting the completion time of the project. The float time for activity C is 26 days and so it may be delayed by a maximum of 26 days without delaying the project as a whole.

Exercise 10D

Example 11

- 1** Sharon's car washing business offers a premium service that involves five activities. These activities, their durations and their immediate predecessors are shown in the table below.

Activity	Description	Duration (minutes)	Immediate predecessors
<i>A</i>	Wash the car	15	-
<i>B</i>	Dry the car	8	<i>A</i>
<i>C</i>	Wax the car	20	<i>B</i>
<i>D</i>	Clean the interior	35	<i>B</i>
<i>E</i>	Polish and shine the car	30	<i>C</i>

- a** Construct an activity network for this project.
- b** Apply the forward scanning technique to determine the shortest time in which the premium service can be expected to be completed.
- c** Apply the backward scanning technique and then answer the following questions.
- i** Construct a table that shows the EST, LFT, LST and float for each activity.
 - ii** Which of the activities will not cause the premium service to take longer than expected if they were delayed?
 - iii** Write down the critical path of this project.
- d** The person responsible for cleaning the interior of the car has been delayed while working on another car. What is the latest time, in minutes after the project begins, that she can start the interior cleaning and still finish cleaning the car on time?

- 2** Anthony is creating a robot for a university project. The activities required to design and build the robot are shown in the table below, along with their duration in days and the immediate predecessor for each activity.

Activity	Description	Duration (weeks)	Immediate predecessors
<i>A</i>	Research robot design and control	5	–
<i>B</i>	Design the internal electronics	8	<i>A</i>
<i>C</i>	Design the remote control	3	<i>A</i>
<i>D</i>	Construct and assemble the robot	15	<i>B</i>
<i>E</i>	Write the code to control the robot	10	<i>B</i>
<i>F</i>	Construct and program the remote control	6	<i>C</i>
<i>G</i>	Debug the code to control the robot	4	<i>E</i>
<i>H</i>	Install the software	1	<i>D, F</i>
<i>I</i>	Test the robot	3	<i>G, H</i>

- a** Construct an activity network for this project.
- b** Apply the forward scanning technique to determine the shortest time in which Anthony can expect to create his robot.
- c** Apply the backward scanning technique and then answer the following questions.
 - i** Construct a table that shows the EST, LFT, LST and float for each activity.
 - ii** Write down the critical path of this project.
- d** Use the information in the table from part **ci** to describe and explain what would happen if Anthony took:
 - i** 3 weeks to research robot design and control instead of 5
 - ii** 10 weeks to construct and program the remote control instead of 6
 - iii** 20 weeks to construct and assemble the robot instead of 15.

10E Altering the duration of an activity

Learning intentions

- ▶ To change the duration of some activities to reduce the completion time of a project.
- ▶ To minimise the cost of reducing the duration of activities and to achieve the maximum reduction.

Altering completion times

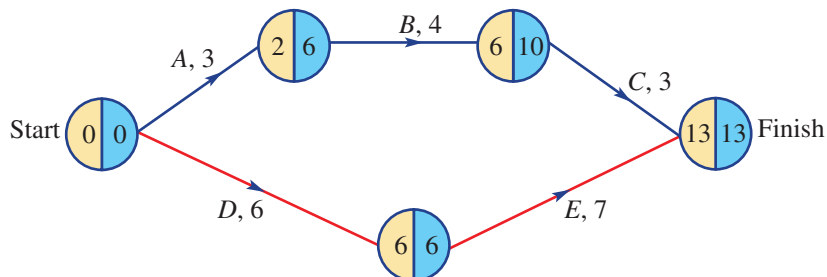
The minimum time it takes to complete a project depends upon the time it takes to complete the individual activities of the project, and upon the predecessors each of the activities have. Critical path analysis can be completed to find the overall minimum completion time.

Sometimes, the managers of a project might arrange for one or more activities within the project to be completed in a shorter time than originally planned. Changing the conditions of an activity within a project, and recalculating the minimum completion time for the project, is called **crashing**.

An individual activity could be crashed by employing more staff, sourcing alternate materials or simply because weather or other factors allow the activity to be completed in a shorter time than usual.

A simple crashing example

A simple activity network is shown in the diagram below. The forwards and backwards scanning processes have been completed and the critical path has been determined. The critical path is shown in red on the diagram.



The minimum time for completion is currently 13 hours. In order to reduce this overall time, the manager of the project should try to complete one, or more, of the activities in a shorter time than normal. Reducing the time taken to complete activity A, B or C would not achieve this goal, however. These activities are not on the critical path and so they already have slack time. Reducing their completion time will not shorten the overall time taken to complete the project.

Activities D and E, on the other hand, lie on the critical path. Reducing the duration of these activities will reduce the overall time for the project. If activity D was reduced in time to 4 hours instead, the project will be completed in 11, not 13, hours.

Crashing with cost

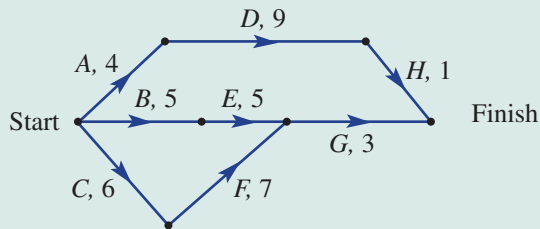
Shortening the completion time for any individual activity could result in an extra cost for the project. In the simple example on the previous page, the cost of reducing the completion time of activity *D* by 1 hour is \$150, while the cost of reducing the completion time of activity *E* by 1 hour is \$18.

Clearly it is best to reduce the completion time, or crash, the activity that will cost the least.



Example 12 Crashing one activity with cost

The directed network below shows the sequence of 8 activities that are needed to complete a project. The time, in days, that it takes to complete each activity is also shown.



- Write down the critical path for this project.
- What is the minimum completion time for the project?

Activity *F* can be reduced by a maximum of 3 days at a cost of \$100 per day.

- What is the new minimum completion time for the project?
- What is the minimum cost that will achieve the greatest reduction in time taken to complete the project?

Solution

a

Path	Duration(days)
<i>A – D – H</i>	14
<i>B – E – G</i>	13
<i>C – F – G</i>	16

The critical path is *C – F – G*.

Explanation

In crashing problems, we first need to identify the critical path, or paths. We will do this by remembering that a critical path is the longest or equal longest path in the activity network. Using this method, set up a table, list all possible paths from *Start* to *Finish* for the directed network and calculate the length of each path. The critical path is the path with the longest time from *Start* to *Finish*.

Solution**b** 16 days

c

<i>Path</i>	<i>Duration (days)</i>	<i>New duration with maximum reduction (F by 3)</i>
$A - D - B$	14	14
$B - E - G$	13	13
$C - F - G$	16	13

The new minimum completion time for the project is 14 days.

d Activity F originally took 7 days to complete. It can be crashed, which means activity F may be reduced by a maximum of 3 days, to result in a completion time of 4 days. It is possible to choose to reduce activity F by 0, 1, 2 or 3 days. Reducing activity F by the maximum 3 days would result in the original critical path being reduced from a total of 16 days, down to 13 days. Considering there is a cost of \$100 per day, this is **not** a desirable outcome; crashing activity F by 3 days results in a new critical path $A - D - H$ with a total completion time of 14 days. If we crash activity F by 2 days only, we create 2 equal critical paths requiring 14 days to complete the project. The minimum cost is \$200.

Explanation

Write the duration of the critical path identified in the previous part.

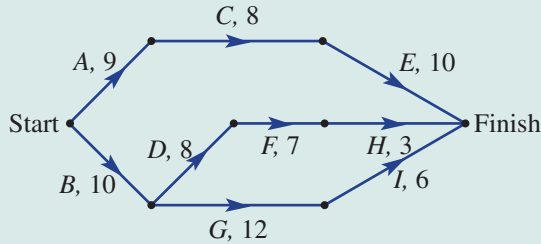
Crash all possible activities by the maximum reduction. Add a new column to the summary table to get an overview of the new duration of each path. This may result in a new critical path. Consider the cost of crashing and whether it is worth applying the maximum reduction.

Reducing activity F by 2 days allows us to reduce the overall completion time of the project at minimum cost.



Example 13 Crashing multiple activities with cost

The directed network below shows the sequence of nine activities that are needed to complete a project. The time, in days, that it takes to complete each activity is also shown.



The minimum completion time for the project is 28 days. It is possible to reduce the completion time for activities *B*, *E*, *G*, *H* and *I*. The completion time for each of these five activities can be reduced by a maximum of two days.

- a** What is the new minimum completion time, in days, that the project could take?

The reduction in completion time for each of these five activities will incur an additional cost. The table opposite shows the five activities that can have their completion times reduced and the associated daily cost, in dollars.

Activity	Daily cost(\$)
<i>B</i>	1500
<i>E</i>	2000
<i>G</i>	700
<i>H</i>	900
<i>I</i>	800

- b** What is the minimum cost that will achieve the greatest reduction in time taken to complete the project?

Solution

- a** List all possible paths from *Start* to *Finish*, including the completion time of each. Crash all activities by their maximum reduction. Identify the new critical path (path with the longest completion time) after the maximum reductions are applied.

Path	Duration (days)	New duration after maximum reduction (<i>B</i> , <i>E</i> , <i>G</i> , <i>H</i> , <i>I</i> by 2)
<i>A</i> – <i>C</i> – <i>E</i>	27	25
<i>B</i> – <i>D</i> – <i>F</i> – <i>H</i>	28	24
<i>B</i> – <i>G</i> – <i>I</i>	28	22

A – *C* – *E* is the new critical path with a duration of 25 days.

Solution

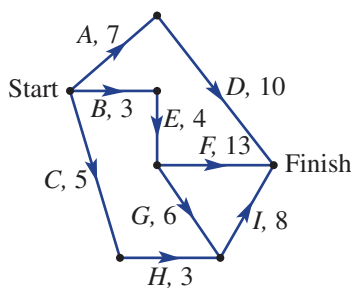
b Steps to determine the minimum cost:

- 1 Begin with the new critical path $A - C - E$. The reduction of activity E must occur to achieve the new minimal completion.
- 2 Consider the path $B - D - F - H$. It has a completion time of 28 days and must be reduced to 25 days to equal the critical path of $A - C - E$. There are two options; reduce B by 2 and H by 1 or reduce B by 1 and H by 2. From the table, it is more expensive per day to reduce B than H ; however, by choosing to reduce B this will also reduce the completion time of the final path $B - G - I$, which is more cost effective. Reducing activity B reduces the completion time of *two* different paths. So reduce activity B by 2 and H by 1 day to reduce the overall completion time of $B - D - F - H$ down to 25 days.
- 3 The final path $B - G - I$ has already been reduced by 2 days due to the reduction of activity B previously chosen. One more day must be reduced for this path to equal the critical path. Activity G has a lower cost of reduction than activity I per day, so include this in your calculation.
- 4 Calculate your total cost of crashing.

$$\begin{aligned}
 \text{Cost of crashing} &= E \text{ by 2 days} + B \text{ by 2 days} + H \text{ by 1 day} + G \text{ by 1 day} \\
 &= 2000 \times 2 + 1500 \times 2 + 900 + 700 \\
 &= \$8600
 \end{aligned}$$

Exercise 10E

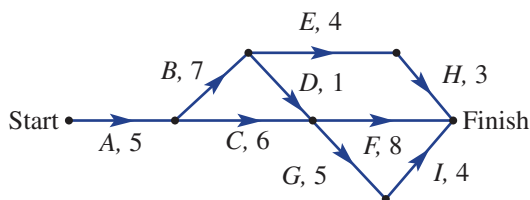
- 1 The activity network for project is shown in the diagram below. The duration for each activity is in hours.



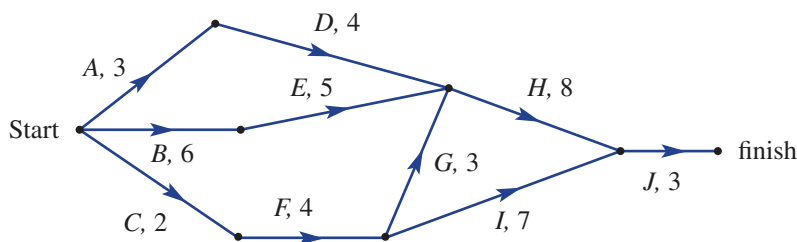
- a List all four paths from the Start to the Finish of the project, with their respective completion times.
- b Identify the critical path and the minimum completion time for the project.
- c If Activity E is reduced by 3 hours, identify the new minimum completion time for this project.

Example 12

- 2** The directed network below shows the sequence of 8 activities that are needed to complete a project. The time, in days, that it takes to complete each activity is also shown.

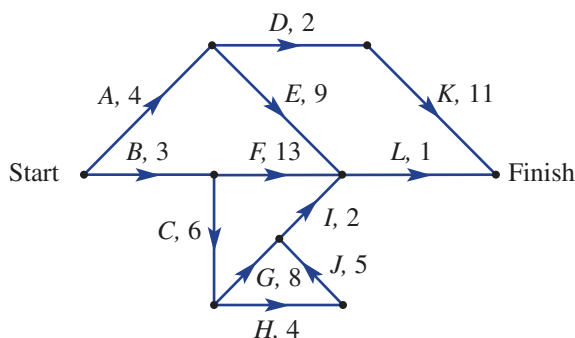


- Write down the critical path for this project.
 - What is the minimum completion time for the project?
Activity *B* can be reduced by a maximum of 3 days at a cost of \$100 per day.
 - What is the new minimum completion time for the project?
 - What is the minimum cost that will achieve the greatest reduction in time taken to complete the project?
- 3** The activity network for project is shown in the diagram below. The duration for each activity is in hours.



- Identify the critical path for this project.
- What is the maximum number of hours that the completion time for activity *E* can be reduced without changing the critical path of the project?
- What is the maximum number of hours that the completion time for activity *H* can be reduced without affecting the critical path of the project?
- Every activity can be reduced in duration by a maximum of 2 hours. If every activity was reduced by the maximum amount possible, what is the minimum completion time for the project?

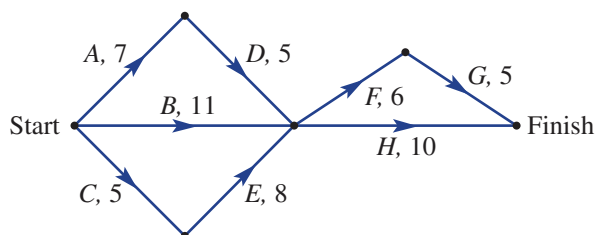
- 4 The activity network for a project is shown in the diagram on the below. The duration for each activity is in hours.



- How many activities could have their completion time increased by 4 hours without altering the minimum completion time?
- If the project is to be crashed by reducing the completion time of one activity only, what is the minimum time, in hours, that the project can be completed in?
- Activity G can be reduced in time at a cost of \$200 per hour. Activity J can be reduced in time at a cost of \$150 per hour. What is the cost of reducing the completion time of this project as much as possible?

Example 13

- 5 The directed network below shows the sequence of 8 activities that are needed to complete a project. The time, in days, that it takes to complete each activity is also shown.



The minimum completion time for the project is 24 days. It is possible to reduce the completion time for activities *D*, *E* and *H*. The completion time for each of these five activities can be reduced by a maximum of two days.

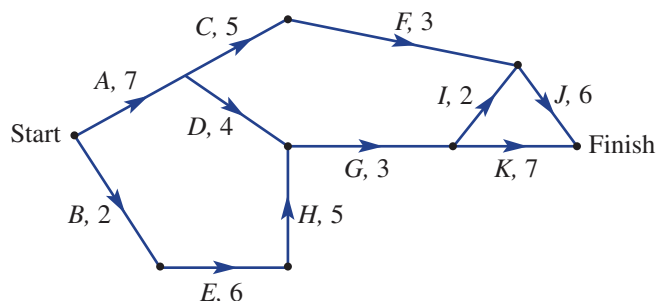
- What is the new minimum completion time, in days, that the project could take?

The reduction in completion time for each of these three activities will incur an additional cost. The table shows the three activities that can have their completion times reduced and the associated daily cost, in dollars.

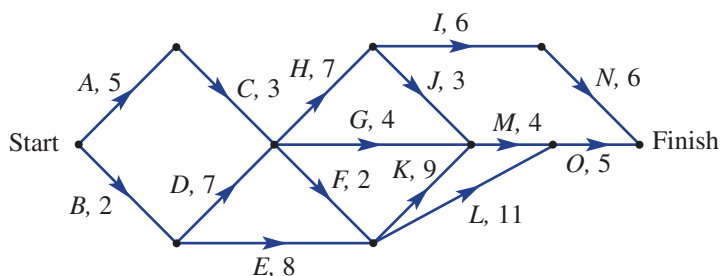
Activity	Daily cost(\$)
<i>D</i>	170
<i>E</i>	350
<i>H</i>	200

- What is the minimum cost that will achieve the greatest reduction in time taken to complete the project?

- 6** The directed network below shows the sequence of 11 activities that are needed to complete a project. The time, in days, that it takes to complete each activity is also shown.

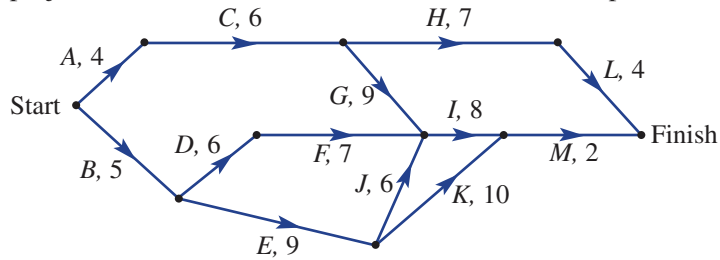


- Which activities are immediate predecessors to activity *G*?
 - Which activities, if crashed, would create more than one critical path?
 - The project could finish earlier if some activities were crashed. Five activities, *B*, *E*, *G*, *H* and *I*, can all be reduced by one hour. The cost of this crashing is \$150 per hour.
 - What is the minimum number of hours in which the project could now be completed?
 - What is the minimum cost of completing the project in this time?
- 7** The directed network below shows the sequence of 15 activities that are needed to complete a project at the MCG. The time, in days, that it takes to complete each activity is also shown.



- What is the minimum completion time?
- How many activities are on the critical path?
- How many paths have a completion time of 28 days?

- 10** The directed graph below shows the sequence of activities required to complete a project. All times are in weeks. There is one critical path for this project.



The total completion time of the project can be reduced by four weeks by reducing

- A** activity *B* by four weeks
- B** activity *F* by four weeks
- C** activity *J* by four weeks
- D** activity *I* by three weeks and activity *J* by one week.

Key ideas and chapter summary

Tree	A tree is a connected graph that contains no cycles, multiple edges or loops. A tree with n vertices has $n - 1$ edges.
Spanning tree	A spanning tree is a tree that connects every vertex of a graph. A spanning tree is found by counting the number of vertices (n) and removing enough edges so that there are $n - 1$ edges left that connect all vertices.
Weighted graph	A weighted graph has numbers, called weights, associated with the edges of a graph. The weights often represent physical quantities as additional information to the edge, such as time, distance or cost.
Minimum spanning tree	A minimum spanning tree is a spanning tree for which the sum of the weights of the edges is as small as possible.
Prim's algorithm	Prim's algorithm is an algorithm for determining the minimum spanning tree of a network.
Shortest path	The shortest path through a weighted graph is the path along edges so that the total of the weights of that path is the minimum for that graph. Shortest path problems involve finding minimum distances, costs or times through a network. Shortest paths can be determined by inspection.
Connector problems	A connector problem is a problem where it is important to minimise the total weight of connections between objects or locations. The weights in connector problems can be length, time, cost or other physical quantity. Connector problems are solved by finding the minimum spanning tree for the graph that represents the problem.
Project	A project is a task that involves a number of individual steps or activities.
Activity	An activity is an individual step in the completion of a project.
Project planning	Project planning involves the analysis and organisation of the activities in a project, taking into account the order in which they must be completed and the time it takes to complete each activity.

Activity network	A directed graph that represents the activities in a project is called an activity network . Edges are used to represent the activities, and these are labelled with the name of the activity and the duration. Vertices of an activity network are not labelled, except for the ‘Start’ and ‘Finish’ vertices.
Immediate predecessor	If activity M must be completed before activity N , then activity M is an immediate predecessor of activity N .
Precedence table	A precedence table is a table that shows all of the activities in a project and their immediate predecessors. Precedence tables may also show the duration of each activity.
Scheduling problem	Scheduling problems involve analysis of the precedence relationship between activities of a project and their durations in order to determine the minimum overall time it would take to complete the project.
Weighted precedence table	A weighted precedence table is a precedence table that also contains the durations of the activities.
Float time	Float time is also called slack time. It is the largest amount of time by which the activity can be delayed without affecting the overall completion time of the project. The float time for an activity is the difference between the latest starting time and the earliest starting time of that activity. Float = LST – EST
Critical path	The critical path for a project is the sequence of activities that cannot be delayed without affecting the overall completion time of the project. Activities on the critical path have float times of zero, that is, the EST and LST are the same. A project may have more than one critical path.
Critical path analysis	Critical path analysis is the process of using knowledge of the precedence and duration for each activity to determine the critical path of a project.
Earliest starting time (EST)	The earliest starting time for an activity is the earliest time after the start of the project that an activity can begin. Earliest starting time is referred to as EST.
Latest finishing time (LFT)	The latest finishing time for an activity is the latest time after the start of the project that an activity can finish without affecting the overall completion time of the project. Latest finishing time is referred to as LFT.

Forward scanning

Forward scanning is the process of determining the EST for each activity in a project. The EST of an activity is added to the duration of the activity to determine the EST of the activities that immediately follow. The EST of any activity is equal to the largest forward scanning value determined from all immediate predecessors.

Backward scanning

Backward scanning is a process of determining the LFT for each activity in a project. The duration of an activity is subtracted from the LFT of activities that immediately follow. The LFT of any activity is equal to the smallest backward scanning value determined from all activities that immediately follow that activity.

Minimum completion time

The overall shortest amount of time in which the project can be completed.

Latest starting time (LST)

The **latest starting time** of an activity is the latest time that activity can start without affecting the overall completion time of the project. Latest starting time is referred to as LST.
For any activity, $LST = LFT - \text{duration}$.

Skills checklist

Download this checklist from the Interactive Textbook, then print it and fill it out to check your skills. 

10A

- 1** I can define what a tree is, identify whether a graph is a tree or not and draw possible trees with a given number of vertices. ☐

See Example 1 and Exercise 10A, Question 1

10A

- 2** I can determine and draw the spanning trees contained in a given graph as subgraphs. ☐

See Example 2 and Exercise 10A, Questions 3 and 4

10A

- 3** I can determine the weight of spanning trees contained in a given graph. ☐

See Example 3, Example 3 and Exercise 10A, Question 5

10A

- 4** I can identify a minimum spanning tree contained in a given graph and determine its weight. ☐

See Example 4 and Exercise 10A, Question 7

10A

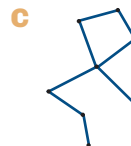
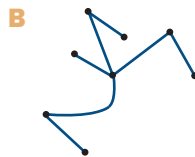
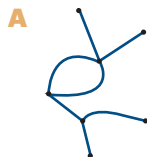
- 5** I can apply the concept of minimum spanning trees to solve practical problems. ☐

See Example 5 and Exercise 10A, Question 6

- 10B** **6** I can identify activities and immediate predecessors from an activity network. ☐
See Example 6 and Exercise 10B, Question 1
- 10B** **7** I can draw an activity network from a precedence table. ☐
See Example 7, Example 8 and Exercise 10B, Question 4
- 10C** **8** I can use forward scanning to determine the earliest starting time of activities in an activity network. ☐
See Example 9 and Exercise 10C, Question 4
- 10C** **9** I can use backward scanning to determine the latest starting time of activities in an activity network. ☐
See Example 9 and Exercise 10C, Question 4
- 10C** **10** I can determine the float time for activities in an activity network. ☐
See Example 9 and Exercise 10C, Questions 4 and 6
- 10C** **11** I can determine the overall minimum completion time for a project using critical path analysis. ☐
See Example 9 and Exercise 10C, Question 4
- 10C** **12** I can determine the critical path for an activity network. ☐
See Example 9 and Exercise 10C, Question 4
- 10D** **13** I can apply critical path analysis to solve problems. ☐
See Example 11 and Exercise 10D, Question 1
- 10E** **14** I can use crashing to reduce the completion time of a project. ☐
See Example 12 and Exercise 10E, Question 2

Multiple-choice questions

- 1** The number of edges of a tree with 12 vertices is
A 9 **B** 10 **C** 11 **D** 12
- 2** Which one of the following graphs is a tree?

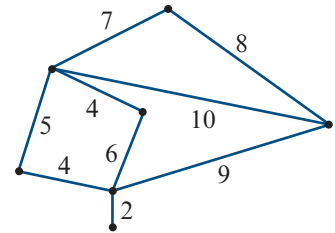


- 3 A graph has 6 vertices and 12 edges. A spanning tree for this graph will have:

A 5 vertices and 12 edges **B** 5 vertices and 5 edges
C 6 vertices and 5 edges **D** 6 vertices and 6 edges

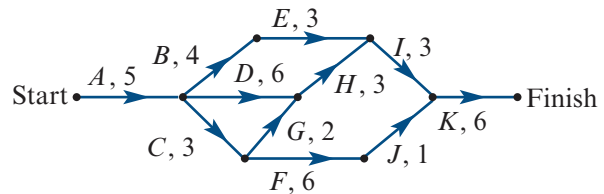
- 4 For the graph shown here, a minimum spanning tree has length:

A 30 **B** 31 **C** 33 **D** 26

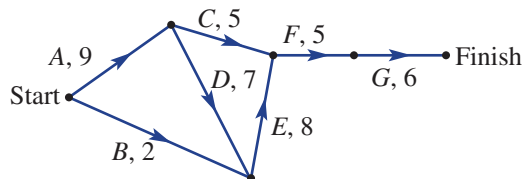


- 5 This activity network is for a project where the component times, in days, are as shown. The critical path for the network of this project is given by:

A A – B – E – I – K **B** A – D – H – I – K
C A – C – G – H – I – K **D** A – C – F – J – K



- 6 The activity network shown represents a project development with activities and their durations (in days) listed on the edges of the graph.

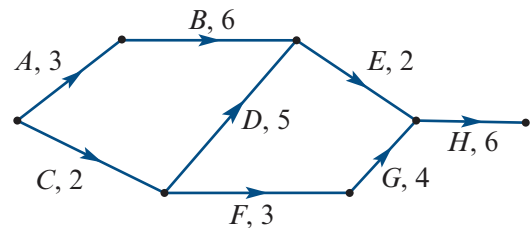


The earliest time (in days) that activity *F* can begin is:

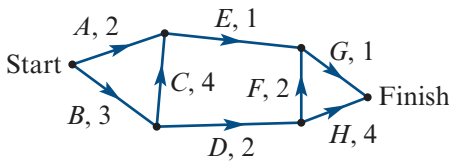
A 12 **B** 14 **C** 22 **D** 24

- 7 The edges in this activity network correspond to the tasks involved in the preparation of an examination. The numbers indicate the time, in weeks, needed for each task. The total number of weeks needed for the preparation of the examination is:

A 14 **B** 15 **C** 16 **D** 17



- 8 This activity network represents a manufacturing process with activities and their duration (in hours) listed on the edges of the graph.
- The earliest time (in hours) after the start of manufacturing that activity G can begin is:

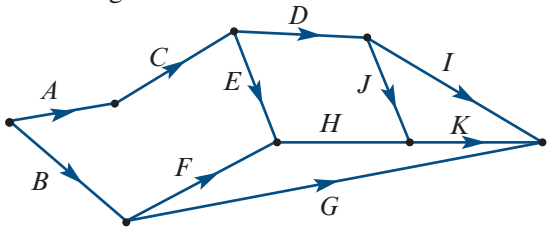


- 9 The table opposite lists the six activities in a project and the earliest start time, in hours, and the predecessor(s) of each task. The time taken for activity F is five hours. Without affecting the time taken for the entire project, the time taken for activity D could be increased by:

A 5 **B** 6 **C** 7 **D** 8

Task	Immediate predecessor	EST
A	–	0
B	–	0
C	A	24
D	B	29
E	C	39
F	D	41
G	E, F	50

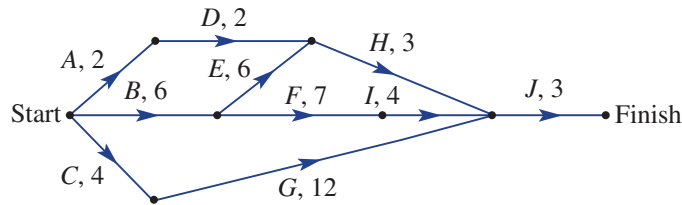
- 10 This network represents a project development with activities listed on the edges of the graph. Which of the following statements must be true?



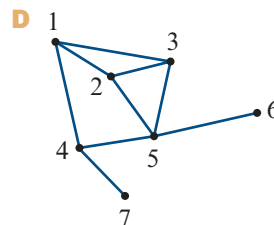
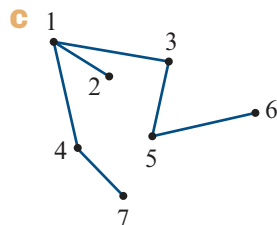
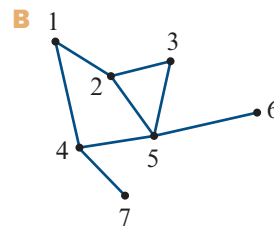
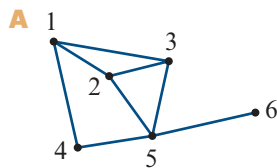
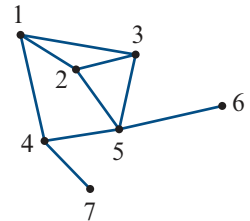
- A** A must be completed before F can start.
B E and F must start at the same time.
C E and F must finish at the same time.
D E cannot commence while A is still taking place.

Use this information to answer Questions 11 to 14.

The activity network below shows the activities required to complete a particular project and the durations, in hours, of those activities.

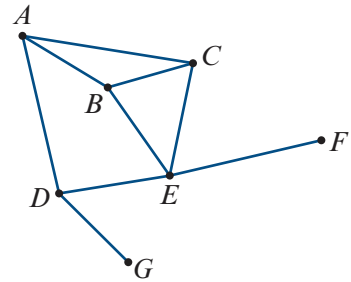


- 11** The earliest time that activity *I* can start is:
A 11 **B** 12 **C** 13 **D** 14
- 12** The latest starting time for activity *E* is:
A 7 **B** 8 **C** 9 **D** 10
- 13** The float time of activity *H* is:
A 2 **B** 3 **C** 4 **D** 5
- 14** The critical path for the project is:
A $A - D - H - J$ **B** $C - G - J$
C $B - E - F - H - J$ **D** $B - F - I - J$
- 15** Which of the following graphs is a spanning tree for the network shown?

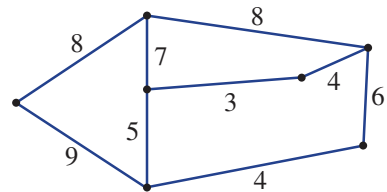


Short-response questions

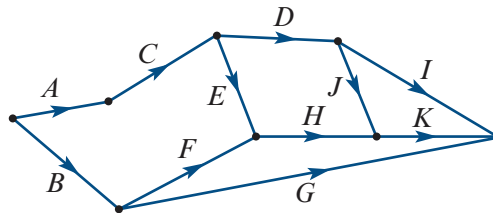
- 1 Consider the graph shown on the right.
- a How many edges must be removed in order to leave a spanning tree?
 - b Two of the edges in this graph must be in every spanning tree. Between which vertices are these edges?
 - c Remove some edges to form two different spanning trees for this graph.



- 2 Determine the minimum spanning tree for the network shown on the right.



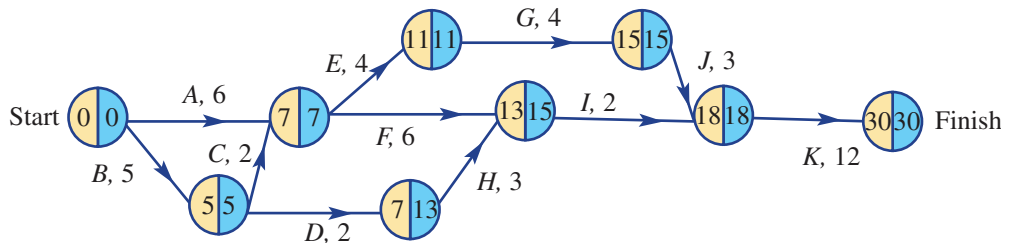
- 3 An activity network is shown below.



- a Which activities are immediate predecessors for activity *D*?
 - b For which activities is activity *B* an immediate predecessor?
 - c How many immediate predecessors does activity *H* have?
- 4 A precedence table for a project is shown below.

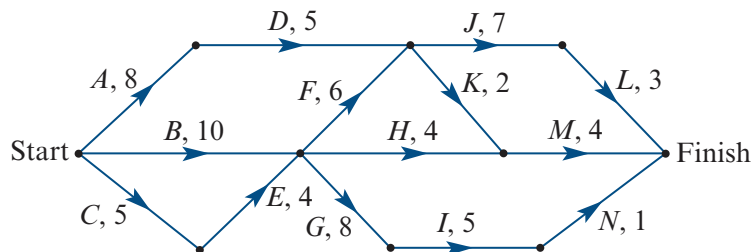
Activity	Immediate predecessors	Duration (weeks)
<i>A</i>	—	5
<i>B</i>	—	6
<i>C</i>	—	3
<i>D</i>	<i>C</i>	8
<i>E</i>	<i>B</i>	2
<i>F</i>	<i>D, E</i>	5
<i>G</i>	<i>A, F</i>	4

- a** Draw an activity network for this project.
- b** Perform the forward scanning process and determine the shortest time in which this project could be completed.
- 5** The activity network for a project is shown below. All durations in this network are in days.



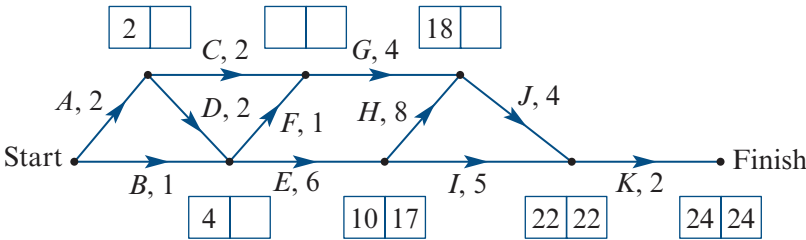
The forward scanning and backward scanning processes have already been completed.

- a** How many days will it take to complete this project?
- b** What is the earliest starting time for activity I?
- c** What is the float time for activity A?
- d** Write down the critical path for this project.
- 6** The activity network for a project is shown below. All durations in this network are in hours.



- a** Construct a precedence table for this project.
- b** Complete the forward scanning process to determine the shortest time in which this project can be completed.
- c** Complete the backward scanning process and write down the critical path for this project.

- 7 a** The assembly of machined parts in a manufacturing process can be represented by the following network. The activities are represented by letters on the arcs and the numbers represent the time taken (in hours) for the activities scheduled.



Activities	A	B	C	D	E	F	G	H	I	J	K
EST	0	0	2	2	4	4		10	10	18	22

- b** The earliest start times (EST) for each activity except *G* are given in the table. Complete the table by finding the EST for *G*.
- c** What is the shortest time required to assemble the product?
- d** What is the float time for activity *I*?
- 8** A precedence table for a project is shown below.

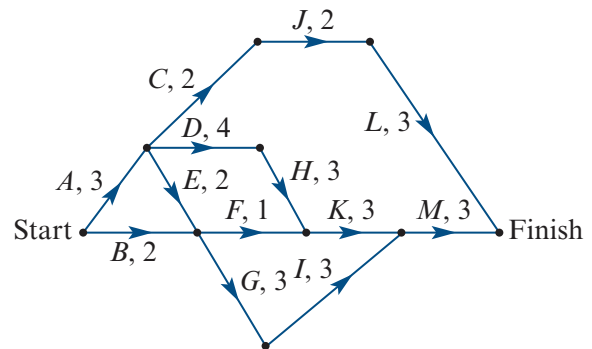
Activity	Immediate
<i>A</i>	—
<i>B</i>	—
<i>C</i>	<i>B</i>
<i>D</i>	<i>A, C</i>
<i>E</i>	<i>A, C</i>
<i>F</i>	<i>E</i>
<i>G</i>	<i>D</i>
<i>H</i>	<i>D</i>
<i>I</i>	<i>H, F</i>
<i>J</i>	<i>G, I</i>

- a** Draw the activity network for this project.
- b** Complete the forward and backward scanning processes to find the minimum number of days in which this project could be completed.
- c** Complete the table of EST, LFT, LST and float values.

Activity	Duration (days)	EST	LFT	LST	Float
A	10				
B	5				
C	3				
D	5				
E	4				
F	6				
G	6				
H	7				
I	5				
J	4				

- d** Explain what it means for an activity to be on the critical path for a project.
- e** Identify the feature of the table above that allows you to write down the critical path for this project.
- f** Write down the critical path for this project.
- g** Explain what would happen to the completion time of the project if activity C started one day later than expected.

- 9** The Bowen Yard Buster team specialises in backyard improvement projects. The team has identified the activities required for a backyard improvement. The network diagram to the right shows the activities identified and the actual times, in hours, needed to complete each activity, that is, the duration of each activity.



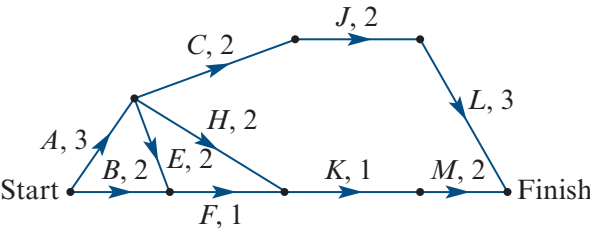
The table on the following page lists the activities, their immediate predecessor(s) and the earliest starting time (EST), in hours, of each of the activities. Activity X is not yet drawn on the network diagram.

- a** Use the information in the network diagram to complete the following table.
- b** Draw and label activity X on the network diagram above, including its direction and duration.

- c The path $A - D - H - K - M$ is the only critical path in this project.
 - i Write down the duration of path $A - D - H - K - M$.
 - ii Explain the importance of the critical path in completing the project.

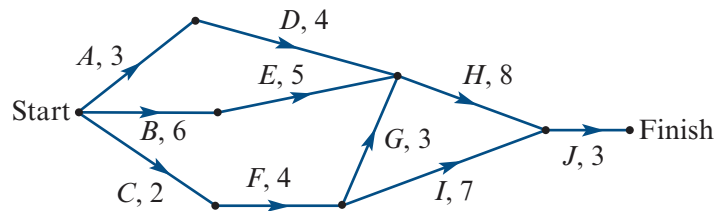
Activity	Immediate predecessor(s)	EST
A	–	0
B	–	0
C	A	3
D	A	3
E		3
F	B, E	5
G	B, E	5
H	D	7
I	G	
J	C, X	8
K	F, H	10
L	J	10
M	I, K	
X	D	7

- d To save money, Bowen Yard Busters decide to revise the project and leave out activities D, G, I and X . This results in a reduction in the time needed to complete activities H, K and M as shown.

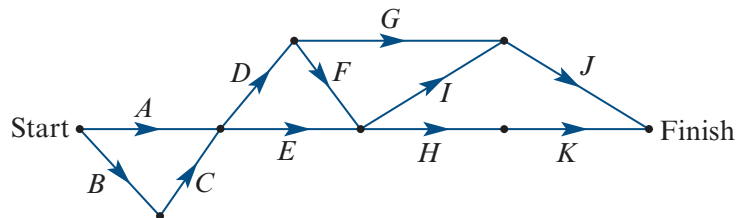


- i For this revised project network, what is the earliest starting time for activity K ?
- ii Write down the critical path for this revised project network.
- iii Without affecting the earliest completion time for this entire revised project, what is the latest starting time for activity M ?

- 10** The activity network for a project is shown in the diagram below. The duration for each activity is in hours.



- Identify the critical path for this project.
 - What is the maximum number of hours by which the completion time for activity *E* can be reduced without affecting the critical path of the project?
 - What is the maximum number of hours by which the completion time for activity *H* can be reduced without affecting the critical path of the project?
 - Every activity can be reduced in duration by a maximum of two hours. If every activity was reduced by the maximum amount possible, what is the new minimum completion time for the project?
- 11** In laying a pipeline, the various jobs involved have been grouped into a set of specific tasks *A – K*, which are performed in the precedence described in the network below.



- List all the task(s) that must be completed before task *E* is started.
- Use the information in the table on the left to complete the second table.

Task	Normal completion time (months)
<i>A</i>	10
<i>B</i>	6
<i>C</i>	3
<i>D</i>	4
<i>E</i>	7
<i>F</i>	4
<i>G</i>	5
<i>H</i>	4
<i>I</i>	5
<i>J</i>	4
<i>K</i>	3

Task	EST	LST
<i>A</i>	0	0
<i>B</i>	0	
<i>C</i>	6	7
<i>D</i>	10	10
<i>E</i>		11
<i>F</i>	14	14
<i>G</i>	14	18
<i>H</i>	18	20
<i>I</i>	18	
<i>J</i>	23	23
<i>K</i>	22	24

- c For this project:
 - i write down the critical path
 - ii determine the length of the critical path (that is, the earliest time the project can be completed).
- d If the project managers are prepared to pay more for additional labour and machinery, the time taken to complete task *A* can be reduced to 8 months, task *E* can be reduced to 5 months and task *I* can be reduced to 4 months.
Under these circumstances:
 - i what would be the critical path(s)?
 - ii how long would it take to complete the project?

Chapter 11

Networks and decision mathematics 2

Chapter questions

UNIT 4 INVESTING AND NETWORKING

Topic 5: Networks and decision mathematics 2

- ▶ How do we define a flow?
- ▶ How do we calculate the maximum flow through a network?
- ▶ How do we draw and use a bipartite graph to solve allocation problems?
- ▶ How do we find the optimal allocation of multiple groups of objects?

In this chapter maximum flow through a network and allocation of tasks are considered.

Flow through a network can be used to represent many physical problems. In this chapter problems of calculating maximum flow through a network from one end to the other are presented.

Bipartite graphs are used to represent the allocation of tasks. The Hungarian algorithm is introduced as a method to help obtain the best matching of tasks

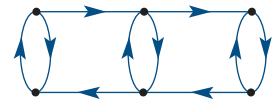
11A Flow networks

Learning intentions

- ▶ To define and describe flow.
- ▶ To calculate the maximum flow through a flow network by observation.
- ▶ To identify cuts and calculate cut capacities.
- ▶ To determine the maximum flow through a flow network by finding the capacity of the minimum cut.
- ▶ To solve flow problems by finding minimum cut capacities.

Directed graphs

In Chapter 9 we introduced **directed graphs (digraphs)**. These graphs record directional information on networks using arrows on the edges. The network on the right shows roads around a city. The vertices are the intersections of the roads and the edges are the possible road connections between the intersections. The arrows show that some of the roads only allow traffic in one direction, while others allow traffic in both directions.



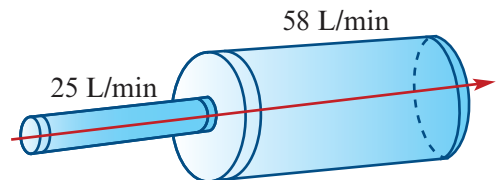
Flow problems

One of the applications of directed graphs to real-life situations is flow problems. Flow problems involve the transfer or **flow** of material from one point, called the **source**, to another point called the **sink**. Examples of this include water flowing through pipes, or traffic flowing along roads.

source $\rightarrow$ flow through network $\rightarrow$ sink

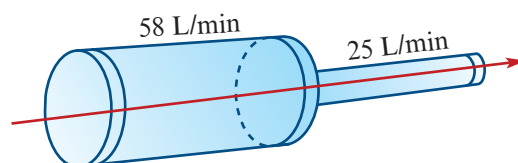
Water flows through pipes in only one direction. In a digraph representing water flow, the vertices are the origin and destination of the water and the edges represent the pipes connecting them. The weights on the edges would be the amount of water that can flow through the pipe in a given time. The weights of flow problem directed graphs are called **capacities**.

The diagram on the right shows two pipes that are joined together. The small pipe has a capacity of 25 litres per minute and this is joined to a larger pipe with capacity 58 litres per minute. Water flows through from the source, into the small pipe, through the large pipe and out to the sink.



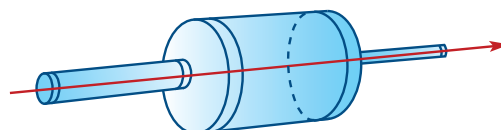
Even though the large pipe has a capacity of 58 litres per minute, the small pipe restricts the flow of water into the large pipe to 25 litres per minute. The flow through the large pipe will never be more than 25 litres per minute.

If the connection is reversed, water will be able to enter the large pipe at the rate of 58 litres per minute, but there will be a ‘bottleneck’ of flow at the junction between the large and small pipe. The large pipe can deliver 58 litres of water every minute to the small pipe, but the small pipe can only allow 25 litres per minute to pass.



In both of the previous situations, the flow through the entire pipe system (both pipes from source to sink) is restricted to a maximum flow of 25 litres per minute. This is the capacity of the smallest pipe in the connection.

If we connect more pipes together, one after the other, we can calculate the overall capacity of **maximum flow** of the pipe system by looking for the smallest capacity pipe in that system.



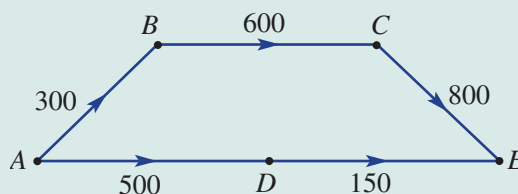
Maximum flow

If pipes of different capacities are connected one after the other, the **maximum flow** through the pipes is equal to the *minimum* capacity of the individual pipes.

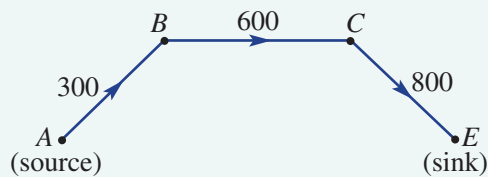


Example 1 Calculating the maximum flow

In the flow network shown on the right, the vertices A , B , C , D and E represent towns. The edges of the graph represent roads and the weights of those edges are the maximum number of cars that can travel on the road each hour. The roads allow only one-way travel, as indicated by the arrow.



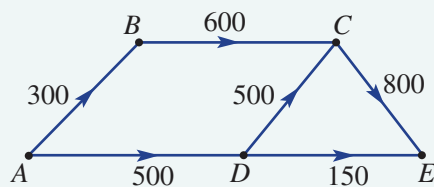
- Determine the maximum traffic flow from A to E through town C .
- Determine the maximum traffic flow from A to E overall.
- A new road is being built to allow traffic from town D to town C . This road can carry 500 cars per hour.
 - Add this road to the flow network.
 - Find the maximum traffic flow from A to E overall after this road is built.

Solution**a**

The maximum flow from A to E through town C is equal to the smallest capacity road along that route. The maximum flow is 300 cars per hour.

b

The maximum flow from A to E overall is: $300 + 150 = 450$ cars per hour.

c i**ii**

The maximum flow through $A - B - C - E$ is 300. But $C - E$ has capacity 800. If another 500 cars per hour come through $D - C$, they will be able to travel from $C - E$. The new maximum flow is now 800 cars per hour.

Explanation

Look at the subgraph that includes town C . The smallest capacity of the individual roads is 300 cars per hour. This will be the maximum flow through town C .

Look at the two subgraphs from A to E . The maximum flow through C will be 300 cars per hour. The maximum flow through D will be 150 cars per hour (minimum capacity). Add the maximum flow through C to the maximum flow through D .

Add the edge from D to C representing the new road to the diagram.

Determine the maximum flow.

There is a method for determining the maximum flow through a network and this is introduced in the following.

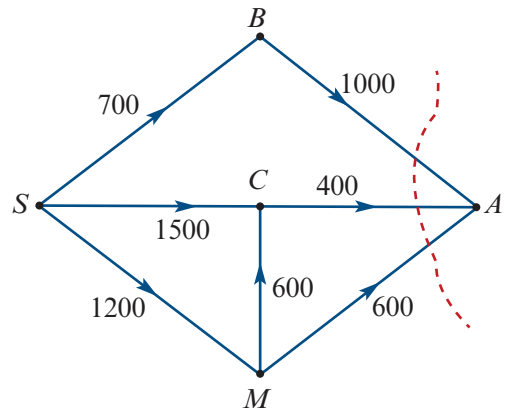
The method is known as the **maximum flow–minimum cut theorem**.

We first consider cuts.

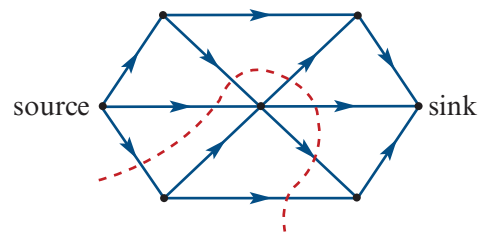
Cuts

For flow networks that contain many vertices and edges, it can be difficult to determine the maximum flow by inspection. We can simplify the search for maximum flow by searching for **cuts** with the network.

A cut divides the flow network into two parts, completely separating the source from the sink. It is helpful to think of cuts as imaginary blocks in the flow that do not allow any flow across them. In the diagram on the right, the dotted line is a cut. It completely blocks the flow from the source (S) to the sink (A).



The second graph shown on the right contains a dotted line that is *not* a cut. It blocks some of the flow through the network but there is still a flow pathway from the source to the sink across the top of the network.



A cut must *completely* block the flow from the source to the sink.

Cut capacity

The **cut capacity** for any cut is the sum of all the weights of the edges that the cut passes through. Only flow from the source side to the sink side is considered in the calculation of cut capacity. Any flow from the sink side across the cut to the source side is ignored.

Cuts and cut capacity

A **cut** is an imaginary line across a flow network that completely blocks all flow from the source to the sink.

The **cut capacity** of a cut is the sum of the capacities for the edges of the flow network that are blocked by that cut. Only edges that flow from the source side of the cut to the sink side of the cut are considered.

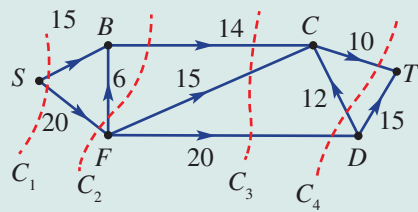


Example 2 Calculating cut capacity

Calculate the capacity of each of the four cuts shown in the flow network on the right.

The cuts are labelled C_1 , C_2 , C_3 and C_4 .

The source is vertex S and the sink is vertex T .



Solution

The capacity of $C_1 = 15 + 20 = 35$

The capacity of $C_2 = 14 + 20 = 34$

The capacity of $C_3 = 14 + 15 + 20 = 49$

The capacity of $C_4 = 20 + 10 = 30$

Explanation

All edges in C_1 are counted as they all flow from S to T across the cut.

The edge from F to B in C_2 is not counted. F is on the sink side of the cut and the flow crosses the cut back to the source side.

All edges in C_3 are counted as they all flow from S to T across the cut.

The edge from D to C in C_4 is not counted. D is on the sink side of the cut and the flow crosses the cut back to the source side.

Maximum flow and cut capacity

The capacity of a cut is important to determine the *maximum flow* through any flow network. Look for the smallest, or **minimum, cut capacity** that exists in the graph. This will be the same as the maximum flow that is possible through that graph.

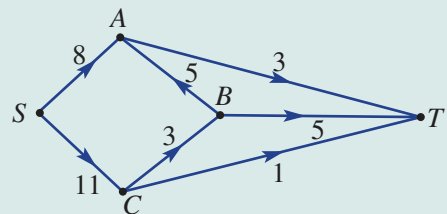
Maximum flow

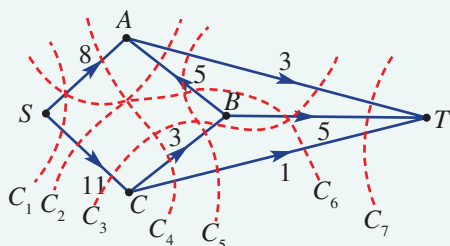
The **maximum flow** that is possible through a flow network is the same as the minimum cut capacity possible for that network.



Example 3 Calculating maximum flow

Calculate the maximum flow from S to T for the flow network shown on the right.



Solution

The capacity of $C_1 = 8 + 11 = 19$

The capacity of $C_2 = 3 + 11 = 14$

The capacity of $C_3 = 3 + 5 + 11 = 19$

The capacity of $C_4 = 8 + 3 + 1 = 12$

The capacity of $C_5 = 3 + 3 + 1 = 7$

The capacity of $C_6 = 8 + 5 + 1 = 14$

The capacity of $C_7 = 3 + 5 + 1 = 9$

The minimum cut capacity is 7 so the maximum flow from S to T is 7.

Explanation

Mark in all possible cuts on the network.

Calculate the capacity of all the cuts.

Identify the minimum cut capacity and write your answer.

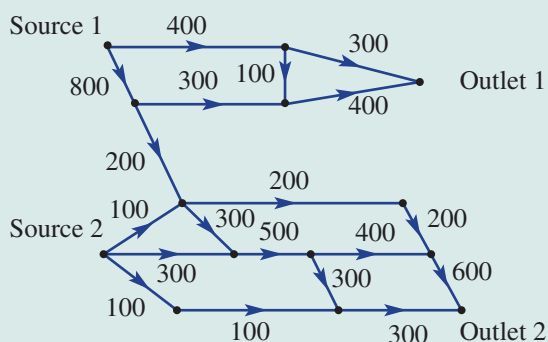
Calculating maximum flow from more than one source



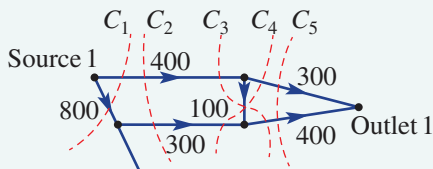
Example 4 Calculating maximum flow from more than one source

Water enters a network of pipes at either Source 1 or Source 2 and flows out at either Outlet 1 or Outlet 2.

The numbers next to the arrows represent the maximum rate, in kilolitres per minute, at which water can flow through each pipe.



Determine the maximum rate, in kilolitres per minute, at which water can flow from these pipes into the ocean at Outlet 1 and Outlet 2.

Solution**Outlet 1**

The capacity of C_1 is: $400 + 800 = 1200$

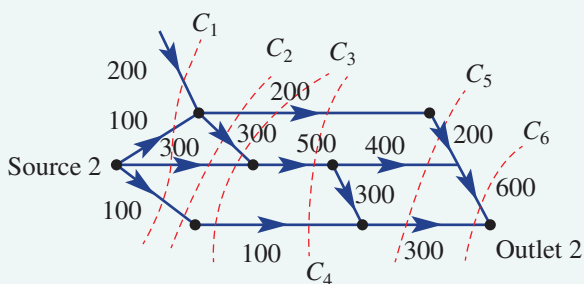
The capacity of C_2 is: $400 + 300 = 700$

The capacity of C_3 is: $400 + 400 = 800$

The capacity of C_4 is: $300 + 100 + 300 = 700$

The capacity of C_5 is: $300 + 400 = 700$

The minimum cut/maximum flow is 700 kilolitres per minute.

Outlet 2

The capacity of C_1 is: $200 + 100 + 300 + 100 = 700$

The capacity of C_2 is: $200 + 300 + 300 + 100 = 900$

The capacity of C_3 is: $200 + 300 + 300 + 100 = 900$

The capacity of C_4 is: $200 + 500 + 100 = 800$

The capacity of C_5 is: $200 + 400 + 300 = 900$

The capacity of C_6 is: $600 + 300 = 900$

The minimum cut/maximum flow is 700 kilolitres per minute.

Explanation

The outlets need to be considered separately.

Look for the minimum cut that prevents water reaching Outlet 1.

Note: The pipe with capacity 200 leading towards Outlet 2 does not need to be considered in any cut because this pipe *always* prevents water from reaching Outlet 1.

Look for the minimum cut that prevents water reaching Outlet 2.

Note: The pipe with capacity 200 leading towards Outlet 2 will need to be considered in any cut because this pipe delivers water towards Outlet 2 and must be 'cut' like all the others. Other cuts are possible, but have not been included in the diagram.

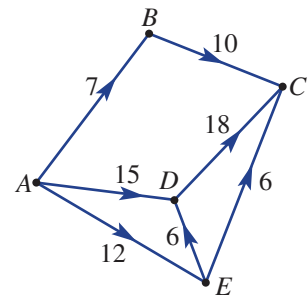
Section Summary

- ▶ A **cut** is an imaginary line across a flow network that completely blocks all flow from the source to the sink.
- ▶ The **cut capacity** of a cut is the sum of the capacities for the edges of the flow network that are blocked by that cut. Only edges that flow from the source side of the cut to the sink side of the cut are considered.
- ▶ The **maximum flow** that is possible through a flow network is the same as the minimum cut capacity possible for that network.

Exercise 11A

Example 1

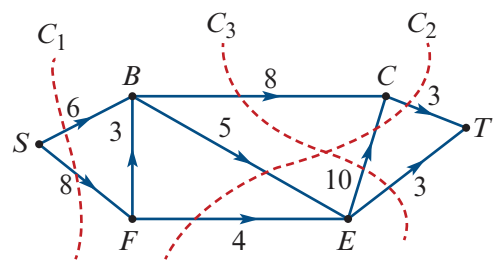
- 1 Flights connect the airports at five cities, A , B , C , D and E . The figures on the network represent the number of passengers, in thousands, that can be carried in a fixed time. Find the maximum number of passengers that can be carried between cities A and C in this fixed time.



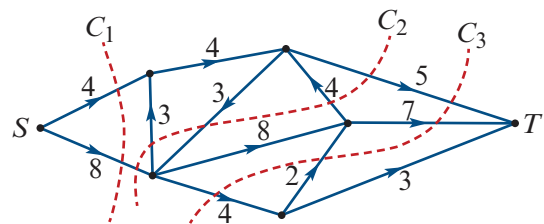
Cuts and cut capacity

Example 2

- 2 Determine the capacity of each of the cuts in the flow network on the right. The source is vertex S and the sink is vertex T .



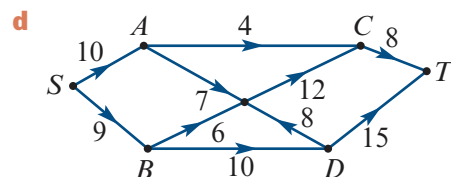
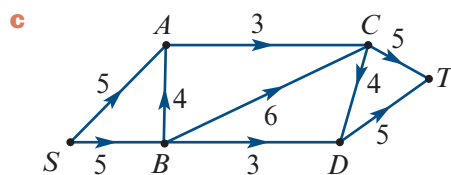
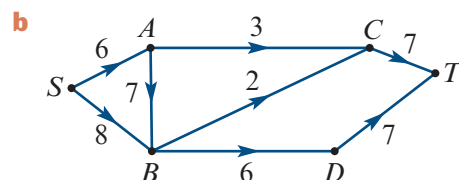
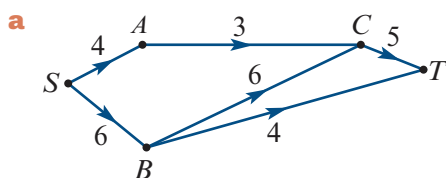
- 3 Determine the capacity of each of the cuts in the flow network on the right. The source is vertex S and the sink is vertex T .



Calculating maximum flow

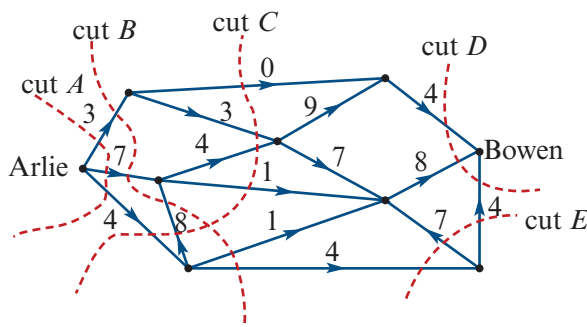
Example 3

- 4 Determine the maximum flow for each of the following flow networks. The source is vertex S and the sink is vertex T .



Applications of maximum flow

- 5** A train has the stages of its journey represented by the edges on the following directed network. The number of available seats for each stage is indicated beside the corresponding edge, as shown on the diagram on below.



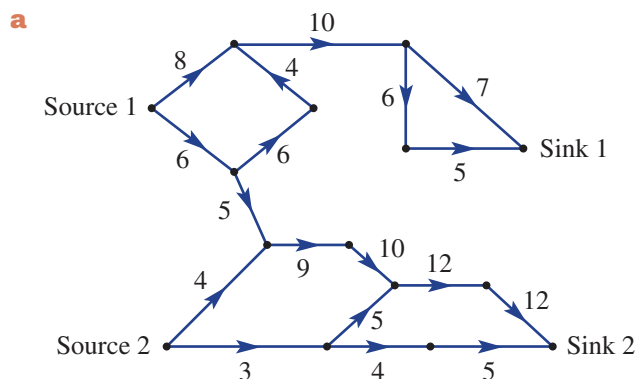
The five cuts, *A*, *B*, *C*, *D* and *E*, shown on the network, are attempts to find the maximum number of available seats that can be booked for a journey from Arlie to Bowen.

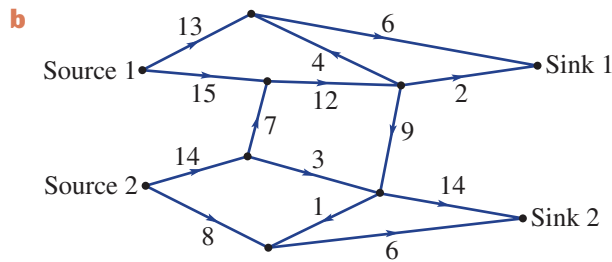
- Write down the capacity of cut *A*, cut *B*, cut *C*, cut *D* and cut *E*.
- Explain why cut *E* is not a valid cut when trying to find the minimum cut between Arlie and Bowen.
- Determine the maximum number of available seats for a train journey from Arlie to Bowen.

Networks with more than one source

Example 4

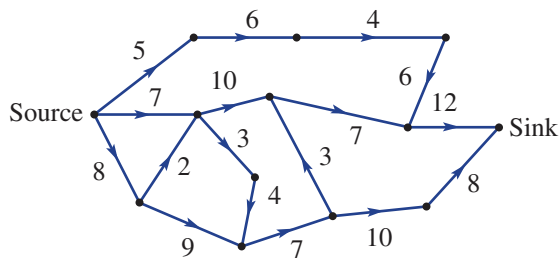
- 6** In each of the following, water pipes of different capacities are connected to two water sources and two sinks. Networks of water pipes are shown in the diagrams below. The numbers on the edges represent the capacities, in kilolitres per minute, of the pipes. For each of the following, find the maximum flow, in kilolitres per minute, to each of the sinks in this network.



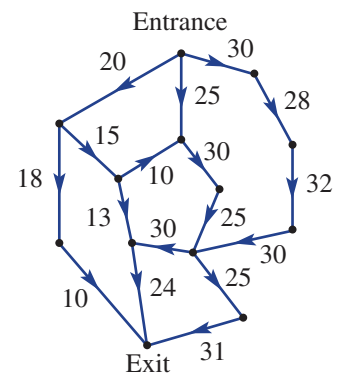


Analysis of maximum-flow problems

- 7** The flow of water through a series of pipes, in litres per minute, is shown in the directed network below.



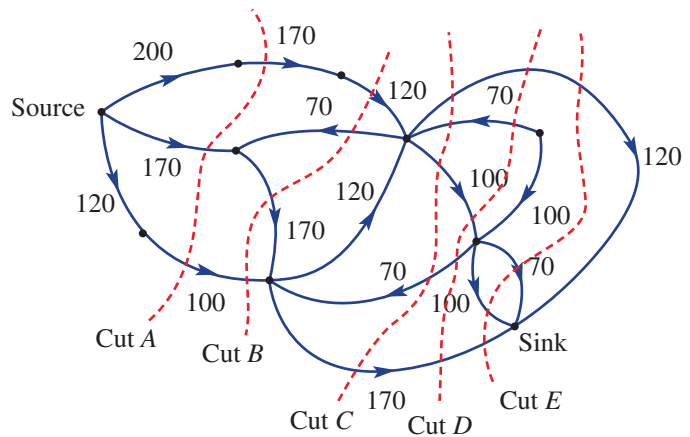
- a** How many different routes from the *source* to the *sink* are possible?
- b** Determine the maximum flow from the *source* to the *sink*.
- 8** The corridors people can walk through to visit different exhibits in a museum is given as a directed network opposite. To avoid congestion around every exhibit, the museum imposes a maximum capacity policy throughout each corridor between exhibits. The numbers on the edges represent the maximum number of people that can walk through each corridor of the museum every 30 minutes.
- a** On the network above, identify a cut that has a capacity of 80.
- b** Determine the maximum flow of people from the entrance to the exit of the museum.
- c** One group of primary school students would like to walk through the museum. The teacher explains that this can happen unsupervised if all students in the group remain together, not separating to explore different routes. What is the maximum number of students that can pass through the museum, from the entrance to the exit every 30 minutes?



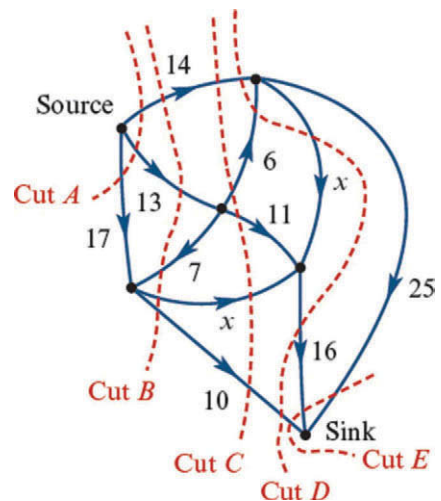
Paper 1-style multiple-choice questions

Questions 7 and 8 refer to the diagram opposite.

The flow of liquid through a series of pipes, in litres per minute, is shown in the directed network opposite. Five cuts labelled A to E are shown on the network.



- 9 The capacity of Cut B is
A 70 **B** 170 **C** 290 **D** 390
- 10 The number of these cuts with a capacity equal to the maximum flow of liquid from the source to the sink, in litres per minute, is
A 1 **B** 2 **C** 3 **D** 4
- 11 The flow of water through a series of pipes, in litres per minute, is shown in the network opposite. The weightings of two edges are labelled x . Five cuts labelled A to E are shown on the network. The maximum flow of water from the source to the sink, in litres per minute, is given by the capacity of



- A** Cut A if $x = 4$
B Cut B if $x = 6$
C Cut C if $x = 8$
D Cut D if $x = 6$

11B Bipartite graphs and assignment problems

Learning intentions

- ▶ To define and describe bipartite graphs.
- ▶ To solve assignment problems including the use of the Hungarian Algorithm.

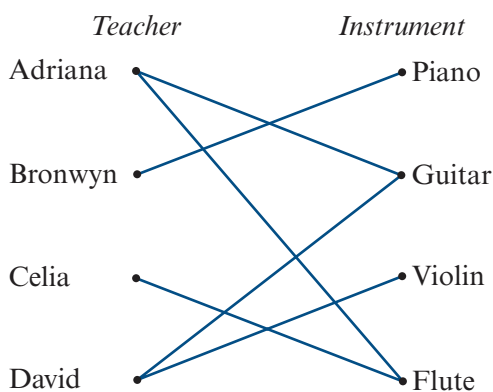
Assignment problems

Assignment problems involve matching two groups of objects to each other based on particular needs or circumstances. For example, a school has particular subjects that need teachers and the school also has teachers that can teach particular subjects. The school would need to assign teachers to subjects to ensure that every class had a teacher. Another example of an assignment problem is a factory that has a number of machines and machine operators. The factory may want to assign an operator to a machine so that the total time it takes a process to occur is minimised.

Bipartite graphs

Assignment problems can be represented graphically using a **bipartite graph**. In a bipartite graph, there are two groups of vertices. The vertices from one group are connected to one or more vertices in the other group by edges.

The bipartite graph below has two groups of vertices, one for the music teachers in a music school and one for the instruments that are taught. Each teacher and instrument are represented by a vertex in the relevant group.



The edge in the bipartite graph connects the teachers to the instruments that they can teach. For example, Adriana can teach both Guitar and Flute because there is an edge connecting Adriana to each of these instruments.

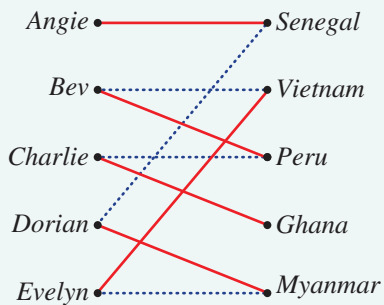
The bipartite graph can help the school assign each teacher to an instrument. For example, Bronwyn is the only teacher that can teach piano and so this assignment is necessary. Celia can only teach Flute and so, even though Adriana can teach both Flute and Guitar, if she teaches Flute then Celia will not be able to teach anything. So, Celia must teach Flute, which in turn means Adriana must teach Guitar.


Example 5 Solving an assignment problem with a bipartite graph

Angie, Bev, Charlie, Dorian and Evelyn are presenters on a TV travel show. Each presenter will be assigned a story to film about one country that they have visited before.

- Angie has visited Senegal.
- Bev has visited Vietnam and Peru.
- Charlie has visited Ghana and Peru.
- Dorian has visited Senegal and Myanmar.
- Evelyn has visited Vietnam and Myanmar.

Construct a bipartite graph of the information above and use it to decide on the assignment of each presenter to one country.

Solution


Angie will visit Senegal. Bev will visit Peru. Charlie will visit Ghana. Dorian will visit Myanmar. Evelyn will visit Vietnam.

Explanation

The two groups of items are: Presenters and Countries. Draw a vertex for each presenter in one column and each country in another. Angie has visited Senegal, so she could be sent there to film her story. Join the vertices for Angie and Senegal with an edge. Bev has visited Vietnam and Peru, so join the vertex for Bev to the vertices for Vietnam and Peru. Join the other presenter vertices to country vertices in a similar way. Definite assignments are shown in red; impossible assignments are shown with dotted lines.

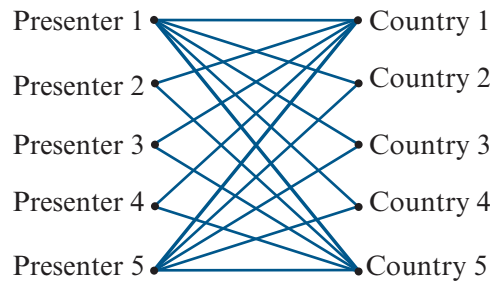
- Angie is only connected to Senegal and so must visit this country. If Angie visits Senegal, Dorian cannot.
- If Dorian cannot visit Senegal, he must visit Myanmar.
- If Dorian visits Myanmar, then Evelyn cannot.
- If Evelyn cannot visit Myanmar, she must visit Vietnam.
- If Evelyn must visit Vietnam, Bev cannot.
- If Bev cannot visit Vietnam, she must visit Peru.
- If Bev must visit Peru, then Charlie cannot.
- Charlie must visit Ghana.

Write the assignments.

Complete bipartite graphs

The assignment problem of presenter to country in Example 4 would be greatly simplified if every presenter could visit all of the countries. Every presenter vertex would be connected to every country vertex and there would be many different assignments that were possible. The bipartite graph for this situation would be a complete graph.

Rather than just assign presenters to countries randomly, the producers could use information about the preferences of the presenters, or perhaps the number of times that they have been to each of the countries, to make the assignment with priority. This information would be weights on the edges of an already very complex graph. Rather than write these weights on the graph, they can be recorded in a table instead.



Section Summary

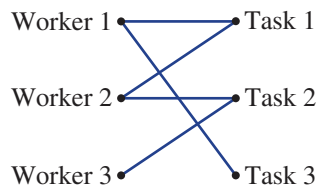
- A **bipartite graph** is a graph whose set of vertices can be split into two distinct groups in such a way that each edge of the graph joins a vertex in the first group to a vertex in the second group

Exercise 11B

Bipartite graphs

Example 5

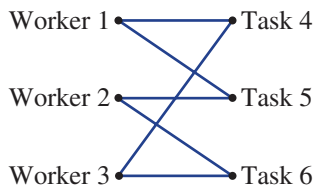
- 1 a On Monday, three workers are each to be allocated one task at work. The bipartite graph below shows which task(s) each person is able to complete.



If each person completes a different task, write down the task each worker must be allocated to on Monday.

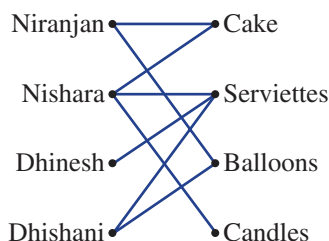
5F

- b** On Tuesday, the same three workers will be allocated to a new set of tasks. The bipartite graph below shows which task(s) each person is able to complete.



Given that worker 2 must complete task 6, write down the new task each worker must be allocated to on Tuesday.

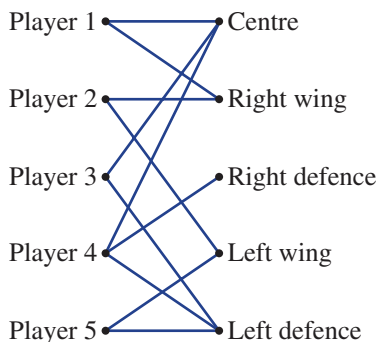
- 2** It is Miko's birthday and his sister Aria has asked some of his friends to assist with the celebrations by purchasing some items for a party. The bipartite graph below shows which item(s) each person is able to purchase on their way to the party.



Each friend must purchase an item. Write down which item each friend must purchase.

- 3** The sport of ice hockey has six player positions: goalie, left defence, right defence, right wing, left wing and centre. A group of six have decided to play. Only one person is happy to play goalie. The other five people must be allocated to the other five positions.

The bipartite graph below shows which positions each of the five players can play.



Each player plays a different position. Write down which position each player must play.

- 4** Gloria, Minh, Carlos and Trevor are buying ice-cream. They have a choice of five flavours: chocolate, vanilla, peppermint, butterscotch and strawberry.

- Gloria likes vanilla and butterscotch, but not the others.
- Minh only likes strawberry.
- Carlos likes chocolate, peppermint and butterscotch.
- Trevor likes all flavours.

a Explain why a bipartite graph can be used to display this information.

b Draw a bipartite graph with the people on the left and flavours on the right.

c What is the degree of the vertex representing Trevor?

The ice-cream shop has no butterscotch ice-cream available. Gloria, Minh, Carlos and Trevor will have only one ice-cream each and will all have a different flavour.

d Who must have the vanilla ice-cream?

e If Carlos chooses peppermint, write down the allocation of ice-cream flavour to these four people.

- 5** Joni, Ian, Dylan and Joshua are teachers in a school. The school has a Maths class, an English class, a Geography class and a Science class, each of which has a teacher. Each teacher can be allocated one class only.

Joni can teach English or Geography.

Ian can teach Maths or Science.

Dylan can teach English or Geography.

Joshua can teach Geography or Science.

a Draw a bipartite graph to show the teachers and the subject that they can teach.

b Explain why Joshua must take the science class.

c Write two different allocations of teachers to subjects.

- 6** The table below shows five people in the rows and five sports in the columns. A tick (✓) in a table cell indicates that the person in that row can coach the sport in that column. A cross (×) indicates that they cannot coach that sport.

Each sport in the table must be coached by only one of the people in the table.

	Hockey	Cricket	Soccer	Rugby	Squash
Rob	✓	×	✓	×	×
Janet	×	✓	✓	✓	✓
Tara	×	✓	×	✓	×
Diana	✓	×	×	✓	×
Jason	×	×	×	✓	×

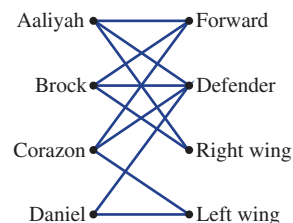
a Draw a bipartite graph to represent the information contained in the table above.

b Explain why Diana must coach hockey.

c Write the allocation of people to sports.

Paper 1-style multiple-choice questions

- 7 The sport of futsal has five player positions: goalkeeper, forward, defender, right wing and left wing. In a group of five friends, Ezekiel will always play goalkeeper, but the other four friends Aaliyah, Brock, Corazon and Daniel will rotate their responsibilities and are able to play a number of positions each.



The bipartite graph on the right shows which positions each of the four friends can play.

Based on the bipartite graph, which one of the following allocations is **not** possible?

A

Friend	Position
Aaliyah	Right wing
Brock	Forward
Corazon	Defender
Daniel	Left wing

B

Friend	Position
Aaliyah	Right wing
Brock	Defender
Corazon	Forward
Daniel	Left wing

C

Friend	Position
Aaliyah	Forward
Brock	Right wing
Corazon	Left wing
Daniel	Defender

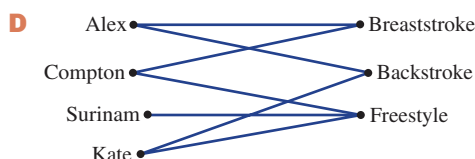
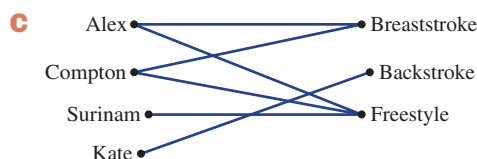
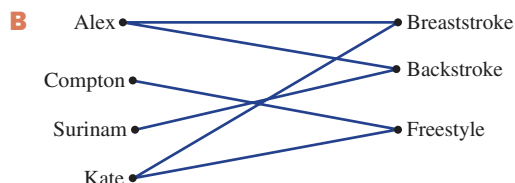
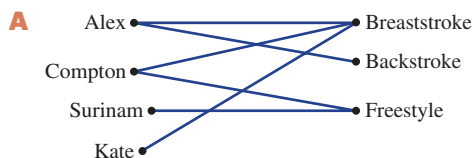
D

Friend	Position
Aaliyah	Forward
Brock	Defender
Corazon	Left wing
Daniel	Right wing

- 8 Four swimmers Alex, Compton, Surinam and Kate each compete in 50 metre races. The strokes, breaststroke(Br), freestyle(F) and backstroke(Ba), they compete in are shown in the table.

Swimmer	Event
Alex	breaststroke, backstroke
Compton	freestyle
Surinam	backstroke
Kate	freestyle, breaststroke

Which bipartite graph represents this information?



11C The Hungarian algorithm

Cost matrix

The table of weights for a bipartite graph is called a **cost matrix**. Even though it is called a cost matrix, the 'cost' does not have to be in terms of money. It could be the time taken to complete a task or the distance that needs to be travelled.

As an example, a factory might need to assign each of four employees to one of four machines.

The table on the below shows the four employees: Wendy, Xenefon, Yolanda and Zelda. The machines in a factory are represented by the letters *A*, *B*, *C* and *D*.

<i>Employee</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
Wendy	30	40	50	60
Xenefon	70	30	40	70
Yolanda	60	50	60	30
Zelda	20	80	50	70

The numbers in the table are the times, in minutes, it takes each employee to finish the task on each machine.

The table is called a **cost matrix**. Even though the numbers do not represent money value, this table contains information about the cost, in terms of time, of employees using each machine. The cost matrix can be used to determine the best way to allocate an employee to a machine so that the overall cost, in terms of the time taken to finish the work, is minimised. The Hungarian algorithm is used to do this.



Performing the Hungarian algorithm

Step 1: Subtract the lowest value in each row from every value in that row.

- 30 has been subtracted from every value in the row for Wendy.
- 30 has been subtracted from every value in the row for Xenefon.
- 30 has been subtracted from every value in the row for Yolanda.
- 20 has been subtracted from every value in the row for Zelda.

<i>Employee</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
Wendy	0	10	20	30
Xenefon	40	0	10	40
Yolanda	30	20	30	0
Zelda	0	60	30	50

Step 2: If the minimum number of lines required to cover all the zeros in the table is equal to the number of allocations to be made and there is a zero in each column, jump to step 6. Otherwise, continue to step 3.

- The zeros can be covered with three lines.
This is less than the number of allocations to be made (4).
- Continue to step 3.

<i>Employee</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
Wendy	0	10	20	30
Xenefon	40	0	10	40
Yolanda	30	20	30	0
Zelda	0	60	30	50

Step 3: If a column does not contain a zero, subtract the lowest value in that column from every value in that column.

- Column *C* does not have a zero.
- 10 has been subtracted from every value in column *C*.

<i>Employee</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
Wendy	0	10	10	30
Xenefon	40	0	0	40
Yolanda	30	20	20	0
Zelda	0	60	20	50

Step 4: If the minimum number of lines required to cover all the zeros in the table is equal to the number of allocations to be made, jump to step 6. Otherwise, continue to step 5a.

- The zeros can be covered with three lines.
This is less than the number of allocations to be made (4).
- Continue to step 5a.

<i>Employee</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
Wendy	0	10	10	30
Xenefon	~~40~~	~~0~~	~~0~~	~~40~~
Yolanda	30	20	20	0
Zelda	0	60	20	50

Step 5a: Add the smallest uncovered value to any value that is covered by two lines. Subtract the smallest uncovered value from all the uncovered values.

- The smallest uncovered element is 10.
- 10 has been *added* to Xenefon–*A* and Xenefon–*D* because these values are covered by two lines.
- 10 has been *subtracted* from all the uncovered values.

<i>Employee</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
Wendy	0	0	0	30
Xenefon	50	0	0	50
Yolanda	30	10	10	0
Zelda	0	50	10	50

Step 5b: Repeat from step 4.

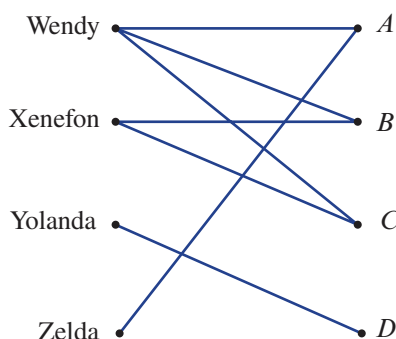
- The zeros can be covered with a minimum of four lines. This is the same as the number of allocations to make.
- Continue to step 6.

Employee	A	B	C	D
Wendy	0	0	0	30
Xenefon	50	0	0	50
Yolanda	30	10	10	0
Zelda	0	50	10	50

Step 6: Draw a bipartite graph with an edge for every zero value in the table.

In the bipartite graph:

- Wendy will be connected to A, B and C
- Xenefon will be connected to B and C
- Yolanda will be connected to D
- Zelda will be connected to A.

**Step 7: Make the allocation and calculate minimum cost**

- Zelda must operate machine A (20 minutes).
- Yolanda must operate machine D (30 minutes).
- Wendy can operate either machine B (40 minutes) or C (50 minutes).
- Xenefon can operate either machine B (30 minutes) or C (40 minutes).

Note: Because Wendy and Xenefon can operate either B or C, there are two possible allocations. Both allocations will have the same minimum cost.

The minimum time taken to finish the work = $20 + 30 + 50 + 30 = 130$ minutes.

**Example 6**

The table below shows the number of hours employees A, B and C require to complete tasks X, Y and Z.

	X	Y	Z
A	21	26	33
B	29	31	26
C	31	48	52

Continued on next page

- a** Allocate the tasks to minimise the total number of hours by
- i** listing each of the six possible sums
 - ii** using the Hungarian algorithm.
- b** Illustrate the result with a bipartite graph.

Solution

- a i** Remember that your choices are restricted so that no two of the numbers can be in the same row or column:

■ $21 + 31 + 52 = 104$	■ $21 + 48 + 26 = 95$
■ $29 + 26 + 52 = 107$	■ $29 + 48 + 33 = 110$
■ $31 + 26 + 26 = 83$	■ $31 + 31 + 33 = 95$

We see that $(31, 26, 26)$ gives the minimum sum. You may have picked this as the most likely already.

- ii** We now follow the Hungarian algorithm:

Subtract the lowest value in each row and then in each column.

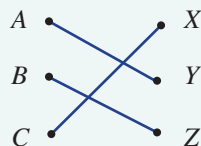
0	5	12
3	5	0
0	17	21

0	0	12
3	0	0
0	12	21

The result becomes evident from the algorithm as drawing lines through rows 1, 2 and 3 covers the zeros. This is the minimum number of lines.

Find the total by identifying the zeros, remembering that you can only have one entry per row. Referring back to the original array the sum is $31 + 26 + 26 = 83$.

b



When you have a 3×3 array as in the example above it is quite practical to list the possibilities. When we go to a 4×4 array there are 24 possibilities and for a 5×5 array there are 120. The Hungarian algorithm can be applied in these cases.

**Example 7**

A company has four sales representatives, Adrienne, Bonnie, Carol and Diva, to allocate to four companies, 1, 2, 3 and 4. The table shows the estimated weekly number of kilometres travelled by each representative when assigned a particular company. How should the sales representatives be allocated to minimise the total number of kilometres travelled?

	1	2	3	4
Adrienne	360	360	340	290
Bonnie	550	560	540	500
Carol	450	470	460	410
Diva	300	330	320	300

Solution

Subtract the lowest value in each row from every value in that row.

	1	2	3	4
Adrienne	70	70	50	0
Bonnie	50	60	40	0
Carol	40	60	50	0
Diva	0	30	20	0

Subtract the lowest value in each column from every value in that column. Draw the lines through the rows and columns that have a zero.

	1	2	3	4
Adrienne	70	40	30	0
Bonnie	50	30	20	0
Carol	40	30	30	0
Diva	--0--	--0--	--0--	0

Add the smallest uncovered value to any value that is covered by two lines. Subtract the smallest uncovered value from all the uncovered values. Draw in new line.

	1	2	3	4
Adrienne	50	20	10	0
Bonnie	--30--	--10--	--0--	0
Carol	20	10	10	0
Diva	--0--	--0--	--0--	20

Repeat: Add the smallest uncovered value to any value that is covered by two lines. Subtract the smallest uncovered value from all the uncovered values.

	1	2	3	4
Adrienne	40	10	0	0
Bonnie	30	10	0	10
Carol	10	0	0	0
Diva	0	0	0	30

For the minimum distance:

Adrienne goes to 4. Bonnie goes to 3. Carol goes to 2. Diva goes to 1.

The minimum total distance travelled = $290 + 540 + 470 + 300 = 1600$ km.

Maximising

The Hungarian algorithm has been used to minimise the cost; however, it can also be used to determine a suitable allocation that will maximise the overall outcome. This might be useful when allocating workers to tasks, where companies may want to maximise productivity or profit.

To solve a maximisation problem using the Hungarian algorithm, subtract all values in the matrix from the largest overall number in the matrix. We work with the array from Example 6.



Example 8

The table below shows the number of hours employees *A*, *B* and *C* require to complete tasks *X*, *Y* and *Z*.

	X	Y	Z
A	21	26	33
B	29	31	26
C	31	48	52

- a Allocate the tasks to **maximise** the total number of hours by
 - i listing each of the 6 possible sums
 - ii using the Hungarian algorithm.
- b Illustrate the result with a bipartite graph.

Solution

- a i Remember that your choices are restricted so that no two of the numbers can be in the same row or column.

■ $21 + 31 + 52 = 104$	■ $21 + 48 + 26 = 95$
■ $29 + 26 + 52 = 107$	■ $29 + 48 + 33 = 110$
■ $31 + 26 + 26 = 83$	■ $31 + 31 + 33 = 95$

We see that $29 + 48 + 33 = 110$ gives the **maximum** sum.

- ii We first form a new array by subtracting all entries from the maximum entry, which is 52. We then use the Hungarian algorithm to find the minimum value of the resulting sum. This will indicate the values to obtain the maximum sum from the original array.

	X	Y	Z
A	$52 - 21$	$52 - 26$	$52 - 33$
B	$52 - 29$	$52 - 31$	$52 - 36$
C	$52 - 31$	$52 - 48$	$52 - 50$

→

	X	Y	Z
A	31	26	19
B	23	21	26
C	21	4	0

We now follow the Hungarian algorithm: Subtract the lowest value in each row from every value in that row.

	X	Y	Z
A	12	7	0
B	2	0	5
C	21	4	0

Subtract the lowest value in each column from every value in that column. You can cover all zeros with only two lines. We need to proceed further with the algorithm.

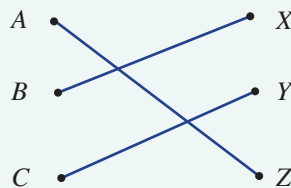
	X	Y	Z
A	10	7	0
B	0	0	5
C	19	4	0

Now subtract the smallest entry from every entry in the table and we have the following. It can be covered with three lines.

	X	Y	Z
A	6	3	0
B	0	0	1
C	15	0	0

Referring back to the original array we have $33 + 29 + 48 = 110$, which is the maximum sum.

b



Section Summary

- The table of weights for a bipartite graph is called a **cost matrix**. Even though it is called a cost matrix, the 'cost' does not have to be in terms of money. It could be the time taken to complete a task or the distance that needs to be travelled.
- The Hungarian algorithm can be used to solve allocation problems.

Exercise 11C

The Hungarian algorithm

Example 6

- 1** The table below shows the number of hours employees A , B and C require to complete tasks X , Y and Z .

	X	Y	Z
A	97	148	160
B	91	97	82
C	67	82	103

- a** Allocate the tasks to minimise the total number of hours by
- listing each of the 6 possible sums
 - using the Hungarian algorithm.
- b** Illustrate the result with a bipartite graph.

Example 7

- 2 a** A cost matrix is shown. Find the allocation(s) by the Hungarian algorithm that will give the minimum cost.

	A	B	C	D
W	110	95	140	80
X	105	82	145	80
Y	125	78	140	75
Z	115	90	135	85

- b** Find the minimum cost for the given cost matrix and give a possible allocation.

	A	B	C	D
W	2	4	3	5
X	3	5	3	4
Y	2	3	4	2
Z	2	4	2	3

Example 8

- 3** The table below shows the number of hours employees A , B and C require to complete tasks X , Y and Z .

	X	Y	Z
A	97	148	160
B	91	97	82
C	67	82	103

Allocate the tasks to maximise the total number of hours by using the Hungarian algorithm.

- 4 A school is to enter four students in four track events: 100 m, 400 m, 800 m and 1500 m. The four students' times (in seconds) are given in the table. The rules permit each student to enter only one event. The aim is to obtain the minimum total time.

<i>Student</i>	100 m	400 m	800 m	1500 m
Dimitri	11	62	144	379
John	13	60	146	359
Carol	12	61	149	369
Elizabeth	13	63	142	349

Use the Hungarian algorithm to select the 'best' student for each event.

- 5 Three volunteer workers, Joe, Meg and Ali, are available to help with three jobs. The time (in minutes) in which each worker is able to complete each task is given in the table opposite. Which allocation of workers to jobs will enable the jobs to be completed in the minimum time?

<i>Student</i>	<i>Job</i>		
	<i>A</i>	<i>B</i>	<i>C</i>
Joe	20	20	36
Meg	16	20	44
Ali	26	26	44

- 6 A company has four machine operators and four different machines that they can operate. The table shows the hourly cost in dollars of running each machine for each operator. How should the machinists be allocated to the machines to minimise the hourly cost from each of the machines with the staff available?

<i>Operator</i>	<i>Machine</i>			
	<i>W</i>	<i>X</i>	<i>Y</i>	<i>Z</i>
<i>A</i>	38	35	26	54
<i>B</i>	32	29	32	26
<i>C</i>	44	26	23	35
<i>D</i>	20	26	32	29

- 7 A football association is scheduling football games to be played by three teams (the Champs, the Stars and the Wests) on a public holiday. On this day, one team must play at their Home ground, one will play Away and one will play at a Neutral ground.

The costs (in \$'000s) for each team to play at each of the grounds are given in the table below.

Determine a schedule that will minimise the total cost of playing the three games and determine this cost.

Note: There are two different ways of scheduling the games to achieve the same minimum cost. Identify both of these.

<i>Team</i>	<i>Home</i>	<i>Away</i>	<i>Neutral</i>
Champs	10	9	8
Stars	7	4	5
Wests	8	7	6

- 8 A roadside vehicle assistance organisation has four service vehicles located in four different places. The table opposite shows the distance (in kilometres) of each of these service vehicles from four motorists in need of roadside assistance.

Service vehicle	Motorist			
	Jess	Mark	Raj	Karla
A	18	15	15	16
B	7	17	11	13
C	25	19	18	21
D	9	22	19	23

Determine a service vehicle assignment that will ensure that the total distance travelled by the service vehicles is minimised. Determine this distance.

- 9 A college holds a meeting for parents of new students. The meeting consists of four talks:

- Talk(1): Welcome
- Talk(2): Facilities of the school
- Talk(3): Expectations of parents
- Talk(4): Expectations of students

Four teachers, A, B, C and D, can deliver these talks and they switch them around for a little variety. The talks are delivered consecutively and there are no breaks between talks. The meeting starts at 10 a.m. and ends when all four talks have been delivered. The time, in minutes, each person takes to deliver each talk is given in the table below.

	Talk 1	Talk 2	Talk 3	Talk 4
A	13	34	28	18
B	14	32	34	14
C	16	32	30	16
D	12	33	34	12

- a i Use the Hungarian algorithm to find the earliest time that the meeting could end.
 ii State the final allocation.
- b Find the latest time that the meeting could end by maximising the sum of the lengths of the talks.
- 10 Four workers, A, B, C and D, are to be assigned to four tasks, P, Q, R and S. Each worker must be assigned to at most one task and each task must be done by just one worker. The amount, in dollars, that each worker would earn while assigned to each task is shown in the table below.

	P	Q	R	S
A	325	325	336	355
B	285	355	315	375
C	355	295	335	365
D	365	305	365	335

Use the Hungarian algorithm to obtain an allocation which **maximises** the total earnings.

Use the following information for Questions 11 and 12.

	Anita	Brad	Carmen	Dexter	Electra
Task 1	1	2	2	5	4
Task 2	4	9	7	11	6
Task 3	5	3	3	9	4
Task 4	8	5	6	6	7
Task 5	5	8	4	6	9

	Job 1	Job 2	Job 3	Job 4
Xena	5	3	7	p
Wilson	1	2	5	6
Yasmine	1	7	1	5
Zachary	4	7	6	p

- Job 1 to Wilson
- Job 2 to Xena
- Job 3 to Yasmine
- Job 4 to Zachary

A 6 **B** 7 **C** 8 **D** 9

Key ideas and chapter summary

Assignment problems	Assignment problems involve matching the objects in one group to objects in another so that the overall cost in terms of time, money or other quantity is minimised. Assignment problems are solved with bipartite graphs and/or the Hungarian algorithm.
Bipartite graph	A bipartite graph is a graph where the vertices exist in separate groups. The edges of a bipartite graph connect vertices in one group with vertices in the other.
Complete bipartite graph	In a complete bipartite graph, every vertex in one group of the bipartite graph is connected by an edge to every vertex in the other group.
Cost matrix	A cost matrix is a table of weights for a complete bipartite graph. It contains the cost, in terms of time, money or other quantity, of assigning objects from one group to objects in another. An example of a cost matrix is a table of the time it will take people (one group) to complete tasks (another group).
Hungarian algorithm	The Hungarian algorithm is an algorithm that is used to determine the best allocation to minimise the total cost.
Flow	Flow is the transfer of material, such as water, gas or traffic, through a directed network.
Flow network	A flow network occurs where the directed edges of the graph represent the flow of material from one vertex to another. The weight of an edge of a flow network is called the capacity of that edge.
Source	The source is the origin of material that flows through a network.
Sink	The sink is the final destination of material that flows through a network.
Flow problems	A flow problem is a problem that involves maximising the amount of material that flows through a network. Flow problems can be solved by finding the minimum cut for the network.
Cut	A cut is an imaginary line dividing a directed graph into two parts, one containing the source and the other containing the sink. It can be imagined that the cut blocks the flow through any edge it crosses.

Cut capacity

The **cut capacity** of a cut is the sum of all the weights on the edges it cuts. Only edges that flow from the source side of the cut to the sink side of the cut are considered.

Maximum flow

The **maximum flow** possible through a directed graph is the same as the smallest cut capacity of the cuts that are possible in that graph.

Skills checklist



Download this checklist from the Interactive Textbook, then print it and fill it out to check your skills. ☒

11A

- 1** I can determine the maximum flow for any section of sequential edges of a directed graph. ☐

See Example 4 and Exercise 11A, Question 4

11A

- 2** I can determine cut capacities. ☐

See Example 2 and Exercise 11A, Question 2

11A

- 3** I can determine the maximum flow as equal to the minimum cut capacity. ☐

See Example 3 and Exercise 11A, Question 5

11B

- 4** I can draw directed and weighted bipartite graphs. ☐

See Example 5 and Exercise 11B, Question 4

11C

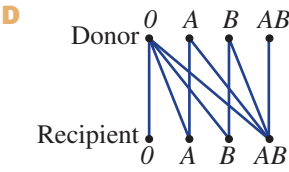
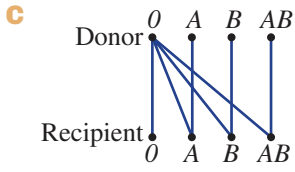
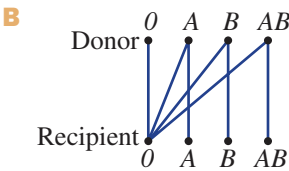
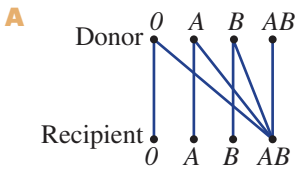
- 5** I can use the Hungarian algorithm to determine an optimum allocation in order to minimise cost. ☐

See Examples 6, 7 and 8 and Exercise 11C, Questions 1, 2 and 3

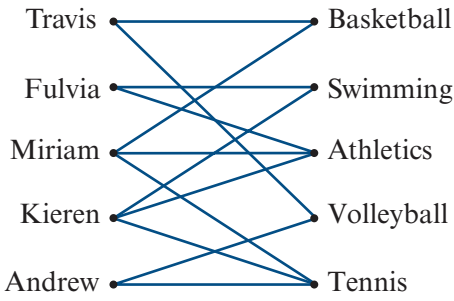
Multiple-choice questions

- 1** There are four different human blood types: O , A , B and AB . Blood can be donated from one human to another using the following rules:
- Type O can donate blood to any type.
 - Type AB can receive blood from any type
 - Each type can donate blood to its own type.
 - Each type can receive blood from its own type.

Which one of the following bipartite graphs correctly represents this information?



2 A group of five students represent their school in five different sports. The information is displayed in a bipartite graph. From this graph we can conclude that:



- A** Travis and Miriam played all the sports between them.
- B** In total, Miriam and Fulvia played fewer sports than Andrew and Travis.
- C** Kieren and Miriam each played the same number of sports.
- D** In total, Kieren and Travis played fewer different sports than Miriam and Fulvia.

3 Five people are to be each allocated one of five tasks (A, B, C, D, E). The table shows the time, in hours, that each person takes to complete the tasks. The total time to complete all the tasks is to be minimised.

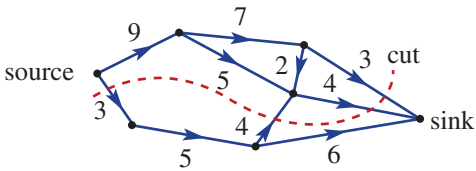
Name	A	B	C	D	E
Francis	12	15	99	10	14
David	10	9	10	7	12
Herman	99	10	11	6	12
Indira	8	8	12	9	99
Natalie	8	99	9	8	11

If no person can help another, Francis should be allocated task:

- A** A
- B** B
- C** D
- D** E

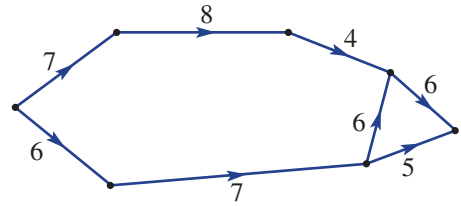
4 For the flow network shown on the right, the capacity of the cut is:

- A** 3
- B** 6
- C** 9
- D** 10



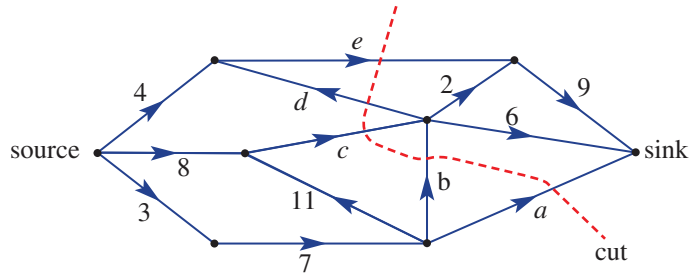
- 5 The maximum flow, from source to sink, in the flow network shown to the right is:

A 10 **B** 11 **C** 12 **D** 13



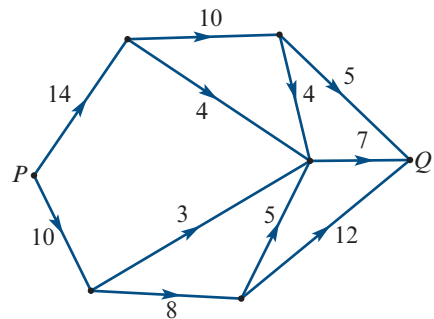
- 6 In the flow network to the right, the weight of each edge is non-zero. The capacity of the cut shown is:

A $a + b + c + d + e$
B $a + c + d + e$
C $a + b + c + e$
D $a + b + c - d + e$



- 7 The flow network to the right shows the capacity of data flow along cables in megabits per second. What is the maximum flow of data, in megabits per second, from server P to server Q ?

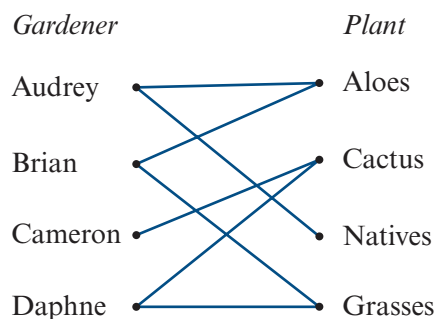
A 20 **B** 22 **C** 23 **D** 24



Short-response questions

- 1 The bipartite graph on the right shows the gardeners in a botanical garden and the plant types that they have experience caring for.

- a** How many gardeners have experience caring for Aloes?
b Which gardener is the only gardener who has experience caring for natives?
c The head gardener will assign Audrey, Brian, Cameron and Daphne to one plant type each. How must the plant types be allocated to gardeners?



2

	Word processing	Editing	Printing	Mailing
David	5	1	0	2
Robyn	4	0	2	0
Linda	2	0	3	2
Anthony	0	3	5	1

Isla has four employees, David, Robyn, Linda and Anthony. She needs to assign each of these people to one of four tasks: Word processing, Editing, Printing, Mailing. Isla has applied the Hungarian algorithm to determine how these tasks should be allocated so that the least overall time is spent completing all the tasks. The result of this is shown in the table above.

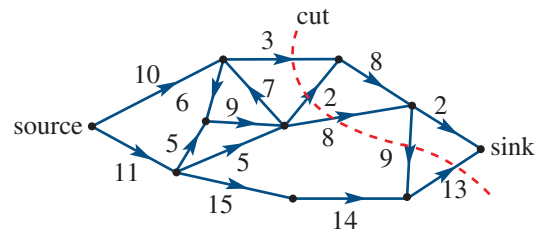
- a Construct a bipartite graph from this table.
- b Determine the allocation of tasks that Isla should make.

3 Steve is a supervisor in a furniture factory. An order for chairs needs to be filled. The table below shows the time, in hours, it would take each of four employees (Julia, Mario, Sylvana, George) to perform each of four tasks (cutting the pieces, assembling the chairs, preparing the chairs for painting, painting the chairs) required to complete the chairs for this order.

	Cutting	Assembling	Preparing	Painting
Julia	8	4	3	4
Mario	6	6	7	5
Sylvana	8	6	4	6
George	5	8	5	6

- a Which employee would take the shortest time to paint the chairs?
- b Which employee would take the longest time to prepare the chairs for painting?
- c Apply the Hungarian algorithm to determine the allocation of employee to task so that the overall minimum time is taken to fill the order.

- 4 In the flow network opposite, the values on the edges give the maximum flow possible between each pair of vertices. The arrows show the direction of flow in the network. Also shown is a cut that separates the source from the sink.

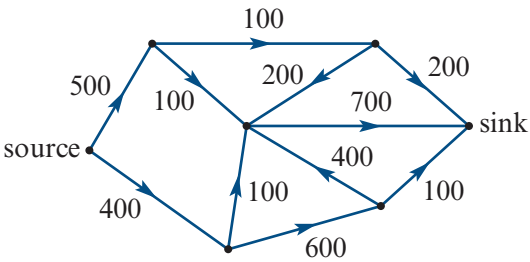


- a Determine the capacity of the cut shown.
- b Determine the maximum flow through this network.
- 5 Bernard, Georgia, Chris and Arthur are student pilots. Their flying instructor, Terry, has four lesson appointments available on a particular Saturday (9 a.m., 10 a.m., 1 p.m. and 3 p.m.).
- Bernard can only fly at 10 a.m.
 - Georgia can fly at any time before midday.
 - Chris can fly at 9 a.m. and then any time after 11 a.m.
 - Arthur can fly any time after 2 p.m.
- a Draw a bipartite graph with the student pilots on the left and the times on the right.
- b Which student pilot will Terry be teaching at 1 p.m.?
- c Write down the appointment times for each of the four student pilots.
- 6 Ann, Bianca, Con and David are four examination supervisors. There are four examination venues: *B*, *C*, *D* and *E*. Each examination venue requires one examination supervisor. The table shows the times (in minutes) that the examination supervisors would take to travel from their home to each examination venue.

Supervisor	B	C	D	E
Ann	25	30	15	35
Bianca	22	34	20	45
Con	32	20	33	35
David	40	30	28	26

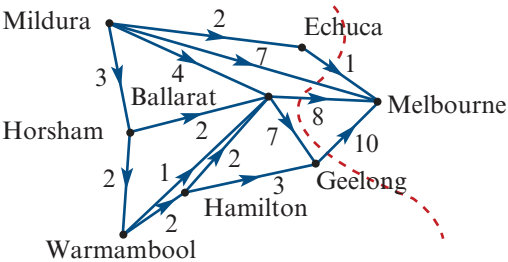
Determine the allocation of examination supervisor to examination venue that will minimise the overall travel time for the supervisors.

- 7** The flow network on the right shows the maximum rate of water flow (in litres per minute) through a system of water pipes. The water flows from the source to the sink.



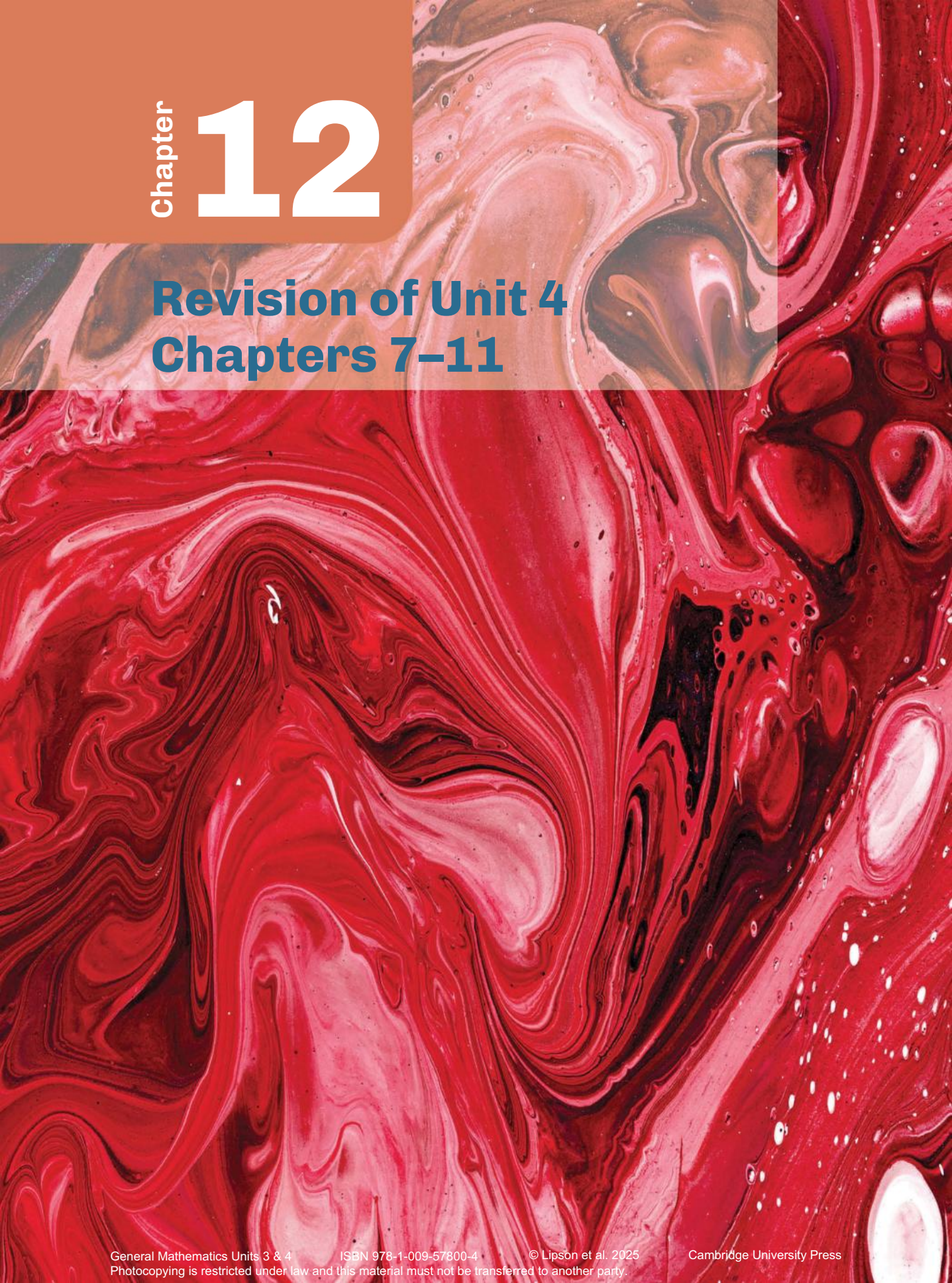
- a** Determine the maximum amount of water that can flow from the source to the sink through this network of pipes.
- b** How many litres of water will flow into the sink over 2 hours?
- c** A tank with capacity 2700 L is placed at the sink. How long will it take this tank to fill?

- 8** WestAir Company flies routes in western Victoria. The network shows the layout of connecting flight paths for WestAir, which originate in Mildura and terminate in either Melbourne or on the way to Melbourne. On this network, the available spaces for passengers flying out of various locations on one morning are shown.



- The network has one cut shown.
- a** What is the capacity of this cut?
- b** What is the maximum number of passengers who could travel from Mildura to Melbourne for the morning?
- 9** A school swimming team wants to select a 4×200 metre relay team. The fastest times of its four best swimmers in each of the strokes are shown in the table below. Which swimmer should swim which stroke to give the team the best chance of winning, and what would be their time to swim the relay?

Swimmer	Backstroke	Breaststroke	Butterfly	Freestyle
Rob	76	78	70	62
Joel	74	80	66	62
Henk	72	76	68	58
Sav	78	80	66	60



Chapter 12

Revision of Unit 4 Chapters 7–11

12A Topic 1: Loans, investments and annuities 1**Multiple-choice questions**

Use the following information to answer Questions 1 and 2.

The balance of a compound interest investment after n quarters, A_n , can be modelled by the recurrence relation $A_0 = 50\,000$, $A_{n+1} = 1.021 \times A_n$.

- 1 The balance of the investment after one year is
A \$50 000
B \$52 122.05
C \$53 216.61
D \$54 334.16
- 2 The annual percentage rate of interest for this investment is
A 0.21% **B** 2.1% **C** 0.84% **D** 8.4%
- 3 A principal of \$14 000 is invested and will earn compound interest at the rate of 2.8% per annum, compounding weekly. The effective annual rate of interest for this investment is closest to
A 2.80% **B** 2.81% **C** 2.83% **D** 2.84%
- 4 Eliana invested \$20 000 for 4 years at an interest rate of 6.6% per annum, compounding monthly. The rule for the balance of her investment, A , after 4 years is
A $A = 20\,000 \times 1.66^4$
B $A = 20\,000 \times 1.066^4$
C $A = 20\,000 \times 1.055^4$
D $A = 20\,000 \times 1.055^{48}$
- 5 Eli borrowed some money. He will be charged compound interest at the rate of 7.08% per annum, compounding monthly. After one year, Eli repaid \$6674.95 as principal and interest. The amount borrowed was closest to
A \$6000 **B** \$6100 **C** \$6200 **D** \$6300
- 6 The balance of a reducing-balance loan after n months, A_n , can be modelled by the recurrence relation $A_0 = 250\,000$, $A_{n+1} = 1.003125 \times A_n - 2300$.
 The total interest that has been paid after one repayment is closest to
A \$781 **B** \$1519 **C** \$1558 **D** \$2300

- 7** A loan of \$180 000 is charged compound interest at the annual percentage interest rate of 3.24% per annum, compounding monthly. The loan is repaid with monthly repayments of \$1200. Let A_n be the balance of the loan after n months. The recurrence relation for the loan is

A $A_0 = 180\,000, A_{n+1} = 1.0324 \times A_n - 1200$

B $A_0 = 180\,000, A_{n+1} = 1.00324 \times A_n - 1200$

C $A_0 = 180\,000, A_{n+1} = 1.027 \times A_n - 1200$

D $A_0 = 180\,000, A_{n+1} = 1.0027 \times A_n - 1200$

- 8** A repayment schedule for the first two repayments of a reducing-balance loan is shown below.

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	40 000.00
1	400.00	160.00	240.00	39 760.00
2	400.00	159.04	240.96	39 519.04

Determine which of the following is the next line of this repayment schedule.

A	3	400.00	158.08	241.92	39 277.12
B	3	400.00	159.04	240.96	39 278.08
C	3	400.00	240.96	159.04	39 360.00
D	3	400.00	241.92	158.08	39 360.96

- 9** Abigail has a loan with monthly repayments of \$2500 over 10 years, and the annual interest rate is 4.8%, compounding monthly. The present value of the loan is closest to
- A** \$200 125
B \$230 750
C \$237 890
D \$275 126
- 10** Tabitha is taking out a loan of \$400 000 with an interest rate of 3.6% annually, compounding quarterly. The loan is to be paid off over 20 years. Her quarterly repayment is
- A** \$6538.40
B \$5975.25
C \$6870.00
D \$7035.68

Short-answer questions

- 1** The following recurrence relation can be used to model a compound interest investment of \$10 000 earning interest at the rate of 7.68% per annum, compounding monthly.

$$A_0 = 10\,000, A_{n+1} = 1.0064 \times A_n$$

In this recurrence relation, A_0 is the balance of the investment after n months.

- a** Apply the recurrence relation to find the balance of the investment after one, two and three months.
 - b** Find how many months it will take for the balance of this investment to first exceed \$10 500.
- 2** Erica has invested \$14 500 into an account that pays compound interest at the rate of 4.8% per annum, compounding monthly.
- a** Construct a recurrence relation model for the balance of Erica's investment after n months.
 - b** Apply the recurrence relation to determine the balance of Erica's investment after three months.
 - c** Using the compound interest formula, determine the balance of Erica's investment after two years.
- 3** Jack has borrowed \$4500 to buy furniture for his home. He will be charged compound interest at the rate of 10.2% per annum, compounding monthly.
Let A_n be the balance of Jack's loan after n months.
- a** State the monthly percentage rate of interest for this loan.
 - b** Construct a recurrence relation that models the balance of Jack's loan.
 - c** Jack pays the principal and all interest charged after one year. Find how much money will he have to repay.
- 4** Hugh has invested \$25 000 in an account that will pay compound interest every month, at the annual percentage rate of interest of 3.84%.
- a** Use the compound interest rule to determine the balance of Hugh's investment after three years.
 - b** Find how much interest has been earned in total after three years.
- 5** Dorothy borrowed some money to pay for a travelling holiday. The annual percentage rate of interest for her loan was 8.76%, compounding monthly. Dorothy repaid the principal and all interest charged, a total sum of \$13 057.59, after travelling for six months.
Use the compound interest rule to determine the principal amount that Dorothy borrowed.

- 6** A reducing-balance loan is modelled using the recurrence relation shown below.

$$A_0 = 3000, A_{n+1} = 1.0024 \times A_n - 150$$
In this recurrence relation, A_n is the balance of the loan after n weekly repayments.
Determine the balance of this loan after five repayments.
- 7** Celia is considering borrowing \$50 000 to buy a caravan. Her bank will charge interest at the rate of 7.08% per annum, compounding monthly. Celia can afford to make monthly repayments of \$500 to repay the loan.
Let A_n be the balance of Celia's loan after n months.
- State the monthly percentage rate of interest for Celia's loan.
 - Determine how many months it will take for Celia's loan to have a balance that is below \$48 000 for the first time.
- 8** Gracie has a loan of \$12 000 that she will repay with monthly repayments of \$450. Interest is charged at the percentage annual interest rate of 6.24%.
- Determine the balance of Gracie's loan after six repayments have been made.
 - Find how much interest has been paid in total after six repayments have been made.
- 9** Leanne currently owes \$138 500 on her home loan. She pays interest at the annual percentage interest rate of 4.32% per annum and repays the loan with monthly repayments of \$1250.
After six months, the interest rate of Leanne's loan increased to 4.44% per annum, compounding monthly. Leanne decided to increase her payments to \$1500 per month. Find how much Leanne will owe on this loan after a further 12 months.
- 10** The details of two different home loans with principal \$320 000 are shown in the table below.

	Interest rate	Term	Compounding period	Extra repayments	Repayment
Loan 1	3.12%	20 years	Monthly	Allowed	\$1800
Loan 2	3.38%	18 years	Fortnightly	Not allowed	\$900

Amanda is trying to decide between the two loans above. She believes that she can afford repayments of \$1900 per month.

- If no extra repayments are made, determine which of the two loans would be best for Amanda. Justify your decision by explaining your mathematical reasoning.
- Amanda may be able to make larger repayments of up to \$1900 per month. Explain whether this changes your answer and provide mathematical reasoning to support your answer.

12B Topic 2: Loans, investments and annuities 2

Multiple-choice questions

- An annuity in withdrawal phase with principal \$256 000 earns interest at an annual percentage rate of interest of 7.6%, compounding quarterly. Quarterly payments of \$7500 is made. The recurrence relation that models the annuity is
A $A_0 = 256\,000, A_{n+1} = 0.76 \times A_n - 7500$
B $A_0 = 256\,000, A_{n+1} = 1.019 \times A_n - 7500$
C $A_0 = 256\,000, A_{n+1} = 1.076 \times A_n - 7500$
D $A_0 = 256\,000, A_{n+1} = 1.19 \times A_n - 7500$
- Elvira has inherited \$100 000 and will invest this money into an annuity from which she will withdraw monthly payments. Interest will be earned at the rate of 4.68% per annum, compounding monthly. The balance of Elvira's investment was \$93 463.21 after five payments have been withdrawn. The value of Elvira's monthly payment is
A \$1024.22 **B** \$1307.36 **C** \$1396.75 **D** \$1687.20
- A payment schedule for the first two payments from an annuity are shown below. The interest compounds monthly and payments are also withdrawn monthly.

Payment number	Payment amount	Interest paid	Principal reduction	Balance of annuity
0	0	0	0	65 000.00
1	1500.00	455.00	1045.00	63 955.00
2	1500.00	447.69	1052.31	62 902.69

The next line of this repayment schedule is

A	3	1500.00	447.69	1052.31	61 850.38
B	3	1500.00	477.69	1052.31	62 902.69
C	3	1500.00	477.69	1052.31	61 850.38
D	3	1500.00	440.32	1059.68	61 843.01

- An annuity can be modelled by the recurrence relation
 Deposit phase: $A_0 = 180\,000, A_{n+1} = 1.005 \times A_n + 800$
 where A_n is the balance of the investment after n monthly payments have been withdrawn or deposited. The future value of the annuity after 2 years is closest to
A \$202 888.76 **B** \$20 345.56 **C** \$223 234.32 **D** \$222 088.76

- 5** Arthur has invested \$40 000 in an annuity. His investment will earn interest at the rate of 7.44% per annum, compounding monthly. Arthur will withdraw \$1200 a month from this annuity.
- The number of payments of \$1200 that Arthur can expect from this annuity is
- A** 37 **B** 38 **C** 39 **D** 40
- 6** Lisa has an annuity with an initial balance of \$200 000. She makes deposits of \$1500 each month for a period of 5 years. After this, \$2000 is withdrawn each month for 2 years. The interest rate over the 7 years is 7.2% per annum, compounding monthly.
- The future value of the annuity after 7 years is closest to
- A** \$320 500 **B** \$372 568
C \$394 305 **D** \$403 718
- 7** Monthly withdrawals of \$350 are made from an account that has an opening balance of \$42 500, invested at 6% per annum, compounding monthly.
- The balance of the account after 1 year is closest to
- A** \$38 300 **B** \$38 967
C \$40 803 **D** \$40 921
- 8** Lachlan invests \$90 000 in an annuity, paying 4.8% per annum, compounding monthly. Lachlan receives a regular monthly payment from the annuity.
- If the value of the annuity after one year is \$72 948.10, the amount of interest earned in the first year is closest to
- A** \$21 000.00 **B** \$17 051.90
C \$3948.10 **D** \$1750
- 9** A perpetuity can be modelled by the following recurrence relation
- A** $A_0 = 85\,000, A_{n+1} = 1.035 \times A_n - 3570$
B $A_0 = 85\,000, A_{n+1} = 1.039 \times A_n - 3570$
C $A_0 = 85\,000, A_{n+1} = 1.041 \times A_n - 3570$
D $A_0 = 85\,000, A_{n+1} = 1.042 \times A_n - 3570$
- 10** A perpetuity will be set up to provide an annual prize of \$800 to the winner of a mathematics competition. Interest will be earned on the principal of the investment at the rate of 4% per annum and this will be used to pay the prize money every year.
- The amount that must be invested is
- A** \$2000 **B** \$3200
C \$20 000 **D** \$80 000

Short-answer questions

- 1** The balance of an annuity, A_n , after n monthly payments have been received is modelled by the recurrence relation below.
- $$A_0 = 250\,000, A_{n+1} = 1.0031 \times A_n - 1800$$
- State the percentage annual rate of interest for this annuity.
 - State the balance of the annuity after five payments have been received.
 - Calculate the total amount of interest that has been earned after five payments have been received.
- 2** Georgina has \$145 000 to invest in an annuity. Interest will be paid at the annual percentage interest rate of 4.08%, compounding monthly. Georgina will withdraw a payment of \$2500 each month from the investment.
- Let A_n be the balance of Georgina's annuity after n monthly payments have been withdrawn.
- Construct a recurrence relation model for the balance of Georgina's investment after n payment withdrawals.
 - Apply the recurrence relation to calculate the amount remaining in Georgina's investment after five payments have been withdrawn.
 - Calculate the total interest that Georgina has earned after five payments have been withdrawn.
- 3** An advertising agency has invested \$140 000 in a perpetuity. The interest earned each month by this investment will pay for a monthly award to a high performing employee.
- The interest on the investment is paid at the rate of 5.16% per annum. Find the value of the prize awarded each month.
 - The prize has value \$800 per month. Find the annual percentage rate of interest for the investment, rounding your answer to two decimal places.
- 4** Brian has \$35 000 to invest. He has two investment options:
- Bank A offers to pay 4.68% per annum, compounding monthly.
 - Bank B offers to pay 4.56% per annum, compounding fortnightly.
- After three years, Brian would like to withdraw the balance of the annuity.
- Determine which of the two investment options would earn Brian the most interest after one year. Explain how you compared the two investment options.
 - Write a letter to Brian explaining the comparison of the two investment options, showing him the calculations for the total amount he could withdraw after three years.

SF

CF

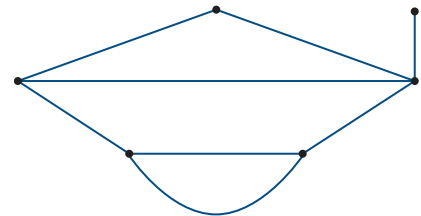
- 5** Fillipe would like to buy an apartment and he will need to borrow \$250 000 to pay for this. Interest will be charged at the annual percentage interest rate of 3.96%.
- Fillipe plans to repay his loan over a period of 25 years.
 - Calculate the monthly repayment amount required to achieve this aim. Round your answer to the nearest cent.
 - Using the rounded repayment amount, calculate the balance of the loan after 25 years.
 - This amount is positive. Explain the significance of this amount.
 - After four years of repayments (48 repayments), Fillipe will make a lump sum repayment of \$50 000. Find how many further repayments will be required to repay the loan. Round your answer to the nearest whole number.
- 6** Luther receives monthly payments of \$5400 from an annuity that is earning interest at the rate of 6.12% per annum, compounding monthly. The balance of Luther's investment is \$326 296.83 after four years of investment.
- State the principal amount of Luther's investment.
 - Find how much interest Luther has earned after four years of investment.
 - Determine how many more payments of \$5400 Luther can withdraw.
- 7** Vusa has recently retired and will invest his superannuation payment into an annuity that will earn interest at the rate of 6.8% per annum, compounding quarterly. The principal amount of this investment will be \$396 000.
- If Vusa withdraws a payment of \$20 000 per quarter, find how many payments in total he can expect to withdraw.
 - Find the balance of Vusa's investment after three years of payments.
- Vusa decided to increase his monthly payment after three years of investment. He will now withdraw monthly payments of \$25 000 until his investment is fully exhausted.
- Find how many payments of \$25 000 he can expect.
 - His final payment will be smaller than \$25 000. Determine how much his final payment will be.
- 8** Marcus's grandparents place \$2000 in an investment account that pays interest at a rate of 4% per annum, compounding annually. For 18 years, they contribute an additional \$1000 to the account.
- When Marcus turns 18, the interest rate increases to 5% and his grandparents stop contributing.
- Calculate how much Marcus will need to add to the account each year if he wishes to have a balance of \$60 000 by the time he is 30. Round your answer to the nearest cent.

- 9** Ethan invests \$126 666 into an annuity from which he receives a regular monthly payment of \$864 for 20 years. The interest rate for the annuity is 5.4% per annum, compounding monthly.
- After two months, the interest rate for this annuity will fall to 4.5%.
- To ensure that Ethan will still receive the same number of \$864 monthly payments, Ethan will add an extra one-off amount into the annuity at this time.
- Determine the value of the one-off addition. Round your answer to the nearest cent.
- 10** Derek invests \$50 000 into a compound interest investment paying 6.3% per annum, compounding annually. Derek invests an additional \$8000 per year immediately after interest is calculated.
- After five years, Derek increases his additional investment to \$10 000 per year.
- Calculate the value of Derek's investment after fifteen years (in total).

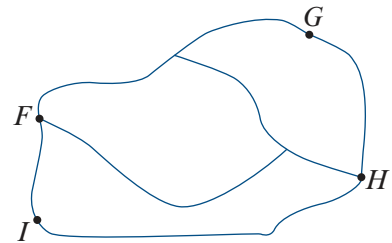
12C Topic 3: Graphs and networks

Multiple-choice questions

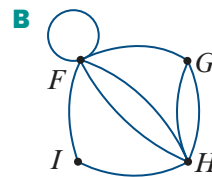
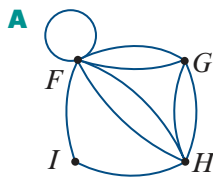
- 1** The sum of the degrees of all the vertices in the network opposite is:
- A** 6
B 7
C 8
D 16

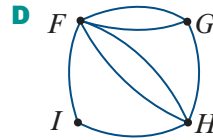
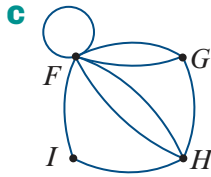


- 2** The diagram shows a map of the roads between four towns: F , G , H and I .

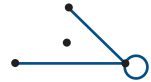


A network diagram that represents the connections between towns on the map is:





- 3** An adjacency matrix for the graph opposite could be:



A

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

B

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

C

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

D

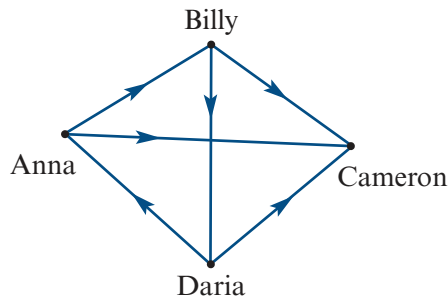
$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

- 4** A connected graph with 12 edges divides a plane into four faces. The number of vertices in this graph will be:

A 6 **B** 10 **C** 12 **D** 13

- 5** The directed graph below shows the results of a chess competition between four competitors, Anna, Billy, Cameron and Daria. The arc on the graph represents the game between the people at the vertices connected by that arc. The arrow points to the loser of the game.

Which one of the following statements about this chess competition is **not** true?

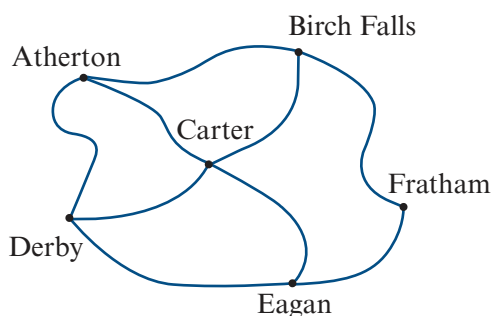


- A** Billy beat Daria.
B Daria lost more than one game.
C Anna won two of the three games she played.
D Nobody won all of their games.

- 6 The graph on the right shows six cities represented by vertices and the railway lines between those cities represented by edges.

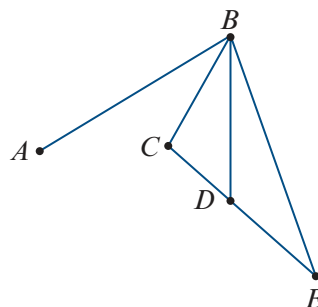
Which one of the following walks along the railway lines is also a trail?

- A Birch Falls – Atherton – Carter – Derby – Atherton – Carter – Eagan
- B Eagan – Carter – Derby – Eagan – Fratham
- C Fratham – Birch Falls – Atherton – Carter – Birch Falls – Fratham
- D Carter – Derby – Carter – Atherton – Derby



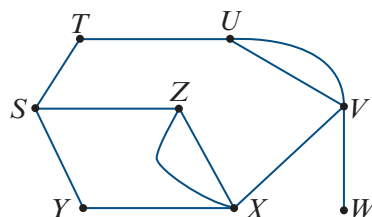
- 7 The graph on the right is best described as:

- A complete
- B Hamiltonian
- C semi-Eulerian
- D Eulerian



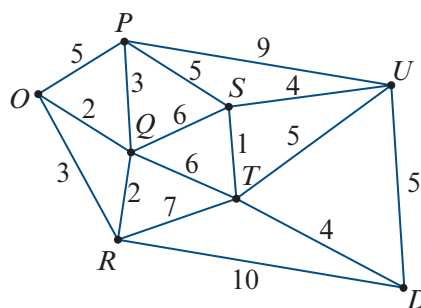
- 8 Adding which one of the following edges to the graph makes a semi-Eulerian trail possible?

- A ST
- B SU
- C SX
- D XW

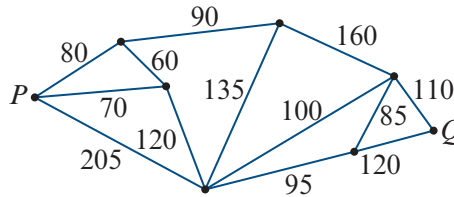


- 9 The length of the shortest path between the origin, O , and destination, D , in the network shown here is:

- A 11
- B 12
- C 13
- D 14



- 10** The graph below shows towns in a particular region represented as vertices and the roads between them represented as edges. The weight on an edge shows the travel time, in minutes, between the vertices connected by that edge. A salesman is about to travel from town P to town Q .

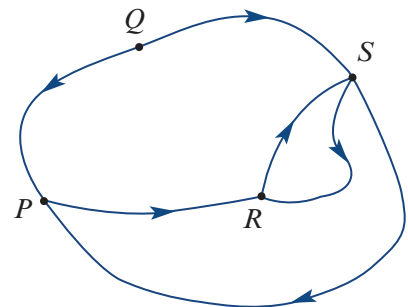
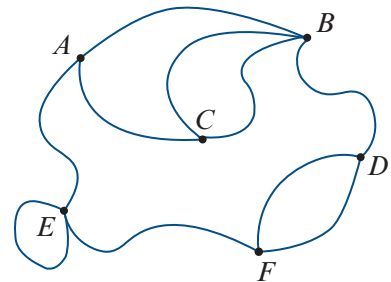


The shortest time it could take him, in minutes, is:

- A** 400 **B** 405 **C** 410 **D** 440
- 11** The number of edges for a complete graph with 20 vertices is:
- A** 10 **B** 20 **C** 21 **D** 190
- 12** The number of vertices for a complete graph with 21 edges is:
- A** 7 **B** 8 **C** 14 **D** 42

Short-response questions

- 1** Consider the graph on the right.
- What is the degree of vertex C ?
 - Which vertex has a loop?
 - How many vertices in this graph have an even degree?
 - Which vertices are immediately adjacent to vertex B ?
 - Which pairs of vertices have multiple edges between them?
- 2** Consider the directed graph shown on the right.
- Which vertex is the only one that can be directly reached from vertex P ?
 - Which vertex cannot be reached from any vertex?
 - Construct an adjacency matrix for this directed graph.



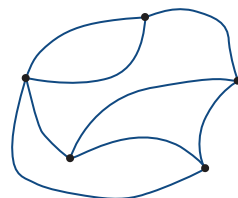
3 The adjacency matrix for a graph is shown on the right.

- a** How many edges are there between vertex A and vertex D ?
b How many loops are there in the graph for this adjacency matrix?
c Draw the graph represented by this adjacency matrix.

	A	B	C	D	E
A	0	1	1	0	2
B	1	1	0	0	1
C	1	0	1	1	0
D	0	0	1	0	0
E	2	1	0	0	0

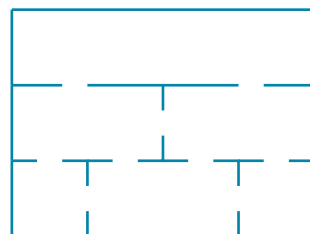
4 Consider the graph on the right.

- a** How many edges does this graph have?
b How many faces does this graph have?
c How many vertices does this graph have?
d Verify Euler's rule for this graph.



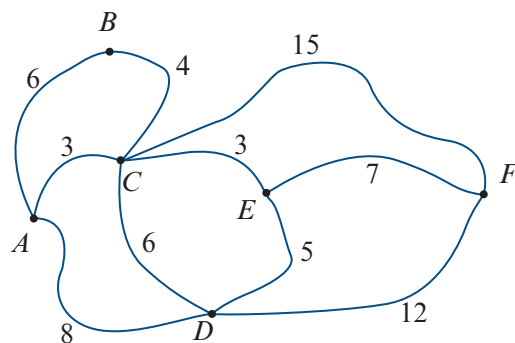
5 Suppose that six rooms in a house are laid out as shown.

The doors are represented by open sections along the walls. Can you walk around the house going through each door exactly once and finishing in the same room where you started?

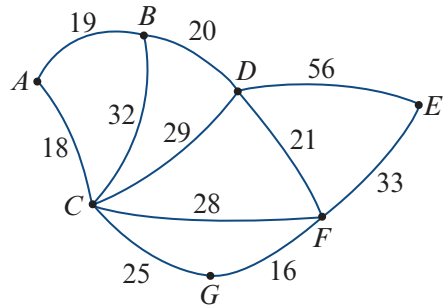


6 The vertices of the network on the right represent camping sites within a national park. The arcs of the network represent the walking tracks between each camping site and the numbers on these arcs show the distances along the tracks, in kilometres.

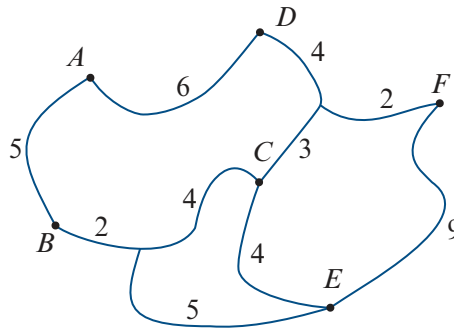
- a** A hiker is at campsite A . If she walks to campsite F , via campsite D , how far would she walk?
b What is the shortest distance from campsite A to campsite F ?



- 7** Seven towns (A, B, C, D, E, F, G) are represented by vertices on the graph to the right. The edges of the graph represent the road connections between the towns. The weights on the edges represent the distance along the road connections, in kilometres.

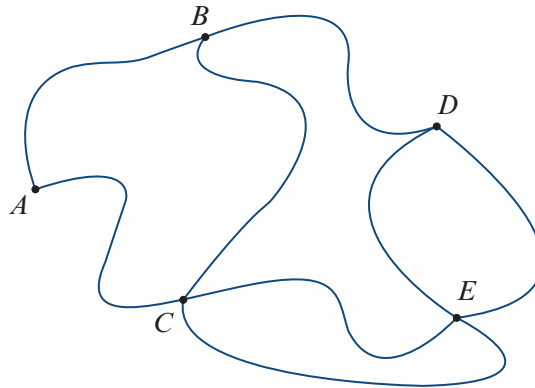


- a** This graph is planar. Explain what this means.
- b** Verify Euler's formula for this graph.
- c** An road inspector will begin at town B and inspect each of the roads only once.
- Where will the inspector end this inspection?
 - What is the name given to the walk that the inspector completes through this graph?
- d** Each of the towns have a branch of a bank. A bank manager from the branch at town C must visit all of the other branches in the region and then return to his own branch at town C .
- Write down one walk that the bank manager could take.
 - What is the name given to this walk?
- 8** The graph below shows six towns, A, B, C, D, E and F , represented by vertices. The towns are connected by roads shown as edges on the graph. The weights on the edges show the length, in kilometres, along each section of road.



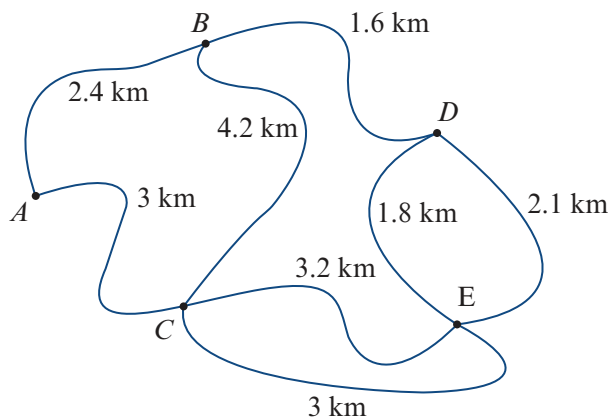
- a** A person can drive from E to B directly. How many kilometres will this journey be?
- b** How many extra kilometres is the journey from E to B via C compared to the direct journey?
- c** Which towns are adjacent to town A on this graph?
- d** Complete the following:
- Draw a graph that shows the shortest direct distances between each of the towns.
 - Construct the adjacency matrix for this graph.

- 9 A family is visiting a theme park and will visit five rides. A map of the theme park is shown below with the vertices representing the rides and the edges representing the paths connecting the rides.



- Determine a Hamilton cycle that the family can follow.
- Make a list of any paths (edges) that the family does not follow.
- Determine a semi-Eulerian trail that the family can follow. Explain how you determined this trail.

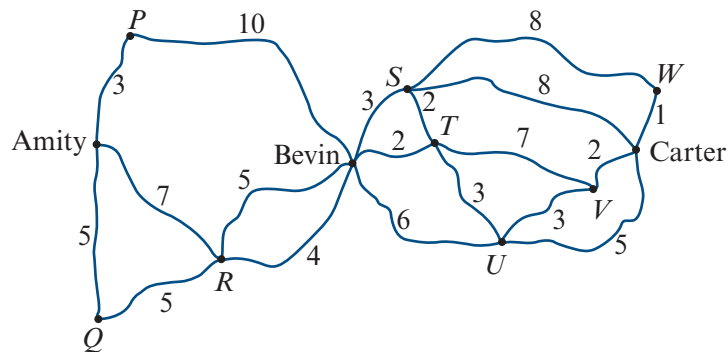
The network below shows the lengths of the paths that join the rides.



The family decide to visit the rides by following a semi-Eulerian trail. Assume that the family can walk at a speed of 3 km/hr. The theme park will close at 5 p.m.

- What is the latest time that the family can enter the theme park?

- 10** The diagram below shows the roads that connect the towns of Amity, Bevin and Carter represented as edges of a network. The vertices of the network, labelled P, Q, R, S, T, U, V and W , are checkpoints for the Amity Cycling Club road race. The numbers on the edges of the network are the lengths, in kilometres, of the roads between the checkpoints and the towns.



- Find, by inspection, the length of the shortest path from Amity to Bevin.
The road race covers the full length of every road on the network in any order or direction chosen by the riders. A rider may pass through each checkpoint more than once, but must travel along each road exactly once.
- One competitor claims this cannot be done. Explain why it *is* possible to travel every road once only during this race.
- If the race begins at Amity, where must this race finish?
- One of the competitors is following this path: Amity– P –Bevin– T – S –Bevin. Which checkpoint should *not* be visited next by this competitor? Explain why.

A road race for junior riders begins at Amity and ends at Carter. Participants are allowed to take any route they prefer.

- Find the shortest path from Bevin to Carter.
- Using your answers to parts **a** and **e**, what is the shortest distance from Amity to Carter?

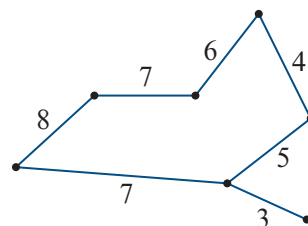
The Water Authority wants to lay water mains along the roads in order to put a fire hydrant at the locations of the checkpoints in the diagram above. A minimal spanning tree will be used for these water mains.

- Draw the minimum spanning tree for the diagram above.

12D Topic 4: Networks and decision mathematics 1**Multiple-choice questions**

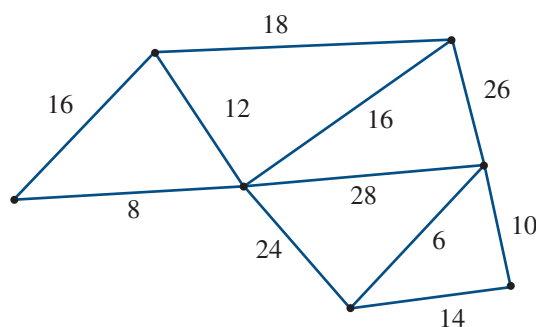
- 1** The sum of the weights of the minimum spanning tree of the weighted graph is:

A 2
B 30
C 32
D 33



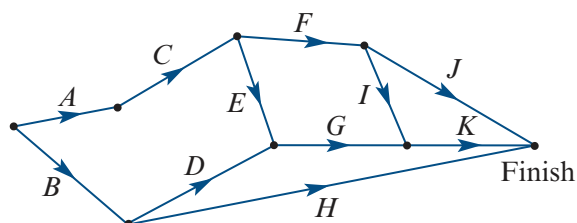
- 2** The minimum spanning tree for the graph opposite has a weight of

A 52
B 72
C 76
D 80



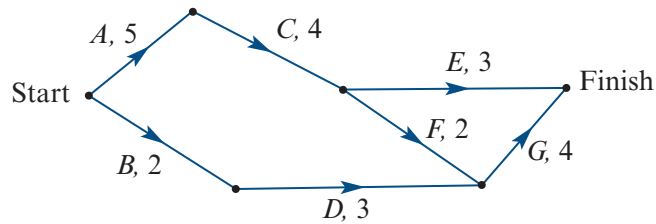
- 3** An activity network for a particular project is shown on the right.
 Activity *C* is an immediate predecessor of:

A activity *E* only
B activity *F* only
C activities *E* and *F*
D activities *E*, *F* and *G*



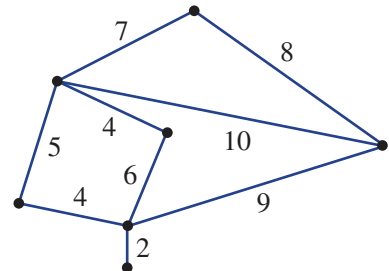
Use the following information to answer Questions 4, 5 and 6.

The activity network for a particular project is shown below. The duration of each activity, in hours, is shown on the network.



- 4 The earliest time that activity *G* can begin is:
- A 4 B 5 C 9 D 11
- 5 What is the shortest number of days it will take to complete this project?
- A 9 B 12 C 15 D 18
- 6 What is the latest possible time that activity *E* can begin without delaying the completion of the entire project?
- A 3 B 9 C 11 D 12
- 7 For the graph shown here, the minimum length spanning tree has length:

- A 30 B 31
- C 33 D 34



- 8 The table shows the information for four events of a project. The duration of an event is measured in minutes.

Activity	Duration	EST	LFT	LST
<i>A</i>	10	0	10	0
<i>B</i>	8	0	10	2
<i>C</i>	3	10	13	10
<i>D</i>	4	6	12	8

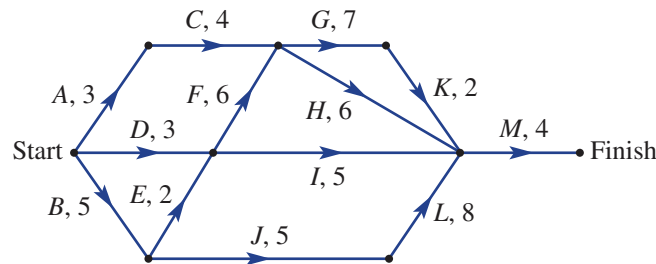
The float time for activity *D* is

- A 0 B 2 C 4 D 6

Activity	A	B	C	D	E	F
Duration	10	28	15	43	33	29

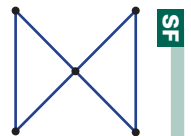
What is the earliest completion time for the project if it starts at 9:00 am?

- 10** The activity network for a particular project is shown below.

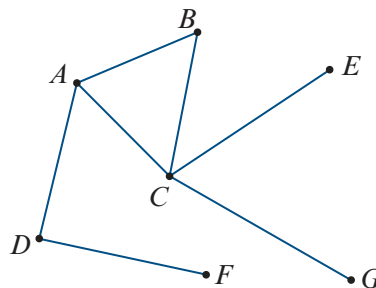


A $A-C-H-M$ **B** $A-C-G-K-M$
C $B-E-F-H-M$ **D** $B-E-F-G-K-M$

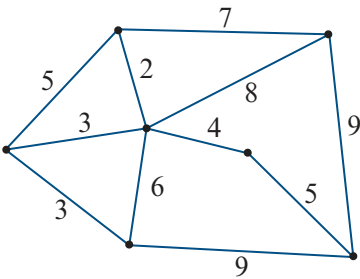
1 Draw two spanning trees for the graph shown here.



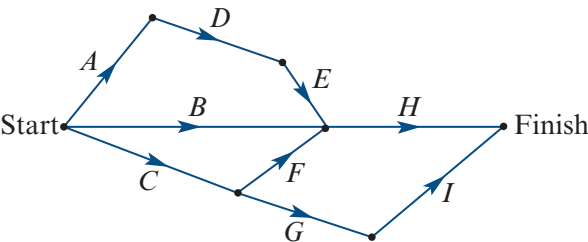
- 2** Consider the graph shown below.
One edge can be removed from this graph in order to leave a spanning tree.
Between which pair(s) of vertices is this edge?



- 3 a Draw the minimum spanning tree for the graph below.
b What is the weight of this minimum spanning tree?



- 4 An activity network for a particular project is shown below.



- a How many activities are required by this project?
b How many immediate predecessors does activity *H* have?
c Write down the activities for which activity *C* is an immediate predecessor.
- 5 A precedence table for a project is shown below.

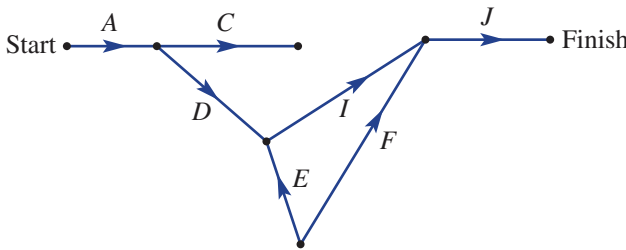
Activity	Immediate predecessors	Duration (hours)
<i>A</i>	–	3
<i>B</i>	<i>A</i>	2
<i>C</i>	<i>A</i>	5
<i>D</i>	<i>B</i>	2
<i>E</i>	<i>C</i>	3
<i>F</i>	<i>E</i>	6
<i>G</i>	<i>D, F</i>	2

- a Draw an activity network for this project.
b Perform the forward scanning process and determine the shortest time in which this project could be completed.

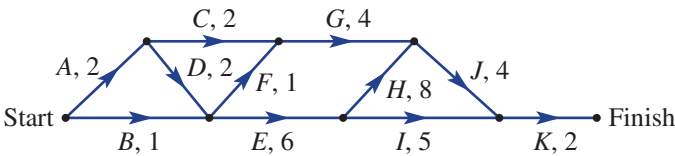
- 6 A project requires nine activities (A – J) to be completed. The duration, in hours, and the immediate predecessor(s) of each activity are shown in the table below.

Activity	Immediate predecessors
A	–
B	–
C	A
D	A
E	B
F	B
G	C
H	C
I	D, E, G
J	H, I, F

The directed network that shows these activities is shown below. Three features are missing. Using the table above, complete the network diagram below.



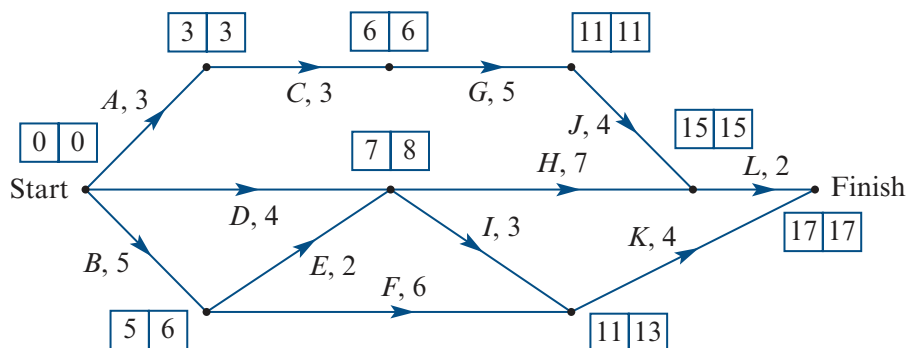
- 7 The assembly of machined parts in a manufacturing process can be represented by the following network. The activities are represented by the letters on the arcs and the numbers represent the time taken (in hours) for the activities scheduled.



Activity	A	B	C	D	E	F	G	H	I	J	K
EST	0	0	2	2	4	4		10	10	18	22

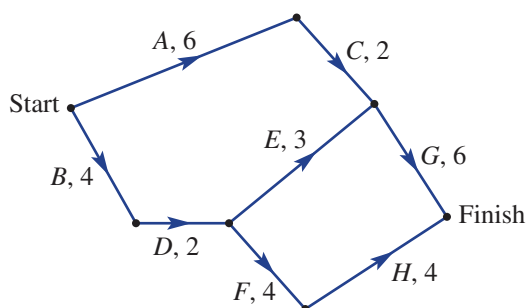
- a The earliest start times (EST) for each activity except G are given in the table. Complete the table by finding the EST for G.
- b What is the shortest time required to assemble the product?
- c What is the float (slack time) for activity I?

- 8 The activity network for a project is shown below. All durations in this network are in hours.



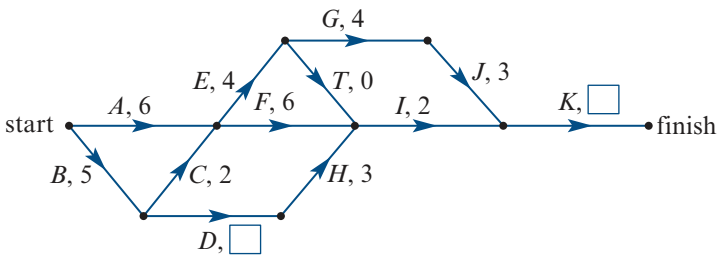
The forward scanning and backward scanning processes have already been completed.

- How many hours should it take for this project to be completed?
 - What is the earliest starting time for activity H ?
 - Calculate the float time for activity E .
 - Write down the critical path for this project.
- 9 Consider the activity network below.



- Determine the earliest start time of activity G .
- Determine the latest start time of activity F .
- Determine the float time for activity C .
- Write down the critical path.
- Determine the minimum completion time for this project.

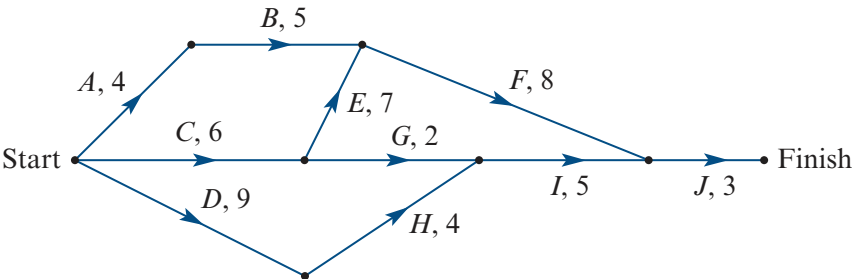
- 10** All the activities and their durations (in hours) in a project at the quarry are shown in the network diagram below. The least time required for completing this entire project is 30 hours.



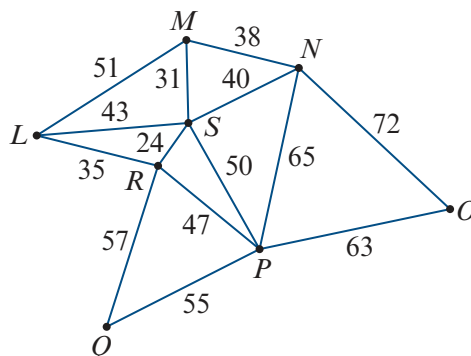
For each activity in this project, the table below shows the completion time, the earliest starting time and the latest starting time.

Activity	Completion time (hours)	Earliest starting time (hours)	Latest starting time (hours)
A	6	0	
B	5	0	0
C	2	5	5
D		5	9
E	4	7	7
F	6	7	
G	4	11	11
H	3	9	13
I	2	13	16
J	3	15	15
K		18	18

- a** Complete the missing times in the table.
- b** Write down the critical path for this project.
- 11** The activity network for a particular project is shown below. All durations in this network are in days.



- a** Construct a precedence table for this project.
- b** Complete the forward scanning process to determine the shortest time in which this project can be completed.
- c** Complete the backward scanning process to determine the critical path for this project.
- 12** A number of towns need to be linked by pipelines to a natural gas supply. In the network shown, the existing road links between towns L, M, N, O, P, Q and R and to the supply point, S , are shown as edges. The towns and the gas supply are shown as vertices. The distances along roads are given in kilometres.

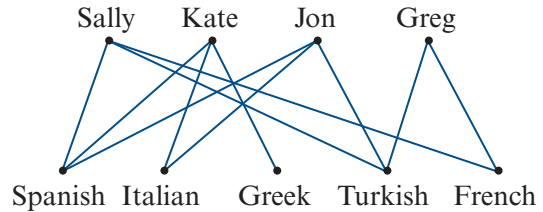


- a** What is the shortest distance along roads from the gas supply point S to the town O ?
- b** The gas company decides to run the gas lines along the existing roads. To ensure that all nodes on the network are linked, the company does not need to place pipes along all the roads in the network.
- What is the usual name given to the network within a graph (here, the road system) which links all nodes (towns and supply) and which gives the shortest total length?
 - Sketch this network.
 - What is the minimum length of gas pipeline the company can use to supply all the towns by running the pipes along the existing roads?

12E Topic 5: Networks and decision mathematics 2**Multiple-choice questions**

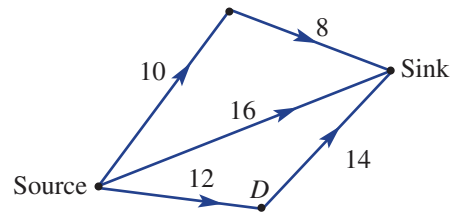
- 1** Which one of the following statements is not implied by this bipartite graph?

- A** There are more translators of French than Greek.
B Sally and Kate can translate five languages between them.
C Kate and Jon can translate more languages between them than can Sally and Greg.
D Sally and Jon can translate more languages between them than can Kate and Greg.



- 2** The directed graph shows the flow of water, in litres per minute, in a system of pipes connecting the source to the sink. The maximum flow is

- A** 40
B 22
C 36
D 38



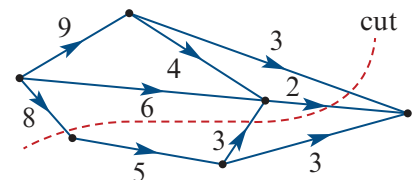
- 3** Five people are to be each allocated one of five tasks (A, B, C, D, E). The table shows the time, in hours, that each person takes to complete the tasks. The tasks must be completed in the least possible total amount of time. If no person can help another, Ismael should be allocated task:

Name	A	B	C	D	E
Francesca	12	15	99	10	14
Daniel	10	9	10	7	12
Harry	99	10	11	6	12
Ismael	8	8	12	9	99
Nathaniel	8	99	9	8	11

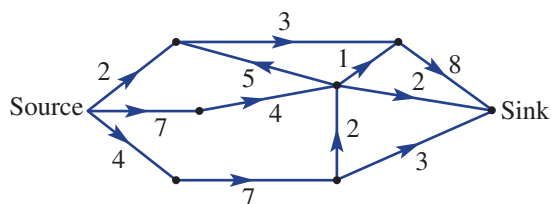
- A** A **B** B **C** C **D** D

- 4** The capacity of the cut in the flow network shown is:

- A** 0 **B** 2 **C** 10 **D** 13

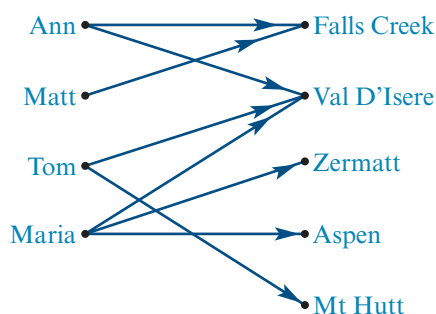


- 5** The flow network on the right shows connected water pipes represented as edges. The arrows show the direction of flow of the water through the pipes. The weights on the edges represent the maximum flow of water, in kilolitres per minute, through each pipe.

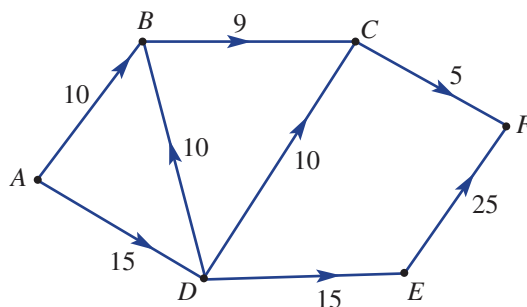


What is the maximum rate of flow of water that is possible from the source to the sink?

- A** 2
B 3
C 9
D 13
- 6** Four students, talking about five ski resorts they have visited, represented their information on the bipartite graph shown here. Which one of the following statements is implied by this bipartite graph?



- A** Matt and Tom have been to four ski resorts between them.
B Maria has visited fewer ski resorts than any of the others.
C Ann and Maria between them have visited all five ski resorts discussed.
D Ann and Tom between them have visited fewer resorts than Matt and Maria between them.
- 7** A source at A pumps water to a sink at F , through pipes, along the flow network shown at right.



What is the maximum rate of flow of water that is possible from the source to the sink?

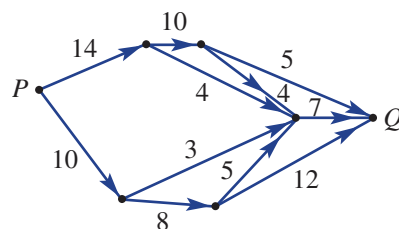
- A** 15 **B** 20 **C** 25 **D** 30

- 8 This matrix shown opposite was obtained after applying the Hungarian algorithm to determine the optimal allocation of three people, Alessandro (*E*), Burkhardt (*B*) and Casandra (*C*), to three tasks: *L*, *M* and *N*.

	<i>L</i>	<i>M</i>	<i>N</i>
<i>A</i>	0	0	6
<i>B</i>	0	2	7
<i>C</i>	1	0	0

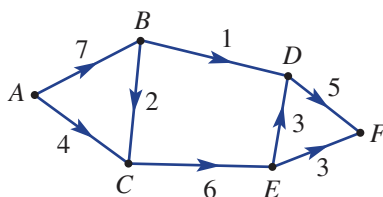
The optimal allocation is

- A** *A* to *N*, *B* to *L*, *C* to *M*
B *A* to *M*, *B* to *L*, *C* to *N*
C *A* to *N*, *B* to *M*, *C* to *L*
D *A* to *M*, *B* to *N*, *C* to *L*
- 9 In the network shown opposite, the numbers represent transmission capacities for information (data) in scaled units. What is the maximum flow of information from station *P* to station *Q*?



- A** 20
B 22
C 23
D 24

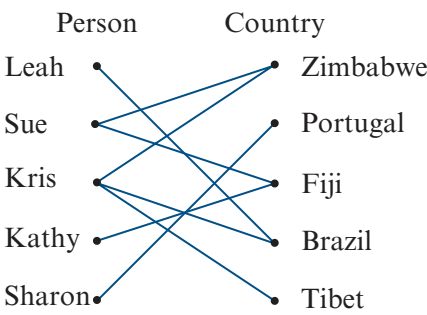
The following graph relates to Questions 10 and 11.



- 10 The maximum flow in the network linking vertex *A* to vertex *F* is:
- A** 5
B 6
C 7
D 8
- 11 The number of ways that vertex *F* can be reached from vertex *A* is:
- A** 2
B 3
C 4
D 5

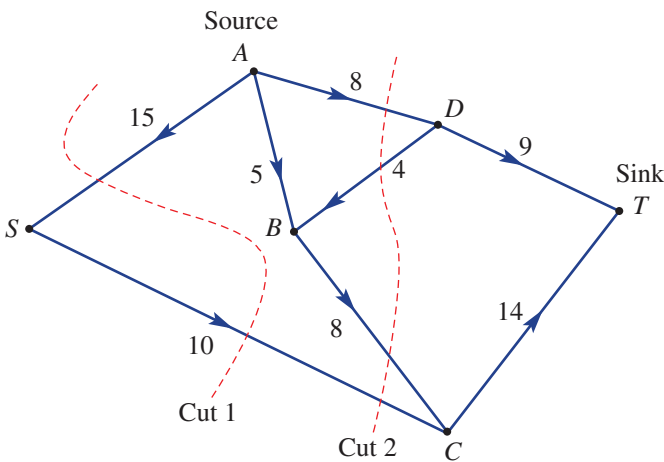
Short-response questions

1 The bipartite graph on the right shows the countries (Zimbabwe, Portugal, Fiji, Brazil and Tibet) that have been visited by five people (Leah, Sue, Kris, Kathy and Sharon).



- a How many countries has Leah visited?
- b Which person is the only one to have visited Portugal?
- c Each of the people will give a speech about their travels to only one country at a travel show. Make a list of each person and the country about which they will speak.

2 Consider the diagram opposite, which represents the number of cars per hour over a suburban road system.



- a Give a reason why Cut 1 is not a valid cut.
- b What is the capacity of Cut 2?
- c What is the maximum flow for this network?

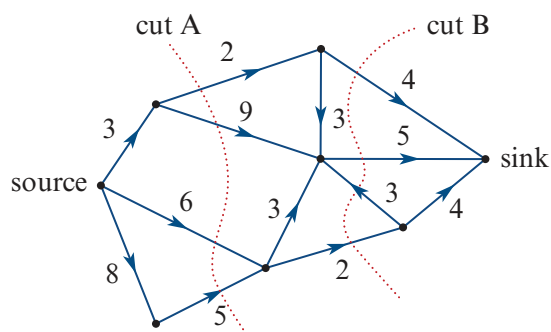
3 Roy runs a catering business with four employees, Ahmet, Beryl, Cynthia and Dorian. Each of these employees will be responsible for preparing one of the courses, canapes, starter, main or desert, for a dinner party. The time that each employee is expected to take to prepare each of the courses is shown in the table below.

	Canapes	Starter	Main	Desert
Ahmet	1	2	4	4
Beryl	5	3	5	3
Cynthia	3	3	3	2
Dario	6	4	4	7

- a Which one of the employees is the quickest to prepare Canapes?

- b** Apply the Hungarian algorithm to determine the allocation of employee to course so that the total time taken to prepare the dinner is as small as possible.
- c** How long will Roy and his employees take to prepare the dinner if:
 - i** the courses are prepared one after the other?
 - ii** preparation of all courses starts at the same time?

- 4** In the flow network on the right, the arcs represent pipes through which water can flow. The numbers on the arcs show the maximum rate at which water can flow through each pipe, in kilolitres per minute. Two cuts are shown on this network.

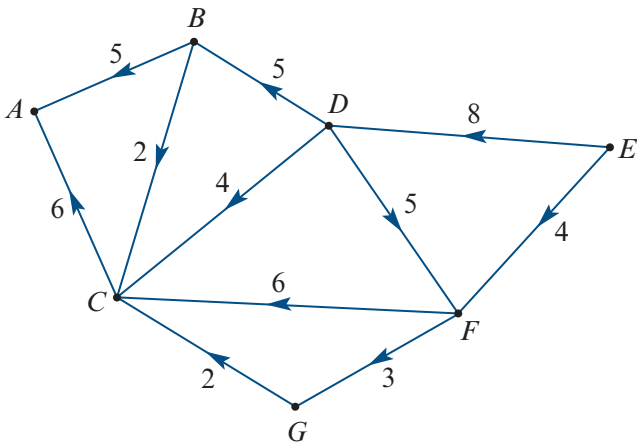


- a** Calculate the capacity of:
 - i** cut A
 - ii** cut B
 - b** The maximum rate at which water can flow from the source to the sink is 8 kilolitres per minute. Draw the cut with this minimum capacity on the network.
- 5** Margaret has four grandchildren, Tyson, Emma, Gregory and Rose. Margaret has four chocolate bars (Flakey, Cherry Chomp, Honey Crunch and Snacker) and will give one to each of her grandchildren.
- Tyson likes Flakey and Snacker.
 - Emma only likes Flakey.
 - Gregory likes all of the chocolates.
 - Rose likes every chocolate except Cherry Chomp.
- a** Draw a bipartite graph with the grandchildren on the left and the chocolate bars on the right.
 - b** Which grandchild will receive the Snacker?
 - c** How will Margaret distribute the chocolate bars so that every grandchild receives one they like?

- 6 Camp sites A, B, C and D are to be supplied with food. Four local residents, W, X, Y and Z , offer to supply one campsite each. The cost in dollars of supplying one load of food from each resident to each campsite is tabulated.

Camp site	W	X	Y	Z
A	30	70	60	20
B	40	30	50	80
C	50	40	60	50
D	60	70	30	70

- a Find the two possible matchings between campsites and residents so that the total cost is a minimum.
- b State this minimum cost.
- 7 A reservoir at E pumps water through pipes along the flow network shown below.

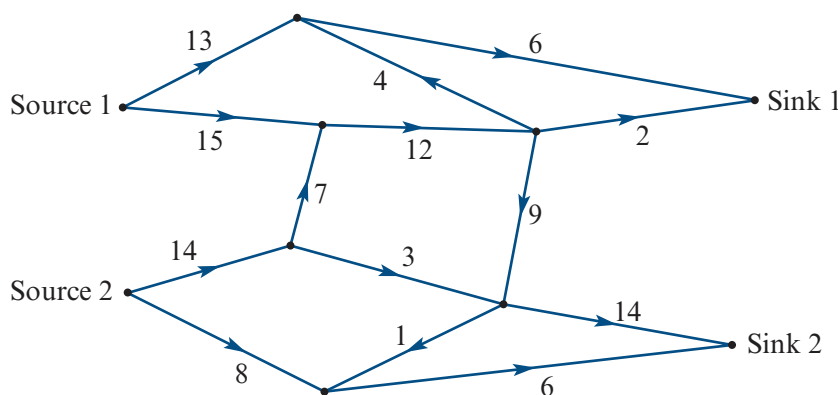


- The capacity of each pipe, in megalitres per day, is shown as weights on the arcs in the network.
- a What is the maximum flow of water that can reach the sink at A from the source at E ?
- b How many litres of water will flow into the sink over a three-day period?
- 8 Four craftsmen A, B, C and D are to be assigned tasks X, Y, Z and W . Each craftsman can only do one task. The table below shows the number of hours each craftsman takes to complete a task.

	X	Y	Z	W
A	48	49	42	42
B	53	49	51	50
C	51	53	48	48
D	47	50	46	43

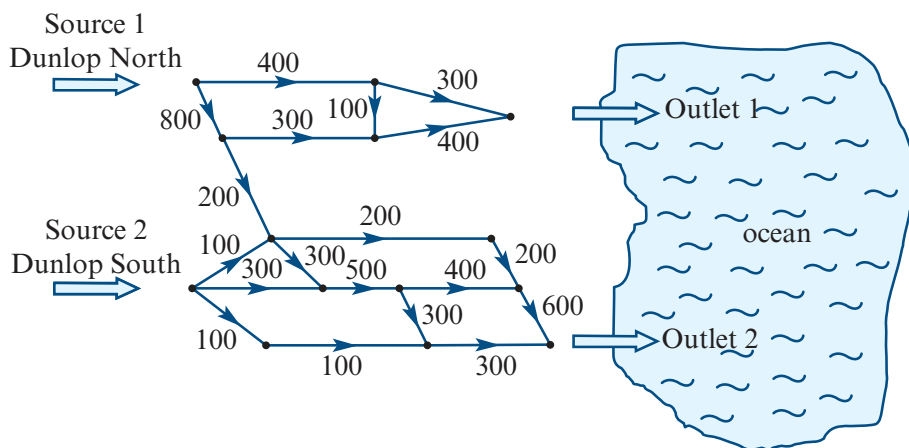
Find the minimum and maximum time to complete the task if each craftsman can only complete one task.

- 9 Water pipes of different capacities are connected to two water sources and two sinks. The network of water pipes is shown in the diagram below. The numbers on the edges represent the capacities, in kilolitres per minute, of the pipes.



Find the maximum flow, in kilolitres per minute, to each of the sinks in this network.

- 10 Storm water enters a network of pipes at either Dunlop North (Source 1) or Dunlop South (Source 2) and flows into the ocean at either Outlet 1 or Outlet 2. On the network diagram below, the pipes are represented by straight lines with arrows that indicate the direction of the flow of water. Water cannot flow through a pipe in the opposite direction. The numbers next to the arrows represent the maximum rate, in kilolitres per minute, at which storm water can flow through each pipe.



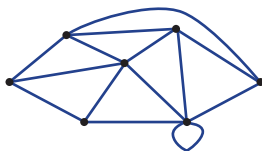
Determine the maximum rate, in kilolitres per minute, at which water can flow from these pipes into the ocean at Outlet 1 and Outlet 2.

Chapter 13

Revision of Units 3&4 Chapters 1–11

13A Paper 1 revision questions**Multiple-choice questions**

- A sequence is generated from the recurrence relation $V_1 = 15$, $V_{n+1} = V_n - 5$. The rule for the value of V_n is
A $V_n = 15 - 5n$ **B** $V_n = n + 10$
C $V_n = 5n + 15$ **D** $V_n = 20 - 5n$
- The third and seventh terms of an arithmetic sequence are 13 and 37 respectively. The rule for the n th term is
A $V_n = 13n + 38$ **B** $V_n = 38 - 13n$ **C** $V_n = -5 + 6n$ **D** $V_n = 13 + 25n$
- City A has latitude 15°N and longitude 50°E . City B has latitude 25°S and longitude 50°E . The shortest distance along the meridian between A and B, in kilometres, is closest to
A 1112 **B** 4448 **C** 111200 **D** 2657
- Location A has latitude 25°S and longitude 50°E . Location B has latitude 25°S and longitude 150°E . The distance along the small circle of the 25°S latitude between A and B, in kilometres, is closest to
A 1112 **B** 5432 **C** 10078 **D** 11 022
- Researchers believe that the level of background music will affect students test scores. They devise an experiment where *test scores* are measured at three different *music levels* (1 = soft music, 2 = moderate music, 3 = loud music). The response variable, and its classification are:
A *music level*, categorical **B** *test score*, categorical
C *music level*, numerical **D** *test score*, numerical
- Consider the graph below.



The number of vertices with a degree greater than or equal to 5 is

- A** 0 **B** 1 **C** 2 **D** 3
- What is the maximum number of edges a bipartite graph with 30 vertices can have?
A 60 **B** 125 **C** 200 **D** 225

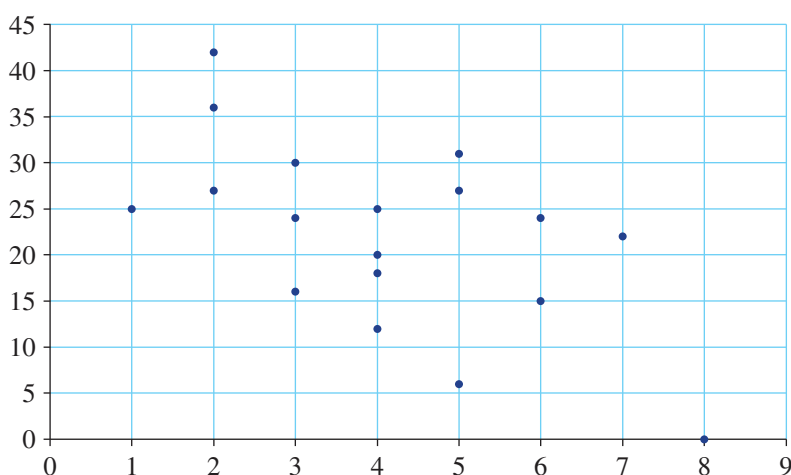
Use the following information to answer Questions 8–10.

The data in the following table were collected when a group of 250 students from Year 9 and Year 10 were asked where they would like to go for their school camp.

	Year level		
Preferred camp	Year 9	Year 10	Total
Beach	78	43	121
Snow	40	66	106
Other	12	11	23
Total	130	120	250

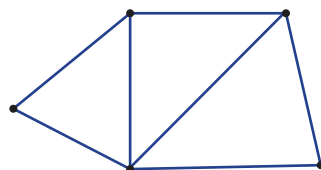
- 8** The percentage of students surveyed who preferred to go to the beach is closest to:
A 48.4% **B** 60.0% **C** 35.8% **D** 65.0%
- 9** Of those students surveyed who are in Year 9, the percentage who chose snow is closest to:
A 16.0% **B** 30.7% **C** 37.7% **D** 42.4%
- 10** The data in the table supports the contention that there is an association between *preferred camp* and *year level* because:
A 35.8% of students in Year 10 preferred to go to the beach, compared to only 55.0% of Year 10 students who preferred to go to the snow.
B 60.0% of students in Year 9 preferred to go to the beach, compared to only 30.8% of Year 9 students who preferred to go to the snow.
C 48.2% of students preferred the beach compared to 42.4% who preferred the snow.
D 60.0% of students in Year 9 preferred to go to the beach, compared to only 35.8% of Year 10 students who preferred to go to the beach.
- 11** Harper invested \$30 000 for 5 years at an interest rate of 5.4% per annum, compounding monthly. The rule for the balance of her investment, A , after 5 years is
A $A = 30\,000 \times 1.54^5$
B $A = 30\,000 \times 1.054^{60}$
C $A = 30\,000 \times 1.045^5$
D $A = 30\,000 \times 1.045^{60}$
- 12** Elijah borrowed some money. He will be charged compound interest at the rate of 8.76% per annum, compounding monthly. After one year, Eli repaid \$5237.78 as principal and interest. The amount borrowed was closest to
A \$4800 **B** \$5000
C \$5979 **D** \$59 792

- 13** The following scatterplot shows a linear association between two numerical variables.



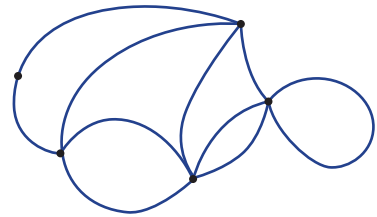
Choose the best description for the direction and strength of the association.

- A** strong positive **B** strong negative
C moderate positive **D** moderate negative
- 14** The equation of the least squares regression line that enables number of mistakes made on a test (y) to be predicted from the time spent studying for the test (x), where:
 $r = -0.515$ $\bar{x} = 25.33$ $s_x = 8.50$ $\bar{y} = 4.08$ $s_y = 1.73$
 is:
- A** $mistakes = -2.53 \times time + 35.67$ **B** $time = 35.67 \times mistakes - 2.53$
C $mistakes = -0.105 \times time + 6.74$ **D** $time = 6.745 \times mistakes - 0.105$
- 15** The coordinates of two points X and Y on the Earth's surface are (15°N ; 15°W) and (25°S , 150°E). Which statement is most likely to be correct about the time difference?
- A** X is 11 hours behind Y. **B** X is 9 hours behind Y.
C Y is 11 hours behind X. **D** Y is 9 hours behind X.
- 16** Point X on the Earth's surface has coordinates (25°N , 32°E), while point Y is at (16°S , 32°E). The distance between X and Y is closest to:
- A** 960 km **B** 4560 km **C** 1000 km **D** 1250 km
- 17** Consider the graph opposite. Euler's formula can be verified for this graph. What values of e , v and f can be used in this verification?
- A** $e = 5, v = 5, f = 3$
B $e = 7, v = 5, f = 4$
C $e = 6, v = 5, f = 3$
D $e = 6, v = 5, f = 4$



- 18** The graph opposite has five vertices and ten edges. How many of the vertices in this graph have an even degree?

A 0 **B** 1 **C** 2 **D** 3



- 19** Noah purchased a motorbike for \$15 000, which is depreciated using the flat-rate depreciation method at a rate of 3.5% per annum. The amount of depreciation after five years is

A \$1575 **B** \$525 **C** \$2625 **D** \$52 500

Use following information to answer Questions 20 and 21.

In a study of the effect of sunlight on the growth of seedlings, a number of seedlings were planted in various locations in a garden bed. The average number of hours of sunlight each plant received each day for 14 days (*sunlight hours*), as well as the amount they grew, in cm, over that period of time (*growth*) were recorded. *Growth* can be predicted from *sunlight hours* from the regression line:

$$\text{growth} = 0.586 \times \text{sunlight hours} + 4.28, \text{ with } r = 0.690$$

- 20** The percentage of variation in *growth* NOT explained by the variation in the *sunlight hours* is closest to:

A 52.4% **B** 31.0% **C** 47.6% **D** 69.0%

- 21** This regression line predicts that, on average, *growth*:

A increases by 4.28 cm for each additional hour of sunlight
B increases by 0.586 cm for each additional hour of sunlight
C decreases by 0.586 cm for each additional hour of sunlight
D decreases by 4.28 cm for each additional hour of sunlight

- 22** The following is a geometric sequence 100, 90, 81, 72.9, ...
 The common ratio r is equal to

A -10 **B** -9 **C** 0.9 **D** 1.1

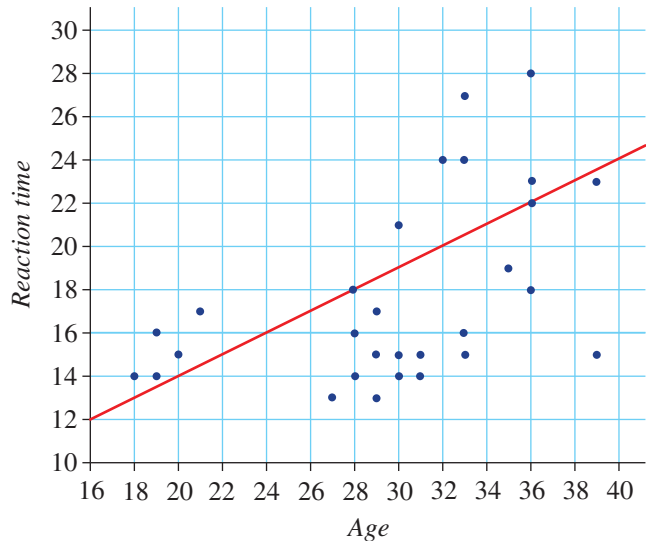
- 23** For a connected graph with five vertices and five edges, the sum of the degrees of the vertices is

A 6 **B** 8 **C** 9 **D** 10

- 24** Cairns is 10 hours ahead of GMT. Delhi in India is 5 hours 30 minutes hours ahead of GMT. What is the time in Delhi when it is 1 p.m. in Cairns?

A 6:30 p.m. **B** 7 p.m. **C** 6:30 a.m. **D** 8:30 a.m.

- 25** The scatterplot shows the reaction time, measured in hundredths of seconds, for a group of 20 people, together with their age. A least squares line has been fitted to the scatterplot with *age* as the explanatory variable, and *reaction time* as the response variable. The equation of the least squares line is closest to:



- A** $\text{reaction time} = 2.0 \times \text{age} + 26$ **B** $\text{reaction time} = 0.5 \times \text{age} + 4.0$
C $\text{reaction time} = 0.5 \times \text{age} + 12.0$ **D** $\text{reaction time} = 0.3 \times \text{age} + 12.0$
- 26** The table below shows the *life expectancy* in years and the percentage of government expenditure which is spent on *health* in eight countries.

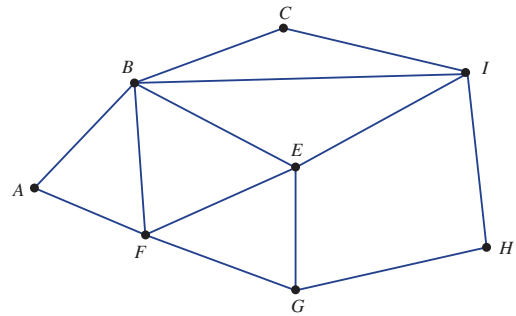
<i>Health (%)</i>	17.3	10.3	4.7	6.0	20.1	6.0	13.2	7.7
<i>Life expectancy (years)</i>	82	76	68	69	83	75	76	76

A least squares line which enables a country's *life expectancy* to be predicted from their expenditure on *health* is fitted to the data. The value of the residual (to the nearest year) when the actual percentage of government expenditure which is spent on health is 6% is closest to:

- A** -3 **B** -2 **C** 2 **D** 3
- 27** A car purchased on 1 June 2024 loses value at a reducing-balance depreciation rate of 15% per year. The original purchase price was \$92 000. The value of the car on 1 June 2027 will be closest to
- A** \$78 200 **B** \$66 470 **C** \$56 500 **D** \$50 600
- 28** The balance of a reducing-balance loan after n months, A_n , can be modelled by the recurrence relation $A_0 = 450\,000$, $A_{n+1} = 1.0069 \times A_n - 2100$. The total interest that has been paid after one repayment is closest to
- A** \$310.50 **B** \$1005 **C** \$3105 **D** \$5205
- 29** Izzy has a loan with monthly repayments of \$3200 over 10 years, and the annual interest rate is 4.86%, compounding monthly. The present value of the loan is closest to
- A** \$8221 **B** \$283 591 **C** \$1 438 325 **D** \$303 656

- 30** If an extra edge is added to this graph it will have an Eulerian circuit. The edge is

- A** AB
B BG
C QG
D IB



- 31** The table below records the monthly electricity cost (in dollars) for an apartment over one calendar year.

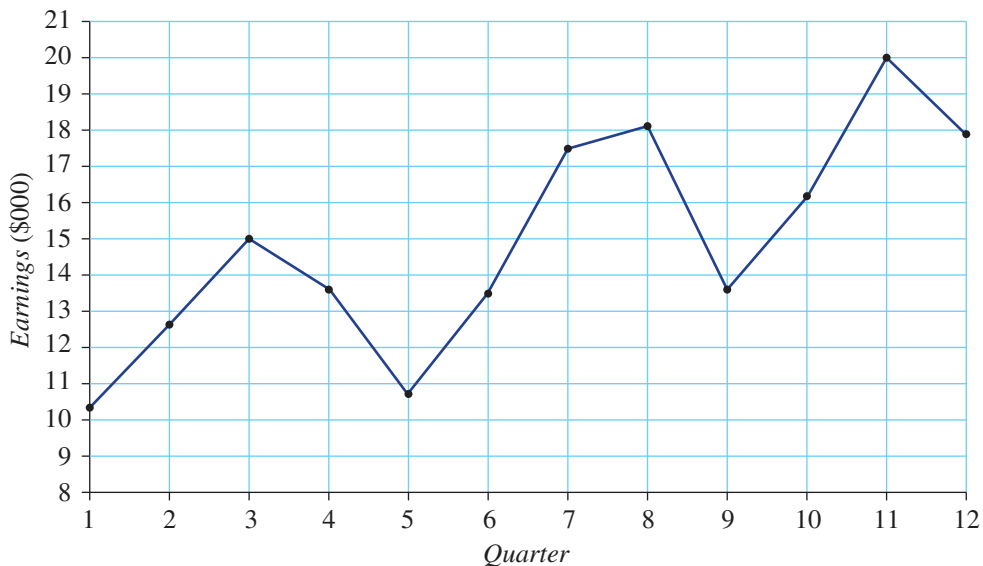
Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
123	90	153	136	101	129	153	143	95	61	85	107

The five-mean smoothed cost of electricity in July is closest to:

- A** \$141.67 **B** \$129.00 **C** \$124.20 **D** \$150.00

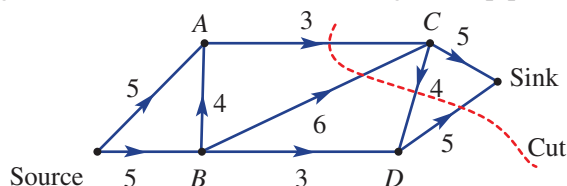
Use following information to answer Questions 32 and 33.

The time series plot below shows earnings per quarter (\$000) for a certain salesperson over a three year period.



- 32** The time series plot is best described as showing:
- A** seasonality only
B seasonality with irregular fluctuations
C an increasing trend with irregular fluctuations
D an increasing trend with seasonality and irregular fluctuations

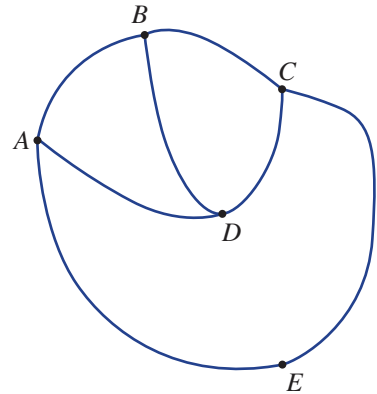
- 33** The nine-median smoothed earnings for the salesperson in Quarter 8, in \$'000s, is closest to:
- A** 14 **B** 16
C 17.5 **D** 18
- 34** A gas pipeline is to be constructed to link several towns in the country. Assuming the pipeline construction costs are the same everywhere in the region, the cheapest network formed by the pipelines and the towns as vertices would form:
- A** a Hamiltonian circuit **B** an Euler circuit
C A minimum length spanning tree **D** a critical path
- 35** The flow of water through a series of pipes is shown in the network below. The numbers on the edges show the maximum flow through each pipe in litres per minute.



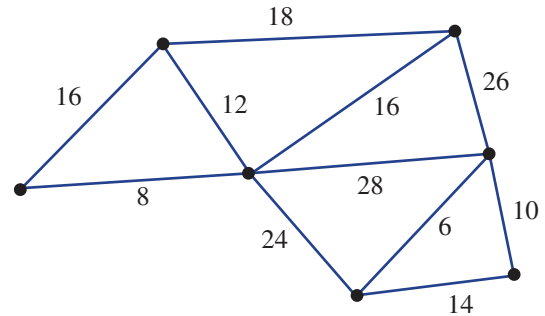
The capacity of the cut in litres per minute, is

- A** 10 **B** 12
C 14 **D** 18
- 36** An annuity in withdrawal phase with principal \$454 000 earns interest at an annual percentage rate of interest of 8.4%, compounding quarterly. Quarterly payments of \$9800 is made. The recurrence relation that models the annuity is
- A** $A_0 = 454\,000, A_{n+1} = 0.84 \times A_n - 9800$
B $A_0 = 454\,000, A_{n+1} = 1.084A_n - 9800$
C $A_0 = 454\,000, A_{n+1} = 1.021 \times A_n - 9800$
D $A_0 = 454\,000, A_{n+1} = 1.21 \times A_n - 9800$
- 37** Margot has inherited \$400 000 and will invest this money into an annuity from which she will withdraw monthly payments. Interest will be earned at the rate of 5.64% per annum, compounding monthly. The balance of Elvira's investment was \$398 384.89 after five payments have been withdrawn. The value of Elvira's monthly payment is
- A** \$323 **B** \$350
C \$2200 **D** \$3495

- 38** The network below shows the cabling between five locations, A, B, C, D and E.
An adjacency matrix for this network is formed.
The number of zeros in this matrix is
- A** 8 **B** 9 **C** 10 **D** 11



- 39** The minimum spanning tree for the graph opposite has a weight of
- A** 52 **B** 72
C 76 **D** 80



- 40** The table below shows the long-term mean monthly sales figures (in \$'000s) for a company, and the associated seasonal indices for the sales. The long-term mean sales figure for January is missing.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Sales		80.0	70.3	62.6	54.6	55.0	52.1	54.2	56.5	52.8	61.8	99.7
SI	0.727	1.289	1.132	1.008	0.880	0.886	0.840	0.874	0.911	0.850	0.996	1.607

The long-term mean sales figure for January is closest to:

- A** 45.1 **B** 58.9 **C** 62.1 **D** 73.4
- 41** The number of job applications received by a large supermarket chain is seasonal. Data has been collected, and a least squares regression line fitted to the deseasonalised data. The equation of the line is

$$\text{deseasonalised job applications} = 12.27 \times \text{month number} + 457.8$$

where *month number* 1 is January 2025.

The monthly seasonal indices for job applications are shown in the following table:

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1.14	1.06	1.22	1.03	0.95	0.95	0.83	0.70	0.78	0.88	1.23	1.23

The actual number of job applications predicted for January 2026 is closest to

- A** 704 **B** 617 **C** 542 **D** 704

- 42** Chloe is taking out a loan of \$500 000 with an interest rate of 4.8% annually, compounding quarterly. The loan is to be paid off over 15 years. Her quarterly repayment is

A \$4752 **B** \$9390 **C** \$11 738 **D** \$11 740

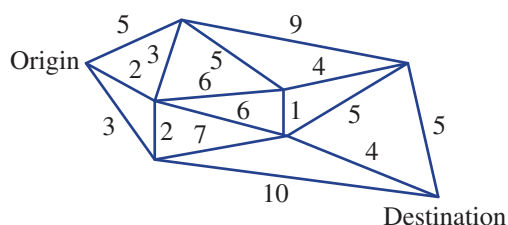
- 43** Adam has invested \$80 000 in an annuity. His investment will earn interest at the rate of 8.16% per annum, compounding monthly. Adam will withdraw \$1950 a month from this annuity.

The number of payments of \$1950 that Adam can expect from this annuity is

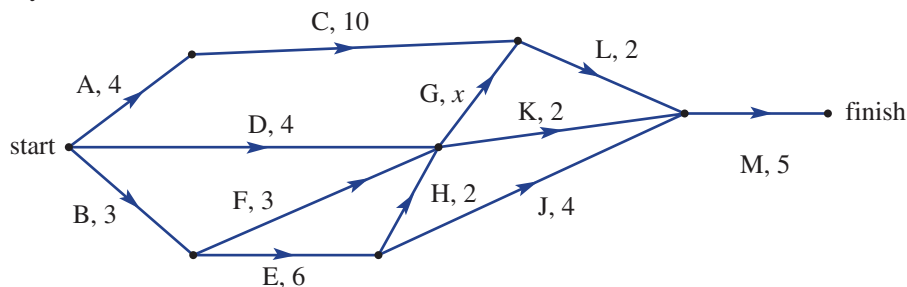
A 46 **B** 47 **C** 48 **D** 49

- 44** The shortest path between the origin and destination in the network shown here is:

A 11 **B** 12
C 13 **D** 14



- 45** The directed graph below shows the sequence of activities required to complete a project. The time taken to complete each activity, in hours, is also shown. The minimum completion time for this project is 21 hours. The time taken to complete activity G is labelled x . The maximum value of x is



A 1 **B** 2 **C** 3 **D** 5

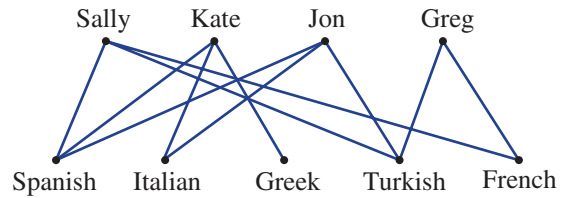
- 46** Petra has an annuity with an initial balance of \$340 000. She makes deposits of \$1750 each month for a period of 5 years. After this, \$1750 is withdrawn each month for 2 years. The interest rate over the 7 years is 6.6% per annum, compounding monthly. The future value of the annuity after 7 years is closest to

A \$326 167 **B** \$596 501 **C** \$629 263 **D** \$635 659

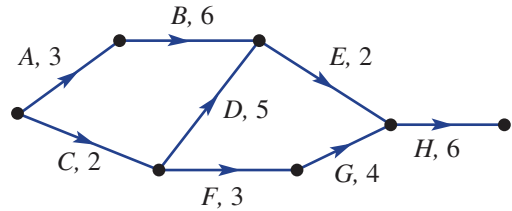
- 47** A perpetuity will be set up to provide an annual prize of \$1000 to the winner of a poetry competition. Interest will be earned on the principal of the investment at the rate of 5.12% per annum and this will be used to pay the prize money every year. The amount that must be invested is

A \$51.20 **B** \$19 531.25 **C** \$5120 **D** \$234 375

- 48 Which one of the following statements is not implied by this bipartite graph?



- A** Sally and Kate can translate five languages between them.
B Jon and Greg can translate four languages between them.
C Kate and Jon can translate more languages between them than can Sally and Greg.
D Sally and Jon can translate more languages between them than can Kate and Greg.
- 49 The edges in the network diagram correspond to the tasks involved in the preparation of an examination. The numbers indicate the time, in weeks, needed for each task. The total number of weeks needed for the preparation of the examination is:



- A** 14 **B** 15 **C** 16 **D** 17

Short-response questions

1 Consider the graph shown here.

a i Explain why this is a planar graph.

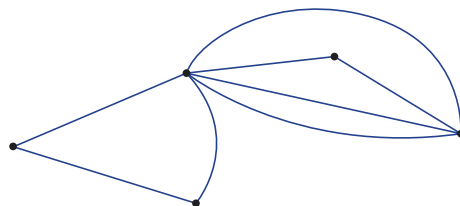
ii For this graph, write down:

- the number of vertices
- the number of edges
- the number of faces.

b i Draw a spanning tree for this graph.

ii For the spanning tree drawn in part i, write down:

- the number of vertices
- the number of edges
- the number of faces.



2 A teacher is concerned that students who spend a lot of time playing video games do not spend enough time reading. The following table shows the data she collected from a group of 10 of her students, who recorded the number of hours they spent reading and the number of hours they spent playing computer games in one week.

<i>Reading</i>	10	4	7	8	6	3	4	1	10	8
<i>Games</i>	7	15	13	15	8	20	10	21	0	2

a Construct a scatterplot of these data, with *games* as the explanatory variable and *reading* as the response variable.

b From the scatterplot, describe the association between *reading* and *games* in terms of direction, form and strength.

3 The *age* in years and percentage *body fat* on food for a group of eight people is given in the following table.

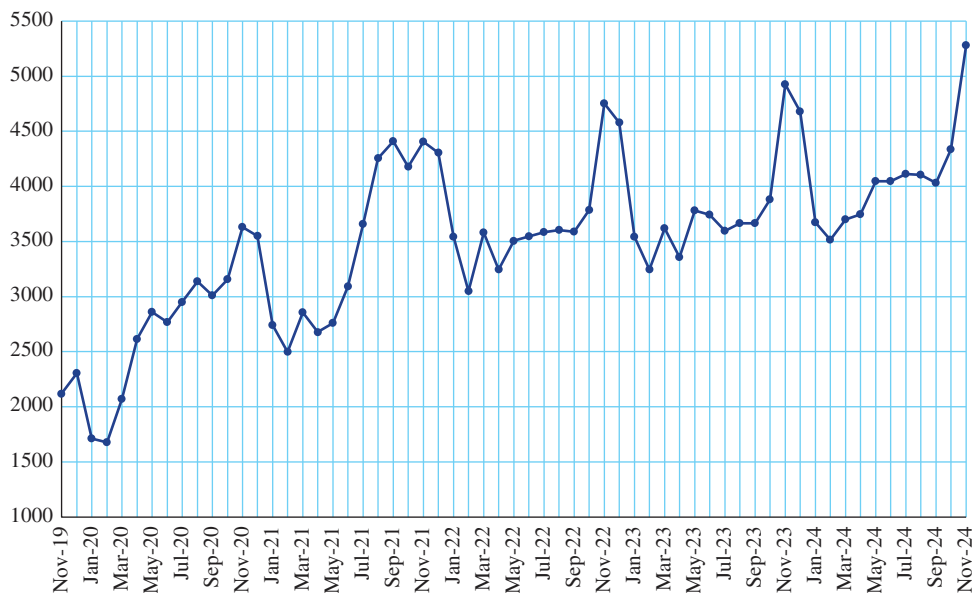
<i>Age</i>	58	72	67	43	51	52	25	35
<i>Body fat</i>	20.1	26.1	25.8	19.5	14.1	27.0	6.1	4.1

Determine the equation of the least squares regression line which will allow *body fat* to be predicted from *age*. Give your answer correct to two decimal places.

4 Two locations lie on the same meridian of longitude. One is 25° south of the other. What is the distance between the two locations, correct to the nearest kilometre?

5 Two locations lie on the same parallel of latitude 10°S . One is 125° east of the other. What is the distance between the two locations around the latitude 10°S small circle, correct to the nearest kilometre?

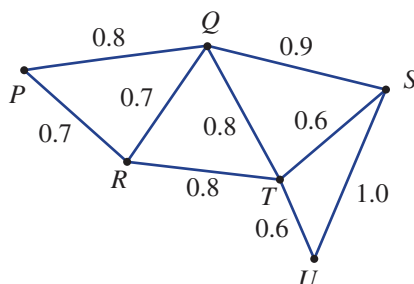
- 6** For each of the following sequences, find the required term.
- a** For an arithmetic sequence with $a = t_1 = 8$ and $d = 3$, determine t_{10} .
 - b** For an arithmetic sequence with $a = t_1 = 300$ and $d = -9$, determine t_{10} .
 - c** For a geometric sequence with $a = t_1 = 20\,000$ and $r = 0.9$, determine t_4 .
 - d** For a geometric sequence with $a = t_1 = 3$ and $r = 2$, determine t_5 .
- 7** Write down the values of t_1 , t_2 and t_3 for each of the following sequences.
- a** An arithmetic sequence with rule $t_n = 6n - 5$.
 - b** A geometric sequence with rule $t_n = 2 \times 5^{n-1}$.
- 8** Write down a rule for the n th term of the sequence for each of the following.
- a** A sequence is defined by the recurrence relation $t_n = t_{n-1} - 10$, with $t_1 = 220$.
 - b** A sequence is defined by the recurrence relation $t_n = 5t_{n-1}$, with $t_1 = 3$.
- 9** The time series plot below shows the amount spent on online retail sales in Australia (\$ millions) each month from November 2019 until November 2024.



Describe the features of the plot.

- 10** Drake has \$500 000 to invest and a bank offers him an interest rate of 6.1% per annum, compounded annually, for the duration of the investment. Find how much his investment will be worth in six years time, giving your answer correct to the nearest dollar.

- 11** A rural town, built on hills, contains a set of roads represented by arcs in the following network. The numbers on the network refer to distances along the roads (in kilometres) and the letters refer to intersections of the roads. The arcs without endpoints refer to the two roads in and out of town.

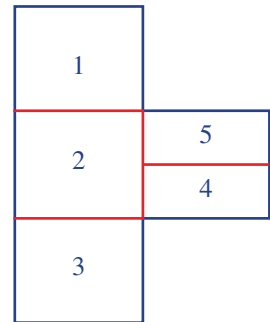


- a**
- i** What is the length of the shortest route through the town from P to U ?
 - ii** A safety officer who enters the town at P needs to examine all intersections in the town before leaving from U to travel on to the next town. To save time she wants to pass through each intersection only once. State a path through the network of roads that would enable her to do this.
- b** A technician from the Electricity Company is checking the overhead cables along each street. The technician elects to follow a semi-Eulerian trail path through the network streets (ignoring the roads in and out of town) starting at R and finishing at S .
- i** Complete the following semi-Eulerian trail.
 $R - Q - P - R - \square - \square - \square - T - U - S$
 - ii** How would the technician benefit from choosing an Euler path?
- 12** The equation of the least squares line that relates the fuel consumption of a certain car, in litres/km, to the speed at which the car is travelling, in km/hr is:

$$\text{fuel consumption} = 0.0218 \times \text{speed} + 6.827$$

- a** Use the equation to predict the *fuel consumption* of the car if it is travelling at 100 km/hr. Round the answer to one decimal place.
- b** Write down the slope of the regression line and interpret in terms of *fuel consumption* and the *speed*.
- c** When the *speed* was 72 km/hr, the actual *fuel consumption* was 8.3 litres/km. Show that, when the least squares line is used to predict the fuel consumption at 72 km/hr, the residual is -0.10 rounded to two decimal places.
- 13** Suppose that the correlation between sales and advertising spend are linearly related, and that $r = -0.68$.
- a** Determine the value of the coefficient of determination, R^2 .
 - b** Interpret R^2 in terms of the variables sales and advertising spend.

- 14** A very large country home garden is divided into five regions labelled 1 to 5 on the diagram opposite. The red lines represent the boundary stonewalls between two regions.



- a** Draw a graph where the five regions of the garden are represented as vertices and the edges of the graph represent the boundary stonewalls between areas.
- b** What is the sum of the degrees of the vertices of this graph?
- 15** Darcy has \$120 000 to invest. For the first five years, she has an interest rate of 4.3% per annum, compounding annually. For the next five years, she has an interest rate of 4.2%, compounding monthly. Find the value of the investment after 10 years.
- 16** The statistical analysis of the set of bivariate data involving variables x and y resulted in the information displayed in the table below:

	x	y
mean	12.5	34.6
standard deviation	3.42	6.84
correlation coefficient	$r = -0.789$	

Use this information to determine the equation of the least squares regression line.

- 17** In a survey of people aged 18 years or more, respondents were asked whether they felt the government was doing enough to reduce homelessness (yes or no), and they were also classified by age group (under 30 years, 30 years or more). The results are summarised in the following two-way frequency table.

<i>Doing enough to reduce homelessness?</i>	Under 30 years	30 years or more
Yes	130	130
No	70	170
Total	200	300

- a** Name the explanatory variable.
- b** Convert the table values to percentages by calculating the column percentages.
- c** Explain whether there is an association between age and agreement with the statement that the government is doing enough to reduce homelessness.

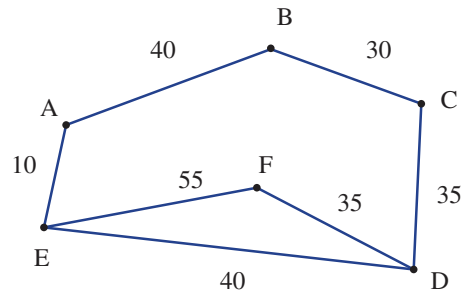
- 18** The following recurrence relation can be used to model a compound interest investment of \$45 000 earning interest at the rate of 6.78% per annum, compounding monthly.

$$A_0 = 45\,000, A_{n+1} = 1.00565 \times A_n$$

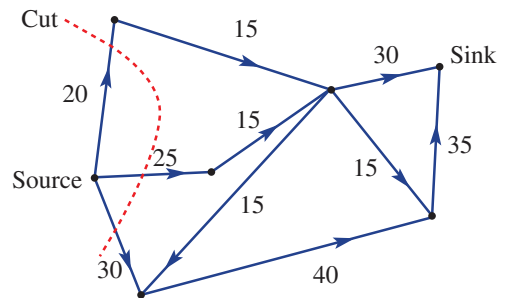
In this recurrence relation, A_0 is the balance of the investment after n months.

- Apply the recurrence relation to find the balance of the investment after one, two and three months.
 - Find how many months it will take for the balance of this investment to first exceed \$46 000.
- 19** Maddy has invested \$17 650 into an account that pays compound interest at the rate of 5.4% per annum, compounding monthly.
- Construct a recurrence relation model for the balance of Maddy's investment after n months.
 - Apply the recurrence relation to determine the balance of Maddy's investment after three months.
 - Using the compound interest formula, determine the balance of Maddy's investment after two years.

- 20** Region 1 of a garden contains 6 circular beds that are labelled A to F , as shown in the graph opposite. The owner wants to have a walk around region 1 visiting each bed on the way. The numbers on the edges joining the vertices give the shortest distance, in metres, between beds.

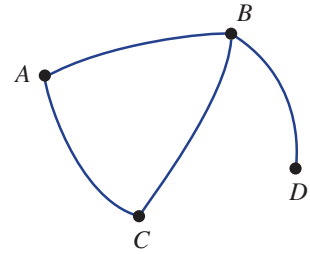


- Explain why the owner could not follow an Eulerian circuit through this network.
 - If the owner follows the shortest Hamiltonian path, name a garden bed at which the owner could start and a garden bed at which the owner could finish.
 - List a semi-Eulerian trail for the graph.
- 21** All areas of a garden require a constant supply of water. The directed graph opposite shows an irrigation system for the garden with the capacity of each section shown in litres per minute. The beginning of the system is labelled source and the end of the system labelled sink.



- What is the capacity of the marked cut in litres per minute?
- Determine the maximum flow of water, in litres per minute, from the source to the sink.

- 22** Construct the adjacency matrix for the graph shown here.



- 23** The table shows the distance each of a group of eight students travels to school, and the time taken when travelling by car.

<i>Distance (km)</i>	3	10	10	15	20	25	30	40
<i>Time (mins)</i>	5	10	18	10	30	25	30	42
<i>Residual</i>								

The equation of the least squares line is:

$$\text{time} = 1.0 \times \text{distance} + 2.8$$

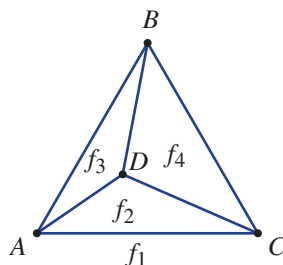
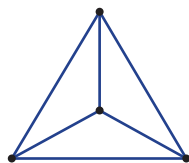
- a** Complete the table of residuals.
b Construct a residual plot.
- 24** The table below shows the quarterly house sales achieved by a real estate company in the years 2024-2025.

<i>Year</i>	Q1	Q2	Q3	Q4
2024	52	59	68	27
2025	57	65	75	29

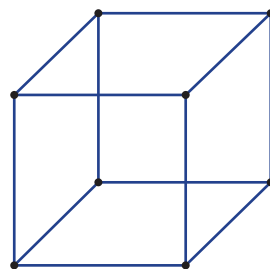
Use the data in the table to find seasonal indices. Give your answers rounded to two decimal places.

- 25** Rob has borrowed \$8200 to buy furniture for his home. He will be charged compound interest at the rate of 6.48% per annum, compounding monthly. The full balance is paid at the end of the loan.
 Let A_n be the balance of Jack's loan after n months.
- a** State the monthly percentage rate of interest for this loan.
b Construct a recurrence relation that models the balance of Jack's loan.
c Jack pays the principal and all interest charged after one year. Find how much money will he have to repay.
- 26** Simon has invested \$65 000 in an account that will pay compound interest every month, at the annual percentage rate of interest of 4.68%.
- a** Use the compound interest rule to determine the balance of Simon's investment after three years.
b Find how much interest has been earned in total after three years.

- 27** Beth borrowed some money to pay for a travelling holiday. The annual percentage rate of interest for her loan was 9.48%, compounding monthly. Beth repaid the principal and all interest charged, a total sum of \$13 627.97, after travelling for six months. Use the compound interest rule to determine the principal amount that Beth borrowed. Give your answer correct to the nearest dollar.
- 28** A tetrahedron may be represented by a connected planar graph as shown.



- a** State the number of edges and vertices.
b Verify Euler's formula.
- 29** Represent this cube by a planar graph and verify Euler's formula.



- 30** The adjacency matrix opposite shows the number of pathways between four points, A , B , C and D . Draw a graph that is represented by this adjacency matrix.

$$\begin{array}{c}
 \begin{array}{c} A \\ B \\ C \\ D \end{array}
 \begin{bmatrix}
 & A & B & C & D \\
 A & 0 & 1 & 0 & 2 \\
 B & 1 & 0 & 1 & 0 \\
 C & 0 & 1 & 0 & 1 \\
 D & 2 & 0 & 1 & 0
 \end{bmatrix}
 \end{array}$$

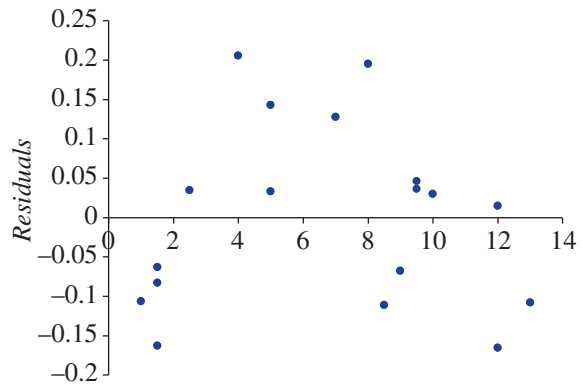
- 31** Lacey's restaurant is open for dinner from Wednesday to Sunday. The number of diners at the restaurant over a two-week period, together with the daily seasonal indices, are shown in the table below:

Week	Wed	Thur	Fri	Sat	Sun
1	45	67	130	154	90
2	57	74	115	150	105
Seasonal index	0.5	0.7	1.3	1.5	1.0

- a** Use the seasonal indices to deseasonalise the data, rounding answers to the nearest whole number.
b Construct a time series plot of the deseasonalised data.

- 32** The table below shows the age in years (*age*) and the length in metres (*length*), for a group of 18 dugongs. A least squares regression line which allows us to predict the length of a dugong from its age was fitted to the data, and the residual plot for this analysis is also shown.

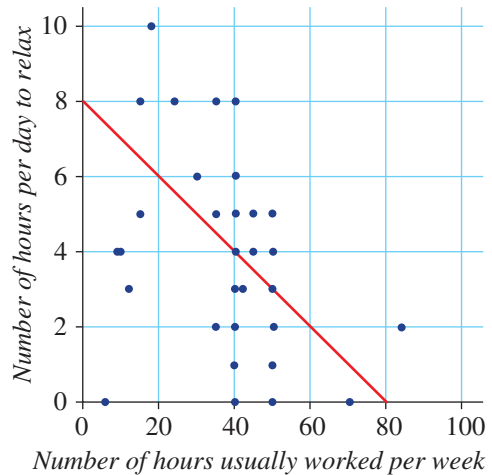
Age	Length	Age	Length
1.0	1.80	8.0	2.47
1.5	1.85	8.5	2.19
1.5	1.87	9.0	2.26
1.5	1.77	9.5	2.40
2.5	2.02	9.5	2.39
4.0	2.27	10.0	2.41
5.0	2.15	12.0	2.50
5.0	2.26	12.0	2.32
7.0	2.35	13.0	2.43



Comment on whether the residual plot supports the assumption that the relationship between the length of a dugong and its age is linear.

- 33** Sandy has \$183 000 to invest in an annuity. Interest will be paid at the annual percentage interest rate of 6.372%, compounding monthly. Sandy will withdraw a payment of \$2650 each month from the investment.
Let A_n be the balance of Sandy's annuity after n monthly payments have been withdrawn.
- Construct a recurrence relation model for the balance of Sandy's investment after n payment withdrawals.
 - Apply the recurrence relation to calculate the amount remaining in Sandy's investment after five payments have been withdrawn.
 - Calculate the total interest that Sandy has earned after five payments have been withdrawn.
- 34** Vancouver, Canada has longitude 123°W while Mackay has longitude 149°E .
- Calculate the difference in longitude between these two places.
 - Calculate the time difference between the two places. (Ignore time zones and daylight saving.)
 - What is the time in Mackay when it is 2:45 a.m. in Vancouver? (Ignore time zones and daylight saving.)
- 35** The position of Honolulu is $(21.3069^\circ\text{N}, 157.8583^\circ\text{W})$. A freighter is 4° to the north of Honolulu and 8° east of Honolulu. Give the latitude and longitude of the freighter correct to the nearest degree.

- 36** The following scatterplot shows the number of hours per day spent in relaxation (*relax*) by a group of people plotted against the number of hours per week they usually work (*work*). Determine the equation of the line which has been fitted to the scatterplot.



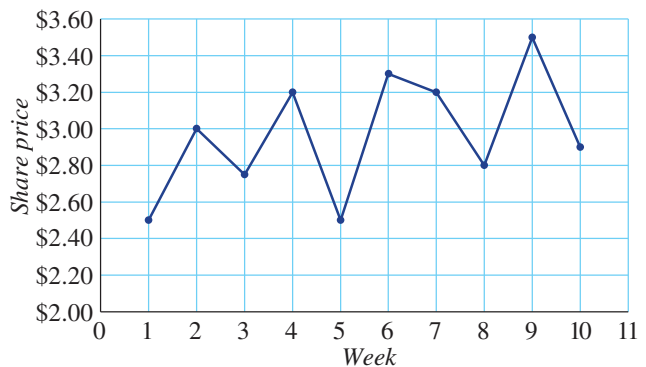
- 37** The following table gives the value of the average price of unleaded petrol in Queensland each year from 2018–2024.

Year	2018	2019	2020	2021	2022	2023	2024
Price (cents/litre)	143.4	141.1	123.9	147.6	184.3	189.7	188.7

- a** Determine the three-mean smoothed value of petrol in Queensland for the year 2022 in cents/litre, rounding your answer to one decimal place.
- b** Determine the three-median smoothed value of petrol in Queensland for the year 2023 in cents/litre, rounding your answer to one decimal place.
- 38** Pontianak, Indonesia has a longitude of 109°E , and a location in Columbia has a longitude of 74°W . Both places lie on the Equator. Find the shortest distance between these two places, to the nearest kilometre.
- 39** Anthony lives in Tully and wants to phone his grandfather in New York. It is 6 p.m. on Saturday in Queensland. What time is it in New York? (New York is 5 hours behind GMT and Rockhampton is 10 hours ahead of GMT.)
- 40** Olive is playing hockey at a tournament in Amsterdam. After her team wins the semi-final at 5 p.m. on Friday she phones her father in Ipswich to tell him the news. What time is it in Ipswich? (Ipswich GMT + 10, Amsterdam GMT +1)
- 41** The balance of an annuity, A_n , after n monthly payments have been received is modelled by the recurrence relation below.
- $$A_0 = 360\,000, A_{n+1} = 1.00875 \times A_n - 2120$$
- a** State the percentage annual rate of interest for this annuity.
- b** State the balance of the annuity after five payments have been received.
- c** Calculate the total amount of interest that has been earned after five payments have been received.

- 42** A consulting firm has invested \$161 500 in a perpetuity. The interest earned each month by this investment will pay for a monthly award to a high performing team within the organisation.
- a** The interest on the investment is paid at the rate of 3.24% per annum. Find the value of the prize awarded each month.
 - b** The agency wants the prize to be at least \$960 per month. Find the minimum annual percentage rate of interest for the investment that will allow this, rounding your answer to two decimal places.
- 43** Tommy invests \$2000 into an annuity, earning interest of 5%, compounding annually. He adds \$2000 a year for five years. He then withdraws \$2000 each year after that.
- a** Find the balance of the annuity after 5 years.
 - b** Find the balance of the annuity after a total of 10 years.
- 44** Jed invests \$5000 into an annuity, earning 6% interest, compounding annually. He deposits \$1000 each year.
- a** Find the balance of the annuity after four years.
 - b** State the total interest earned over the four years.

- 45** The time series plot opposite shows the maximum weekly price of shares in a company over a period of 10 weeks. Use three-median smoothing to graphically smooth the plot and comment on the smoothed plot.



- 46** The number of guests each month at a beach resort is seasonal, with seasonal indices as shown below:

Season	Summer	Autumn	Winter	Spring
Seasonal index	2.1	0.8	0.2	0.9

It is also known that over a specific time period the number of guests each season has generally been declining, according to the following equation which was determined from deseasonalised data:

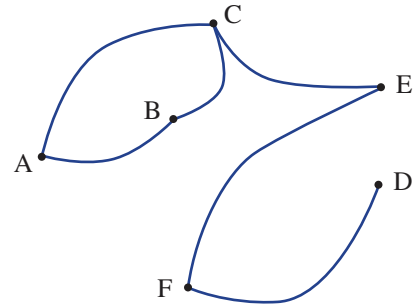
$$\text{number of guests} = -5 \times \text{season} + 335$$

where Season 1 is Summer 2025. Show that the actual number of guests predicted for the following Summer is 651.

13B Paper 2 revision questions

Short-response questions

- 1 a** The Penvale swimming club has six new members, A, B, C, D, E and F. The graph opposite shows the members who have competed together before joining the club. For example, the edge between A and B shows that they have previously competed together.



- i** How many of these players had E competed with before joining the club?
 - ii** Who had competed with both A and B before joining the club?
- b** The swimming club has a medley relay team. Three of the new club members, A, B and C, can complete the following sectors of the medley race: backstroke, butterfly and breaststroke. The table below shows the average times in seconds for 100 m for these sectors for each of the three swimmers. The freestyle swimmer has been chosen and has much better freestyle times than the three new members. How should the swimmers be allocated to minimise the team's time?

Swimmer	Backstroke	Breaststroke	Butterfly
A	72	74	66
B	68	72	62
C	70	76	62

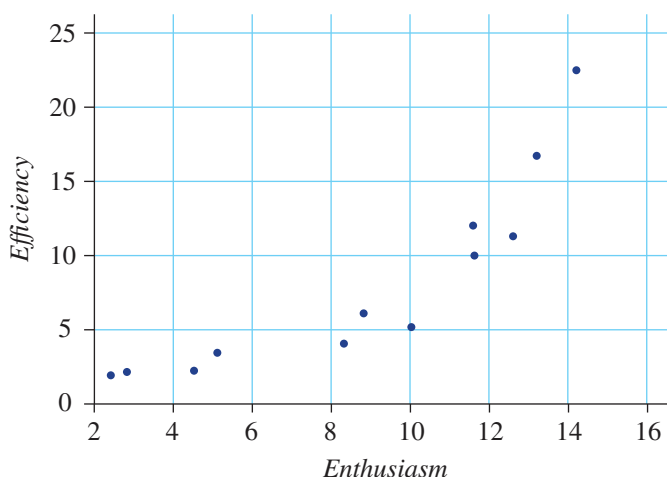
- c** The Penvale swimming club rooms are to undergo renovation. This project involves eight activities: A to I. The table opposite shows the earliest start time (EST) and duration, in months, for each activity. The immediate predecessor(s) is also shown. The duration for activity D is missing. The information in the table above can be used to complete an activity network.

Activity	EST	Duration	Immediate predecessor(s)
A	0	3	–
B	3	6	A
C	3	2	A
D	5	...	C
E	5	9	C
F	9	6	B
G	15	4	D
H	15	15	F
I	30	2	E, G, H

- i What is the duration, in months, of activity D?
 - ii Draw the associated activity network for this renovation.
 - iii Name the four activities that have a float time.
 - iv The project is to be crashed by reducing the completion time of one activity only. What is the minimum time, in months, that the project can be completed in?
- 2** The cost of hiring a car for a year involves a flat rate of \$3000 and then a cost of \$1.20 per kilometre travelled.
- a Write down the cost of hiring the car for a year where n kilometres are travelled.
 - b Find how much it will cost if the car travels 250 km in the month.
 - c Determine how many kilometres can be travelled in a month if no more than \$4500 can be spent on hiring the car.
- 3** Foxes are causing a problem in a nature reserve. Under normal conditions, the fox population grows at a rate of 15% per year. When counted at the start of the year, there were 680 foxes in the nature reserve.
- Write down a mathematical model of the form:
- $$N_{n+1} = rN_n \quad \text{where} \quad N_1 = a$$
- that can be used to describe the growth of the fox population in the nature reserve under normal conditions. (N_n represents the number of foxes in the nature reserve at the end of the n th year.)
- Using your model, plot a graph of fox numbers against the year for the first five years. Determine whether the fox population will exceed 1500 over this time period and whether the model is realistic.
- 4** A reducing-balance loan is modelled using the recurrence relation shown below.
- $$A_0 = 6000, A_{n+1} = 1.0024 \times A_n - 680$$
- In this recurrence relation, A_n is the balance of the loan after n weekly repayments. Determine the balance of this loan after five repayments.
- 5** Cindy is considering borrowing \$145 000 to buy a granny flat. Her bank will charge interest at the rate of 6.12% per annum, compounding monthly. Celia can afford to make monthly repayments of \$2720 to repay the loan.
- Let A_n be the balance of Cindy's loan after n months.
- a State the monthly percentage rate of interest for Cindy's loan.
 - b Determine how many months it will take for Cindy's loan to have a balance that is below \$100 000 for the first time.

- 6** In a study of the association between a person's *enthusiasm* for their job (a numerical variable measured on a scale from 0 to 15), and their *efficiency* when performing job (a numerical variable measured on a scale from 0 to 25), data was collected from a group of 12 employees.

The following scatterplot was constructed, with *enthusiasm* as the explanatory variable, and *efficiency* as the response variable.



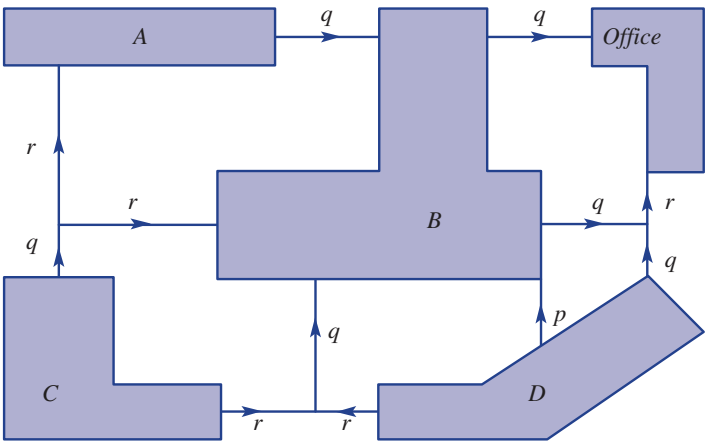
- a** Describe the association between *efficiency* and *enthusiasm* in terms of form and strength.
- b** Evaluate the appropriateness of calculating the value of the correlation coefficient r for this data.
- 7** To investigate the association between the weight of a certain species of fish in grams (*weight*) and its length in centimetres (*length*) data was collected from a sample of 35 fish. It was determined from the data that the mean length of the sample of fish is 30.306 cm with a standard deviation of 3.594 cm, and the mean weight of the sample of fish is 617.829 grams with a standard deviation 209.206 grams. The value of the correlation coefficient $r = 0.937$. Use this information to predict the weight of a fish which is 50 cm long. Give your answer rounded to the nearest gram.
- 8** Daniel will travel from Mt Gambier in South Australia (38°S , 141°E) to Sapporo (43°N , 141°E) on Wednesday on a charter flight. The flight will leave Mt Gambier at 11:20 a.m., and will take 11 hours and 40 minutes to reach Sapporo. Mt Gambier is thirty minutes ahead of Sapporo.
- a** On what day and at what time will Daniel arrive in Sapporo?
- b** What is the great circle distance between Mt Gambier and Sapporo?

- c** Ignoring time zones and summertime, what time is it in Marseille when it is 11:00 p.m. in Sapporo?
- d** Both cities are on the 43°N latitude. Find the shortest distance between the two cities following the 43°N small circle.

-

- b i** Modify your planar weighted graph from part **a ii** above to show only the shortest direct pathway between adjacent buildings.
- ii** How much will the covered walkways on these pathways cost to build?
- c** It has been decided that covering all of these walkways is too expensive. Only the minimum number of pathways that are necessary to allow students to walk from one building to any other while remaining under cover will be built.
 - i** Draw the graph that shows the pathways that should be covered so that the overall cost of making and installing the covering structure is a minimum.
 - ii** Calculate the cost of the covering structure in part **i**.

- d** In emergency situations, some of the doors in building *B* are locked and students are directed to evacuate the university via other pathways. The diagram below shows the locations of these evacuation pathways.

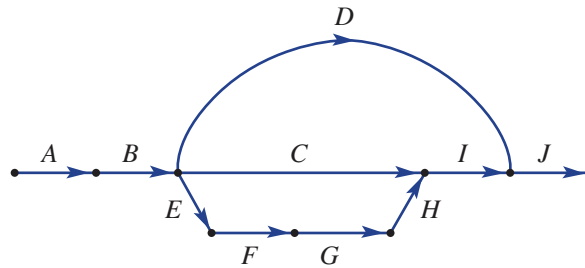


The pathways and doors allow different rates of students to flow along per minute. On the diagram:

- p = flow rate of 80 students per minute
 - q = flow rate of 120 students per minute
 - r = flow rate of 150 students per minute
- i** If there are 875 students in building *D* when the alarm bell rings, what is the minimum time it could take all students to leave this building? Assume there are no students in the other buildings.
- ii** Evening school is held in building *C*. On a particular evening there were 840 students in building *C*. Assume that there were no other students in any of the other buildings. How long would it take to evacuate all of these students through the office building?
- 11** The HiHo construction company builds a particular type of house using the project plan given in the table below.

Activity	Description	Duration (days)	Immediate predecessors
<i>A</i>	Build foundation	5	-
<i>B</i>	Constructing frame	8	<i>A</i>
<i>C</i>	Construct roof	12	<i>B</i>
<i>D</i>	Electrical wiring	5	<i>B</i>
<i>E</i>	Windows	4	<i>B</i>
<i>F</i>	Insulation	1	<i>E</i>
<i>G</i>	Installing plumbing	1	<i>F</i>
<i>H</i>	put on siding	6	<i>G</i>
<i>I</i>	paint house	3	<i>C, H</i>
<i>J</i>	add fixtures/fitings	3	<i>D, I</i>

A network diagram for this project is shown here.



- a Determine the earliest start time for activity *E*
 - b Determine the latest start time for activity *D*?
 - c Determine the minimum time to complete this project.
 - d List the critical paths for this project.
 - e Determine the float time for activity *D*.
- 12** Given the following information about two sequences, find the required term.
- a For an arithmetic sequence $t_6 = 37$ and $t_{10} = 65$, find t_{20} .
 - b For a geometric sequence $t_4 = 16$ and $t_{10} = 1024$, find t_7 .
- 13** The quarterly electricity bills(\$) for a household over a 2 year period are given in the following table.

Year	Quarter 1	Quarter 2	Quarter 3	Quarter 4
1	602	584	537	589
2	612	594	543	596

Construct a time series plot the deseasonalised data.

- 14** The equation of the least squares line that relates the fuel consumption of a certain car, in litres/km, to the speed at which the car is travelling, in km/hr is:

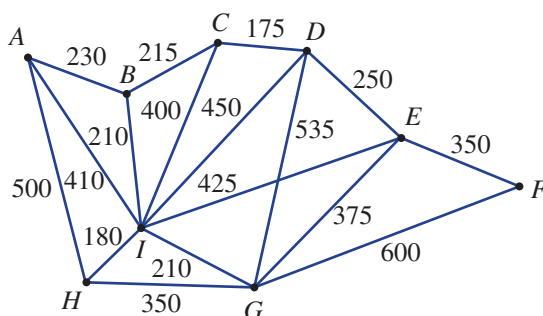
$$\text{Fuel consumption} = 0.0218 \times \text{speed} + 6.827$$

- a Use the summary statistics shown to determine the coefficient of determination as a percentage, rounded to one decimal place.

	Fuel consumption	Speed
mean	8.7556	88.444
standard deviation	0.52941	22.367

- b Interpret the value of the coefficient of determination in terms of *fuel consumption* and *speed*.
- 15** A caravan was purchased for \$88 000. It depreciates in value at the rate of 9% per year, using a flat-rate depreciation method.
- a Write down a rule for the value of the caravan, V_n , after n years.
 - b Use this rule to find the value of the caravan after six years.
 - c Find the total depreciation of the caravan over five years.

- 16** Yixuan borrows \$30 800 from a bank and is charged simple interest at the rate of 12% per annum. Let t_n be the value of the loan after n years.
- Write down a rule for the value of the loan after n years.
 - Find how much Yixuan need to pay the bank after 3 years.
 - Determine how many years it takes for the value of the loan to first exceed \$60 000.
- 17** A company has constructed a new industrial complex with 9 buildings in a layout as shown below. The minimum distances in metres between adjacent buildings in the complex are also shown.

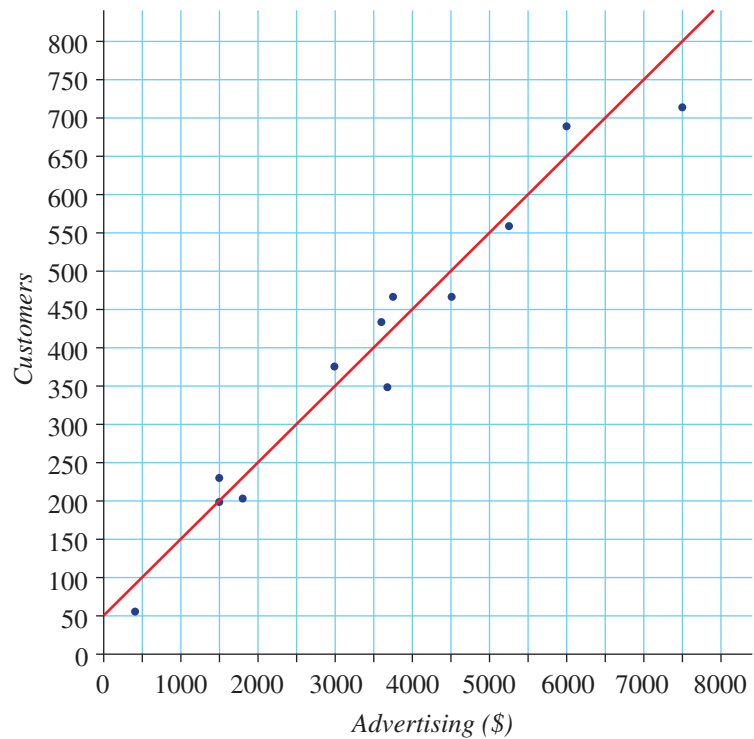


An electrical network is to be built to serve all the buildings.

- Draw a network that will ensure that all the buildings are connected to the network but that also minimises the amount of cable used. Label each node in the network.
 - Determine the minimum length of cable required.
- 18** Brisbane is 10 hours ahead of Coordinated Universal Time (UTC +10) and Miami is 5 hours behind Coordinated Universal Time (UTC−5). Elizabeth travels from Brisbane to Miami. She leaves Brisbane 7:10 p.m. on Thursday local time and travels to Miami. The flights takes 19 hours and 34 minutes. What time and day is it in Miami when the plane lands?
- 19** For a certain species of tree the correlation between tree growth and rainfall is 0.53, whilst the correlation between tree growth and temperature is −0.63.
- Interpret each of these correlation coefficients in terms of the variables in the study.
 - Explain with reasons which of the variables, *rainfall* or *temperature*, is the better predictor of tree growth.

- 20** Felix would like to buy an apartment and he will need to borrow \$620 000 to pay for this. Interest will be charged at the annual percentage interest rate of 4.92%, compounding monthly.
- a** Felix plans to repay his loan over a period of 25 years.
 - i** Calculate the monthly repayment amount required to achieve this aim. Round your answer to the nearest cent.
 - ii** Using the rounded repayment amount, calculate the balance of the loan after 25 years.
 - iii** This amount is positive. Explain the significance of this amount.
 - b** After four years of repayments (48 repayments), Felix will make a lump sum repayment of \$100 000. Find how many further **full** repayments will be required to repay the loan.
- 21** Luke receives monthly payments of \$6100 from an annuity that is earning interest at the rate of 6.24% per annum, compounding monthly. The balance of Luke's investment is \$386 694.12 after four years of investment.
- a** State the principal amount of Luke's investment.
 - b** Find how much interest Luke has earned after four years of investment.
 - c** Determine how many more payments of \$6100 Luke can withdraw.
- 22** Milo's grandparents place \$5000 in an investment account that pays interest at a rate of 5% per annum, compounding annually. For 18 years, they contribute an additional \$1000 to the account.
- When Milo turns 18, the interest rate increases to 6% and his grandparents stop contributing.
- Calculate how much Milo will need to add to the account each year if he wishes to have a balance of \$120 000 by the time he is 30. Round your answer to the nearest cent.
- 23** **a** A plane flies from point A(26°S, 145°E) for 4000 km due north along the 145°E meridian. What will be the location of the plane at the end of this flight?
- b** A plane flies from point A(45°N, 145°E) for 4000 km due south along the 145°E meridian. What will be the location of the plane at the end of this flight?
- 24** **a** A plane flies from point A(26°S, 145°E) for 4000 km east along the 26°S parallel of latitude. What will be the location of the plane at the end of this flight?
- b** A plane flies from point A(45°N, 5°E) for 4000 km west along the 45°N parallel of latitude. What will be the location of the plane at the end of this flight?
- 25** Greg has a loan of \$22 000 that he will repay with monthly repayments of \$610. Interest is charged at the percentage annual interest rate of 5.88%.
- a** Determine the balance of Greg's loan after six repayments have been made.
 - b** Find how much interest has been paid in total after six repayments have been made.

- 26** Data were collected to investigate the association between the number of customers served at a coffee shop each week (*customers*) and the amount of money the shop owner spent in advertising (*advertising*) in dollars over the previous week. The following scatterplot shows this data with a fitted line added.



- a** Determine how much on average the coffee shop needs to spend on advertising to attract one additional customer.
- b** The coffee shop owner finds a new advertising option, which promised to cost only \$2 to attract each additional customer. Determine the amount the owner would save in attracting 100 customers using this advertising option.
- 27** Lexi currently owes \$226 800 on her home loan. She pays interest at the annual percentage interest rate of 4.26% per annum and repays the loan with monthly repayments of \$6540. After six months, the interest rate of Lexi's loan increased to 4.62% per annum, compounding monthly. Lexi decided to increase her payments to \$7000 per month. Find how much Lexi will owe on this loan after a further 12 months.
- 28** The details of two different home loans with principal \$320 000 are shown in the table below.

	Interest rate	Term	Compounding period	Extra repayments	Repayment
<i>Loan 1</i>	3.48%	20 years	Monthly	Allowed	\$1853
<i>Loan 2</i>	3.64%	18 years	Fortnightly	Not allowed	\$933

Alex is trying to decide between the two loans options. He believes that he can afford repayments of \$1900 per month.

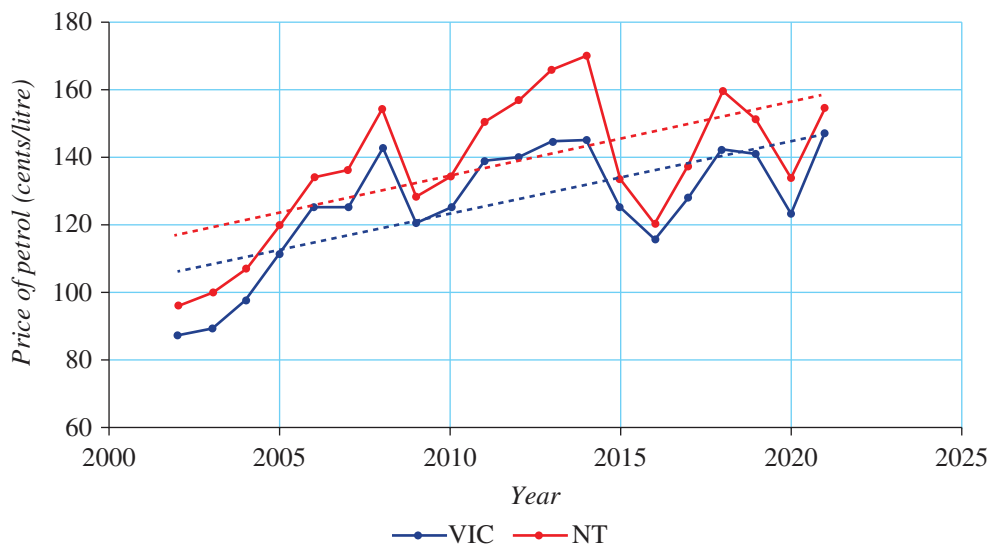
- a** If no extra repayments are made, determine which of the two loans would be best for Alex. Justify your decision by explaining your mathematical reasoning.
- b** Alex may be able to make larger repayments. Explain whether this changes your answer and provide mathematical reasoning to support your answer.

- 29** The follow table gives the value of the average price of unleaded petrol in Victoria each year from 2015–2021.

Year	2015	2016	2017	2018	2019	2020	2021
Price (cents/litre)	126.3	116.4	128.7	143.4	141.1	123.9	147.6

- a** Find the five-mean smoothed value of petrol in Victoria for the year 2017 in cents/litre, rounding your answer to one decimal place.
- b** Find the five-median smoothed value of petrol in Victoria for the year 2019 in cents/litre.

The following time series plot shows the price of petrol in Victoria in cents/litre, and the price of petrol in the Northern Territory (NT) in cents/litre, over the years 2002–2021.



Least squares regression lines (shown on the plot) have been fitted to both sets of time series data, and the following equations determined:

Victoria: $\text{petrol price} = -4071.53 + 2.08699 \times \text{year}$

NT: $\text{petrol price} = -4290.07 + 2.20128 \times \text{year}$

- c i** Write down the slope of the least squares line for Victoria rounded to one decimal place, and interpret.
- ii** Write down the slope of the least squares line for the NT rounded to one decimal place, and interpret.

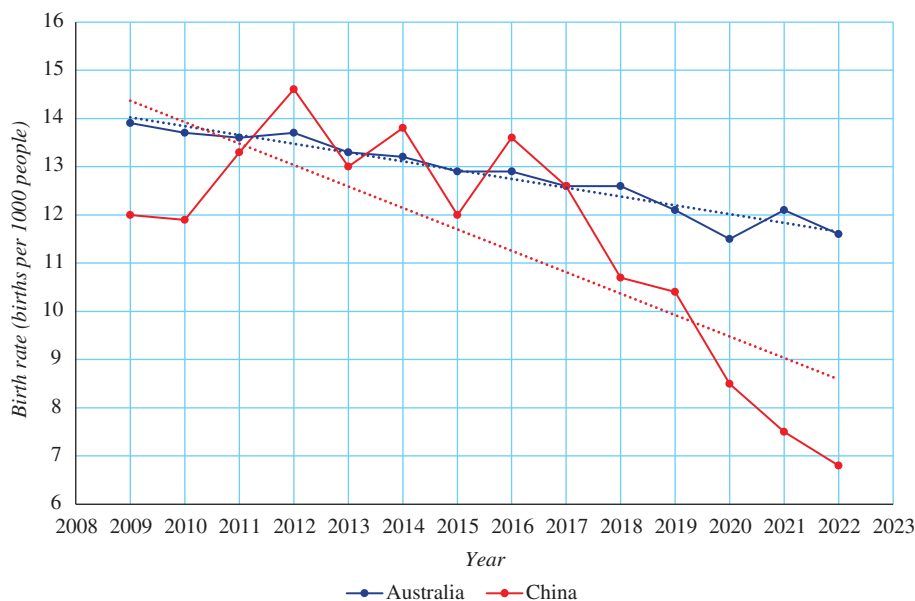
- d** Use the least squares regression lines to predict the price of petrol in 2026:
- i** in Victoria
 - ii** in the NT
- e** Do the equations predict that the difference in petrol prices between Victoria and the NT will decrease, stay the same, or increase? Explain your answer, quoting appropriate statistics.

- 30** The following table gives the minimum temperature in a city over a 10 day period.

Day	1	2	3	4	5	6	7	8	9	10
Temperature	12.7	12.9	12.8	a	b	12.0	11.6	11.8	11.4	11.0

Given that the five-mean smoothed value for Day 3 is 12.6, and the three-mean smoothed value for Day 6 is 11.9, determine the values of a and b .

- 31** The following time series plot shows the birth rate (in live births per 1000 people) for Australia and for China over the years 2009–2022.



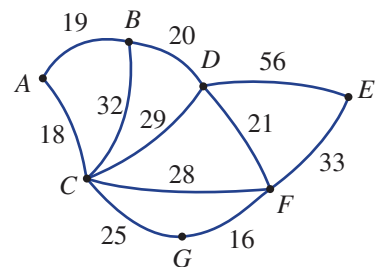
Least squares regression lines (shown on the plot) have been fitted to both sets of time series data, and the following equations determined:

Australia : $\text{birth rate} = -0.182 \times \text{year} + 380.06$

China : $\text{birth rate} = -0.444 \times \text{year} + 906.71$

Use the equations of the least squares regression lines to predict the first year in which the birth rate for Australia will be double the birth rate for China.

- 32** Benjamin has \$75 000 to invest. He has two investment options:
- Bank A offers to pay 5.28% per annum, compounding monthly
 - Bank B offers to pay 4.68% per annum, compounding fortnightly
- After three years, Benjamin would like to withdraw the balance of the annuity.
- a** Determine which of the two investment options would earn Benjamin the most interest after one year. Explain how you compared the two investment options.
- b** Write a letter to Benjamin explaining the comparison of the two investment options, showing him the calculations for the total amount he could withdraw after three years.
- 33** **a** An aeroplane flies from Cairns (17°S , 146°E) to Bourke in NSW (31°S , 146°E) by flying along the 146°E meridian. How far is this, correct to the nearest km?
- b** It now flies an equal distance west along the 31°S parallel of latitude. What is the longitude of the point it arrives at?
- 34** **a** A plane leaves A(21°S , 146°E) and flies along the 146°E meridian until it reaches the 35°S parallel of latitude. It then flies west along the 35°S parallel of latitude until it reaches the 135°E meridian. What is the total distance flown? (Final location (35°S , 135°E)).
- b** A plane leaves A(21°S , 146°E) and flies west along the 21°S parallel of latitude until it reaches the 135°E meridian. It then flies south along the 146°E meridian until it reaches the 35°S meridian parallel of latitude. What is the total distance flown? (Final location (35°S , 135°E)).
- c** Find the difference of the two total distances.
- 35** Two locations X and Y have the same latitude 35°N . The longitude of X is 125°E and the longitude of Y is 146°E .
- a** Find the radius of the small circle of the 35°N parallel of latitude.
- b** Find the distance between X and Y around the parallel of latitude 35°S .
- 36** The graph shows the road network between seven towns. Distance between the towns is in kilometres.
- a** A cable network is to be established between the towns. The cable is going to be laid along the sides of roads. What is the minimal length of cable required here if back-up links are not considered necessary; that is, there are no loops in the cable network?



- b** Treating the towns as vertices and roads as edges in a graph, what is the distance of a journey that forms a Hamiltonian cycle in the graph?
 - c** An inspection of roads starts from town *B*. Every road needs to be driven along once and once only. Determine a path that the driver must take to accomplish this and determine the distance.
- 37** Elijah invests \$140 000 into an annuity from which he receives a regular monthly payment of \$5300 for 28 months. The interest rate for the annuity is 5.52% per annum, compounding monthly.
- After two months, the interest rate for this annuity will fall to 4.44%.
- To ensure that Elijah will still receive the same number of \$5300 monthly payments, Elijah will add an extra one-off amount into the annuity at this time.
- Determine the value of the one-off addition. Round your answer to the nearest cent.

13C List of Unit 4 and Units 3 & 4 assessment and examination practice online items

These assessment practice items can be found in the interactive textbook and in the teacher resources of the online teaching suite.

Interactive Textbook

For student and teacher access:

- 1** IA3: A practice internal examination on Unit 4
- 2** EA: A practice external examination on Units 3&4

Online Teaching Suite

For teacher access:

- 1** IA3: An internal examination on Unit 4
- 2** EA: A practice external examination on Units 3&4

Assessment items for Unit 3 are listed at the end of Chapter 6.

Appendix A

The problem-solving and modelling task

Joel Speranza

Chapter contents

- ▶ **A1** About the problem-solving and modelling task
- ▶ **A2** A content guide for a PSMT report

A1 About the problem-solving and modelling task

Mathematisation is the process of taking a real-world problem, translating it to a mathematically purposeful representation, and solving that problem. In General Mathematics you will be assessed on your ability to do this through an assessment item called a problem-solving and modelling task (PSMT).

This chapter outlines how to plan, solve and present your PSMT at a high level and provides real examples of high-level student work. Each section of this chapter is accompanied by a video lesson with additional advice, accessible via the included QR codes.

How is the PSMT marked?

Before you begin any assessment, you should consider how teachers will make judgements of your work. Your teacher will use the Instrument-specific Marking Guide (ISMG) from the syllabus to determine your mark. The ISMG is broken into four criteria (Formulate, Solve, Evaluate and Communicate) and each criterion is assessed using three to five descriptors. Throughout this chapter, we will be focusing on the top descriptor in each criterion. These top descriptors will be displayed in this appendix where needed. The full ISMG with all descriptors can be accessed through the Interactive Textbook.

A2 A content guide for a PSMT report

A suggested set of headings for writing the PSMT report has been provided in a downloadable Word document in the Interactive Textbook. The rest of this appendix provides notes on what to include under these headings, which are reproduced here in black ('Introduction to the task', 'Formulating a solution', 'Developing a solution', and so on). Extracts from the ISMG display the criteria and descriptors covered by each heading of the report.

You also need to adhere to the required word limit and conventions of the mathematical report genre. You may have seen another report genre, the scientific report, and the two have similarities.

Word length guides of each section of the report are indicative only and will be dependent on your specific PSMT. Setting out your PSMT in this way is not compulsory but does help you to achieve this highlighted descriptor from the ISMG:

Communicate	Marks
The student response has the following characteristics:	
- correct use of appropriate mathematical language - logical organisation of the response, which can be read independently of the task sheet - justification of decisions using mathematical reasoning	3–4



Video A1

1. Introduction to the task

(200 words)

These are the criteria and descriptors to be covered by the introduction:

Formulate	Marks
The student response has the following characteristics:	
- justified statements of important assumptions - justified statements of important observations - justified mathematical translation of important aspects of the task	3–4



Video A2

Communicate	Marks
The student response has the following characteristics:	
- correct use of appropriate mathematical language - logical organisation of the response, which can be read independently of the task sheet - justification of decisions using mathematical reasoning	3–4

Note that, as long as you understand the task, and have written down what you need to do, you do not have to finalise the introduction at the beginning of the process. You may find it better to write the introduction alongside writing the conclusion, ensuring that the conclusion addresses the goal of the task which was stated in the introduction.

The purpose of the introduction is twofold:

- 1 To introduce a report which ‘can be read independently’ of the task sheet. This means that a person who has never seen the task sheet will be able to understand what the task is, simply by reading your introduction. No reference to the task sheet should be made here. You should also begin with providing some context on what the problem is to be solved, and why solving the problem is important.
- 2 To show ‘justified mathematical translation of important aspects of the task’. Mathematical translation is the process of taking a real-world problem and moving it into the mathematical world – a process known as **mathematisation**. In this part of an introduction, you should give a brief description of the mathematical techniques you will use to solve the problem. This description can include:
 - Applicable mathematical or statistical principles
 - Mathematical concepts and techniques
 - Technology that you will use throughout the task

The student sample below shows:

- how the task can be read independently of the task sheet
- justified mathematical translation of important aspects of the task.

Introduction

As the population of Australia continues to grow, it is crucial that we plan ahead. Important aspects in life such as resources, housing, and medical products and services need to account for this growing population before the demand for necessary requirements of life have such a demand that cannot be catered for (Sommerfeld, 2018). By knowing predictions for a future population count, we can plan ahead and make choices that will benefit the future now.

The purpose of this task is to determine a prediction for the rate of population growth in Australia for 2061 by exploring multiple mathematical models. A multitude of mathematical techniques will be utilised throughout this task, including; modelling of exponential, logarithmic, logistic, and sinusoidal functions, as well as deriving to find a rate of change, determining percentage error, and using technology such as GeoGebra, a TI-84 Plus CE Graphics Calculator and Excel to model equations and manage data.

Student sample taken from the 2020 Mathematical Methods subject report

Using mathematical language

In the introduction and throughout the PSMT, you must demonstrate correct use of mathematical language.

These are the criteria and descriptors to be covered:

Communicate	Marks
The student response has the following characteristics:	
• correct use of appropriate mathematical language • logical organisation of the response, which can be read independently of the task sheet • justification of decisions using mathematical reasoning	3–4



Video A3

The syllabus objectives elaborate on mathematical language as ‘terminology, symbols, conventions and representations’. What follows is a non-exhaustive list of ways to demonstrate each aspect of mathematical language.

Terminology

- Procedural mathematical language (mathematical verbs) e.g. determine, solve, verify, calculate integrate, derive.
- Technical mathematical language (mathematical nouns) e.g. search the unit outline for terminology specific to the content being assessed.

Conventions

- Equations have a left and right hand side
- Equal signs are aligned
- Define variables before using them

Symbols

- Use mathematical symbols correctly
- If typing mathematics, use equation editor

Representations: Graphs

- An appropriate title
- Axes labelled with appropriate units
- An appropriate scale for each axis
- A legend if appropriate

Representations: Diagrams

- An appropriate title
- Labelled
- Drawn to scale or a label indicating otherwise
- Vertices of shapes labelled with capital letters
- Angles labelled with Greek letters

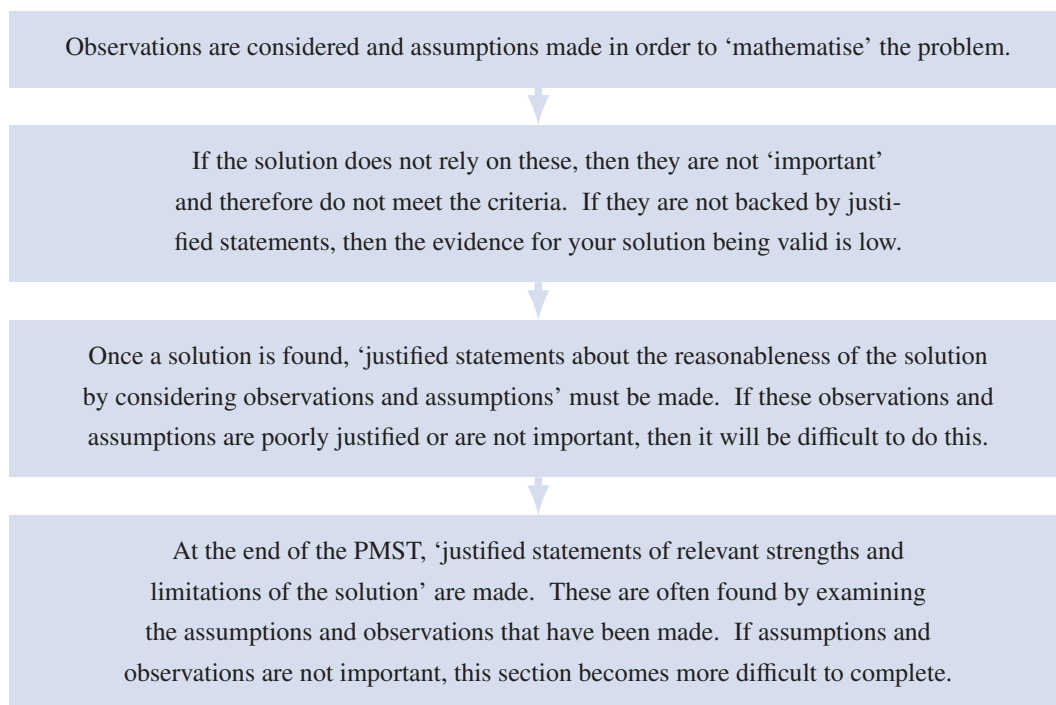
2. Formulating a solution

(400 words)

In the section titled ‘Formulating a solution’, you are aiming to demonstrate the descriptors of the ISMG shown below.

Formulate	Marks
The student response has the following characteristics:	
- justified statements of important assumptions - justified statements of important observations - justified mathematical translation of important aspects of the task	3–4

It is vital that your assumptions are both important and expressed as justified statements. This is the most important section of the PSMT, as many sections of the PSMT cannot be completed to a high-standard without them. This is demonstrated in the following flowchart.



Observations vs assumptions

Students can often be confused about the difference between an observation and an assumption. Both are vital to creating a mathematical model but serve different purposes. The table below outlines the key differences between them.

Aspect	Observations	Assumptions
Definition	Data or information required to solve a mathematical problem and/or develop a mathematical model.	Conditions that are stated to be true when beginning to solve a mathematical problem and/or develop a mathematical model.

Aspect	Observations	Assumptions
Nature and verifiability	Factual, based on actual data or information, and can be empirically verified.	While not directly verifiable, are rational and necessary for solving the problem. They are based on plausible reasoning or existing knowledge.
Role in modelling	Provides data and empirical evidence for creation of the model.	Simplifies complexity and fills data gaps. Dictates the strengths, limitations, and applicability of the model.
Flexibility	Generally rigid; they are facts that can't be altered. However, the interpretation of observations can evolve with new data.	More flexible and can be adjusted or replaced as new information becomes available or as the model evolves.
Examples	1 Measuring the temperature and humidity levels in various regions over a year. 2 Recording the frequency and intensity of rainfall in a specific area. 3 The observable fact that warm air rises and cool air sinks, affecting weather patterns.	1 Assuming that future weather patterns will reflect past trends due to climate consistency. 2 Assuming a certain level of accuracy in satellite data used for cloud cover analysis. 3 Assuming that ocean currents will remain relatively stable over the short-term forecasting period.
Role in validation	Used for validating the model by comparing its predictions with actual observations.	Validated indirectly through the model's performance and its ability to make accurate predictions within the defined framework of these assumptions.

2.1 Observations

To ensure that your observations reach the level of 'justified statements' they should include:

- a discussion of how the observations affect the mathematical model/solution
- in-text referencing to a reputable source (while it is possible to justify statements without this, consider the inclusion of a reference the 'gold-standard').

The following student sample demonstrates justified statements of observations. Each observation contains:

- a statement of what the observation is
- a reference that provides justification for the observation being made
- justification for how this observation impacts the mathematical model/solution.

Structuring your observations in this way provide a high likelihood of them being considered justified statements on the ISMG.



Video A4

Observations

It was observed that the male-to-female ratio in Western Australia is 102 males for every 100 females (McCrindle, 2014). This observation directly impacts the mathematical model as the initial total population must divide by this ratio only to consider the female Western Australian population. This is relevant as males do not reproduce any offspring and cannot be factored into the Leslie matrix.

It was observed that Western Australia takes 30% of immigrants into Australia each year (Australian Bureau of Statistic, 2021). Further, 12,706 to 18,200 immigrants settled in Australia during 2018 and 2020. (Lawrence, 2018). These observations impacted the mathematical model as an increase in immigrants will impact the projected populations, causing an increase or decrease in total population growth rate.

Humans live to approximately 100 years (Vaupel, 2010); therefore, a 21×21 matrix is reasonable. From this, the age classes for the Leslie matrix were split into 5-year categories. Hence, each new generation produced by the Leslie matrix represents a 5-year gap. This observation is relevant as it impacts the scope mathematical model by investigating only eight new generations.

Student sample taken from the 2022 Specialist Mathematics subject report

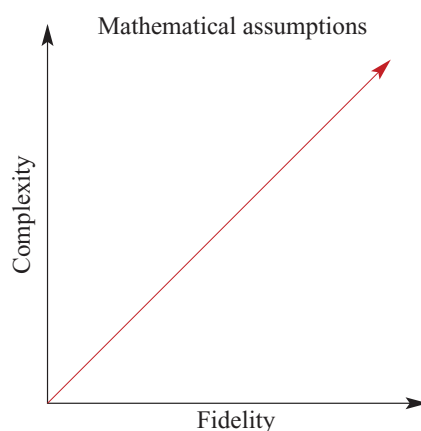
2.2 Assumptions

As with observations, to justify the assumptions you make you should include discussion of how the assumptions affect the mathematical model/solution; and in-text referencing to a reputable source. While it is possible to justify statements without this, it is considered best practice to include one.

Referencing should be used to justify an assumption and it can be done in two different ways:

- use a source that brings your mathematical model closer to representing the real world
- use a source that takes your mathematical model further from representing the real world but is required to reduce the complexity of the model.

To illustrate this further, each assumption can be thought of as existing on the graph shown below, with mathematical complexity increasing as fidelity (how closely the assumption models the real world) increases. Therefore, each assumption can be thought of as a trade-off between complexity and fidelity. Adopting this approach from the beginning makes future sections of the PSMT (evaluating the reasonableness of the solution and strengths and limitations) easier to complete.



Video A5

In the student sample which follows, each assumption contains:

- a statement of what the assumption is
- a reference that provides justification for the assumption being made
- justification for how the assumption impacts the mathematical model/solution.

Assumptions

It was assumed that birth rates would only impact women aged 15–44 as in Australia. The average woman's reproductive years are between ages 15 and 44 (Watson, 2018). This assumption restricts birth rates to only six of twenty-one age classes. The assumption was made to reduce the anomalies to develop clean data.

It was assumed that the new immigrant population introduced into Western Australia was equally divided into the age classes from 18 to 34. Most migrants to Australia are young adults, with 61.2% aged between 18 and 34 years (abs.gov.au, 2018). Further, 12 706 to 18 200 immigrants settled in Australia during 2018 and 2020. (Lawrence, 2018). This assumption was made to create a realistic data spread that included the possible impact immigrants' survival and or birth rate would have on the total population.

It was assumed that the investigation started during 2016 as the female population data collected was from 2016 (abs.gov.au, 2016); therefore, a 21×21 matrix is reasonable. This assumption impacts the investigation as the potential increasing or decreasing growth rate and total population can be compared to secondary data to determine the model's validity.

Student sample taken from the 2022 Specialist Mathematics subject report.

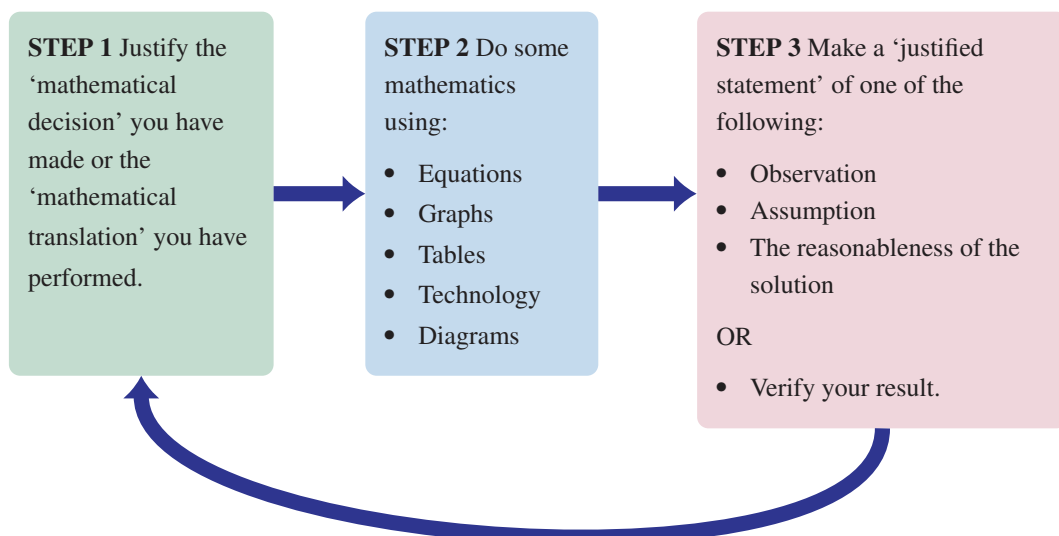
3. Developing a Solution

(700 words)

In this section of the PSMT, you are attempting to demonstrate almost all the descriptors of the ISMG, as shown in the following example. This is also a section where students can get a little confused about how to set things out. Below is a flowchart you can use throughout this section to ensure that you demonstrate the full range of ISMG descriptors.



Video A6



The sample following is a simple example of how this flowchart can be put into practice. It is from a PSMT in which the task is to develop a flying fox. The mathematics has been simplified and the example annotated the flow chart steps and matching ISMG descriptors.

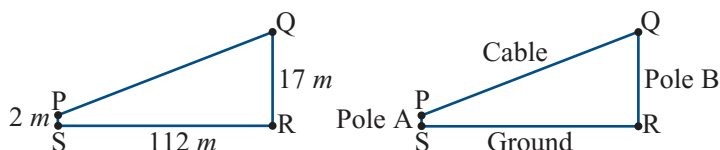
Use them to guide your solution, but do not annotate your own report with them.

3.1 Calculating the length of the cable

STEP 1

The flying fox cable, two supporting poles and the ground can be modelled as the quadrilateral, $PQRS$, as shown below.

STEP 2



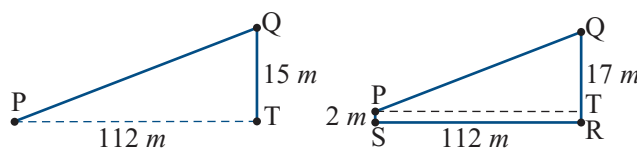
STEP 3

It is assumed that the cable is perfectly taut and has no sag. While the cable in a real flying fox will have a sag (Evans, 2017), this assumption allows the cable's length to be calculated.

STEP 1

A horizontal line drawn through point P creates a right triangle PQT . Creating a right triangle will allow the length of the cable to be calculated using Pythagoras' theorem.

STEP 2

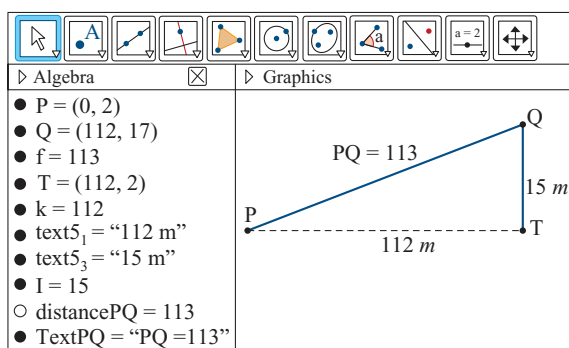


$$\begin{aligned}\text{Length of cable} &= \sqrt{112^2 + 15^2} \\ &= \sqrt{12544 + 225} \\ &= \sqrt{12769} = 113\end{aligned}$$

STEP 3

The length of the cable is calculated to be 113 metres.

This result can be verified using a scale diagram drawn in Geogebra, as shown in this screenshot.



Formulate

- Justified mathematical translation of important aspects of the task

Communicate

- Correct use of appropriate mathematical language

Formulate

- Justified statements of important assumptions

Communicate

- Justification of decisions using mathematical reasoning

Solve

- Accurate use of mathematical knowledge for important aspects of the task

Solve

- Efficient use of technology

STEP 1**3.2 Calculating the total cost of the cable**

The total cost of cable can now be calculated. In construction, it is common to assume that you require an additional 10% of materials to allow for wastage (Jones, 2017). 8 mm aircraft-grade galvanised cable is perfect for ziplines up to 150 metres in length and has a current cost of \$8 per metre. (cable-ride.com, n.d)

STEP 2

$$\text{Total cost} = 113 \times 8 \times 1.1 = \$994.40$$

Adding an additional 10% to materials is a reasonable solution when considering the assumption made that there is no sag in the cable. In reality there will be sag, and this will increase the amount of the cable required.

STEP 3

The \$994.40 total cost for the cable can be verified by comparing it to a 90 metre zipline kit available online for \$1387. (cable-ride.com, n.d). While this is \$392.60 more expensive than our cable, it is for a complete kit rather than just a cable. These two prices are close enough to support our result.

Formulate

- Justified statements of important assumptions
- Justified statements of important observations

Solve

- Accurate use of mathematical knowledge for important aspects of the task

Evaluate

- Justified statements about the reasonableness of the solution by considering the assumptions

Evaluate

- Verified results

Efficient use of technology

Throughout your assignment, you should be looking for opportunities to demonstrate the efficient use of technology,

as required by the ISMG descriptor shown. The key descriptor here is 'efficient', which can be understood as using technology in a way that minimises wasted effort and/or time. An example of efficiency would be using Excel formulas to perform repeated calculations, rather than performing each calculation by hand or on a calculator.

It is often the case that while students use quite a bit of technology throughout their PSMT, they often don't provide the evidence that they have used it. Here are some types of technology students use and the evidence that they can provide to demonstrate that they have used it.

Solve

The student response has the following characteristics:

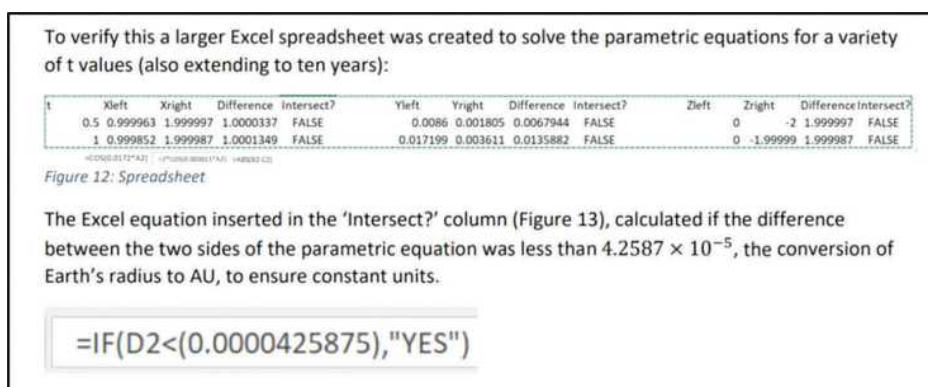
- Efficient use of technology



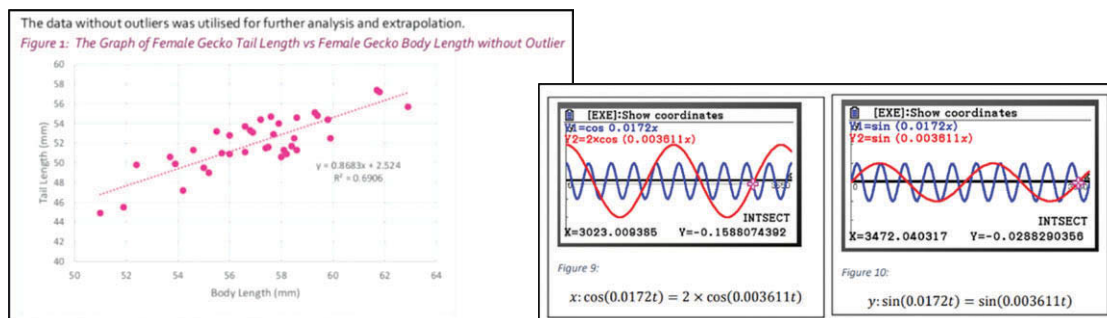
Video A7

Types of technology	Evidence provided
Spreadsheet software (e.g. Microsoft Excel)	■ Screenshots of graphs or tables ■ Samples of spreadsheet formulas used
Calculator	■ Photos or screenshots of the calculator screen ■ Descriptions of how the calculator was used
Graphing software (e.g. Desmos or Geogebra)	■ Screenshots ■ Descriptions of how the software was used
Logging software (e.g. data collectors)	■ Sample of the data collected ■ Screenshots

The following examples which evidence efficient use of technology are taken from the 2022 Mathematics subject reports.



Student sample taken from the 2022 Specialist Mathematics subject report



Student sample taken from the 2022 General Mathematics subject report

Student sample taken from the 2022 Specialist Mathematics subject report

4. Evaluating and verifying the solution

(600 words)

4.1 Verifying results

In this section you are aiming to demonstrate the descriptor shown here, verifying the overall solution. If you have

been following the advice from the previous section, you will have been verifying some of your results in the process of coming to your solution. Below, we look at four techniques students can use for verifying results.

- 1 Verifying through estimation:** This method simplifies the problem before calculating an approximate solution. For instance, if you have calculated the area of a composite shape, you can verify your solution by simplifying it into a basic rectangle and recalculating the area. If the estimated area is close to your calculated area, it helps confirm the accuracy of your solution.

Evaluate

The student response has the following characteristics:

- Verified results



Video A8

- 2 Verifying through research:** This method involves comparing results with reliable sources. This method is particularly useful for tasks related to historical data or correlation studies. Findings can be matched against data from textbooks, academic journals, or credible online sources. For example, in a task exploring the correlation between car weight and fuel consumption, you can verify the solution against existing research on this topic.
- 3 Verifying through technology:** Using technology to redo algebraic calculations can aid in verifying results. For example, when calculating the area under a curve, employing graphing calculators or software for an approximation and comparing it to manual calculations can serve as a verification method.
- 4 Verifying through an alternative method:** This method is effective in problems with multiple solutions and involves employing various techniques to solve the same issue. For example, you may initially solve an algebraic equation by factoring, then recheck the solution using the quadratic formula. If both methods produce identical results, it strongly suggests that the solution is correct.

4.2 Evaluating reasonableness of the solution

In this section you are aiming to demonstrate the content descriptor below:

Evaluate	Marks
The student response has the following characteristics:	
• verified results • justified statements about the reasonableness of the solution by considering the assumptions • justified statements about the reasonableness of the solution by considering the observations • justified statements of relevant strengths of the solution • justified statements of relevant limitations of the solution	4–5



Video A9

When evaluating the reasonableness of your solution by considering assumptions and observations you should:

- consider the solution found
- consider how it is affected by the observation or assumption
- consider how the solution might be different if the observation or assumption was altered (often with some mathematical working included).

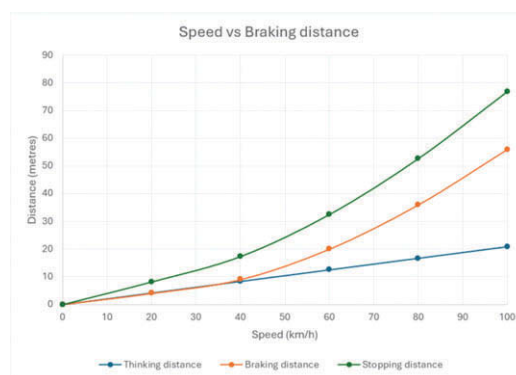
On the next page is an example of evaluating the reasonableness of a solution by considering an assumption from a PSMT investigating the braking distance of a car dependent on the speed at which it is travelling.

Example

The distance calculated for a car to come to a complete stop is underestimated because of the assumption that driver's response times were instantaneous. If driver response time were factored in, the braking distance would be greater than the distances calculated in this report.

Assuming average driver response time of 1.5 seconds (Muttart, 2004), we can see how the solution would change if this

time was taken into account in the graph pictured. At the top speed of 100 km/h, braking distance is increased from the initial solution of 56 m to 77 m, an increase of 37.5%.



Note: Students often make the mistake of evaluating the reasonableness of their assumption, rather than evaluating the reasonableness of their solution by considering their assumption. The distinction is subtle but important. The justified statements made must refer to the solution and how it is affected by the assumption, not just the assumption itself.

Correct: The solution is reasonable because...

Incorrect: The assumption is reasonable because...

A subtle but important difference.

4.3 Strengths and limitations of the solution

In this section you are aiming to demonstrate the content descriptor below:

Evaluate	Marks
The student response has the following characteristics:	
• verified results • justified statements about the reasonableness of the solution by considering the assumptions • justified statements about the reasonableness of the solution by considering the observations • justified statements of relevant strengths of the solution • justified statements of relevant limitations of the solution	4–5



Video A10

Strengths and limitations of the solution can be thought of in the following way.

- Strengths – aspects of the model that make it useful
- Weaknesses – aspects of the model that limit its usefulness

A series of questions to help identify these strengths and limitations is below. Do not aim to answer all these questions but use them as prompts to generate ideas.

Strengths	Limitations
What assumptions were made that closely align to the real world?	What assumptions were made that do not align closely with the real world?
What aspects of the real world does the solution consider?	What aspects of the real world does the solution not consider?
Could the method used to create this solution be easily adapted and used to solve other, related problems?	Are there other, related problems that the method used could not be easily adapted to solve?
What aspects of the solution can be verified using other observational data?	What aspects of the solution cannot be verified using observational data?
What are the potential, positive consequences of using this solution in the real world?	What are the potential, negative consequences of using this solution in the real world?
What aspects of the solution will continue to be accurate into the future?	What aspects of the solution will cease to be accurate into the future?

To make justified statements of strengths/limitations:

- state the strength/limitation
- justify why it is a strength/limitation.

The student sample of limitations below provides examples of this.

The following limitations were observed There was a limited amount of data points that were used as a sample. This means the findings were less reliable as it may not be an accurate representation of all rugby games. Another limitation is that when using extrapolation with regards to the regression line, it may not be accurate to predict further outcomes because the prediction is outside the sample data range. One final limitation is that the R^2 value found is not considerably strong, therefore a smaller percentage of the points scored per game can be attributed to the line breaks achieved per game, decreasing the reliability of the study.

Student sample taken from the 2022 General Mathematics subject report

5. Conclusion

(100 words)

The conclusion is another opportunity to show logical organisation of your response. In the conclusion you should:

- restate the purpose of the mathematical report
- provide a summary of your solution, stating an appropriate answer to the task.

Conclusion

The purpose of this report was to use functions and derivatives to create a reasonable prediction for the rate of change for the Australian population in 2061. It was found that the most reasonable model was solution three as it produces a reasonable population and somewhat reasonable rates of change. Therefore, using this logistic function, it is predicted that the population will reach approximately 39.6 million in 2061, with a percentage rate of change of 0.95% and an instantaneous rate of change of approximately 370 000 addition people per year.

Student sample taken from the 2022 Mathematical Methods subject report

6. Reference List

(Not included in page or word count)

Use a standard referencing style. Ask your teacher for guidance on this if you need it.

7. Appendix

(Not included in page or word count)

An appendix is for supporting material such as data, diagrams, calculations and screenshots or print-outs from technology, that don't form a direct part of the solution or evaluation. The appendix is not marked so don't include important items that you want a mark for. If you haven't already included your use of technology in the report, you could put a small sample into the body of your assignment to get marks for efficient use of technology.

Glossary

A

Activity (CPA) [p. 473] A task to be completed as part of a project. Activities are represented by the edges in the project diagram.

Activity network [p. 473] An activity network is a weighted directed graph that shows the required order of completion of the activities that make up a project. The weights indicate the durations of the activities they represent.

Adding to the principal See annuity investment.

Adjacency matrix [p. 399] A square matrix showing the number of edges joining each pair of vertices in a graph.

Algorithm [p. 441] A step-by-step procedure for solving a particular problem that involves applying the same process repeatedly. Examples include Prim's algorithm and the Hungarian algorithm.

Allocation [p. 451] Allocation is the process of assigning a series of tasks to different members of a group in a way that enables the tasks to be completed for the minimum time or cost.

Amortisation [p. 382] Amortisation is the repayment of a loan or an investment with regular payments made over a period of time.

Amortisation table [p. 382] An amortisation table charts the amortisation (repayment) of a reducing balance loan or annuity on a step-by-step (payment-by-payment) basis or the payment of a compound interest investment with additional payments.

Annuity [p. 351] An annuity is a compound interest investment from which regular payments are made.

Annuities formula [p. 348] Given the periodic payment d , decimal rate of interest per compounding period, i , and the number of compounding periods, n , the present value of an annuity is $A_{FV} = d \times \frac{1 - (1 + i)^n}{i}$.

Arc [p. 219] The part of a circle between two given points on the circle. The length of the arc of a circle is given by $s = r \left(\frac{\theta}{180} \right) \pi$, where r is the radius of the circle and θ is the angle in degrees subtended by the arc at the centre of the circle.

Arithmetic sequences [p. 166] A sequence is arithmetic if it satisfies the recurrence relation $t_1 = \text{starting value}$, $t_{n+1} = t_n + d$, where d is the common difference. Arithmetic sequences are used to model linear growth and decay situations. The rule for the n th term of an arithmetic sequence is $t_n = t_1 + (n - 1)d$, where t_1 is the starting value.

B

Backward scanning [p. 485] Backward scanning is the process of determining the LST for each activity in a project activity network.

Balance [p. 324] The balance of a loan or investment is the amount owed or accrued after a period of time.

Bipartite graph [p. 446] A graph whose set of vertices can be split into two subsets, X and Y , in such a way that each edge of the graph joins a vertex in X and a vertex in Y .

Bivariate data [p. 2] Data in which each observation involves recording information about two variables for the same person or thing. An example would be data recording the height and weight of the children in a preschool.

Break-even analysis Finding the point where the revenue of a business first equals the costs of running the business. Past this point, the business is running at a profit:
profit = revenue – costs.

Bridge [p. 394] An edge in a **connected graph** that, if removed, would leave the graph no longer connected.

C

Capacities (flow network) [p. 455] The weights of the directed edges in a flow network are called **capacities**. They give the maximum amount that can move between the two points in the flow network represented by these vertices in a particular time interval. This could be, for example, the maximum amount of water in litres per minute or the maximum number of cars per hour.

Categorical variable [p. 2] Categorical variables generate data values that are names or labels, such as gender (male, female) or coffee size (small, medium, large).

Causal relationship [p. 84] When a change in the **explanatory variable** leads to a change in the **response variable**, this is known as a causal relationship.

Coefficient of determination (R^2) [p. 45] A coefficient which gives a measure of the predictive power of a regression line. It gives the percentage of variation in the RV that can be explained by the variation in the EV.

Complete graph [p. 394] A graph with edges connecting all pairs of vertices.

Compound interest [p. 199] Where the interest paid on a loan or investment is added to the principal and subsequent interest is calculated on the total.

Compound interest investments with additional payments A compound interest investments with periodic payments is an investment to which additions are made to the

principal on a regular basis. Also known as ‘adding to the principal’.

Compounding period [p. 290] The compounding period is the time period for the calculation of interest for an investment or loan. Typical compounding periods are yearly, quarterly, monthly or daily.

Connected graph [p. 394] A **connected graph** is a graph that has no isolated vertices and no separate parts.

Continuous variable [p. 3] A variable representing a quantity that is measured rather than counted, for example the weights of people in kilograms.

Coordinated Universal Time (UTC) [p. 241] A measure of time used to regulate time across the world. Equivalent to **GMT**.

Correlation coefficient r [p. 36] A statistical measure of the strength of the linear association between two numerical variables.

Cost matrix [p. 448] A cost matrix is a table that contains the cost of allocating objects from one group, such as people, to objects from another group, such as tasks. The cost can be money, or other factors such as the time taken to complete the project.

Critical path [p. 483] The project path that has the longest completion time.

Critical path analysis [p. 483] A project planning method in which activity durations are known with certainty.

Cut [p. 457] A line dividing a directed (flow) graph into two parts in a way that separates all ‘sinks’ from their ‘sources’.

Cut capacity [p. 457] The capacity of a cut is the sum of the capacities of the cuts passing through the cut that represents flow from the source to the sink. Edges that represent flow from the sink to the source do not contribute to the capacity of the cut.

Cycle (graphs) [p. 413] A **walk** with no repeated vertices that starts and ends at the same vertex.

Cycle (time series) [p. 115] Periodic movement in a time series but over a period greater than a year.

D

Degree of a vertex ($\deg(A)$) [p. 391] The number of edges attached to the vertex. The degree of vertex A is written as $\deg(A)$.

Depreciation [p. 195] The reduction in value of an item over time.

Deseasonalise [p. 130] The process of removing seasonality in time series data.

Directed graph (digraph) [p. 395] A graph or network in which directions are associated with each of the edges.

Discrete variable [p. 3] A variable representing a quantity that is determined by counting, for example, the number of people waiting in a queue.

E

Earliest starting time (EST) [p. 484] The earliest time an activity in a project can be started.

Edge [p. 391] A line joining one vertex in a graph or network to another vertex or itself (a loop).

Effective rate of interest [p. 303] Used to compare the interest paid on loans (or investments) with the same annual nominal interest rate r but with different compounding periods (daily, monthly, quarterly, annually, other).

Elements [p. 450] The numbers or symbols displayed in a matrix.

Equivalent graph [p. 405] *See isomorphic graphs.*

Eulerian graph and Eulerian trails [p. 455] A connected graph is **Eulerian** if it has a closed trail (starts and ends at the same vertex) and includes every edge. Such a trail is called an **Eulerian trail** or **Eulerian circuit**.

Euler's formula [p. 407] The formula $v - e + f = 2$, which relates the number of vertices, edges and faces in a connected planar graph.

Explanatory variable [p. 4] When investigating associations in **bivariate data**, the explanatory variable (EV) is the variable used to explain or predict the value of the **response variable** (RV).

Extrapolation [p. 75] Using a mathematical model to make a prediction *outside* the range of data used to construct the model.

F

Face [p. 406] An area in a graph or network that can only be reached by crossing an edge. One such area is always the area surrounding a graph.

Finance solver A finance solver is a computer/calculator application that automates the computations associated with analysing a reducing balance loan, an annuity or an annuity investment.

Flat-rate depreciation [p. 180] Depreciation where the value of an item is reduced by the same amount each year. Flat-rate depreciation is equivalent, but opposite, to simple interest.

Float (slack) time [p. 483] The amount of time available to complete a particular activity that does not increase the total time taken to complete the project.

Flow [p. 455] Flow is the transfer of material, such as water, gas or traffic through a directed network.

Flow network [p. 455] A **flow network** occurs where the directed edges of the graph represent the flow of material from one vertex to another. The weight of an edge of a flow network is called the **capacity** of that edge.

Forward scanning [p. 484] Forward scanning is the process of determining the EST for each activity in a project activity network.

Future value [p. 389] The future value (A) of a compound interest loan or investment is the balance of that loan or investment after some number of compounding periods.

G

Geometric decay [p. 188] When a recurrence rule involves multiplying by a factor less than one, the terms in the resulting sequence are said to decay geometrically.

Geometric growth [p. 188] When a recurrence rule involves multiplying by a factor greater than one, the terms in the resulting sequence are said to grow geometrically.

Geometric sequences [p. 187] A sequence is geometric if it satisfies the recurrence relation: $t_{n+1} = r \times t_n$ and a starting point usually t_1 . Geometric sequences are used to model geometric growth and linear decay situations. The rule for the n^{th} term of a geometric sequence is $t_n = r^{n-1}t_1 = r^{n-1}a$, where $a = t_1$ is the starting value.

Graph or network [pp. 391, 500] A collection of points called vertices and a set of connecting lines called edges.

Great circle [p. 223] A circle on a sphere whose plane passes through the centre of the sphere. The shortest distance between two points on a sphere is along an arc of the great circle passing through the two points. *See also* **small circle**.

Greenwich Mean Time (GMT) [p. 241] Equivalent to **UTC**, this is a measure of time centred around Greenwich, England and is used across the world.

H

Hamiltonian cycle [p. 420] a **Hamiltonian cycle** is a path that starts and finishes at the same vertex and visits every other vertex exactly once.

Hamiltonian graph [p. 420] A Hamiltonian graph is a graph that contains a **Hamiltonian cycle**.

Hamiltonian path [p. 420] A path through a graph or network that passes through each vertex exactly once. It may or may not start and finish at the same vertex.

Hungarian algorithm [p. 448] An algorithm for solving allocation (assignment) problems.

I

Immediate predecessor [p. 473] An activity that must be completed immediately before another one can start.

Intercept (of a straight line) [p. 60] The y-intercept is where the regression line cuts across the y-axis.

Interest [p. 178] The amount of money paid (earned) for borrowing (lending) money over a period of time.

Interest rate [p. 115] The rate at which interest is charged or paid. Usually expressed as a percentage of the money owed or lent.

International Date Line [p. 242] An imaginary line through the Pacific Ocean that corresponds to 180° **longitude**.

Interpolation [p. 75] Using a regression line to make a prediction *within* the range of values of the explanatory variable.

Irregular (random) fluctuations [p. 117] Unpredictable fluctuations in a time series. Always present in any real world time series plot.

Isolated vertex [p. 393] A vertex that is not connected to any other vertex. Its degree is zero.

Isomorphic graphs [p. 405] Equivalent graphs. Graphs that have the same number of edges and vertices that are identically connected.

Iteration Each application of a recurrence rule to calculate a new term in a sequence is called an iteration.

L

Latest start time (LST) [p. 487] The latest time an activity in a project can begin, without affecting the overall completion time for the project.

Latitude [p. 224] The angle or angular distance north or south of the equator.

Least squares method [p. 56] A way of finding the equation of a regression line by minimising the sum of the squares of the residuals.

Linear decay [p. 169] When a recurrence rule involves subtracting a fixed amount, the terms in the resulting sequence are said to decay linearly.

Linear growth [p. 169] When a recurrence rule involves adding a fixed amount, the terms in the resulting sequence are said to grow linearly.

Linear regression [p. 58] The process of fitting a straight line to bivariate data.

Longitude [p. 224] The angle or angular distance east or west of the **prime meridian**.

Loop [p. 391] An edge in a graph or network that joins a vertex to itself.

M

Matrix [p. 399] A rectangular array of numbers or symbols set out in rows and columns within square brackets (*pl*: matrices).

Maximum flow (graph) [p. 455] The capacity of the 'minimum' cut.

Mean ($\bar{x}$) [p. 60] The balance point of a data distribution. The mean is given by $\bar{x} = \frac{\sum x}{n}$, where $\sum x$ is the sum of the data values and n is the number of data values. Best used for symmetric distributions.

Median [p. 89] The median (M) is the middle value in a data distribution. It is the midpoint of a distribution dividing an ordered data set into two equal parts. Can be used for skewed or symmetric distributions.

Meridian [p. 224] Semi-great circles that pass through north and south poles.

Meridians of longitude [p. 224] Semi-great circles which pass through the north and south poles.

Minimum cut (graph) [p. 458] The cut through a graph or network with the minimum capacity.

Minimum spanning tree [p. 441] The spanning tree of minimum length. For a given connected graph, there may be more than one minimum spanning tree.

Modelling [pp. 198, 285] Mathematical modelling is the use of a mathematical rule or formula to represent real-life situations.

Moving mean smoothing [p. 129] In three-moving mean smoothing, each original data value is replaced by the mean of itself and the value on either side. In five-moving mean smoothing, each original data value is replaced by the mean of itself and the two values on either side.

Moving median smoothing [p. 137] Moving median smoothing is a graphical technique for smoothing a time series plot using moving medians rather than moving means.

Multiple edge [p. 393] Where more than one edge connects the same two vertices in a graph.

N

Network [pp. 424, 526] A set of points called vertices and connecting lines called edges, enclosing and surrounded by areas called faces.

Nominal interest rate [p. 290] The annual interest rate for a loan or investment that assumes the compounding period is 1 year. If the compounding period is less than a year, for example monthly, the actual or **effective interest rate** will be greater than r .

Numerical variable [p. 2] A variable used to represent quantities that are counted or measured. For example, the number of people in a queue, the heights of these people in cm. Numerical variables come in types: discrete and continuous.

O

Outliers [pp. 26, 117] Data values that appear to stand out from the main body of a data set.

P

Parallels of latitude [p. 224] Small circles whose planes are parallel to that of the equator.

Path [p. 412] A **path** is a walk with no repeated edges and no repeated vertices.

- An **open path** is a path that starts and finishes at different vertices.
- A **closed path** or **cycle** is a path that starts and finishes at the same vertex.

See also **trail**.

Percentage frequency [p. 11] Frequency expressed as a percentage.

Perpetuity [p. 376] An investment where an equal amount is paid out on a regular basis forever.

Planar graph [p. 405] A graph that can be drawn in such a way that no two edges intersect, except at the vertices.

Precedence table [p. 474] A table that records the activities of a project, their immediate predecessors and often the duration of each activity.

Prim's algorithm [p. 441] An algorithm for determining a minimum spanning tree in a connected graph.

Prime meridian [p. 225] The meridian located at 0° which passes through Greenwich, England.

Principal (P) [p. 178] The initial amount borrowed, lent or invested.

R

Radius [p. 219] The distance from the centre to any point on the circle (sphere). Half the diameter.

Recurrence relation [pp. 167, 281] A relation that enables the value of the next term in a sequence to be obtained by one or more current terms. Examples include ‘to find the next term, add two to the current term’ and ‘to find the next term, multiply the current term by three and subtract five’.

Reducing-balance depreciation [p. 201] When the value of an item is reduced by the same percentage each year. Reducing-balance depreciation is equivalent to, but opposite to, compound interest.

Reducing-balance loan [p. 324] A loan that attracts compound interest, but where regular repayments are also made. In most instances the repayments are calculated so that the amount of the loan and the interest are eventually repaid in full.

Reseasonalise [p. 130] The process of converting seasonal data back into its original form.

Residual [p. 59] The vertical distance from a data point to a straight line fitted to a scatterplot is called a residual:

residual = actual value – predicted value

Residual plot [p. 68] A plot of the residuals against the explanatory variable. Residual plots can be used to investigate the linearity assumption.

Response variable [p. 4] The variable of primary interest in a statistical investigation.

Rule for the future value [p. 318] The future value of a compound interest loan or investment after n periods where P is the principal and i is the interest rate per compounding period is

$$P = \frac{A}{(1+i)^n}.$$

S

Scatterplot [p. 20] A statistical graph used for displaying bivariate data. Data pairs are represented by points on a coordinate plane, the EV is plotted on the horizontal axis and the RV is plotted on the vertical axis.

Seasonal indices [p. 130] Indices calculated when the data shows seasonal variation. Seasonal indices quantify seasonal variation. A seasonal index is defined by the formula:

$$\text{seasonal index} = \frac{\text{value for season}}{\text{seasonal average}}$$

For seasonal indices, the average is 1 (or 100%).

Seasonality [p. 115] The tendency for values in the time series to follow a seasonal pattern, increasing or decreasing predictably according to time periods such as time of day, day of the week, month, or quarter.

Semi-Eulerian graph and semi-Eulerian trails [p. 456] A connected graph is **semi-Eulerian** if it has an open trail that includes every edge. Such a trail is called a **semi-Eulerian trail**.

Semi-Hamiltonian graph [p. 462] A semi-Hamiltonian graph contain an open path called a **Hamiltonian path** that involves every vertex of the graph

Sequence [p. 163] A list of numbers or symbols written down in succession, for example 5, 15, 25, ...

Shortest path [p. 424] The path through a graph or network with minimum length.

Simple graph [p. 394] A graph with no loops or multiple edges.

Simple interest [pp. 178, 285] Interest that is calculated for an agreed period and paid only on the original amount invested or borrowed.

Sink [p. 455] See **sink and source**.

Sink and source [p. 455] In a flow network, a **source** generates flow while a **sink** absorbs the flow.

Slope (of a straight line) [p. 60] The slope of a straight line is defined to be: $\text{slope} = \frac{\text{rise}}{\text{run}}$. The slope is also known as the gradient.

Small circle [p. 223] Any circle on a sphere whose plane does not pass through the centre of the sphere. See also **great circle**.

Smoothing [p. 121] A technique used to eliminate some of the variation in a time series plot so that features such as seasonality or trend are more easily identified.

Source [p. 455] See **sink and source**.

Spanning tree [p. 439] A subgraph of a connected graph that contains all the vertices of the original graph, but without any multiple edges, circuits or loops.

Standard deviation (s) [p. 60] A summary statistic that measures the spread of the data values around the mean. The standard deviation is given by $s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$

Strength of a linear relationship [p. 36] Classified as weak, moderate or strong. Determined by observing the degree of scatter in a scatterplot or calculating a correlation coefficient.

Structural change (time series) [p. 116] A sudden change in the established pattern of a time series plot.

Subgraph [p. 395] Part of a graph that is also a graph in its own right.

T

Time series data [p. 107] A collection of data values along with the times (in order) at which they were recorded.

Time series plot [p. 107] A line graph where the values of the response variable are plotted in time order.

Time zone [p. 241] A region of the Earth that has a uniform standard time or local time. There are 24 time zones in total.

Total interest [pp. 338, 354] The total interest paid on a reducing-balance loan after n repayments is the difference between the total repayments made and the principal.

Trail [p. 412] A **trail** is a walk with no repeated edges.

- An **open trail** is a trail that starts and finishes at different vertices.
- A **closed trail** or **circuit** is a trail that starts and finishes at the same vertex.

See also **path**.

Tree [p. 439] A **connected graph** with no circuits, multiple edges or loops.

Trend [p. 113] The tendency for values in the time series to generally increase or decrease over a significant period of time.

Trend line forecasting [p. 143] Using a line fitted to an increasing or decreasing time series to predict future values.

Two-way frequency table [p. 8] A frequency table in which subjects are classified according to two categorical variables. Two-way frequency tables are commonly used to investigate the associations between two categorical variables.

U

Unit-cost depreciation [p. 180] Depreciation based on how many units have been produced or consumed by the object being depreciated. For example, a machine filling bottles of drink may be depreciated by 0.001 cents per bottle it fills.

V

Variable [p. 2] A symbol used to represent a number or group of numbers.

Vertex (graph) [p. 391] The points in a graph or network (p vertices).

W

Walk [p. 411] Any continuous sequence of edges, linking successive vertices, that connects two different vertices in a graph. See also **trail** and **path**.

Weighted graph [p. 424] A graph in which a number representing the size of some quantity is associated with each edge. These numbers are called weights.

Chapter 1

Section 1A

- 1 a** Categorical **b** Categorical
c Numerical **d** Numerical
e Categorical **f** Numerical
g Numerical **h** Categorical
i Categorical
- 2 a** Two categorical variables
b One categorical and one numerical
c Two numerical
d Two categorical
- 3 a** EV: colour; RV: toxicity
b EV: type of diet; RV: weight loss
c EV: age; RV: price
d EV: fuel; RV: cost
e EV: location; RV: house price
- 4 a** Age **b** Years of education
c Temperature **d** Time of year
e Age group **f** State of residence
- 5 a** Sex - categorical, EV; attitude to lowering the drinking age - categorical, RV
b Hours of study - numerical; hours spent using social media - numerical. Either variable could be the EV or RV, it would depend on the question asked.
c Gestation time - numerical, EV; birth weight - numerical, RV
d Sex - categorical, EV; hours spent using social media - numerical, RV
e Voting preference - categorical, support for tax cuts - categorical. Either variable could be the EV or RV, it would depend on the question asked.
- 6 C** **7 A** **8 C**

Section 1B

- 1 a** Gender is the EV, university is the RV.

b

		Gender		Total
		F	M	
University	No	4	4	8
	Yes	8	4	12
Total		12	8	20

- 2 a** Age group is the EV, reduce fees is the RV.

b

Reduce fees	Age group			Total
	17–18	19–25	26 or more	
No	3	3	4	10
Yes	8	6	6	20
Total	11	9	10	30

c

Reduce fees	Age group		
	17–18	18–25	26 or more
No	27.3%	33.3%	40.0%
Yes	72.7%	66.7%	60.0%
Total	100.0%	100.0%	100.0%

- 3 a** Enrolment status

b 81.8%

c 242, 81.8%, 80.5%, not.

- 4 a** Satisfaction with job is the EV, satisfaction with life is the RV.

b

Satisfaction with life	Satisfaction with job	
	Dissatisfied	Satisfied
Dissatisfied	75.0%	22.6%
Satisfied	25.0%	77.4%
Total	100.0%	100.0%

c 77.4%, 25.0%.

- 5 a** Gender is the EV, are you satisfied with your life overall? is the RV.

Satisfied with life?	Gender	
	Female	Male
Yes	86.4%	84.7%
No	13.6%	15.3%
Total	100.0%	100.0%

There is no association between satisfaction with life overall and the gender of the respondent. Very similar percentages of males and females report that they are satisfied with their life overall (Females: 86.4%, Males: 84.7%).

6 a

Handedness	Gender	
	Male	Female
Left	9.0%	9.8%
Right	91.0%	90.2%
Total	100.0%	100.0%

- b** There is no association between handedness and the gender of the respondent. Very similar percentages of males and females report that they are left handed (Females: 9.8%, Males: 9.0%).

- 7 a** Gender

- b** 54.9%

- c** There are several ways you can answer the question.

There is an association between gender and level of exercise. The percentage of males who rarely exercised (28.8%) was much lower than the percentage of females who rarely exercised (39.2%).

or

There is an association between gender and level of exercise. The percentage of males who exercised regularly (18.6%) was much higher than the percentage of females who exercised regularly (5.9%).

Note: For the category 'sometimes', there is no association between level of exercise and sex.

- 8** Class is the EV, exam grade is the RV.

Exam grade	Class	
	Dr Evans	Dr Smith
Fail	11.1%	9.4%
Pass	61.1%	62.5%
Credit or above	27.8%	28.1%
Total	100.0%	100.0%

The data does not support the suggestion that Dr Evans's maths class achieves better grades than Dr Smith's maths class. Both classes achieved similar percentages in each group at all exam grade levels, with 61.1% of Dr Evans's class achieving a pass, which is very close to the 62.5% who achieved a pass in Dr Smith's class.

- 9** Type of treatment is the EV, outcome of treatment is the RV.

Outcome of treatment	Type of treatment	
	Drug	Pillow
Complete cure	9.8%	31.3%
Partial cure	26.8%	37.5%
No improvement	63.4%	31.3%
Total	100.0%	100.0%

The data supports the contention that the special pillow would be more effective in the treatment of snoring than the treatment with drugs. A much higher percentage of those using the pillow experienced a complete cure (31.3%) compared to those using the drug treatment (9.8%).

- 10 a** 11.9% **b** 52.3% **c** ordinal

- c** There is an association between marital status and attitude to life. The highest percentage of people who found life exciting were those who had never married (52.5%). The percentages of people who find life exciting who are married (47.6%), divorced (46.7%) and separated (45.7%) are all quite similar. Those who are widowed were least likely to report they find life exciting (33.8%).

- 11** A

- 12** B

- 13** C

Section 1C

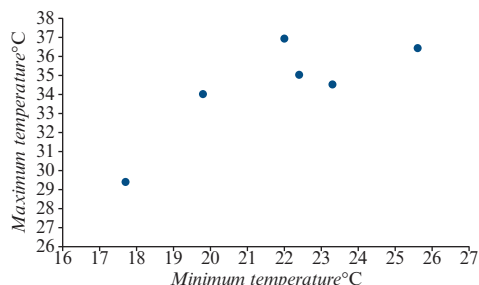
- 1 a** Number of seats

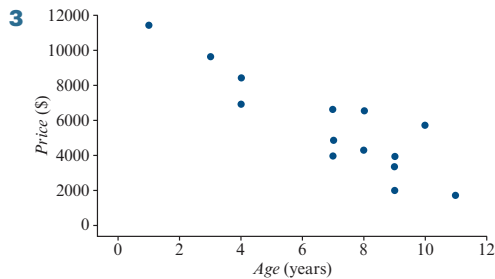
- c** 8 aircraft

- b** Numerical

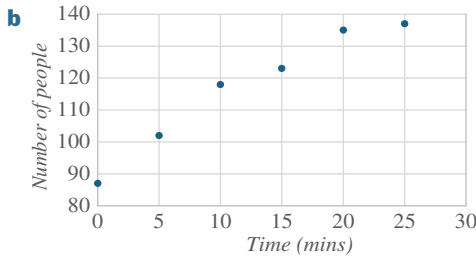
- d** Around 800 km/h

- 2**

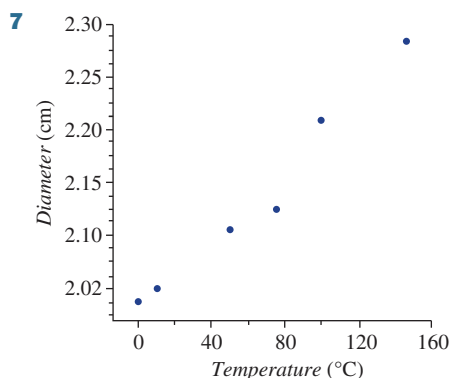
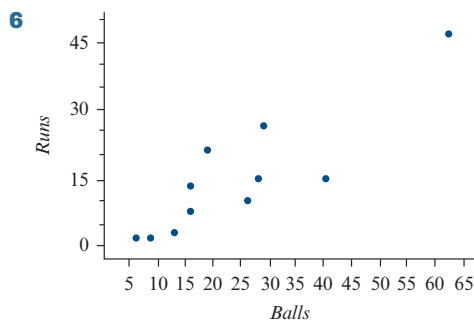
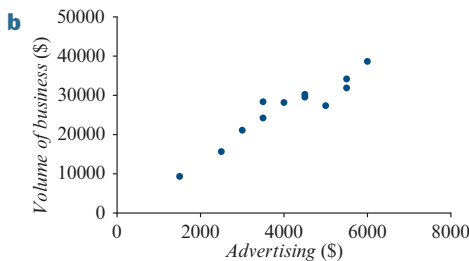




4 a Time is the EV and number in the theatre is the RV.



5 a Advertising is the EV and volume of business is the RV.



8 D

9 B

Section 1D

1 Note: There are no absolute right or wrong answers to these questions as answering them requires a degree of personal judgement.

- a** No association **b** Yes, positive
c Yes, positive **d** Yes, positive
e Yes, negative **f** Yes, negative

2 a People who undertake higher levels of daily exercise tend to have higher levels of fitness.

b People who run faster tend to take less time to run a marathon.

- 3 a** Positive, linear, moderate
b Negative, linear, weak
c Positive, linear, strong
d No association

4 a Positive **b** Linear
c Hard to classify with so little data

5 a Positive **b** Linear **c** Strong

6 strong, negative, linear, lower

7 a There is a strong, positive, linear association between time spent studying and mark. Those students who spent more time studying tended to receive higher marks.

b There is a strong, negative, linear association between the age of a secondhand sailboat and its price. Older boats tended to be cost less.

c There is a strong, positive, non-linear association between time spent practising and performance level. Performance level tended to increase as more time was spent practising up until about 8 hours of practice, after which performance level tended to level out.

d There is moderate, negative, linear association between exam room temperature and scores on a test. When the room temperature was higher students tended to achieve lower scores.

8 A

Section 1E

1 a A: strong, positive, non-linear relationship with no outliers

B: strong, negative, linear relationship with an outlier

C: weak, negative, linear relationship with no outliers

b A: non-linear B: outlier

2 Estimates could vary by ± 0.1

- a** 0.9 **b** 0.7 **c** -0.6 **d** 0

- 3 0.73 4 0.972
 8 0.570 9 0.722 10 -0.396
- 11 a None b Weak, negative
 c Strong, positive d Weak, positive
 e Strong, positive f Moderate, negative
 g Moderate, positive h None
 i Weak, negative j Weak, positive
 k Perfect positive or strong positive
 l Perfect negative or strong negative
- 12 D 13 D

Section 1F

- 1 a 45.6% b 11.9% c 32.1% d 45.3%
 e 1.5% f 0.04%
- 2 a The coefficient of determination is $r^2 = (-0.611)^2 = 0.373$ or 37.3%; that is, 37.3% of the variation observed in hearing test scores can be explained by variation in age.
- b The coefficient of determination is $r^2 = (0.716)^2 = 0.513$ or 51.3%; that is, 51.3% of the variation observed in mortality rates can be explained by variation in smoking rates.
- c The coefficient of determination is $r^2 = (-0.807)^2 = 0.651$ or 65.1%; that is, 65.1% of the variation observed in life expectancies can be explained by variation in birth rates.
- d The coefficient of determination is $r^2 = (0.818)^2 = 0.669$ or 66.9%; that is, 66.9% of the variation observed in daily maximum temperature is explained by the variability in daily minimum temperatures.
- e The coefficient of determination is $r^2 = (0.8782)^2 = 0.771$ or 77.1%; that is, 77.1% of the variation in the runs scored by a batsman is explained by the variability in the number of balls they face.
- 3 a $r = 0.906$ b $r = -0.353$
- 4 a $R^2 = 36\%$, thus 36% of variation in the weekly expenditure on food can be explained by the variation in income.
- b $R^2 = 25\%$, thus 25% of variation in the weekly expenditure on leisure can be explained by the variation in income.
- c People's incomes are associated with their expenditure on both leisure and food, but it is a better predictor of expenditure on food.
- 5 D 6 A 7 D
 8 C 9 C

Chapter 1 review

Multiple-choice questions

- 1 A 2 D 3 B 4 D
 5 C 6 B 7 D 8 D
 9 C 10 D 11 C 12 C
 13 B 14 A 15 D

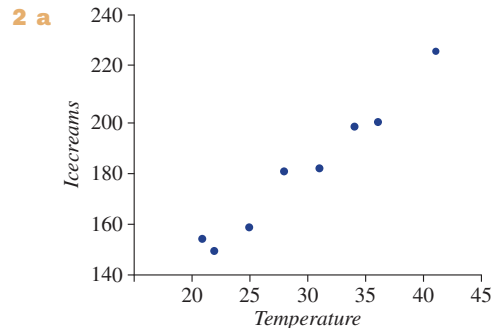
Short-response questions

- 1 a Number of accidents and age; both categorical variables
 b RV: Number of accidents; EV: age
 c 470

d

Number of accidents	Age < 30	Age ≥ 30
At most one accident	21.7%	42.5%
More than one accident	78.3%	57.5%

- e The statement is correct. Of drivers aged less than 30, 78.3% had more than one accident compared to only 57.5% of drivers in the older category.



- b There is a strong, positive, linear association between temperature and the number of ice-creams sold. When the temperature is higher more ice-creams tend to be sold.
- c $r = 0.984$
- d strong
- 3 a There is a moderate, negative, linear relationship between government expenditure on health and infant mortality. Those countries which spend more on health tend to have lower infant mortality rates. There is one country (14, 36) which is a possible outlier - this country seems to have a higher infant mortality than is indicated by the health expenditure.
- b Yes, because both variables are numerical, and relationship is linear

- 4 a EV; Gender, RV: Part-time job

Part-time job	Male	Female
Yes	67%	77%
No	33%	23%
Total	100%	100%

- c In this sample of 300 year 12 students there is an association between gender and having a part-time job. Females are more likely to work part-time than males (77% of females, 67% of males).
- 5 a $R^2 = 59.3\%$
- b 59.3% of the variation in hearing test scores is explained by the variation in age
- 6 a Years experience is the EV, which is always plotted on the horizontal axis, hourly rate is the RV, which is always plotted on the vertical axis.
- b There is a strong, positive, linear relationship; that is, people with more experience tend to be paid a higher starting pay rate. There are no apparent outliers.
- c 0.973
- d Coefficient of determination $R^2 = 0.947$; that is, 94.7% of variation in pay rate is explained by the variation in experience.
- 7 a There is a moderate, negative, linear relationship between time taken and the number of mistakes - those who take longer tend to make fewer mistakes.
- b Because both variables are numerical, and the relationship is linear.
- c $R^2 = 40.3\%$. 40.3% of the variation in the number of mistakes made can be explained by the variation in time taken.
- 8 a There is a moderate, positive, linear relationship between hours of sunlight and height of the seedlings. Those seedlings which get more sunlight tend to be taller. There is one plant (4.8, 11.5) which is a possible outlier - this plant seems to be taller than would be indicated by the number of hours of sunlight.
- b The value of r will be closer to 1, indicating a stronger correlation.
- 9 a There is a moderate, positive linear relationship between years employed and salary.
- b There is a strong, positive linear relationship between years employed and salary for those with a tertiary education. There is a

weak, positive linear relationship between years employed and salary for those with a secondary education

- c The number of years for which an employee has worked for the company explains 90.3% of the variation in salary for tertiary educated employees, and is thus a powerful predictor of salary for this group. However, the number of years for which an employee has worked for the company explains only 17.5% of the variation in salary for secondary educated employees, and so is not so useful for this group.

Chapter 2

Section 2A

- 1 a RV: *pollution level*; EV: *traffic volume*
b $\text{pollution level} = 49 \times \text{traffic volume} - 330$
- 2 a RV: *life expectancy*; EV: *birth rate*
b $\text{life expectancy} = -1.5 \times \text{birth rate} + 107.2$
- 3 a RV: *distance travelled*; EV: *age*
b $\text{distance travelled} = 11.1 \times \text{age} + 15.6$
- 6 a Answer given in question.
b $\text{runs} = 0.729 \times \text{balls faced} - 2.649$
- 7 a RV: *number of TVs*
c $\text{number of TVs} = 0.93 \times \text{number of cars} + 61.20$
- 8 0.451
- 9 a r is also negative.
b Slope is zero: regression line is horizontal.
c Intercept = $\bar{y}$ (mean of RV)
- 10 C 11 A 12 B

Section 2B

- 1 $\text{mark} = -4.3 \times \text{days absent} + 80$
- 2 a 0.33, 1
b 0, 2.9
- 3 a 575: On average, the company will achieve \$575 in sales when their online advertising expenditure is \$0.
b 4.85: On average sales will increase by \$4.85 for each additional \$1.00 spent on online advertising.
- 4 a 80 cm, extrapolating
b 92 cm, interpolating
c 98 cm, extrapolating

- 5 a** \$487.50, extrapolating
b \$1023.50, interpolating
c \$1224.50, extrapolating
6 a 173 cm, reliable, interpolating
b 189 cm, unreliable, extrapolating
c 165 cm, reliable, interpolating

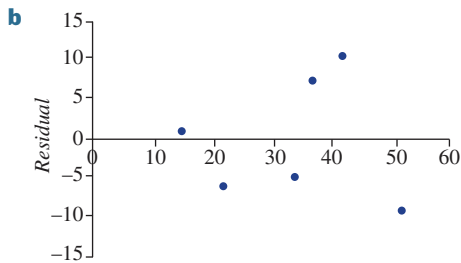
7 Answers given in question.

- 8 a** 9.7 **b** -0.8

- 9 a** 2 **b** -1 **c** 2

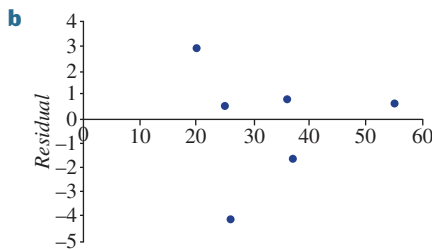
10 a

Resid	-9.1	1.0	-6.0	10.6	-4.8	7.5



11 a

Resid	-4.08	2.90	0.82	0.55	-1.61	0.65



- 12 A:** clear curved pattern in the residuals (not random), C: curved pattern in the residuals (not random).

- 13 a** 20.3%
b 42.3%
c The number of *hours* is a better predictor as it explains 42.3% of the variation in *exam score*, much more than *IQ* which explains only 20.3% of the variation in *exam score*.

- 14 a** 27.8: On average a packet of chips with 0 gm of fat contains 27.8 calories.
b 14.7: On average, the calorie content increases by 14.7 for each one additional gram of fat included.
c 75.7% of the variation in calorie content of the chips is explained by the variation in fat content.
d 145.4 **e** -13.4

- 15 a** -0.278: On average, for each additional one metre the golfer is from the hole the success rate decreases by 27.8%.

- b** 73.5%
c 3.54 m
d -0.705
e 49.7%: 49.7% of the variation in success rate in putting is explained by the variation in the distance the golfer is from the hole.

- 16 a** $\text{pay rate} = 0.289 \times \text{experience} + 8.56$

- b** 93.5%

- c** The pay rate for a worker with no experience is \$8.56 per hour.

- d** On average, the pay rate increases by 29 cents per hour for each additional one year of experience.

- e i** \$10.87 **ii** \$0.33

- f** The residual plot indicates a slight curve in the scatterplot, indicating the relationship may not be linear.

- 17 a** $r = -0.608$

- b** 37% of the variation in the hearing test score is explained by the variation in age.

- c** $\text{score} = -0.043 \times \text{age} + 4.9$

- d** -0.043; the hearing test score, on average, decreases by 0.043 for each one additional year of age.

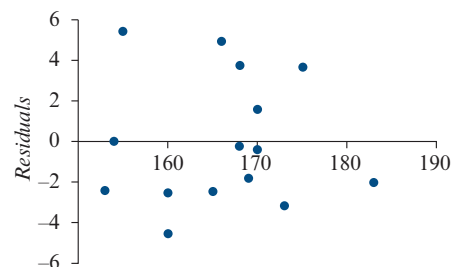
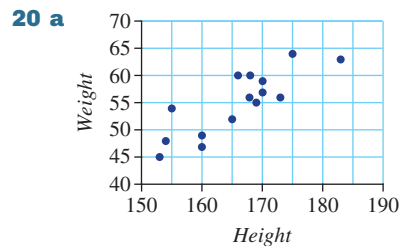
- e i** 4.04 **ii** -2.04

- f i** 0.3 **ii** -0.4

- g** Yes; no clear pattern in the residual plot.

- 18** negative, drug dose, -0.855; -10.19; 57.05; decreases, 10.19; 57.05; 73.1, response time, drug dose; curved pattern

- 19** positive, $r = 0.951$, $\text{height} = 0.807 \times \text{arm span} + 32.97$, increase, 0.81 cm, arm span, 90.4%, height, arm span.



- b** From the scatterplot, we can see that there is a strong, positive, linear relationship between height and weight: $r = 0.840$. There are no obvious outliers. The equation of the least squares regression line is $weight = 0.587 \times height - 42.35$. The slope of the regression line predicts an increase of 0.59 kg in weight for each 1 cm increase in height. The coefficient of determination indicates that for this sample 70.5% of the variation in weight is explained by the variation in height.

21 68.6 **22** D **23** B **24** A

Section 2C

Note: These answers are for guidance only. Alternative explanations for the source of an association may be equally acceptable as the variables suggested.

- 1** Not necessarily. In general, older children are taller and have been learning mathematics longer. Therefore they tend to do better on mathematics tests. Age is the probable common cause for this association.
- 2** Not necessarily. While one possible explanation is that religion is encouraging people to drink, a better explanation might be that towns with large numbers of churches also have large populations, thus explaining the larger amount of alcohol consumed. Town size is the probable common cause for this association.
- 3** Probably not. The amount of ice cream consumed and the number of drownings would both be affected by weather conditions. Weather conditions are the probable common cause.
- 4** Maybe but not necessarily. Bigger hospitals tend to treat more people with serious illnesses and these require longer hospital stays. A common cause could be the type of patients treated at the hospital.
- 5** Possible confounding variables include personality type (a highly anxious person may both smoke and be susceptible to heart disease), diet, alcohol consumption, level of exercise.
- 6** Coincidence
- 7** The size of the fire.
- 8** **a** Moderate

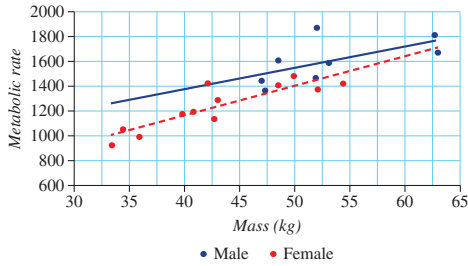
- b i** A confounding variable is a variable, other than the explanatory variable, which may also explain any observed changes in the response variable.
- ii** Nurses who have more years experience may have also continued to study and have achieved higher qualifications. It may be the possession of higher qualifications that explains the higher salaries, not the years of experience.
- iii** Restrict the study to include only nurses who have the same level of qualifications.

9 D **10** C

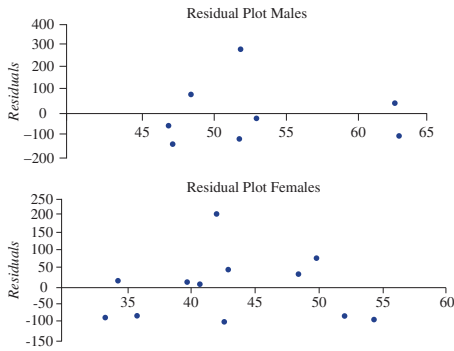
Section 2D

- 1** **a** percentaged two-way frequency table
b scatterplot **c** parallel box plots
d scatterplot **e** scatterplot
f percentaged two-way frequency table
g percentaged two-way frequency table
h parallel box plots or back-to-back stem plots
- 2** In 1990 there is an association between agreement with this statement and ethnicity for males, with low but quite similar for levels of agreement for males in Australia (23.0%) and in the UK 29.5%), but a higher level of agreement for males in Europe (37%). However, at this time there is no association between agreement with this statement and ethnicity for females, with low and quite similars for level of agreement for females in Australia (25.1%), the UK (27.0%), and Europe (27.5%).
 In 2010 there is no association between agreement with this statement and ethnicity for males. Levels of agreement with the statement are similar for males in Australia (13.9%), the UK (16.7%) and Europe (19.3%). There is also no association between agreement with this statement and ethnicity for females. Levels of agreement with the statement are similar for females in Australia (10.0%), the UK (11.0%) and Europe (13.7%).
 Between 1990 and 2010 agreement with the statement has lowered considerably for both sexes across all three ethnicities.
- 3** A study was conducted to investigate the hypothesis that lean body mass would have a strong association with metabolic rate. Data were collected from a sample of 20 people,

12 women and 8 men. The results are displayed in the following scatterplot.



There are no obvious outliers, and the assumption that the relationship between lean body mass and metabolic rate is linear for males and females is confirmed by the following residual plots.



For males there is a moderate, positive, linear relationship between lean body mass and metabolic rate: $r = 0.614$. The equation of the least squares regression line is: $\text{metabolic rate} = 17.16 \times \text{lean body mass} + 689.1$. The slope of the regression line predicts an increase of 17.61 in metabolic rate for each 1 kg increase in lean body mass.

For females there is a strong, positive, linear relationship between lean body mass and metabolic rate: $r = 0.876$. The equation of the least squares regression line is: $\text{metabolic rate} = 23.92 \times \text{lean body mass} + 206.6$

The slope of the regression line predicts an increase of 23.92 in metabolic rate for each 1 kg increase in lean body mass.

Comparing the values of the coefficient of determination for each group we can see that for this sample 76.7% of the variation in metabolic rate is explained by the variation in lean body mass for females, while only 37.7% of the variation of the variation in metabolic rate is explained by the variation in lean body mass for males.

In conclusion, the researchers hypothesis only partially supported, in that there is a strong, positive association between lean body mass and metabolic rate for females, but only a moderate, positive association between lean body mass and metabolic rate for males. Scatterplot shows relationship is linear. For the total data set, $r = 0.871$. For the females $r = 0.876$, for the males $r = 0.614$, so the relationship is stronger for females than males.

Chapter 2 review

Multiple-choice questions

1 C	2 A	3 B	4 D
5 C	6 B	7 B	8 D
9 A	10 A	11 A	12 D
13 D	14 A	15 C	16 A
17 C	18 B	19 C	20 C
21 B	22 D	23 C	

Short-response questions

- $c = 91.725$, $m = 2.625$
- Number of ice creams = $58.2 + 4.1 \times \text{temperature}$
- Residual = 600
- On average, price decreases by \$5675 per year.
 - On average, the price of a new caravan is \$87 500.
- 3.91 secs
 - 3.19 secs
 - Predicting for a person 55 years of age is interpolation, and we can be reasonably confident that this prediction will be reliable. Predicting for a person 35 years of age is extrapolation, and we cannot be confident of this prediction.
- score
 - slope is 0.34, intercept is 32
 - 70.1
 - 7.9
- days of rain
 - 6.88, 2850
 - 2024
 - decrease, 6.88
 - 0.696
 - 48.4, days of rain
 - 1873
 - 483
 - interpolation

8 a Cost

b strong, positive, linear

c i \$307.30

ii extrapolating

d i 222.48: The fixed costs of preparing meals is \$222.48.

ii \$4.039: The slope of the regression line predicts that, on average, meal preparation costs increase by \$4.04 for each additional meal produced.

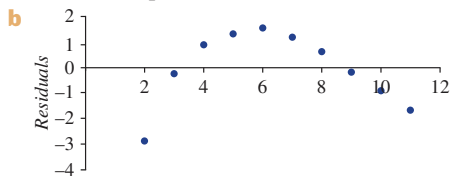
9 a RV: height; EV: femur length

b $height = 36.3 + 5.35 \times femur\ length$

c On average, height increases by 5.35 cm for each 1 cm increase in femur length.

d $r^2 = 0.988$; that is, 98.8% of the variation in height is explained by the variation in femur length.

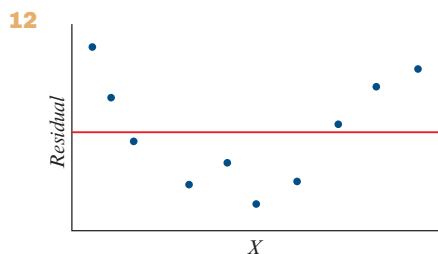
10 a 83 cm, extrapolation



c No, the residual plot shows a clear curve, indicating that this relationship is not linear.

11 a There is a weak positive correlation between educational attainment and the amount spent on education ($r = 0.43$). Those countries which spend more on education also tend to have higher educational attainment. There is a moderate negative correlation between educational attainment and the student: teacher ratio ($r = -0.64$). Those countries with a higher student: teacher ratio tend to have lower educational attainment.

b Student: teacher ratio explains 41.0% of the variation in educational attainment, making it a much more important explanatory variable than amount spent on education, which explains only 18.5%.



13 Correlation implies that two variables have been observed to vary together, either one increasing as the other increases (positive correlation) or one increasing as the other decreasing (negative correlation). It may be that this is because there is a causal relationship between the variables (for example, the further we drive the less fuel there will be in the tank of the car) or it could be because there is a third variable which is affecting both (e.g. age will affect both height and scores on an IQ test, meaning that although there might be a high correlation between height and IQ scores in children, we do not think taller people are smarter).

14 No a causal relation cannot be assumed. It may be that only mature people are likely to be prepared to have children. Age is also a confounding variable here. Older people are likely to both have children and have higher maturity levels.

15 a i 142 beats/min

ii extrapolating

b -6.3

c i linearity

ii the lack of a clear pattern in the residual plot supports the linearity assumption.

d In a study of the association between heart rate before exercise and heart rate after exercise data was collected from 13 students. There is a moderate, positive linear association between heart rate before exercise and heart rate after exercise ($r = 0.699$). Those people with a higher resting heart rate tend to have a higher heart rate after exercise. There are no obvious outliers, and the linearity assumption is confirmed by the residual plot. The least square regression equation is:

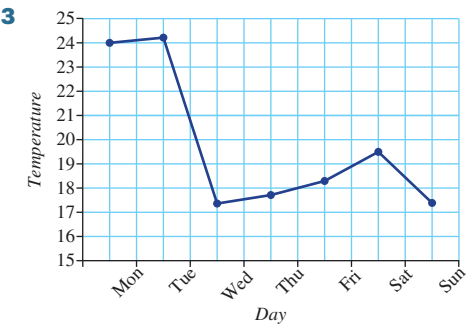
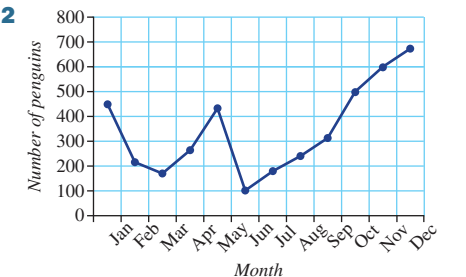
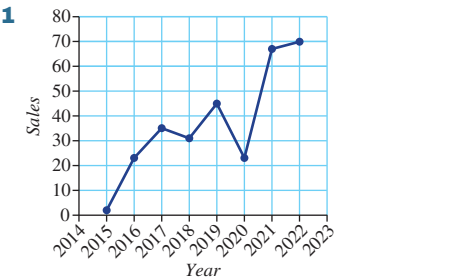
$heart\ rate\ after\ exercise =$

$0.561 \times heart\ rate\ before\ exercise + 85.671$

The slope of the regression line predicts an increase of 0.561 beats/min after exercise each 1 beat/min before exercise. The coefficient of determination indicates that 48.9% of the variation in heart rate after exercise can be explained by the variation in heart rate before exercise.

Chapter 3

Section 3A



4

Feature	A	B	C
Irreg fluct	✓	✓	✓
Inc			✓
Dec	✓		

5

Feature	A	B	C
Irreg fluct	✓	✓	✓
Inc trend			✓
Dec trend	✓		
Cycles		✓	
Seasonality	✓		✓

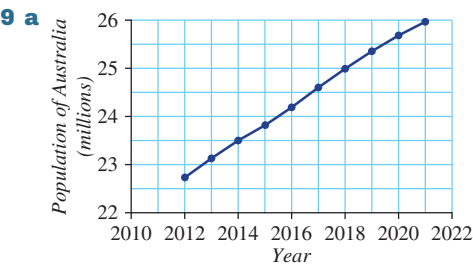
6

Feature	A	B	C
Irreg fluct	✓		✓
Struct change	✓		
Inc trend	✓		
Dec trend	✓		
Seasonality			✓

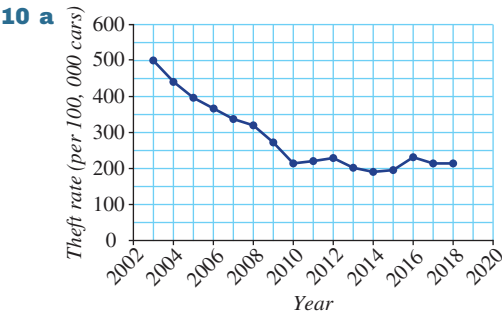
7

Feature	A	B	C
Irreg fluct	✓	✓	✓
Struct change		✓	
Inc trend	✓		
Dec trend			
Outlier			✓

8 The number of mobile phones per 100 people increases rapidly over the years 2000–2008. The number continues to increase from 2009 until 2019, but the increase in the number of phones is at much lower level than in the preceding years.



b The plot shows a steady increase in the population of Australia over the years 2012–2021.



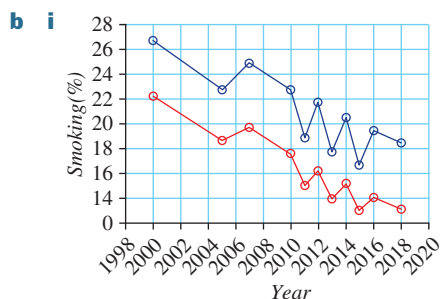
b The plot shows a steady decline in the number of vehicle thefts over the years 2003–2010, after which the number of vehicle thefts has remained reasonably steady, showing only irregular variation.

11 The number of cases of measles show an increasing trend between 1989 and 1992. In 1993–1994 there is a rapid increase in the number of measles cases, followed by a rapid decrease in 1995–1996. The number of cases continued to decrease until 2000, since then have remained low, showing only irregular variation over the years 2001–2023.

12 The number of overseas arrivals (millions people per month) in Australia increased steadily from November 2011 until April 2020. The number of arrivals is clearly seasonal, with the peak time for arrivals in the January quarter each year. The number of arrivals dropped suddenly to almost zero in August 2020, due to the COVID 19 pandemic, and remained at this level until October 2021. Since then, arrivals have again steadily increased, but are yet not back to the pre-COVID levels.

13 a i The percentage of males who smoke has consistently decreased since 1945, while the percentage of females who smoke increased slightly from 1945 to 1975 but then decreased at a similar rate to males over the period 1975–1992.

ii The difference in smoking rates between males and females has decreased steadily over these years. In 1945 the smoking rate for males was more than 40% more than the smoking rate more females, but by 1995 this difference was less than 10%.



ii Whilst both plots show irregular fluctuation, overall the percentages of male and females who smoke have declined substantially over the years 2000–2018.

iii Although the smoking rate for males continues to be higher than the smoking rate for female, this difference has remained almost the same over these years, and is less than 5%.

14 D

15 C

Section 3B

- 1 a i** 3 **ii** 1 **iii** 4
b i 3.2 **ii** 1.2 **iii** 2.2
c i 2.6 **ii** 2.0
d 2.3

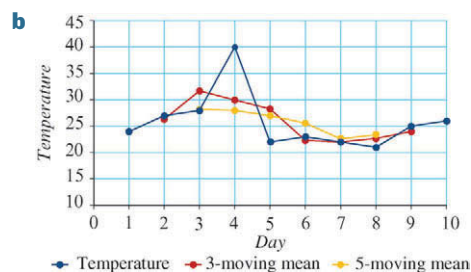
- 2 a** 24.4 **b** 20.0 **c** 23.2

3

t	1	2	3	4	5	6	7	8	9
y	10	12	8	4	12	8	10	18	2
3-mean	–	10	8	8	8	10	12	10	–
5-mean	–	–	9.2	8.8	8.4	10.4	10	–	–

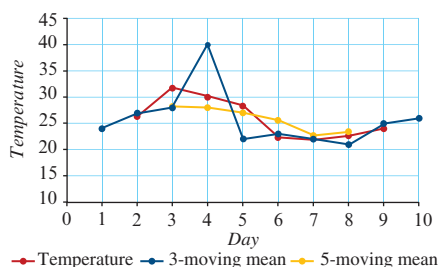
4 a

Day	Temperature (C°)	3-moving mean	5-moving mean
1	24	–	–
2	27	26.3	–
3	28	31.7	28.2
4	40	30.0	28.0
5	22	28.3	27.0
6	23	22.3	25.6
7	22	22.0	22.6
8	21	22.7	23.4
9	25	24.0	–
10	26	–	–



The smoothed plots show that the 'average' maximum temperature changes relatively slowly over the 10-day period (the 5-day average varies by only 5°) when compared to the daily maximum, which can vary quite widely (for example, nearly 20° between the fourth and fifth day) over the same period of time.

5 a, c



The exchange rate has a downward trend over the 10-day period. This is most obvious from the smoothed plots, particularly the 5-moving mean plot.

Day	Exchange rate	3-moving mean	5-moving mean
1	0.743	—	—
2	0.754	0.745	—
3	0.737	0.747	0.742
4	0.751	0.737	0.738
5	0.724	0.733	0.730
6	0.724	0.720	0.729
7	0.712	0.724	0.722
8	0.735	0.721	0.720
9	0.716	0.721	—
10	0.711	—	—

6 C

7 A

8 D

9 C

Section 3C

1 a (3, 3)

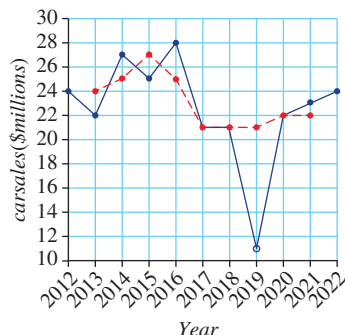
b (2, 2)

c (3, 2)

d (3, 3)

2 a 30°C b 25°C

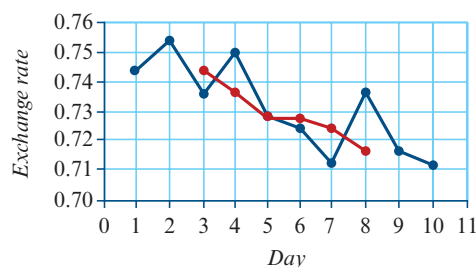
3



The smoothed plot shows that sales were quite consistent up to 2017 when they dropped, and have remained at this lower level from 2017–2022.

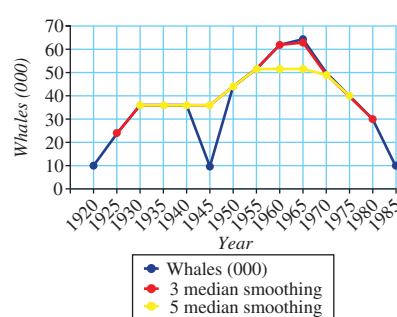
4 a 25°C b 25°C

5



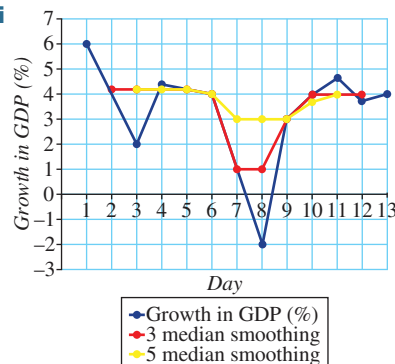
The smoothed plot shows that there was a general decreasing trend in the exchange rate over this period.

6



7 a about 4%

b i, ii



c The plot of GDP growth over 1 year, shows a great deal of variability, with no clear trend is apparent. When smoothed over a 3-year period, GDP growth is still variable but to a lesser extent. No clear trend is apparent, but GDP appears to be going through a period of below average growth during the time period Year 7 to Year 9. When smoothed over a 5-year period, GDP growth is much less variable but clearly shows the period of below average growth during the period Year 7 to Year 9.

8 D

9 C

10 D

Section 3D

- 1 a** 1.0
b Sales in February are typically 30% higher than sales in the average month.
c Sales in September are typically 10% lower than sales in the average month.

2 a 7.8 **b** 6.7

3 a 3.9 **b** 6.9

- 4 a** Increase by 42.9%.
b Decrease by 23.1%

5 a 1.2 **b** 1514 **c** 1437 **d** 1005

6 Number of students: 56 125 126 96
 Deseasonalised numbers: 112 125 97 80
 Seasonal index: 0.5 1.0 1.3 1.2

7 a, c

Deseasonalised: 152 142 148 153
 Seasonal index: 1.30 1.02 0.58 1.1

- b** In quarter 1 the restaurant chain employs 30% more waiters than the number employed in an average quarter.

8

Q1	Q2	Q3	Q4
0.89	0.83	1.12	1.16

9

Jan	Feb	Mar	Apr	May	Jun
0.89	0.96	1.04	1.26	1.33	1.11

Jul	Aug	Sep	Oct	Nov	Dec
0.67	0.74	0.59	0.81	1.11	1.48

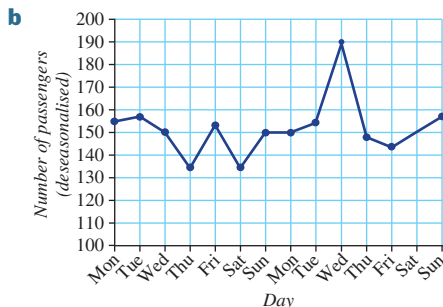
10

Jan	Feb	Mar	Apr	May	Jun
1.06	0.96	1.18	1.18	1.03	0.86

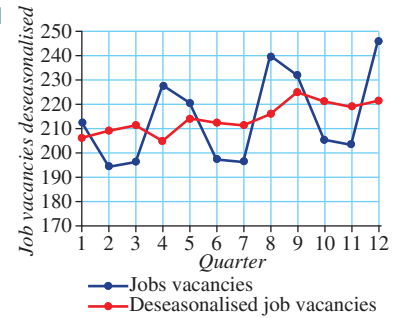
Jul	Aug	Sep	Oct	Nov	Dec
0.96	0.79	0.74	0.54	1.18	1.50

11 a

Week	Mon	Tue	Wed	Thu	Fri	Sat	Sun
1	155	157	150	134	153	134	150
2	150	154	190	148	143	150	157



12 a, d



b

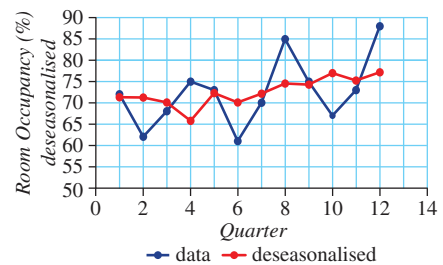
Q1	Q2	Q3	Q4
1.03	0.93	0.93	1.11

c

Year	Q1	Q2	Q3	Q4
1	206	209	211	205
2	214	212	211	215
3	224	220	218	221

- e** The number of job vacancies has increased steadily over this three year period.

13 a, d



b

Sum	Aut	Win	Spr
1.01	0.87	0.97	1.14

c

Season	2022	2023	2024
summer	71	72	74
autumn	71	70	77
winter	70	72	75
spring	66	75	77

- e** The percentage occupancy rates increased steadily over this three year period.

14 D 15 C 16 D 17 B

Section 3E

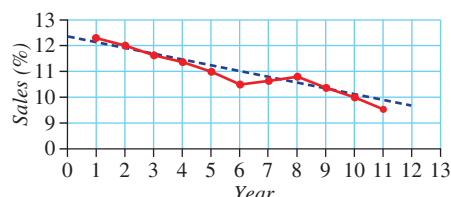
- 1 a** The population of Australia has been increasing steadily over the years 2014–2023.
b $population = 0.365 \times year - 712.540$
c 28 400 000 people
2 a The number of commencing university students in Australia increased steadily from 2015–2019. From 2020–2022 the

numbers decreased rapidly, following by a large increase between 2022–2023.

b *commencing students* =
 $6.958 \times \text{year} - 13\,430.467$

c 690 000

3 a, d



b General decreasing trend in the percentage of retail sales made in department stores

c $\text{sales} = -0.258 \times \text{year} + 12.5$. The percentage of total retail sales that are made in department stores is decreasing by approximately 0.3% per year.

e 8.6%

4 a $\text{age} = 0.088 \times \text{year} - 147$; On average, the average age of mothers increased by 0.088 years (equivalent to 1 month) each year between 2010 and 2020.

b 31.6 years; Unreliable as we are extrapolating 10 years beyond the period in which the data were collected.

5 a $\text{earnings} = 42.07 \times \text{year} - 83\,280$; On average, average weekly earnings increased by \$42.07 each year between 2014 and 2021.

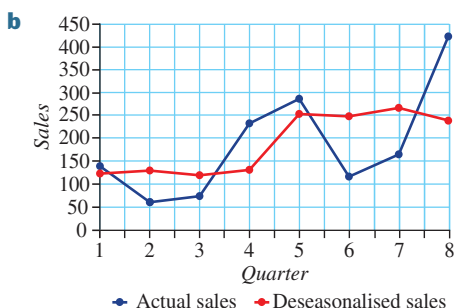
b \$2122.10; Unreliable as we are extrapolating 9 years beyond the period in which the data were collected.

6 a *deseasonalised number* =
 $1.59 \times \text{quarter number} + 50.9$

b deseasonalised number = 76.34
 reseasonalised (actual) number = 90 (to the nearest whole number)

7 a

Year	Quarter 1	Quarter 2	Quarter 3	Quarter 4
1	122	128	118	130
2	250	245	263	236



The deseasonalised sales appear to show an increasing trend over time.

c *deseasonalised sales* = $80.8 + 23.5 \times \text{quarter}$

d forecasted actual sales = $386.3 \times 1.13 = 437$

8 B

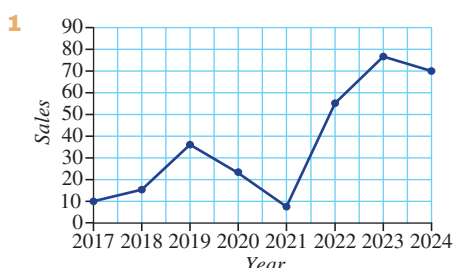
9 A

Chapter 3 review

Multiple-choice questions

1 D	2 D	3 A	4 A
5 B	6 C	7 D	8 C
9 C	10 C	11 B	12 C
13 B	14 A	15 D	16 D
17 D	18 A	19 B	20 A

Short-response questions



2 The time series plot shows that the sales of houses is seasonal, with most houses sold in Quarter 3. The least numbers of houses are sold in Quarter 1. The plot also shows increasing trend, indicating that house sales have increased over this three year period. The plot also shows random fluctuations.

3 a 22.3

b 32.8

4 a 0.708

b 0.704

5 a 1.86

b On Saturday the ice cream sales are 86% higher than they are on an average day.

c 244

6

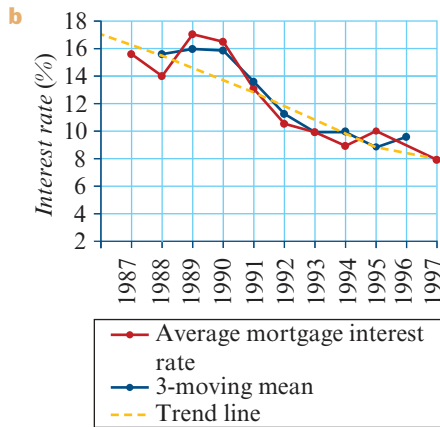
Q1	Q2	Q3	Q4
0.63	1.77	1.18	0.42

7 a $\text{visitors} = 273.89 \times \text{year} - 546\,236.46$

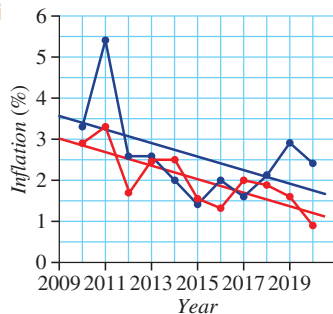
b 9 760 000

c May be unreliable as we are extrapolating.

8 a 3-mean smoothed value for 1992 = 11%



- c** Both the raw data and the smoothed data reveal a steadily decreasing trend.
- 9 a** $CO_2 = -3.33 \times \text{year} + 7\,124.33$
b 371
c May be unreliable as we are extrapolating.
- 10** 205
- 11 a**
- | Sum | Aut | Win | Spr |
|------|------|------|------|
| 0.29 | 0.35 | 1.36 | 2.01 |
- b** $\text{dolphins} = -8.51 \times \text{season} + 380.40$, where Summer, 2022 is season 1.
- c**
- | Sum | Aut | Win | Spr |
|-----|-----|-----|-----|
| 68 | 80 | 297 | 423 |
- 12 a i** $\text{inflation} = 332 - 0.164 \times \text{year}$
b i $\text{inflation} = 339 - 0.167 \times \text{year}$
a ii, b ii



- c** The trend lines are approximately parallel. As such, they are unlikely to cross, so the inflation rate for China will remain higher than the inflation rate for Australia.

Chapter 4

Section 4A

- 1 a** 5, Add 5
c 10, Subtract 1
e 101, Add 4
b 2, Add 2
d 22, Subtract 3
f 87, Add 5

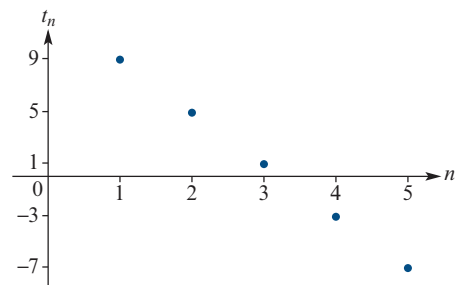
- 2 a** 35
e 127
b 18
f 78
c 60
d 11
- 3 a** 31, 37
d 21, 17
b -21, -27
e 37, 41
c 44, 49
f 80, 69
- 4 a** 2, 8, 14, 20, 26
c 10, 22, 34, 46, 58
b 5, 2, -1, -4, -7
d 56, 45, 34, 23, 12
- 5 a** 4, 6, 8, 10, 12
c 24, 20, 16, 12, 8
e 2, 10, 50, 250, 1250
b 50, 10, 2, 0.4, 0.08
d 5, 15, 45, 135, 405
- 6 a** 2
e -10
b 14
f 2
c 7
d 56
- 7 a** 1, 8, 15, 22, 29
c 7, 16, 25, 34, 43
e 70, 87, 104, 121, 138
f 27, 24, 21, 18, 15
b 14, 12, 10, 8, 6
d 69, 58, 47, 36, 25
- 8** C
9 A
10 D

Section 4B

- 1** 18
- 2**
- | n | 8 | 12 | 16 | 20 | 24 |
|-------|-------|-------|-------|-------|-------|
| t_n | t_1 | t_2 | t_3 | t_4 | t_5 |
- 3 a** 15
b 24
c 27
- 4 a** 27
b 33
c 9
- 5 a** 3, 3
b 6, 6
c -4, -4
d -11, -11
- 6 a** 7, 7, 7, 7, arithmetic
b $t_6 = 39, t_7 = 46$
- 7 a** -8, -4, -2, -1 - Not arithmetic
b N/A
- 8 a** $d = 5, t_5 = 23$
c $d = -7, t_5 = 161$
b $d = -4, t_5 = -1$
d $d = 98, t_5 = 605$
- 9** 3, 7, 11, 15, 19
- 10** 15, 11, 7, 3, -1
- 11 a** 4
c 7 times, $t_8 = 48$
b 40
d 12 times, $t_{13} = 68$
- 12 a** -2
d -13
b -5
e -93
c 6, $t_7 = -7$

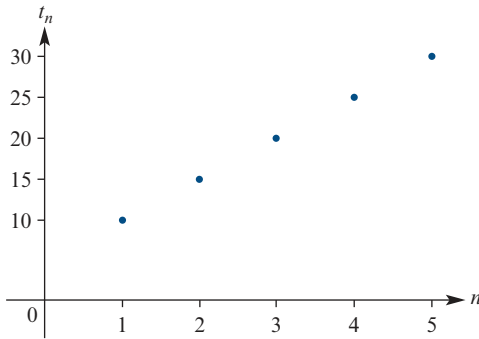
13

n	1	2	3	4	5
t_n	9	5	1	-3	-7



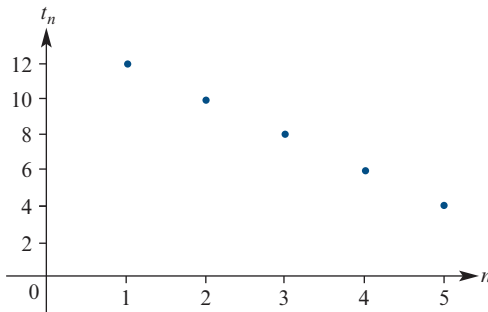
14

n	1	2	3	4	5
t_n	10	15	20	25	30



15

n	1	2	3	4	5
t_n	12	10	8	6	4



- 16 a** $t_1 = 7, t_{n+1} = t_n + 3$
b $t_1 = 19, t_{n+1} = t_n + 15$
c $t_1 = 62, t_{n+1} = t_n - 8$
- 17 a** $t_1 = 19, t_{n+1} = t_n + 4$
b $t_1 = 26, t_{n+1} = t_n - 6$
c $t_1 = 11, t_{n+1} = t_n + 9$
d $t_1 = 4, t_{n+1} = t_n - 11$

18 t_9

- 19** $t_1 = 370, t_{n+1} = t_n - 12$
 After 31 weeks all the wine glasses will be broken.

20 a -299 **b** -599

21 B **22** C **23** C

Section 4C

- 1 a** 220 **b** 290 **c** 1700
2 a 55 **b** -16 **c** 0.36 **d** 101
e -5.5 **f** -528 **g** -1
3 a $t_n = 7n - 1$ **b** $t_n = 59 - 11n$
c $t_n = 11n - 3$ **d** $t_n = 1020 - 20n$

4 4

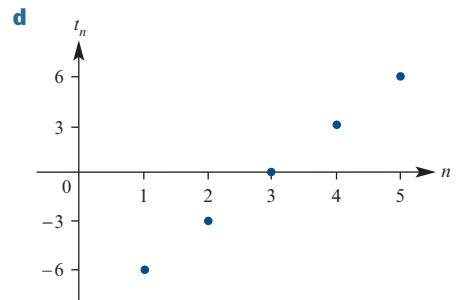
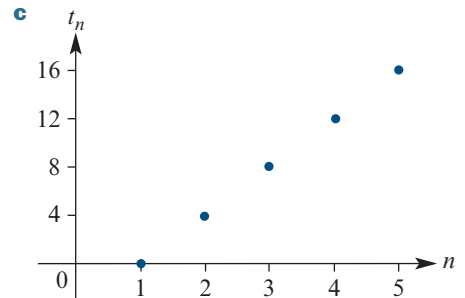
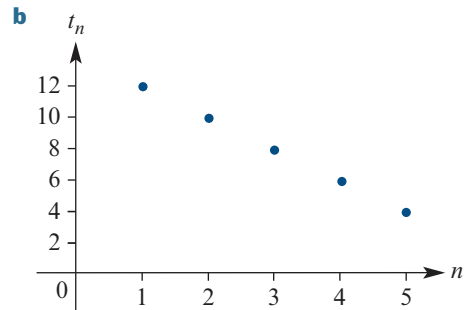
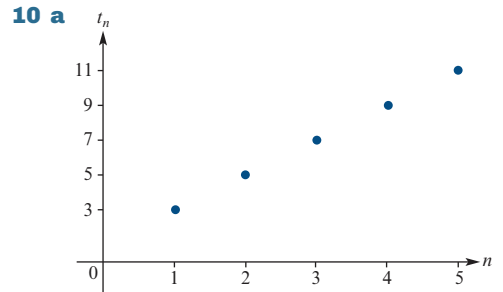
5 8

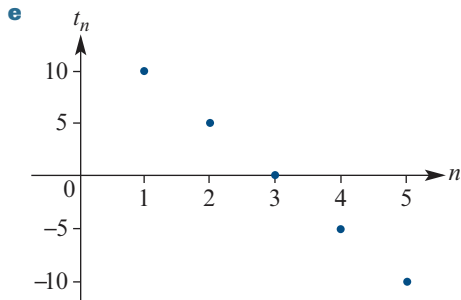
6 11

7 70

- 8 a** 7, 12, 17 **b** 1, 4, 7
c 20, 16, 12 **d** 99, 104, 109
e 20, 16, 12 **f** 120, 110, 100
g -25, -20, -15

- 9 a** 12 **b** 19 **c** 9 **d** 8
e 7 **f** 10 **g** 27





11 A : $a = 25, d = -4, t_n = 29 - 4n$

B : $a = 5, d = 5, t_n = 5n$

C : $a = 18, d = -1, t_n = 19 - n$

D : $a = 1, d = 6, t_n = 6n - 5$

12 a $t_n = 2n$ **b** 2, 4, 6, 8

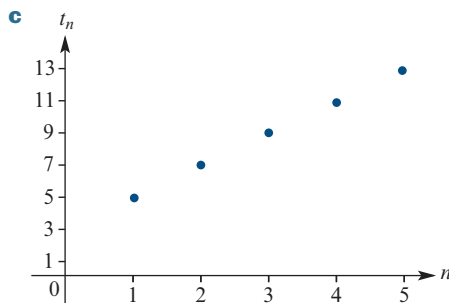
13 a $t_n = 6n + 5$ **b** 11, 17, 23, 29

14 a $t_n = 59 - 3n$ **b** 56, 53, 50, 47

15 a $t_n = 2n + 3$

b

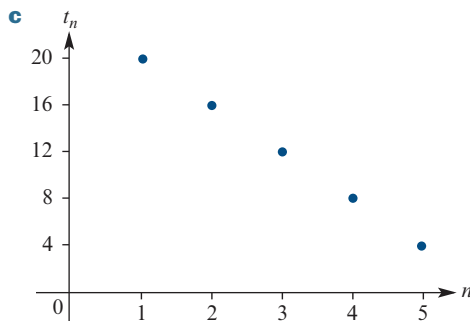
n	1	2	3	4	5
t_n	5	7	9	11	13



16 a $t_n = 24 - 4n$

b

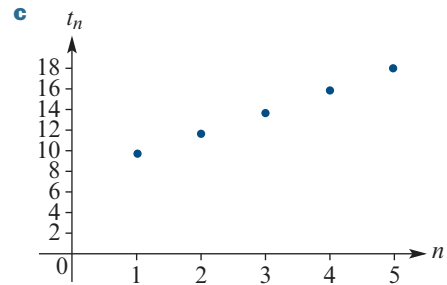
n	1	2	3	4	5
t_n	20	16	12	8	4



17 a $t_n = 2n + 8$

b

n	1	2	3	4	5
t_n	10	12	14	16	18



18 $t_1 = 7, d = 5$

19 $t_1 = 2, d = -2$

20 a $t_1 = 18, d = 7; t_1 = 18, t_{n+1} = t_n + 7$

b $t_1 = 3, d = -5; t_1 = 3, t_{n+1} = t_n - 5$

c $t_1 = 9, d = -3; t_1 = 9, t_{n+1} = t_n - 3$

d $t_1 = 47, d = 6; t_1 = 41, t_{n+1} = t_n + 6$

21 B **22** C **23** D

Section 4D

1 a \$375

b

Year	1	2	3	4	5
Value(\$)	5375	5750	6125	6500	6875

c $A = 5000 + 375n$ dollars

d i \$10 625

ii \$14 375

2 a \$4500

End of year	1	2	3	4	5
Amount (\$)	54 500	59 000	63 500	68 000	72 500

c $A = 50\,000 + 4500n$

d i \$117 500

ii \$162 500

3 a $A = 60\,000 + 2700n$

b \$73 500

c 8 years

4 a $A = 2000 + 76n$ **b** \$2456 **c** 14 years

5 a $A = 7000 + 518n$ **b** \$10 108 **c** 6 years

6 a $C = 2.90 + 1.94k$ **b** \$35.88

7 a $R = 3.10 + 1.86k$ **b** $G = 3.65 + 1.72k$

c 4 kilometres

8 a $A = 100\,000 - 4000n$ **b** \$80 000

9 a $A = 235\,000 - n \times 11\,985$ **b** In 2031

10 a $V = 48\,000 - \frac{n}{5}$

b \$45 000

c 25 000 km

11 a $V = 82\,000 - n \times 350$

b \$77 100

c 100 000 km

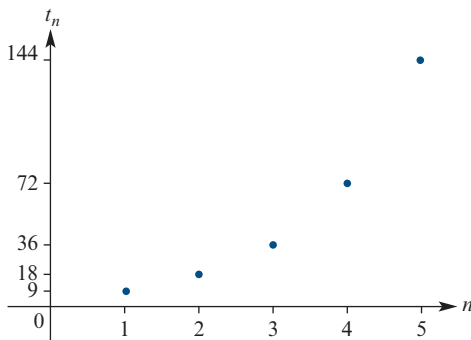
- 12 a** $C_n = 2.25 + 0.02n$
b i \$2.45 **ii** \$2.95
c i 80 pages **ii** 200 pages
- 13 a** $g_n = 86 + 40n$
b i 686 kg **ii** 1166 kg
c 35
- 14 a** $a_n = 15 - 0.2n$
b i 10.4 litres **ii** 6 litres
c 67
- 15 a** $s_n = 20\,500 + 450(n - 1)$
 or $s_n = \$450n + \$20\,050$
b i \$22 300 **ii** \$23 650
c Start of 67th year
- 16 a** $m_n = 430 - 25.75(n - 1)$
 or $m_n = \$455.75 - \$25.75n$
b i \$275.50 **ii** \$121
c 17 days
- 17 C** **18 D** **19 B**

Section 4E

- 1** 3, 12, 48, 192, 768
2 15, 30, 60, 120, 240
3 a 10 **b** 20 000 **c** 4 **d** 14
e 2.46×10^{14}
4 a $\frac{1}{4}$ **b** 4 **c** 6 **d** $\frac{1}{4}$
e 0.00390625 **f** 3.23×10^{-27}
5 a -5
b -625
c 5
d 3125
e 14
f -6 103 515 625

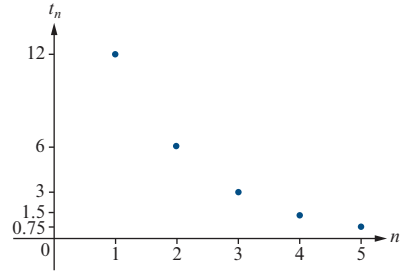
6

n	1	2	3	4	5
t_n	9	18	36	72	144



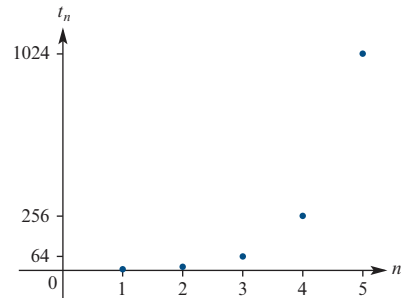
7

n	1	2	3	4	5
t_n	12	6	3	1.5	0.75



8

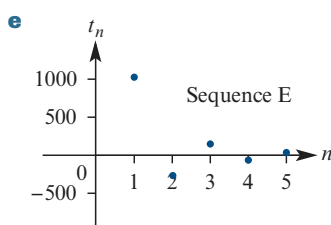
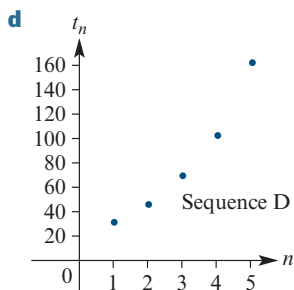
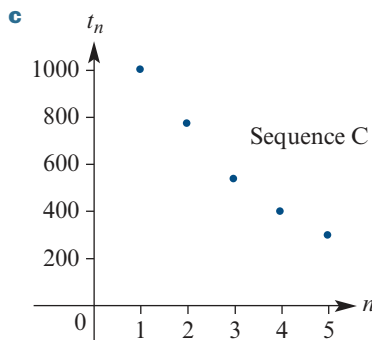
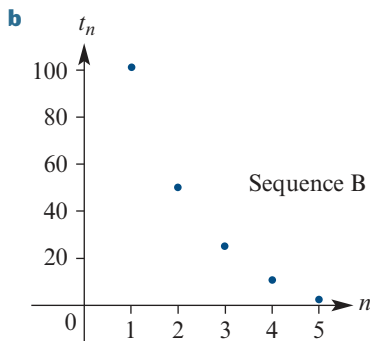
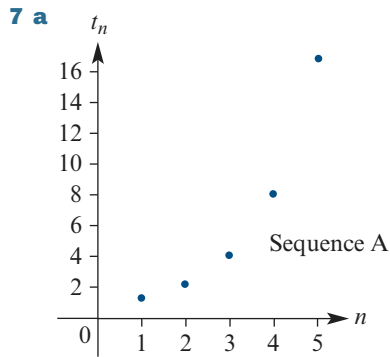
n	1	2	3	4	5
t_n	4	16	64	256	1024



- 9 a** $t_1 = 4, t_{n+1} = 2 \times t_n$ **b** $t_1 = 10, t_{n+1} = 7 \times t_n$
c $t_1 = 16, t_{n+1} = \frac{1}{2} \times t_n$
10 a $t_1 = 3, t_{n+1} = 3 \times t_n$ **b** $t_1 = 2, t_{n+1} = 4 \times t_n$
c $t_1 = 200, t_{n+1} = \frac{1}{2} \times t_n$
11 t_9
12 a $t_1 = 2, t_{n+1} = 2 \times t_n$ **b** Day 13
13 a 4, 8, 16, 32, 64, 128, 256
b $t_1 = 4, t_{n+1} = 2 \times t_n$ **c** Day 6
14 $t_1 = 20, t_{n+1} = 2 \times t_n$
15 B **16 B** **17 A** **18 C**

Section 4F

- 1 a** $t_n = 2 \times 5^{n-1}$ **b** $t_n = 5000 \times 0.8^{n-1}$
2 a 16 **b** 4096
c 268 435 456
3 a $t_n = 8 \times 5^{n-1}$ **b** $t_n = 5 \times 0.6^{n-1}$
c $t_n = 32 \times 0.25^{n-1}$
4 a 15 625 **b** 0.128
c 512 **d** -131 220
e 0.015625 **f** 161.051
g 0.015625
5 a 1, 3, 9
b 20 000, 10 000, 5000
c 10, -20, 40
d 128, 160, 200
6 a 8 **b** 9 **c** 12 **d** 10
e 22 **f** 16 **g** 15



- 8** Sequence A
a 25 **b** $0 < r < 1$ **c** $r \approx 0.6$
 Sequence B
a 10 **b** $r > 1$ **c** $r \approx 1.2$
 Sequence C
a 10 **b** $0 < r < 1$ **c** $r \approx 0.5$
9 a 2, 4, 8, 16 **b** 2
10 C **11 C** **12 B**

Section 4G

- 1 a** 160 000, 320 000, 640 000, 1 280 000
b $B_n = 80\,000 \times 2^n$ **c** 2 560 000
2 a 200 000, 400 000 **b** $N = 100\,000 \times 2^n$
c 6 400 000
3 a $A = 2700 \times 1.05^n$ **b** 3446 **c** 15 years
d 971 **e** 18 years
4 a 20

0	1	2	3
10	20	40	80

b $b_n = 10 \times 2^n$
d i 160 **ii** 640 **iii** 40 960
5 a 1400

0	1	2	3	3
1000	1400	1960	2744	3842

b $F_n = 1000(1.4)^n$
d i 7530 **ii** 28 925 **iii** 836 683
6 \$5150, \$5304.50, \$5463.64
7 a $A = 6000 \times (1.042)^n$
b \$7073.30
8 a \$15 810, \$16 663.74, \$17 563.58
b $A_n = 15\,000 \times 1.054^n$ **c** 9 years
9 a $A = 20\,000 \times (1.063)^n$
b \$27 145.40 **c** 7 years
d $A = 18\,000 \times (1.094)^n$
10 a $A = 8000 \times (1.125)^n$
b \$11 390.63 **c** \$3390.63
d \$1265.63
11 a $A = 3300 \times (1.075)^n$
b \$6801.40 **c** \$3501.40
d \$474.52
12 a $A = 9800 \times (0.965)^n$ **b** \$8201
c \$319
13 a $A = 1200 \times (0.88)^n$ **b** \$490.41 **c** 6 years
14 a \$18 500

0	1	2	3
20 000	18 500	17 113	15 829

b $V_n = 20\,000(0.925)^n$

- d** i \$12 528
 ii \$10 719
 iii \$6211
- 15 a** \$27 000
- | | | | | |
|----------|--------|--------|--------|------------|
| | 0 | 1 | 2 | 3 |
| b | 12 000 | 27 000 | 60 750 | 136 687.50 |
- i $V_n = 12\,000(2.25)^n$
- c** i \$307 547 (to the nearest dollar)
 ii \$1 556 956 (to the nearest dollar)
 iii \$39 903 081 (to the nearest dollar)
- 16 a** 3754 (to the nearest share)
b $B = 30\,000 \times 0.965^n$
- 17 a** $S = 570\,000 \times 1.0615^n$
b $N = 651\,765.23 \times 1.0494^n$
- 18 a** Yes. Next value calculated by multiplying by 0.75.
 i 113.91 cm ii $H = 360 \times 0.75^n$
 iii 4.81 cm
- 19 C** **20 B** **21 C**

Chapter 4 review

Multiple-choice questions

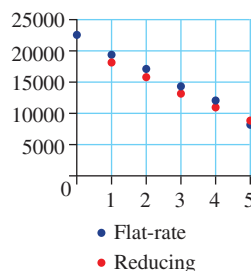
- 1 C** **2 A** **3 B** **4 B**
5 C **6 A** **7 D** **8 D**
9 C **10 C** **11 D** **12 B**
13 D **14 A** **15 C** **16 C**
17 B **18 B** **19 A** **20 D**

Short-response questions

- 1 a** 46 **b** 56
2 a 320 **b** 7.8125
3 72
4 a 6, 12, 18 **b** 1000, 500, 250
5 a $t_n = 13 - 5n$ **b** $t_n = 3^{n-1}$
6 a $t_n = 120 \times 0.25^{n-1}$
b 7.5
c Approximately 0.000002 (1.788×10^{-6})
7 a $A = 38\,500 \times 0.905^n$
b \$23 372.42
c \$15 127.58

- 8 a** $A = 20\,000 + 1880n$ **b** \$29 400
c 11 years
9 a 20 cents **b** $A = 1650 - 0.2n$
c \$1250

10 $A = 22\,500 - 2700n$; $B = 22\,500 \times 0.84^n$



- 11** A: \$6575; B: \$6777.89; 7.1%
- 12 a** See table * at bottom of page
b 162.5 kg
c i 0.75
 ii See table ** at bottom of page
d Day 8
e 3350.67kg
f 20.1%

Chapter 5

Section 5A

- 1 a** $32^\circ 27'$ **b** $43^\circ 12'$ **c** $122^\circ 27.6'$
d $91^\circ 7.2'$ **e** 45°
- 2 a** 32.75° **b** 15.58° **c** 7.37°
d 142.73° **e** 67.25°
- 3 a** 50.27 cm **b** 87.96 m **c** 282.74 mm
d 37.70 mm **e** 43.98 m
- 4 a** $\frac{1}{4}$ **b** $\frac{3}{4}$ **c** $\frac{1}{12}$
d $\frac{1}{3}$ **e** $\frac{1}{6}$ **f** $\frac{5}{12}$
- 5 a** 7.85 cm **b** 10.47 cm **c** 26.18 cm
d 23.56 cm **e** 36.65 cm **f** 57.60 cm
- 6 a** 13.09 cm **b** 5.24 cm **c** 78.54 cm
d 37.70 cm **e** 122.17 cm **f** 109.96 cm
- 7** 45.81 cm
- 8 a** 95.5° **b** 47.75° **9** 6.20 cm

* Day	1	2	3	4	5	6	7	8	9	10
Emission each day	1500	1370	1240	1110	980	850	720	590	460	330

** Day	1	2	3	4	5	6	7	8	9	10
Emission	1500.00	1125.00	843.75	632.81	474.61	355.96	266.97	200.23	200.00	200.00

Section 5B

- 1 a** Bucharest, Romania
b Oslo, Norway **c** Houston, USA
d Singapore
e Cape Town, South Africa
- 2 a** 30.04°N, 31.24°E **b** 19.43°N, 99.13°W
c 34.60°S, 58.38°W **d** 19.08°N, 72.88°E
e 6.52°N, 3.38°E **f** 17.83°S, 31.05°E
- 3 a** 28.04°S, 148.58°E **b** 26.80°S, 153.13°E
c 21.14°S, 149.19°E **d** 12.65°S, 141.85°E
e 24.87°S, 152.35°E **f** 20.08°S, 146.26°E
g 138°E
- 4 a** (65°N, 0°) **b** (0°, 75°E)
- 5 a** (18°N, 24°E) **b** (38°N, 44°E)
c (17°S, 23°E) **d** (43°N, 37°W)
- 6 a i** Mexico City, Yangon
ii Brisbane
b i Brisbane
ii Plymouth
iii London
iv Zurich
c Marseilles, Mexico City, Plymouth, Zurich, London, Yangon
d Plymouth, Mexico City, Lima, London (just)
e Mexico City, Yangon
f Cooktown, Melbourne
g London
h Wellington
- 7** 6338 km
8 2313 km
9 1779 km
10 34.27°S
- 11 a** 3892 km **b** 3002 km **c** 4893 km
d 4448 km **e** 3336 km
- 12** 3558 km
13 4003 km
- 14 a** 5560 km **b** 7784 km **c** 3336 km
d 6672 km **e** 16 902 km
- 15** 5.4°
- 16 a** 3336 km **b** 6672 km **c** 13 344 km
17 a 4226 km **b** 5782 km **c** 14 234 km
18 a 4176 km **b** 14 184 km **c** 5832 km
19 a Equator 4670 km; North Pole 5338 km; South Pole 14 678 km
b Equator 6116 km; North Pole 3892 km; South Pole 16 124 km
c Equator 1668 km; North Pole 11 676 km; South Pole 8340 km

d Equator 1557 km; North Pole 11 565 km; South Pole 8451 km

- 20 a** 6154 km **b** 5517 km
c 4505 km **d** 3186 km
- 21** 60°N, 60°S
- 22 a** 1134 km **b** 2829 km **c** 5439 km
- 23** 10 045 km
- 24 a** 1371 km **b** 6783 km
c 4689 km **d** 2391 km
- 25 a** 9129 km **b** 12 485 km
c 13 232 km **d** 11 078 km
- 26 a** 445 km **b** 645 km **c** 1090 km
- 27 a** 2335 km **b** 2340 km **c** 4675 km
- 28** Answers will vary
- 29** A **30** B **31** A

Section 5C

- 1 a** X is 2 hours ahead of Y
b X is 5 hours ahead of Y
c Y is 5 hours ahead of X
d Y is 1 hour 48 minutes ahead of X
- 2 a i** 1 hour 28 minutes behind
ii 9 hours 52 minutes ahead
iii 3 hours 48 minutes ahead
iv 10 hours 12 minutes ahead
v 16 hours 48 minutes ahead
- b i** 7:28 a.m.
ii 8:08 p.m. the day before
iii 2:12 a.m.
iv 7:48 p.m. the day before
v 1:12 p.m. the day before
- 3 a** 150°W **b** 2 p.m.
c 7:30 p.m. the day before
- 4 a** 5:30 p.m. **b** 7:30 a.m.
- 5 a** 45° **b** 3 hours
c 2:45 p.m.
- 6 a** 3:26 a.m. Sunday
b 5:34 p.m. Saturday
c 6:30 p.m. Saturday
d 3:22 a.m. Sunday
e 10:26 a.m. Saturday
f 7:22 a.m. Saturday
- 7 a** 3 a.m. the same day
b 3 p.m. the next day
- 8 a** 3 a.m. the same day
b 5 a.m. the next day
- 9 a** 4 p.m. Monday
b 14 hours
c 13 hours 55 minutes

- 10 a** 12 midday same day
b 11:30 a.m. same day
c 10 a.m. same day
d 12 midday same day
e 2 p.m. same day
f 1 p.m. same day
g 2 p.m. same day
h 4 p.m. the day before
i 1 p.m. same day

11 D **12** C **13** B

Chapter 5 review

Multiple-choice questions

- 1** C **2** B **3** A **4** D
5 D **6** B **7** C

Short-response questions

- 1** 3336 km
2 3892 km
3 a 52°
b 3 hours 28 minutes
c 6:13 a.m.
4 (0°, 160°W)
5 19 905 km
6 3 hours 8 minutes
7 8 a.m.
8 5 a.m.
9 7 p.m.
10 6 a.m. Saturday
11 a 4 p.m. Tuesday
b 2 a.m. Thursday
12 a 15 hours 28 minutes
b 11:28 a.m. Tuesday
c 8:32 a.m. Thursday
13 a 5543 km
b 1451 km
14 a 5226.4 km
b 5842.58 km
c 16 622.26 km

Chapter 6

Section 6A

Topic 1: Bivariate data analysis 1

Multiple-choice questions

- 1** B **2** A **3** D **4** C **5** A
6 D **7** B **8** D **9** D **10** B
11 D **12** D **13** B **14** A **15** C

Short-response questions

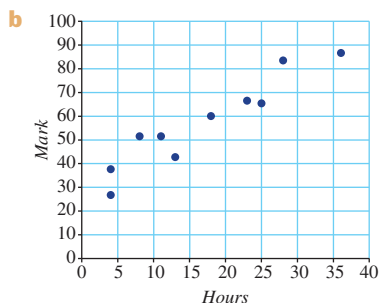
- 1 a** Numerical & numerical, scatterplot
b Categorical & numerical, parallel boxplots
c Categorical & categorical, two-way frequency table

2 a age, distance

b

Study mode	Gender	
	Female	Male
On campus	3	3
Online	4	2
Total	7	5

3 a EV: hours, RV: mark.



- c** There is a strong, positive linear relationship between hours of study and mark. Those who studied more hours obtained higher marks in the examination.
- 4 a** 0.966. There is a strong, positive linear relationship between male and female life expectancies. Those countries that have high male life expectancies also tend to have high female life expectancies.
- b** The variables are both numerical and the relationship is linear.
- 5 a** EV: time of observation, RV: activity
b 38%
c There is an association between the dolphin's activity and time of day. Only 6.7% of dolphins were observed feeding in the afternoon, less than the percentage observed feeding in the morning 38%, and very much less than the percentage observed feeding in the evening 70.9%.
- 6 a** Birth rate 76.7%, births attended by skilled staff 51.0%, exclusive breastfeeding 2.9%, health expenditure 24.5%, literacy rate female 64.0%, literacy rate male 72.1%, safe sanitation 40.4%, safe drinking water 66.7%.
b Of the variables listed, only exclusive breastfeeding does not appear to be related to infant mortality. The variation in infant mortality is strongly related to birth rate,

and parent's educational levels (as indicated by literacy rate, particularly their fathers). Also important are access to clean water, skilled health staff, and safe sanitation (in that order). Expenditure on health is also indicated to be able explain a reasonable percentage of the variation in mortality rate.

- 7** There is an association between preference for type of instruction and faculty. A higher level of preference for on campus study is seen in Arts and Science students (63.8% and 64% respectively), reducing to 41.4% for Business students.
- 8** The value of the correlation coefficient would be closer to -1 . The value of the slope would become closer to -2 .
- 9** There is a relationship between political affiliation and attitude to a republic for males, with 28.2% of those who identify as Liberal in favour of retaining the King, compared to only 8.1% of those who identify as Labor wanting to retain the King. There is a relationship between political affiliation and attitude to a republic for females, with 37.7% of those who identify as Liberal in favour of retaining the King, compared to only 16.3% of those who identify as Labor wanting to retain the King. For both males and females, the relationship between support for a republic and political affiliation is the same, with both groups showing a stronger preference for retaining the King in Liberal supporters than in Labor supporters.

Section 6B

Topic 2: Bivariate data analysis 2

Multiple-choice questions

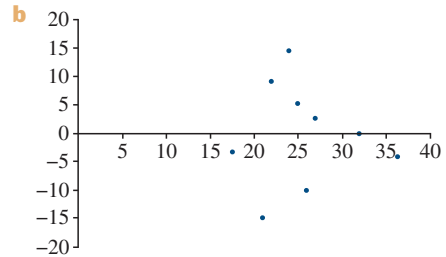
- 1** C **2** C **3** A **4** D
5 B **6** A **7** B **8** D
9 B **10** C **11** A **12** D

Short-response questions

- 1 a** EV: mother height, RV: daughter height
b 0.612, moderate
c $\text{daughter height} = 0.715 \times \text{mother height} + 50.23$
 Intercept: mother's height of 0 cm predicts a daughter's height of 50.23 cm, which is meaningless. Slope: An increase of 1 cm in mother's height predicts an increase of 0.715 cm in daughter's height.

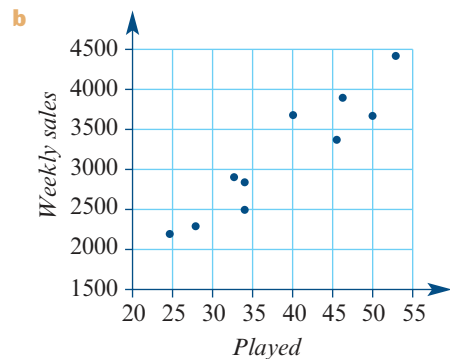
- d** $R^2 = 37.5\%$. 37.5% of the variation in daughter's height can be explained by the variation in mother's height.
e 171.8 cm
f Interpolation, as 170 cm is within the range of the data used to determine the equation of the regression line.

- 2 a** $-15.5, 2.7$



The residual plot shows only random variation, so the linearity assumption holds.

- 3 a** EV: played, RV: weekly sales



- c** $r = 0.9458$
d Strong, positive, linear relationship
e $\text{weekly sales} = 74.3 \times \text{played} + 293$
f Slope: on average, the number of downloads increases by 74.3 for each additional time the song is played on the radio in the previous week. Intercept: predicts 293 downloads of the song if it is not played on radio in the previous week.
g 7723
h Extrapolating so may be unreliable.
4 a $\text{cost} = 5.97 \times \text{number of meals} + 165.63$
b Base cost = \$165.63
c Additional cost per meal = \$5.97
5 A study was conducted to investigate the relationship between femur length and radius length. Data were collected from a sample of 10 people.
 From the scatterplot of radius versus femur, we can see that there is a strong, positive, linear relationship between femur length and

radius length: $r = 0.988$. That is, those people with a longer femur also tended to have a longer radius. There are no obvious outliers, and the linearity assumption is confirmed by the residual plot. The equation of the least squares regression line is:

$$\text{radius} = 0.739 \times \text{femur} - 7.25$$

The slope of the regression line predicts an increase of 0.739 cm in radius for each 1 cm increase in femur.

The coefficient of determination tells us that 97.5% of the variation in radius is explained by the variation in femur.

- 6 a** On average, height increases by 0.815 cm for each additional 1 cm increase in arm span.

b i Females: $R^2 = 64.6\%$

ii Males: $R^2 = 69.9\%$

iii Since the value of the coefficient of determination for males (69.9%) is higher than the value for females (64.6%), then we can say that arm span is a better predictor of height for males than for females.

- c i** The models predict that when both have arm span measurements of 160 cm, a male will be 1.8 cm taller than a female.

ii The models predict that when both arm span measurements of 190 cm, a female will be 4.6 cm taller than a male.

iii The differences predicted is not reliable for a height of 160 cm as this is value is outside the range of height data for males. The prediction is not reliable for a height of 190 cm as this value is outside the range of height data for females.

- 7 a** There is a moderate strength, non-linear association between *expenditure* and *score*.

b i The linearity assumption.

ii No, there is a clear structure in the residual plot. If the linearity assumption had been met the residuals would have been randomly scattered around a horizontal line at $y = 0$.

- 8 a** \$22.90 per pair of jeans

b \$147.90

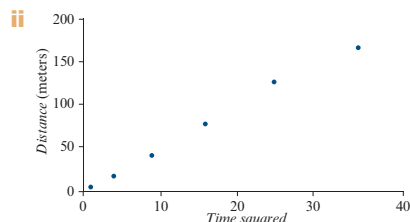
c i \$258.83

ii \$83.83

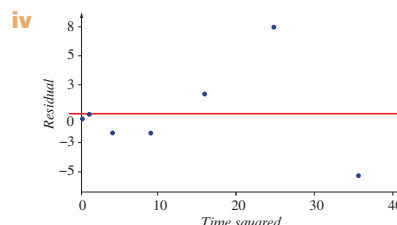
- 9 a** The residual plot shows clear curved structure indicating that the relationship is not linear.

b i

Time (seconds)	(Time) ²	Distance (metres)
0	0	0
1	1	5.2
2	4	18
3	9	42
4	16	79
5	25	128
6	36	168



iii $\text{distance} = 4.803 \times \text{time}^2 + 0.45$



The residual plot shows less structure indicating that the linearity assumption is better met.

v 235.8 metres

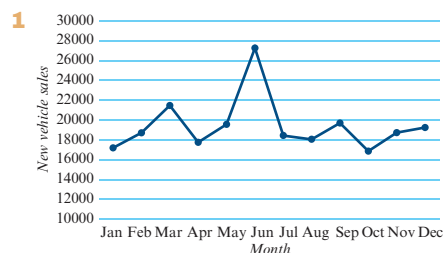
Section 6C

Topic 3: Time series analysis

Multiple-choice questions

1 B	2 A	3 A	4 B
5 C	6 A	7 A	8 D
9 C	10 D	11 D	12 B
13 B	14 B	15 D	16 A

Short-response questions



New vehicle sales in Queensland in this year varied between about 16 500 and 19 500 in most months, with a small increase in sales in March, and a much larger increase in sales in June (27 270).

2 a Summer

b The time series plot shows that the number of visitors is seasonal, with the most visitors in Summer, and the least in Winter. The plot also shows a decreasing trend.

3 a 0.6134 b 0.6144

4 a 0.5

b Overnight there are only 50% of the average number of cars in the car park for each time period.

5

Quarter	Q1	Q2	Q3	Q4
SI	1.13	1.06	0.49	1.33

6 a \$4 600 b \$5 080

c i 0.8 ii \$3 500

7 a

Quarter	Sum	Aut	Win	Spr
SI	1.04	1.15	1.30	0.52

b

	Sum	Aut	Win	Spr
2022	1327	1415	1415	1385
2023	1492	1539	1582	1394
2024	1874	1727	1654	1904

c number of purchases =
 $45.0 \times \text{quarter} + 1266.4$

d Gradient = 45.0. On average, the number of purchases increases by about 45 each quarter.

h Seasonalised predicted purchases for summer 2026 = 2212.7

8 a i \$91 000 ii \$99 000

b i On average, the price of bitcoin increases by \$1536.04 each month.
ii \$137 566

c i On average, the price of bitcoin increases by \$4776.66 each month.
ii \$163 158

iii The model based on the 2024 data would be preferred, as it more recent data, and extrapolation over a shorter period of time. The change in the rate of increase of bitcoin during 2024 is also confirmed by the time series plot.

9 a The average weekly earnings for both males and females have increased steadily, and at similar rates, between 2013 and 2023.

b Since the rate of increase for females is \$24.08 each 6 months, higher than the rate of increase for males of \$21.75 each six months, then we expect the difference

between the average weekly earnings males and females to decrease over time. If the current trends continue, the trend lines would eventually cross, meaning the average weekly earnings for females would be higher than the average weekly earnings for males. This is predicted to first occur in November, 2071.

7 Answers will vary.

8 Answers will vary.

Section 6D

Topic 4: Growth and decay in sequences

Multiple-choice questions

- 1 C 2 D 3 C 4 C
5 D 6 D 7 A 8 C
9 D 10 D

Short-response questions

1 a 51 b 310 c 4096 d 162

2 a $t_1 = 5, t_2 = 13, t_3 = 21$

b $t_1 = 2, t_2 = 10, t_3 = 50$

3 a $t_n = 815 - 15n$ b $t_n = 2 \times 6^{n-1}$

4 \$249 445

5 a -15 b 20

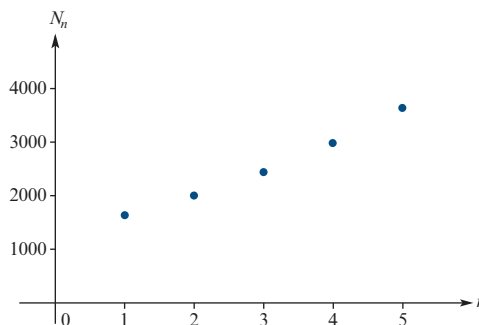
6 a $V_n = 48\,000 \times 0.92^n$ b \$29 105
c \$16 364

7 a $V_n = 25\,550 + 4080n$ b \$37 740
c 7 years

8 a $C_n = 5000 + 0.02n$ b \$7400 c 250 000

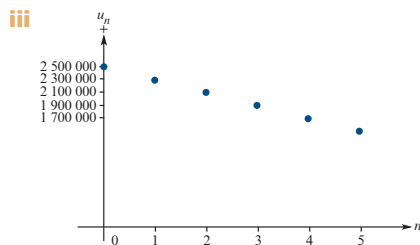
9 $N_{n+1} = 1.22N_n, N_1 = 1654$

Start of year	1	2	3	4	5
Number of deer	1654	2018	2462	3004	3665

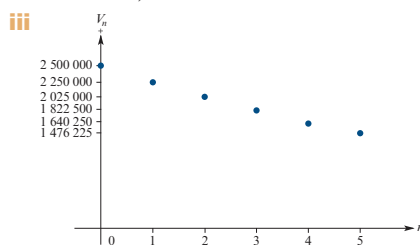


The population will not exceed 5000 in the first five years. Realistic in the short term but in the long term, there may be a lack of food and shelter to support a large and growing population.

- 10 a i** $V_n = 2\,500\,000 - 200\,000n$
ii 2 300 000, 2 100 000, 1 900 000,
 1 700 000, 1 500 000



- b i** $U_n = 2\,500\,000 (0.9)^n$
ii 2 250 000, 2 025 000, 1 822 500,
 1 640 250, 1 476 225



- c i** Method 2 **ii** Method 1

Section 6E

Topic 5: Earth geometry and time zones

Multiple-choice questions

- 1** B **2** B **3** D **4** D
5 C **6** A **7** B **8** C

Short-response questions

- 1** 4114 km
2 15 038 km
3 a 264°
b 17 hours 36 minutes
c 11:36 a.m. the next day
4 a 5727 km
b 4281 km
c 15 735 km
5 140°E, 22°N
6 4838 km
7 a 35 670 km
b 25 190 km
8 a 11 120 km
b 15 330 km
c 4210 km
9 0°, 26.5°W
10 a 556 km
b 481.5 km
c 26°N or 26°S

- 11 a** 7691 km
b 7761 km
c 70 km
d 3 hours 40 minutes. A is ahead of B

Chapter 7

Section 7A

- 1** $A_0 = 5000$, $A_{n+1} = 1.084 \times A_n$
2 $A_0 = 8500$, $A_{n+1} = 1.042 \times A_n$
3 a \$2050, \$2101.25, \$2153.78
b \$153.78 **c** \$52.53 **d** 10 years
4 a \$26 600, \$28 302.40, \$30 113.75
b \$5113.75 **c** \$1811.35 **d** 8 years
5 a 0.006 **b** 0.0093 **c** 0.0031
d 0.0052 **e** 0.019 **f** 0.0261
6 a 3.75% **b** 6.4% **c** 9.6%
d 11.28% **e** 6.12% **f** 10.92%
7 $J_0 = 2000$, $J_{n+1} = 1.007 \times J_n$
8 $R_0 = 34\,000$, $R_{n+1} = 1.0018 \times R_n$

9 a

Principal: \$12 000	
Annual interest rate: 5.2%	
Compounds per year	Balance after 1 year
1	\$12 624.00
2	\$12 632.11
4	\$12 636.27
12	\$12 639.09
52	\$12 640.18

b

Principal: \$25 000	
Annual interest rate: 9.4%	
Compounds per year	Balance after 1 year
1	\$27 350.00
2	\$27 405.23
4	\$27 434.14
12	\$27 453.94
52	\$27 461.66

- 10 a** 1 month: \$6028.80, 2 months: \$6057.74, 3 months: \$6086.82
b \$86.82
c \$28.94
d 7 months
11 a 1 year: \$2076, 2 years: \$2152, 3 years: \$2228
b 14 years

- 12 a** 1 year: \$7518, 2 years: \$8036,
3 years: \$8554 **b** 6 years
- 13 a** 1 month: \$8488.20, 2 months: \$8577.33, 3
months: \$8667.39
b \$88.20
c \$89.13
d \$267.39
e 11 months
- 14 a** 0.5%
b $A_0 = 7600$, $A_{n+1} = 1.005A_n$
c \$7830.87
- 15 a** 2% **b** $A_0 = 3500$, $A_{n+1} = 1.02A_n$
c \$3788.51 **d** \$312.30
- 16 a** Weekly **b** \$1.46
- 17 a** Quarterly
b \$9.13
- 18 C** **19 D** **20 C**

Section 7B

- 1 a** $H_n = 2500 \times 1.035^n$
b \$2771.79
- 2 a** $G_n = 15\,000 \times 1.028^n$
b \$16 751.89
- 3** \$6377.17
- 4** \$32 709.21
- 5** \$9785.98
- 6** \$16 563.11
- 7 a** $B_n = 10\,000 \times 1.052^n$
b \$6601.88
- 8** \$6491.04
- 9** \$816.24
- 10** \$1252.06
- 11** First year: \$372.12, Second year: \$399.82,
Third year: \$429.57, Fourth year: \$461.54,
Fifth year: \$495.89
- 12** Bank A results in the least amount of interest
(\$3015.27) over the life of the loan.
- 13** At least 4 years
- 14 D** **15 D** **16 B**

Section 7C

- 1 a** 5.54% **b** 8.76% **c** 4.91%
d 13.10% **e** 7.64%
- 2 a** 7.02% **b** 5.76% **c** 11.84%
d 4.58% **e** 2.31%
- 3 a** More compounds earn more interest
b 4.68%
c 4.70%
d More frequent compounds (monthly) has a
higher effective interest rate.

- 4 a** Fewer compounds charge less interest
b 8.25%
c 8.24%
d Less frequent compounds (monthly) has a
lower effective interest rate.
- 5 a** A: 8.62% B: 8.11%
b A: \$3018.10 B: \$2837.08
c Luke should choose loan B as he will pay
less interest than with loan A.
- 6 a** A: 5.43% B: 5.61%
b A: \$7602.92 B: \$7860.27
c Sharon should choose investment B as
she will earn more interest than with
investment A.
- 7 B** **8 D** **9 D**

Section 7D

- 1** \$8500 **2** \$24 000 **3** \$35 000 **4** \$6400
5 \$31 000 **6** 4.14% **7** 6.12% **8** 6.12%
9 5.28% **10** 4.92% **11** 4.90% **12** 6.00%
13 6 years **14** 50 years **15** 48 quarters
16 4.00% **17** 6.25%
18 D **19 A** **20 C** **21 A**

Section 7E

- 1** $A_0 = 6500$, $A_{n+1} = 1.14A_n - 1800$
- 2 a** $A_0 = 14\,000$, $A_{n+1} = 1.028A_n - 2000$
b \$12 392, \$10 738.98, \$9039.67
- 3 a** $A_0 = 22\,000$, $A_{n+1} = 1.018A_n - 1000$
b \$20 155.19
- 4 a** $A_0 = 85\,000$, $A_{n+1} = 1.0067A_n - 1800$
b \$81 283.71
- 5** $A_0 = 150\,000$, $A_{n+1} = 1.0054A_n - 1700$;
\$147 315.56
- 6** $A_0 = 245\,000$, $A_{n+1} = 1.0016A_n - 1200$;
\$242 572.12
- 7 a** \$600 **b** \$1201.44 **c** \$501.44
8 a \$860 **b** \$868.19 **c** \$168.19
9 a 8.64% **b** \$10 282.98
c \$282.98
- 10 a** 11.4% **b** \$4717.02 **c** \$217.02
- 11 a** $A_0 = 20\,000$, $A_{n+1} = 1.0059A_n - 600$
i \$17 561.39 **ii** \$561.39
b $A_0 = 20\,000$, $A_{n+1} = 1.0059A_n - 800$
i \$16 549.52 **ii** \$549.52
c \$11.87
- 12** Balance A: \$47 132.08,
Balance B: \$47 116.39, Interest A: \$3132.08,
Interest B: \$3116.39
- 13** \$18.09
- 14 B** **15 D** **16 C**

Section 7F

1 a

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	8400.00
1	500	79.80	420.20	7979.80
2	500	75.81	424.19	7555.61
3	500	71.78	428.22	7127.39
4	500	67.71	432.29	6695.10
5	500	63.60	436.40	6258.70

b \$358.70

2 a

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	2000.00
1	100.00	7.00	93.00	1907.00
2	100.00	6.67	93.33	1813.67
3	100.00	6.35	93.65	1720.02
4	100.00	6.02	93.98	1626.04
5	100.00	5.69	94.31	1531.73

b i \$20.02
ii \$26.04

3

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	4600.00
1	300.00	18.40	281.60	4318.40
2	300.00	17.27	282.73	4035.67
3	300.00	16.14	283.86	3751.81
4	450.00	15.01	434.99	3316.82
5	450.00	13.27	436.73	2880.09

Balance after five repayments = \$2880.09

4

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	7500.00
1	\$300.00	71.25	228.75	7271.25
2	\$300.00	69.08	230.92	7040.33
3	\$600.00	66.88	533.12	6507.21
4	\$600.00	61.82	538.18	5969.03
5	\$600.00	56.71	543.29	5425.74

Balance after five repayments = \$5425.74

5

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	6000.00
1	\$6000.00	78.00	500	5578.00
2	\$5578.00	72.51	500	5150.51
3	\$5150.51	66.96	700	4517.47
4	\$4517.47	58.73	700	3876.20

Balance interest = \$276.20

6 a i \$14 680.68

ii Too much would be paid with the final payment at the end of the 11th month. It must be less than normal at \$2086.82

iii \$516.82

b i \$1194.53

ii \$494.53

iii \$22.29

c Better to be made as repayment number 1. It would save the most interest because the payment was made earlier in the term of the loan.

7 D **8 B**

Section 7G

- 1** \$36 800.44 **2** \$156 864.12
3 \$322 456.27 **4** \$78 723.82.40
5 \$107 124.54 **6** \$186 933.16
7 \$310 413.73 **8** \$161 810.14
9 \$164 173.43 **10** \$269 343.29
11 a \$174 447.18 **b** \$151 095.40
c Loan B
12 a \$191 683.94 **b** \$170 839.20
c Loan B
13 Loan B
14 B **15 A**

Section 7H

- 1** \$59 700.50 **2** \$5373.23
3 \$17 947.68 **4** \$4635.06
5 \$15 877.22 **6** \$7718.16
7 \$1480.29
8 a \$2613.32 **b** \$470 397.60
c \$170 397.60
9 a \$530 338.60 **b** \$840 000
c \$309 661.40
10 a \$4772.95 **b** \$572 753.78
c \$122 753.78
11 a \$240 466.22 **b** \$316 800
c \$76 333.78

- 12** \$345 004.72 **13** \$443 236.28
14 \$320 791.59 **15** \$121 279.20
16 \$533 746.99 **17** Option A
18 \$251 676.44
19 C **20** B **21** B **22** C

Chapter 7 review

Multiple-choice questions

- 1** B **2** C **3** D **4** C
5 A **6** C **7** A **8** D
9 A **10** D **11** C **12** B
13 C **14** B **15** A **16** D
17 B **18** C **19** B **20** C

Short-response questions

- 1 a** $A_0 = 12\,500$, $A_{n+1} = 1.0065A_n$
b \$12 828.18
2 a 1 month: \$5832.48
 2 months: \$5865.14
 3 months: \$5897.99
b 7 months
3 a \$1600
b 18.2%
c One week: \$1605.60, two weeks: \$1611.22,
 three weeks: \$1616.86
d 9 weeks
4 a 3.8%
b $A_0 = 2200$, $A_{n+1} = 1.038A_n$
c \$2553.95
5 a \$1528.33
b \$28.33
6 a \$9500 **b** \$250.00 **c** \$8188.07
7 a 0.49%
b $A_0 = 250\,000$, $A_{n+1} = 1.0049A_n - 2400$
c \$242 863.07 **d** 9
8 a \$16 411.74 **b** \$1411.74
9 a \$3500
b \$600.00
c i 0.92% **ii** 11.04%
d \$613.87

10 a

Repayment number	Repayment amount	Interest paid	Principal reduction	Balance of loan
0	0	0	0	125 000.00
1	1000.00	387.50	612.50	124 387.50
2	1000.00	385.60	614.40	123 773.10
3	1000.00	383.70	616.30	123 156.80
4	1000.00	381.79	618.21	122 538.59

b \$109 763.84

- 11 a** \$2323.06 **b** \$83 630.16
c \$8630.16

12 18.6%

13 3.95%

14 a Bank A: 6.45% per annum Bank B: 6.27% per annum

b Bank A. It has the largest effective rate of interest and will earn more interest in one year than Bank B.

c \$9

15 a Bank A. It has the largest effective rate of interest (8.37% per annum compared to 8.35% per annum) and will earn more interest in the time of the investment.

b \$7

16 a i \$15 701.99

ii \$15 704.73

iii \$15 705.80

b Effective annual rate of interest

c The weekly compounding investment has the highest effective annual rate of interest. It will earn more interest than the other compounding frequencies and so will have the most benefit to Eva. This can be seen by the fact that her account will be the highest after one year if she has weekly compounding.

17 a \$2300.00

b 17.68%

c i \$2428.36

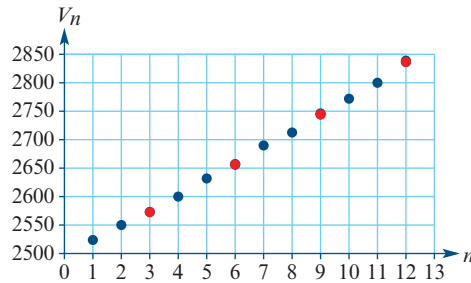
ii \$128.36

iii \$443.98

18 a

Month	Quarterly compounds	Monthly compounds
1		\$2526.00
2		\$2552.27
3	\$2578.00	\$2578.81
4		\$2605.63
5		\$2632.73
6	\$2658.43	\$2660.11
7		\$2687.78
8		\$2715.73
9	\$2741.38	\$2743.97
10		\$2772.51
11		\$2801.35
12	\$2826.91	\$2830.48

b i



ii The table shows that Lucille will pay more with monthly compounds compared to quarterly compounds. The graph of monthly compounds is higher on the axes than that for quarterly compounds.

c \$3.57

Chapter 8

Section 8A

1 a i $A_0 = 0$, $A_{n+1} = 1.025 \times A_n + 5000$
ii \$15 378.13

b i $A_0 = 0$, $A_{n+1} = 1.016 \times A_n + 6500$
ii \$19 813.66

c i $A_0 = 320\,000$, $A_{n+1} = 1.009 \times A_n + 8000$
ii \$352 934.64

d i $A_0 = 460\,000$,
 $A_{n+1} = 1.0058 \times A_n + 4200$
ii \$480 723.73

e i $A_0 = 845\,000$,
 $A_{n+1} = 1.0041 \times A_n + 7500$
ii \$878 028.55

f i $A_0 = 1\,250\,000$,
 $A_{n+1} = 1.0012 \times A_n + 2700$
ii \$1 262 615.13

2 a i $A_0 = 120\,500$, $A_{n+1} = 1.028 \times A_n - 8000$
ii \$106 229.79

b i $A_0 = 276\,000$,
 $A_{n+1} = 1.0126 \times A_n - 4600$
ii \$272 590.20

c i $A_0 = 358\,000$,
 $A_{n+1} = 1.0143 \times A_n - 25\,000$
ii \$297 501.26

d i $A_0 = 440\,000$,
 $A_{n+1} = 1.0036 \times A_n - 5000$
ii \$429 715.06

e i $A_0 = 845\,000$,
 $A_{n+1} = 1.0067 \times A_n - 9600$
ii \$833 105.16

f i $A_0 = 1\,360\,000$,
 $A_{n+1} = 1.0015 \times A_n - 2900$
ii \$1 357 416.13

3 a The recurrence relation has a value subtracted at the end.

b \$1030.00 **c** \$2030.50
d \$120.50

4 a The recurrence relation has a positive value added at the end. **b** \$300
c \$7666.58 **d** \$166.58

5 a i 2.16% **ii** \$62 942.44
b i \$7500 **ii** \$62 942.44
iii \$25 875.72
c i \$442.44 **ii** \$433.28 **iii** \$875.72

6 a i 3.72% **ii** \$269 729.13
b i \$1800 **ii** \$269 729.13
iii \$266 828.64
c i \$2479.13 **ii** \$2499.51
iii \$4978.64

7 a $A_0 = 125\,000$, $A_{n+1} = 1.0013 \times A_n + 695$
b 6 **c** \$1502.75

8 a $A_0 = 32\,000$, $A_{n+1} = 1.0036 \times A_n - 3500$
b 4 **c** 9 **d** \$1094.43
e \$1098.37

9 a $A_0 = 54\,000$, $A_{n+1} = 1.0064 \times A_n + 1500$
b \$65 252.30
c $A_0 = 65\,252.30$, $A_{n+1} = 1.0063 \times A_n - 1800$
d \$4586.48

10 B **11 C**

Section 8B

1 a \$164 000.00
b 5.8%
c \$155 822.47
d \$747.95
e \$2765.26
f i 3805.16 **ii** \$3805.16

g	6	3500.00	721.46	2778.54	147 526.62
	7	3500.00	708.13	2791.87	144 734.75

2

Payment number	Payment withdrawn	Interest	Principal reduction	Balance of annuity
0	0	0	0.00	50000.00
1	4000.00	400.00	4400.00	54 400.00
2	4000.00	435.20	4435.20	58 835.20
3	4000.00	470.68	4470.68	63 305.88
4	4000.00	506.45	4506.45	67 812.33
5	4000.00	542.50	4542.50	72 354.83

3

Payment number	Payment withdrawn	Interest	Principal reduction	Balance of annuity
0	0	0	0.00	135 000.00
1	1200.00	486.00	1686.00	136 686.00
2	1200.00	492.07	1692.07	138 378.07
3	1200.00	498.16	1698.16	140 076.23
4	1200.00	504.27	1704.27	141 780.50
5	1200.00	510.41	1710.41	143 490.91

4

Payment number	Payment withdrawn	Interest	Principal reduction	Balance of annuity
0	0	0	0	25 000.00
1	1000.00	143.75	856.25	24 143.75
2	1000.00	138.83	861.17	23 282.58
3	1000.00	133.87	866.13	22 416.45
4	1000.00	128.89	871.11	21 545.34
5	1000.00	123.89	876.11	20 669.23

5

Payment number	Payment withdrawn	Interest	Principal reduction	Balance of annuity
0	0	0	0	380 000.00
1	12 000.00	4560.00	7440.00	372 560.00
2	12 000.00	4470.72	7529.28	365 030.72
3	12 000.00	4380.37	7619.63	357 411.09
4	12 000.00	4288.93	7711.07	349 700.02
5	12 000.00	4196.40	7803.60	341 896.42

6

Payment number	Deposit amount	Interest	Principal increase	Balance of annuity
0	0	0	0	10 000.00
1	1000.00	510.00	1510.00	11 510.00
2	1000.00	587.01	1587.01	13 097.01

Payment number	Payment withdrawn	Interest	Principal reduction	Balance of annuity
1	1000.00	667.95	332.05	12 764.96
2	1000.00	651.01	348.99	12 415.97

7

Payment number	Deposit amount	Interest	Principal increase	Balance of annuity
0	0	0	0	30 000.00
1	1500.00	360.00	1860.00	31 860.00
2	1500.00	382.32	1882.32	33 742.32
3	1500.00	404.91	1904.91	35 647.23

Payment number	Payment withdrawn	Interest	Principal reduction	Balance of annuity
1	1500.00	427.77	1072.23	34 575.00
2	1500.00	414.90	1085.10	33 489.90
3	1500.00	401.88	1098.12	32 391.78

8 D **9** C

Section 8C

- 1** **a** $d = 1000, n = 5, i = 0.05$
b \$5525.62
- 2** \$6477.82
- 3** \$101 846.82
- 4** \$11 058.58
- 5** **a** $P = 5000, d = 1000, n = 5, i = 0.05$
b \$11 907.04
- 6** \$29 691.25
- 7** \$130 065.76
- 8** **a** $P = 10\ 000, d = 1000, n = 5, i = 0.05$
b \$7237.18
- 9** \$144 226.52
- 10** \$327 661.26
- 11** \$435 539.53
- 12** D **13** B

Section 8D

- 1** \$3618.96 is positive so it will last.
- 2** **a** \$6115.35 **b** 40 payments
c \$6188.74 is positive so it will last.
- 3** 21
- 4** **a** 58 **b** \$430.43 **c** \$10 380.43
- 5** **a** 120
b **i** \$474.81 **ii** \$474.29
c \$16 976.68
- 6** **a** \$412 000.00 **b** \$65 662.50
- 7** **a** 5.27% **b** \$396 300.89
c 487 **d** 5
- 8** **a** 9.24% **b** 58 (57 full plus one smaller)
- 9** **a** 97 (96 full plus one smaller)
b \$4279.64
- 10** \$22 965.28
- 11** C **12** B

Section 8E

- 1** \$3075
- 2** \$25 480
- 3** \$4791.67
- 4** \$3350
- 5** **a** \$9790.50 **b** \$642 000
- 6** \$7000
- 7** \$10 080
- 8** \$11 880
- 9** Yes, it is a perpetuity because the payments are equal to the interest earned.
- 10** D **11** B **12** A

Section 8F

- 1 a** Verified $P = \frac{d}{i}$ **b** 3.125%
- $$= \frac{2500}{0.025} = 100\,000$$
- 2** \$720 000
3 \$1 100 000
4 \$200 000
5 a \$12 500 **b** 6.25%
6 6.3%
7 6.25%
8 3%
9 \$4200
10 \$991 866.88
11 A **12** B

Chapter 8 review

Multiple-choice questions

- 1** A **2** C **3** D **4** B
5 C **6** A **7** C **8** C
9 D **10** B **11** C **12** B
13 C **14** D **15** D **16** D
17 D **18** A **19** D **20** C

Short-response questions

- 1 a** \$624 000 **b** \$2500
c \$613 834.21
2 a $A_0 = 345\,000$, $A_{n+1} = 1.0115 \times A_n - 12\,000$
b \$295 397.96 **c** \$22 397.96
3 a 58 months **b** \$58 291.53 **c** \$1941.53
4 a \$71 318.97 **b** \$7208.47 **c** 20 quarters
5 \$474.81
6 a \$84 000.00
b \$14 500.00
c i 0.47% **ii** 5.64%
d \$12 888.23
7 \$16 528.65
8 a \$150 000 **b** 9%
9 a \$825
b \$147 835.30
c 39 months
d i 43 months **ii** \$453
10 a \$1 175 244.58
b 295 months plus one smaller payment
c \$3525
11 a \$5250
b \$441

- c i** \$5335.88 **ii** \$57 934.44
d \$5014.44

- 12 a** \$2460
b \$191.88
c i \$5083.55 **ii** \$27 519.87
d i \$7732.89 **ii** \$2649.34

- 13 a** \$720.00
b i $A_0 = 150\,000$,
 $A_{n+1} = 1.0048 \times A_n - 2000$
ii \$142 227.25 **iii** 16
c i 36 **ii** \$1850.89

- 14 a** \$496 000.00 **b** \$69 623.14
c 86 **d** \$2047.67

- 15 a** 89 **b** \$1554.55

- 16 a** 8% of \$1800 is more than 8.2% of \$1600 and so Byron will have more deposited into his superannuation account each fortnight with Option 1.

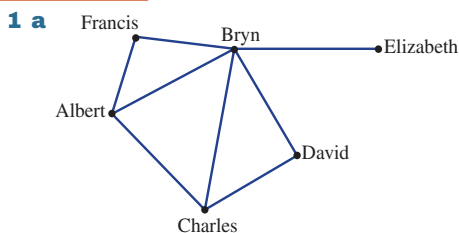
- b i** \$102 107.91 **ii** \$22 275.91

- 17** Quarterly payment = \$7029.52

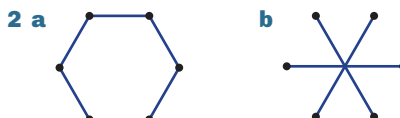
Total interest = \$107 475

Chapter 9

Section 9A

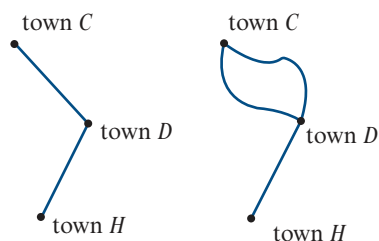


- b** 3
c i Elizabeth **ii** Bryn
d Albert, Bryn, Charles, Elizabeth
e 1



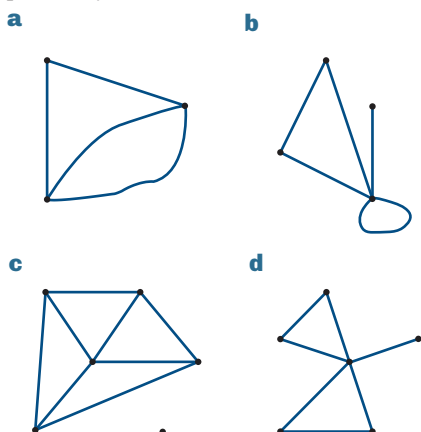
- 3 a i** 5 **ii** 6 **iii** 0 **iv** 2
v 3 **vi** 4 **vii** 1
b i 4 **ii** 7 **iii** 1 **iv** 6
v 2 **vi** 2 **vii** 2
4 a i 3 **ii** 2 **iii** 1
b i 7 **ii** 14 **iii** Verified as 14

c Two answers are possible, as shown below.

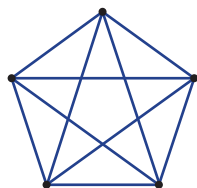


d H and D

5 Many different answers are possible. One possibility for each is shown.



6 a K_5



b 10

7 14

8 15

9 A

10 C

11 C

12 A

13 C

14 C

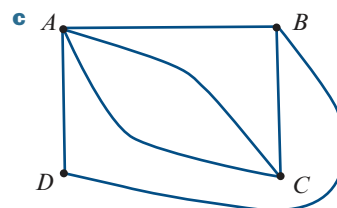
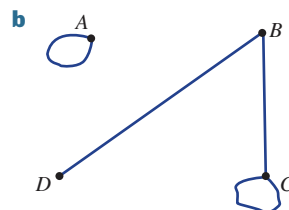
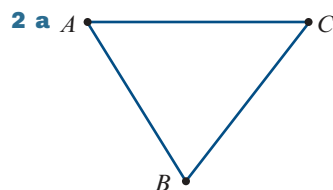
Section 9B

1 a	$\begin{matrix} A & B & C & D \\ A & \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix} \\ B & \begin{bmatrix} 1 & 0 & 1 & 1 \end{bmatrix} \\ C & \begin{bmatrix} 1 & 1 & 0 & 0 \end{bmatrix} \\ D & \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \end{matrix}$	b	$\begin{matrix} A & B & C & D \\ A & \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix} \\ B & \begin{bmatrix} 1 & 0 & 0 & 1 \end{bmatrix} \\ C & \begin{bmatrix} 1 & 0 & 0 & 1 \end{bmatrix} \\ D & \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix} \end{matrix}$
c	$\begin{matrix} A & B & C & D \\ A & \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \\ B & \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \\ C & \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \\ D & \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \end{matrix}$	d	$\begin{matrix} A & B & C & D \\ A & \begin{bmatrix} 0 & 1 & 1 & 1 \end{bmatrix} \\ B & \begin{bmatrix} 1 & 0 & 1 & 1 \end{bmatrix} \\ C & \begin{bmatrix} 1 & 1 & 0 & 1 \end{bmatrix} \\ D & \begin{bmatrix} 1 & 1 & 1 & 0 \end{bmatrix} \end{matrix}$

e

$$\begin{matrix} A & B & C & D & E & F \\ A & \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \\ B & \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \\ C & \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \\ D & \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix} \\ E & \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \\ F & \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{matrix}$$

f

$$\begin{matrix} A & B & C & D \\ A & \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix} \\ B & \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \\ C & \begin{bmatrix} 0 & 0 & 0 & 2 \end{bmatrix} \\ D & \begin{bmatrix} 0 & 1 & 2 & 0 \end{bmatrix} \end{matrix}$$


3 C is an isolated vertex.

4 Leading diagonals will all be '1'.

5

$$\begin{matrix} A & B & C & D & E \\ A & \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \end{bmatrix} \\ B & \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \end{bmatrix} \\ C & \begin{bmatrix} 1 & 1 & 0 & 1 & 1 \end{bmatrix} \\ D & \begin{bmatrix} 1 & 1 & 1 & 0 & 1 \end{bmatrix} \\ E & \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \end{bmatrix} \end{matrix}$$

6 B

7 A

8 C

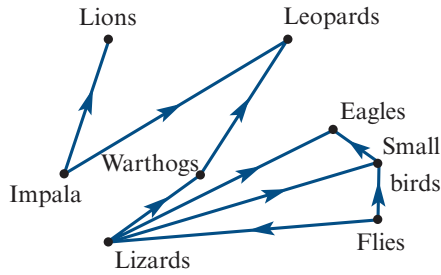
9 C

10 D

Section 9C

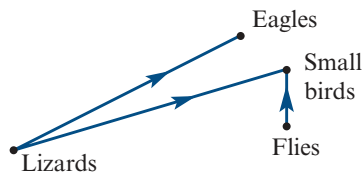
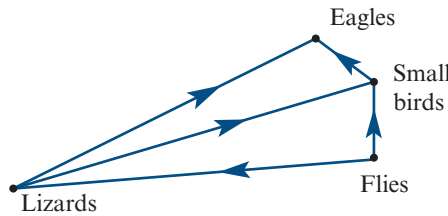
- 1 a** Chelsea
b Wayne
c Samantha
d i deg (Samantha) = 4
 ii deg (Eli) = 4

2 a



b $\deg(\text{Warthog}) = 2$

c Multiple answers are possible. Two of these are shown below.



3 a

$$\begin{matrix} A & B & C \\ A & \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \\ B & \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \\ C & \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} \end{matrix}$$

b

$$\begin{matrix} A & B & C & D \\ A & \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \\ B & \begin{bmatrix} 1 & 0 & 0 & 1 \end{bmatrix} \\ C & \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \\ D & \begin{bmatrix} 0 & 1 & 1 & 0 \end{bmatrix} \end{matrix}$$

c

$$\begin{matrix} M & N & P & Q \\ M & \begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix} \\ N & \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \\ P & \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \\ Q & \begin{bmatrix} 0 & 0 & 1 & 1 \end{bmatrix} \end{matrix}$$

d

$$\begin{matrix} E & F & G & H \\ E & \begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix} \\ F & \begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix} \\ G & \begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix} \\ H & \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \end{matrix}$$

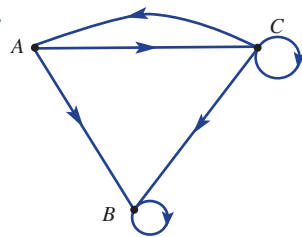
e

$$\begin{matrix} A & B & C & D & E \\ A & \begin{bmatrix} 0 & 1 & 1 & 0 & 0 \end{bmatrix} \\ B & \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \end{bmatrix} \\ C & \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \end{bmatrix} \\ D & \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix} \\ E & \begin{bmatrix} 0 & 0 & 1 & 1 & 0 \end{bmatrix} \end{matrix}$$

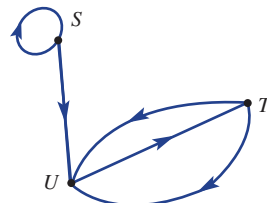
f

$$\begin{matrix} A & B & C & D & E & F \\ A & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \\ B & \begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} \\ C & \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \\ D & \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ E & \begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} \\ F & \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{matrix}$$

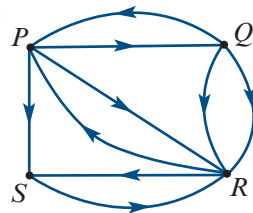
4 a



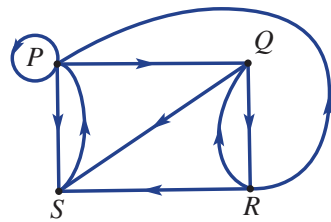
b



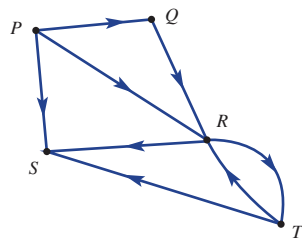
c



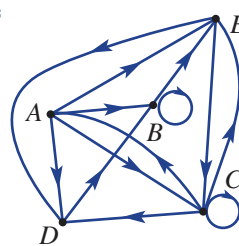
d



e



f



5 A

6 D

7 D

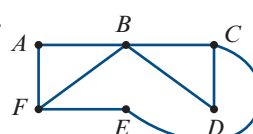
Section 9D

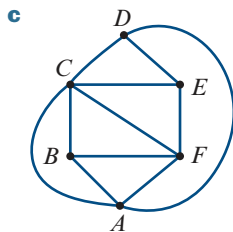
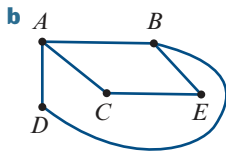
1 a i

b ii

c ii

2 a





d not possible

3 a $v = 8, e = 12, f = 6$

b $v = 6, e = 12, f = 8$

c $v = 7, e = 12, f = 7$

d $v = 5, e = 6, f = 3$

e $v = 5, e = 9, f = 6$

f $v = 6, e = 8, f = 4$

4 a $f = 4$ **b** $v = 12$ **c** $f = 11$ **d** $e = 19$

5 7

6 D **7** C **8** B **9** C **10** D

Section 9E

1 a path **b** trail **c** path
d closed walk **e** trail **f** path

2 a trail - not a path

b walk only

c path

d path

3 a closed path or cycle

b open walk only

c open walk only

d trail

e closed walk

f closed path or cycle

4 C **5** C

Section 9F

1 a i Semi-Eulerian
ii $E - A - B - E - D - B - C - D - A$

b i Neither

c i Semi-Eulerian

ii $A - C - E - C - B - D - E - F$

d i Eulerian

ii $A - B - C - E - D - C - A$

e i Eulerian

ii $E - F - D - E - A - B - D - C - B - E$

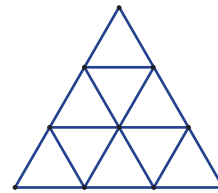
2 a $A : 4, B : 2, C : 5, D : 2, E : 4,$
 $F : 4, G : 3$

b i C or G **ii** G or C

c i C and G

ii Eulerian trail

3 a



b Each vertex has even degree.

4 a $1 \times m$ where $m > 1$

$m \times 1$ where $m > 1$

$2 \times 3, 3 \times 2.$

b 2×2

5 C

6 D

7 A

8 D

9 D

Section 9G

1 a i $A - B - C - F - I - H - E - G - D$

ii $E - G - D - A - B - C - F - I - H - E$

b i $A - B - C - D - E - F$

ii $E - F - A - B - C - D - E$

c i $A - B - D - C - E$

ii $E - A - B - D - C - E$

2 (Other examples may be given)

a $v_1 - v_2 - v_3 - v_4$

b $v_3 - v_5 - v_6 - v_4 - v_1 - v_2$

c $v_1 - v_3 - v_2 - v_4$

d $v_1 - v_2 - v_3 - v_4 - v_5 - v_6 - v_7$

3 (Other examples may be given)

a $v_1 - v_4 - v_3 - v_2 - v_1$

b $v_1 - v_2 - v_3 - v_5 - v_6 - v_4 - v_1$

c $v_1 - v_5 - v_2 - v_3 - v_4 - v_1$

d $v_1 - v_2 - v_5 - v_7 - v_6 - v_4 - v_3 - v_1$

4 a $A - F - G - B - C - H - E - D$

b $F - A - B - C - D - E - H - G$

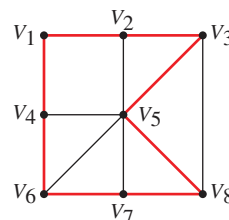
5 We seek a Hamiltonian cycle starting and

finishing at at v_1 . One such cycle is

$v_1 - v_2 - v_3 - v_5 - v_8 - v_7 - v_6 - v_4 - v_1$,

and this is shown in red below. There are

many others.



6 a 5

b i Hamilton path

ii $E - W - D - C - B - A - F$

7 D

8 B

Section 9H

- 1 **a** D and E **b** 17 minutes
c 8 minutes
d 36 minutes ($A - B - C - D - E$)
- 2 11
- 3 **a** 34 km
b 56 km
c Two answers are possible: $A - E - F - G - I$
or $A - C - F - G - I$
- 4 **a** $S - B - D - F$, 12
b $S - A - C - D - F$, 10
c $S - B - D - F$, 15
d $S - A - E - G - F$, 19
- 5 19 km
- 6 **B** 7 **C** 8 **A** 9 **D**

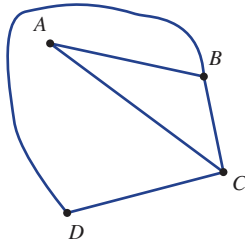
Chapter 9 review

Multiple-choice questions

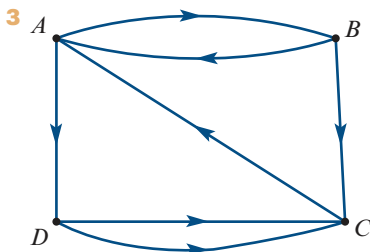
- 1 **C** 2 **A** 3 **C** 4 **D**
5 **A** 6 **D** 7 **C** 8 **A**
9 **B** 10 **B** 11 **B** 12 **A**
13 **C** 14 **B** 15 **B** 16 **D**
17 **A** 18 **C** 19 **B** 20 **B**
21 **B**

Short-response questions

- 1 **a** 3 **b** 2 **c** A and D
d



- 2
- | | A | B | C | D |
|---|---|---|---|---|
| A | 0 | 1 | 1 | 1 |
| B | 1 | 0 | 1 | 1 |
| C | 1 | 1 | 0 | 1 |
| D | 1 | 1 | 1 | 0 |



- 4 **a** Mackay Masters and Gladstone Gladiators
b 2 **c** Bundaberg Braves

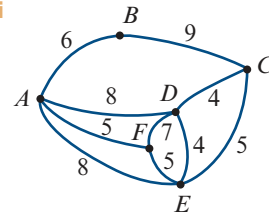
- 5 **a** 4 **b** 6 **c** 8
d $4 + 6 - 8 = 2$

- 6 **a** C and D **b** 12 hours
c 13 hours

- 7 **a** It can be drawn so that none of the edges in the graph cross over each other.
b $v + f + e = 9 + 7 - 14 = 2$
c 750 m

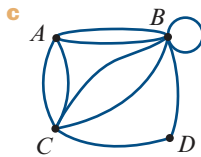
- d** **i** Yes, all vertices have even degrees and so the graph is Eulerian.
ii Multiple answers possible.
One is: Office – $C5 - C7 - C8 - C6 - C5 - C4 - C3 - C2 - C4 - C1 - C2 - C8 - C1$ – office
e **i** Hamiltonian cycle
ii $C7$ to park office. Other answers possible.
iii Office – $C1 - C2 - C3 - C4 - C5 - C6 - C8 - C7$ – office, or same route in reverse order.
Other answers possible.

- 8 **a** **i**



- ii**
- | | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| A | 0 | 1 | 0 | 1 | 1 | 1 |
| B | 1 | 0 | 1 | 0 | 0 | 0 |
| C | 0 | 1 | 0 | 1 | 1 | 0 |
| D | 1 | 0 | 1 | 0 | 1 | 1 |
| E | 1 | 0 | 1 | 1 | 0 | 1 |
| F | 1 | 0 | 0 | 1 | 1 | 0 |

- b** **i** 45 km (at minimum)
ii Some vertices are visited more than once.
iii $F - E - D - C - B - A - F$
iv 33 km (for route above; other answers possible)
c F and C
- 9 **a** 7
b 2



- c** Vertices are not all even.
- 10 a** Park Entrance - Information Centre - Boathouse - Campsite - Lookout - Park Entrance.
Park Entrance - Lookout - Campsite - Boathouse - Information Centre - Park Entrance.
- b** Yes. There are exactly two vertices that have an odd degree.
- c** **i** 30 minutes
ii 2 hours and 45 minutes
- 11 a** 4
b 18
c $v + f = e + 2$; $6 + 5 = 9 + 2$
d No, the salesperson has visited a vertex more than once (Melville), therefore this is not a Hamiltonian cycle.
e Hamiltonian path, because all vertices were visited without repeating any vertices and the starting vertex is different to the ending vertex.
f Melville-Croghon-Bartow-Stratmoore-Osburn-Kenton-Melville. The shortest distance is 58 kilometres.
g An Eulerian circuit is possible if all vertices have an even degree. The vertices that represent Croghon and Stratmoore both have an odd degree.
h This walk described is an Eulerian trail. The inspector could start their route at either Croghon or Stratmoore, because these are the only two vertices with an odd degree. One option is: Stratmoore - Osburn - Kenton - Melville - Osburn - Croghon - Bartow - Stratmoore - Melville - Croghon (the reverse is also acceptable)
i 3 : 29 pm

Chapter 10

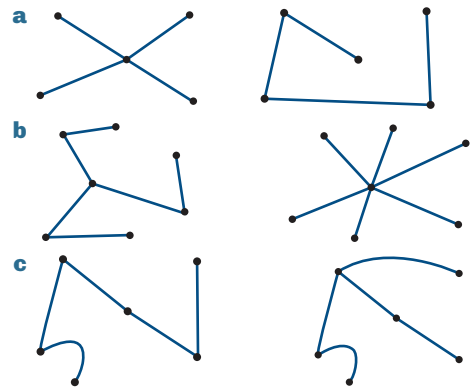
Section 10A

- 1 a** 11
b 9
c Multiple answers possible

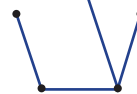


2 A, B, D

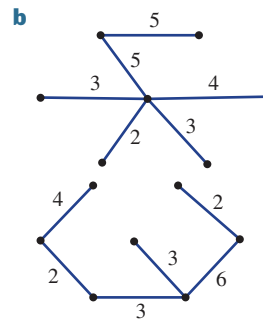
3 Multiple answers possible.



4



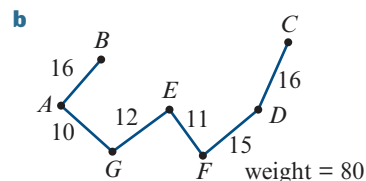
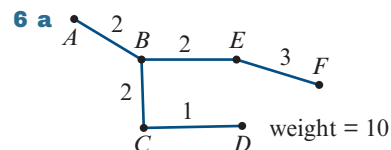
5 a 6

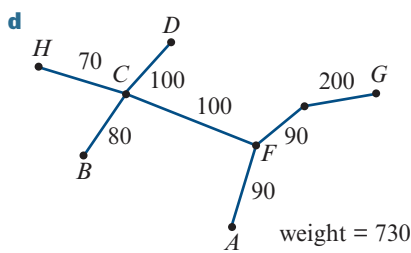
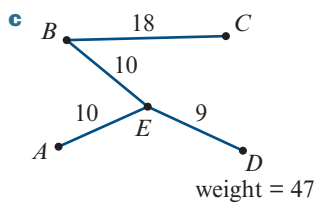


Note: other answers are possible

c 22, 20

Note: other answers are possible

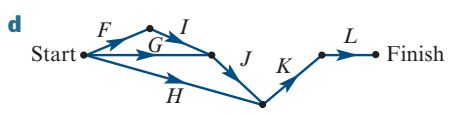
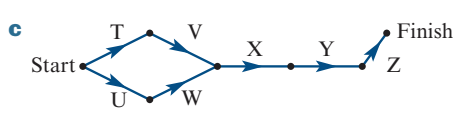
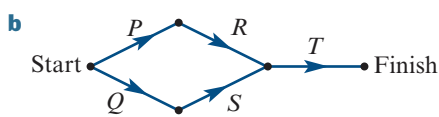
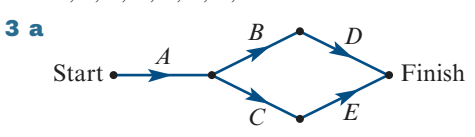




- 7** 179 km **8** 94 km **9** 490 m
10 D **11** A **12** C **13** C
14 D **15** D

Section 10B

- 1 a** 8 **b** C and E
c H **d** 1
2 a F, D, H
b A, B, C, D, E, F, G, H



4 a

Activity	Immediate Predecessors
A	–
B	–
C	A
D	A
E	B, C
F	D
G	E

b

Activity	Immediate Predecessors
P	–
Q	P
R	P
S	Q
T	Q
U	S, V
V	R
W	R
X	T, U

c

Activity	Immediate Predecessors
J	–
K	–
L	J
M	N
N	K
O	K
P	N
Q	L, M
R	P
S	O, R
T	Q

d

Activity	Immediate Predecessors
A	–
B	–
C	A
D	A
E	D
F	C, E
G	D
H	B

e

Activity	Immediate Predecessors
P	–
Q	–
R	P
S	P
T	Q
U	R
V	S
W	T
X	U
Y	W
Z	V, X, Y

f

Activity	Immediate Predecessors
A	—
B	A
C	A
D	A
E	B
F	C
G	D
H	E, F
I	G
J	I
K	H

5 a

Activity	Immediate predecessors
A	—
B	—
C	—
D	A
E	B, F
F	C
G	B, F
H	D, E
I	H
J	I, K
K	G
L	G
M	H
N	J, L
O	N

- b** A - D - H - M
A - D - H - I - J - N - O
c B - E - H - M
B - E - H - I - J - N - O
B - G - K - J - N - O
B - G - L - N - O
d C - F - E - H - M
C - F - E - H - I - J - N - O
C - F - G - K - J - N - O
C - F - G - L - N - O
6 D **7 B** **8 D** **9 C**

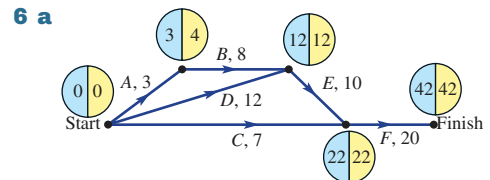
Section 10C

- 1 a** $p = 12$ **b** $w = 10$ **c** $m = 8, n = 8$
d $a = 10, b = 18, c = 11$
2 a 3 **b** 5 **c** 2
d 13 **e** A - C
3 a 12 **b** 10 **c** 9
d 1 **e** 3 **f** 9
4 a 40 **b** 8 **c** 11
d 28
e i 20 **ii** EST is not the same as LST
f A

5 a

Activity	Duration	EST	LFT	LST	Float
A	10	0	11	1	1
B	9	0	9	0	0
C	3	0	17	14	14
D	2	10	13	11	1
E	4	9	13	9	0
F	1	13	14	13	0
G	1	13	15	14	1
H	3	14	17	14	0
I	2	14	17	15	1
J	1	17	18	17	0

b B - E - F - H - J



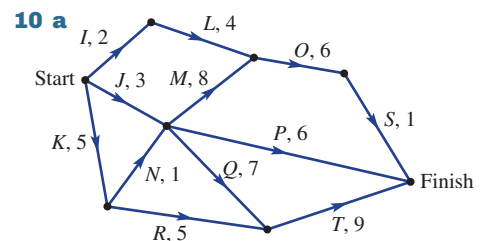
- b** 42
c D - E - F
d i A : 1, B : 4, C : 15
ii B : 1, A : 1, C : 15

- 7 a** 8 **b** 11
c C - F - G **d** B

- 8 a** D, F, G **b** 13
c Activity H lies on the critical path and if delayed, the completion time of the project will be extended.

- d** 15 **e** F

- 9 a i** 25 **ii** 29 **iii** 30
b 5 **c** 2



- b**
-
- ```

graph LR
 Start((0|0)) -- "I, 2" --> I((2|11))
 Start -- "J, 3" --> J((3|6))
 Start -- "K, 5" --> K((5|5))
 I -- "L, 4" --> L((6|15))
 J -- "M, 8" --> M((11|19))
 K -- "N, 7" --> N((12|19))
 L -- "O, 6" --> O((17|23))
 M -- "P, 6" --> P((17|23))
 N -- "Q, 7" --> Q((19|26))
 O -- "S, 1" --> S((23|24))
 P -- "T, 9" --> T((26|35))
 Q -- "R, 5" --> R((24|29))
 S --> Finish((22|22))
 T --> Finish
 R --> Finish

```
- c** 22 weeks  
**d** K - N - Q - T

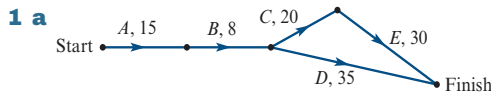
**e**

| Activity | Duration | EST | LFT | LST | Float |
|----------|----------|-----|-----|-----|-------|
| I        | 2        | 0   | 11  | 9   | 9     |
| J        | 3        | 0   | 6   | 3   | 3     |
| K        | 5        | 0   | 5   | 0   | 0     |
| L        | 4        | 2   | 15  | 11  | 9     |
| M        | 8        | 6   | 15  | 7   | 1     |
| N        | 1        | 5   | 6   | 5   | 0     |
| O        | 6        | 14  | 21  | 15  | 1     |
| P        | 6        | 6   | 22  | 16  | 10    |
| Q        | 7        | 6   | 13  | 6   | 0     |
| R        | 5        | 5   | 13  | 8   | 3     |
| S        | 1        | 20  | 22  | 21  | 1     |
| T        | 9        | 13  | 22  | 13  | 0     |

**f** Verified

- 11** A    **12** B    **13** C    **14** D  
**15** D    **16** C    **17** A    **18** D

### Section 10D



**b** 73 minutes

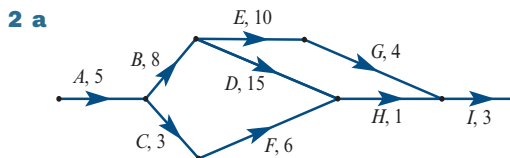
**c i**

| Activity | Duration | EST | LFT | LST | Float |
|----------|----------|-----|-----|-----|-------|
| A        | 15       | 0   | 15  | 0   | 0     |
| B        | 8        | 15  | 23  | 15  | 0     |
| C        | 20       | 23  | 43  | 23  | 0     |
| D        | 35       | 23  | 73  | 38  | 15    |
| E        | 30       | 43  | 73  | 43  | 0     |

**ii** Cleaning the interior

**iii**  $A - B - C - E$

**d** 38 minutes



**b** 32 weeks

**c i**

| Activity | Duration | EST | LFT | LST | Float |
|----------|----------|-----|-----|-----|-------|
| A        | 5        | 0   | 5   | 0   | 0     |
| B        | 8        | 5   | 13  | 5   | 0     |
| C        | 3        | 5   | 22  | 19  | 14    |
| D        | 15       | 13  | 28  | 13  | 0     |
| E        | 10       | 13  | 25  | 15  | 2     |
| F        | 6        | 8   | 28  | 22  | 14    |
| G        | 4        | 23  | 29  | 25  | 2     |
| H        | 1        | 28  | 29  | 28  | 0     |
| I        | 3        | 29  | 32  | 29  | 0     |

**ii**  $A - B - D - H - I$

**d i** The project would be completed in a minimum of 30 weeks.

**ii** Nothing. This activity has a float of 14 and so it could be extended in duration by 14 weeks.

**iii** The project would be completed in a minimum of 37 weeks.

### Section 10E

**1 a**  $A - D$ : 17 hours;  $B - E - F$ : 20 hours;  
 $B - E - G - I$ : 21 hours;  $C - H - I$ :

16 hours

**b**  $B - E - G - I$ , 21 hours

**c** 18 hours

**2 a**  $A - B - D - G - I$     **b** 22 days

**c** 20 days    **d** \$200

**3 a**  $B - E - H - J$     **b** 2 hours

**c** 6 hours    **d** 14 hours

**4 a** 5    **b** 17 hours    **c** \$1200

**5 a** 22 days    **b** \$870

**6 a** C, D, H

**b** B, E, H, I, J

**c i** 21 days    **ii** \$450

**7 a** 29 days

**b** 6

**c** 4

**d** Two answers possible:

H, 2 J, 0 K, 2 L, 1 M, 1

H, 2 J, 0 K, 1 L, 1 M, 2

**8** B    **9** B    **10** D

## Chapter 10 review

### Multiple-choice questions

**1** C    **2** B    **3** C    **4** A

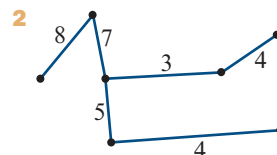
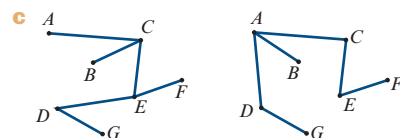
**5** B    **6** D    **7** D    **8** D

**9** D    **10** D    **11** C    **12** B

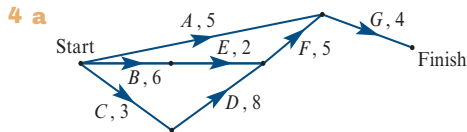
**13** A    **14** D    **15** C

### Short-response questions

**1 a** 3    **b** D and G, E and F



3 a C b F and G c 2



b 20 weeks

5 a 30 days b 13 days  
c 1 day d  $B - C - E - G - J - K$

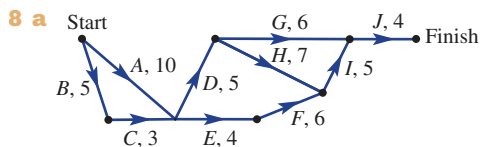
6 a

| Activity | Immediate predecessors |
|----------|------------------------|
| A        | —                      |
| B        | —                      |
| C        | —                      |
| D        | A                      |
| E        | C                      |
| F        | B, E                   |
| G        | B, E                   |
| H        | B, E                   |
| I        | G                      |
| J        | D, F                   |
| K        | D, F                   |
| L        | J                      |
| M        | H, K                   |
| N        | I                      |

b 26 hours

c  $B - F - J - L$

7 a 5 hours b 24 hours c 7 hours



b 31 days

c

| Activity | Duration | EST | LFT | LST | Float |
|----------|----------|-----|-----|-----|-------|
| A        | 10       | 0   | 10  | 0   | 0     |
| B        | 5        | 0   | 7   | 2   | 2     |
| C        | 3        | 5   | 10  | 7   | 2     |
| D        | 5        | 10  | 15  | 10  | 0     |
| E        | 4        | 10  | 16  | 12  | 2     |
| F        | 6        | 14  | 22  | 16  | 2     |
| G        | 6        | 15  | 27  | 21  | 6     |
| H        | 7        | 15  | 22  | 15  | 0     |
| I        | 5        | 22  | 27  | 22  | 0     |
| J        | 4        | 27  | 31  | 27  | 0     |

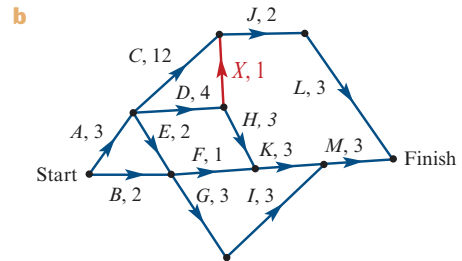
d If an activity is on the critical path, it is an activity that cannot be delayed or extended in duration without affecting the overall minimum completion time of the project.

e The activities that have float equal to zero.

f  $A - D - H - I - J$

g Nothing. C has a float time of 2, which means it can be delayed by up to 2 days without affecting the overall completion time of the project.

9 a Immediate predecessor of E is A. EST for I is 8. EST for M is 13.



c i 16 hours

ii Critical path is the sequence of activities that cannot be delayed without delaying the entire project.

d i 6 hours

ii  $A - C - J - L$

iii 8 hours

10 a  $B - E - H - J$  b 2 hours c 6 hours

d 14 hours

11 a A, B, C

b LST for B is 1, EST for E is 10, LST for I is 18

c i  $A - D - F - I - J$

ii 27 months

d i  $B - C - D - F - I - J$

ii 25 months

## Chapter 11

### Section 11A

1 31 000

2 Cut  $C_1 - 14$ , Cut  $C_2 - 12$ , Cut  $C_3 - 21$

3 Cut  $C_1 - 12$ , Cut  $C_2 - 16$ , Cut  $C_3 - 16$

4 a 9 b 11 c 8 d 18

5 a Cut A - 14, Cut B - 23, Cut C - 12, Cut D - 16, Cut E - not a cut

b It does not completely separate the source from the sink.

c 12

6 a sink 1 = 10, sink 2 = 11

b sink 1 = 8, sink 2 = 18

7 a 9 b 18

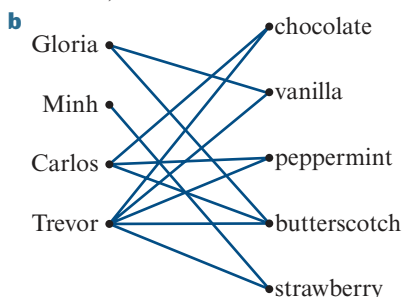


- 8 a** Cut passes through edges with weights 20, 10, 30, 30  
**b** 59      **c** 25  
**9** D      **10** C      **11** B

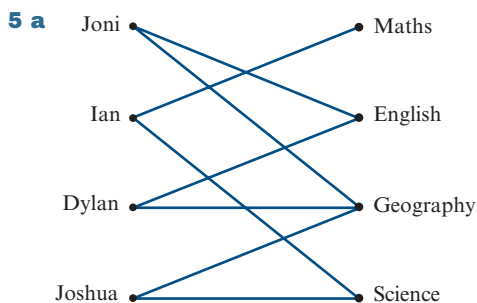
### Section 11B

- 1 a** Worker 1 – Task 3; Worker 2 – Task 1; Worker 3 – Task 2  
**b** Worker 1 – Task 5; Worker 2 – Task 6; Worker 3 – Task 4  
**2** Niranjan – Cake; Nishara – Candles; Dhinesh – Serviettes; Dhishani – Balloons  
**3** Two answers possible  
 Player 1 – Right Wing; Player 2 – Left wing; Player 3 – Centre; Player 4 – Right Defence; Player 5 – Left Defence  
 Or  
 Player 1 – Centre; Player 2 – Right wing; Player 3 – Left Defence; Player 4 – Right Defence; Player 5 – Left Wing

- 4 a** two distinct groups of vertices (people and flavours)

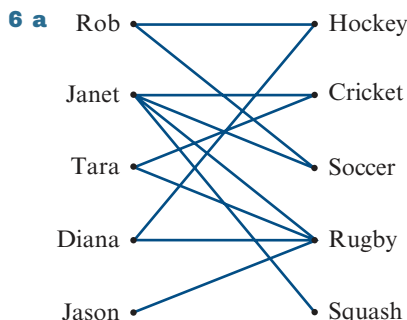


**c** 5



- b** Ian is the only teacher who can teach Maths and so he cannot teach Science. Joshua is the only other teacher who can teach Science and so he must take this class.  
**c** Ian – Maths, Joshua – Science, Dylan – English, Joni – Geography  
 Ian – Maths, Joshua – Science, Dylan – Geography, Joni – English

- d** Gloria  
**e** Gloria – vanilla, Minh – strawberry, Carlos – peppermint, Trevor – chocolate



- b** Jason can only coach Rugby and so Diana cannot. The only other sport Diana can coach is Hockey.  
**c** Jason – Rugby, Diana – Hockey, Rob – Soccer, Janet – Squash, Tara – Cricket

**7** D      **8** B

### Section 11C

- 1 a i**  $97 + 97 + 103 = 297$ ,  
 $97 + 82 + 82 = 261$ ,  
 $91 + 148 + 103 = 342$ ,  
 $91 + 82 + 160 = 333$ ,  
 $67 + 148 + 82 = 297$ ,  
 $67 + 97 + 160 = 324$   
**ii**  $97 + 82 + 82 = 261$ ,  $A - X$ ,  $B - Z$ ,  $C - Y$   
**b**  $A - X$



- 2 a**  $W - D$ ,  $X - A$ ,  $Y - B$ ,  $Z - C$   
**b** Minimum cost is 11; Many allocations possible, e.g.  $W - A$ ,  $X - B$ ,  $Y - D$ ,  $Z - C$   
**3**  $A - Y$ ,  $B - X$ ,  $C - Z$ . Maximum is 342  
**4** Dimitri 800 m, John 400 m, Carol 100 m, Elizabeth 1500 m  
**5** Joe C, Meg A, Ali B  
**6**  $A - Y$ ,  $B - Z$ ,  $C - X$ ,  $D - W$   
**7** Champs – Home, Stars – Away, Wests – Neutral; or Champs – Neutral, Stars – Away, Wests – Home.  
 Cost = \$20 000  
**8** A Mark, B Karla, C Raj, D Jess; or A Karla, B Raj, C Mark, D Jess; 55 km  
**9 a i** 11:26 am  
**ii** A – talk 3, B – talk 1, C – talk 2, D – talk 4  
**b** 11:41 am

- 10 The maximum value is 1430.  
The allocation is:  $C - P, B - Q, D - R, A - S$   
11 A    12 A    13 D

## Chapter 11 review

### Multiple-choice questions

- 1 D    2 C    3 D    4 D  
5 A    6 C    7 A

### Short-response questions

- 1 a 2    b Audrey  
c Audrey – natives, Brian – Aloes,  
Cameron – Cactus, Daphne – grasses
- 2 a
- 
- b David – printing, Robyn – mailing,  
Linda – editing, Anthony – word processing
- 3 a Julia    b Mario  
c Julia – assembling, Mario – painting,  
Sylvana – preparing, George – Cutting
- 4 a 26    b 15
- 5 a
- 
- b Chris  
c Bernard – 10 a.m., Georgia – 9 a.m.,  
Chris – 1 p.m., Arthur – 3 p.m.
- 6 Ann – D, Bianca – B, Con – C, David – E
- 7 a 600 litres per minute  
b 72 000 L  
c 4.5 min
- 8 a 26  
b 15
- 9 Rob – breaststroke, Joel – backstroke,  
Henk – freestyle, Sav – butterfly or Rob –  
breaststroke, Joel – butterfly, Henk –  
backstroke, Sav – freestyle. Time = 276

## Chapter 12

### Section 12A

#### Topic 1: Loans, investments and annuities 1

#### Multiple-choice questions

- 1 D    2 D    3 D    4 D  
5 C    6 A    7 D    8 A  
9 C    10 D

#### Short-response questions

- 1 a  $A_1 = \$10\ 064.00$ ;  $A_2 = \$10\ 128.41$ ;  
 $A_3 = \$10\ 193.23$     b 8
- 2 a  $A_0 = 14\ 500$ ,  $A_{n+1} = 1.004 \times A_n$   
b \$14 674.70    c \$15 957.95
- 3 a 0.85%  
b  $A_0 = 4500$ ,  $A_{n+1} = 1.0085 \times A_n$   
c \$4981.08
- 4 a \$28 047.29    b \$3047.29
- 5 \$12 500.00
- 6 \$2282.56
- 7 a 0.59%    b 10
- 8 a \$9643.96    b \$343.96
- 9 \$121 649.93
- 10 a Loan 2. Amanda will pay less interest in total with this loan.  
b No. Even with payments of \$1900, Loan 2 still results in less interest overall.

### Section 12B

#### Topic 2: Loans, investments and annuities 2

#### Multiple-choice questions

- 1 B    2 D    3 D    4 C  
5 A    6 D    7 C    8 C  
9 D    10 C

#### Short-response questions

- 1 a 3.72%  
b \$244 843.13  
c \$3843.13
- 2 a  $A_0 = 145\ 000$ ,  $A_{n+1} = 1.0034 \times A_n - 2500$   
b \$134 896.53  
c \$2396.53
- 3 a \$602    b 6.86%
- 4 a Bank A, it has the highest effective rate of interest

**b** Multiple answers possible.

Total amount withdrawn from Bank A is  
\$40 264.69

Total amount withdrawn from Bank B is  
\$40 126.15

**5 a i** \$1314.08 **ii** \$1.62

**iii** This is the amount that has been overpaid.  
The bank must refund this.

**b** 175

**6 a** \$485 000

**b** \$100 496.83

**c** 72

**7 a** 25 **b** \$221 020.56 **c** 9

**d** \$16 584.37

**8** \$418.94 will give just under \$60 000 or  
\$418.95 will give just over \$60 000

**9** \$9786.81

**10** \$342 272.21

### Section 12C

#### Topic 3: Graphs and networks

##### Multiple-choice questions

**1** D **2** A **3** B **4** B

**5** B **6** B **7** C **8** B

**9** B **10** A **11** D **12** A

##### Short-response questions

**1 a** 3 **b** E **c** 2 **d** A, C, D

**e**  $B - C, F - D$

**2 a** R

**b** Q

**c** P Q R S

P  $\begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}$

Q  $\begin{bmatrix} 1 & 0 & 0 & 1 \end{bmatrix}$

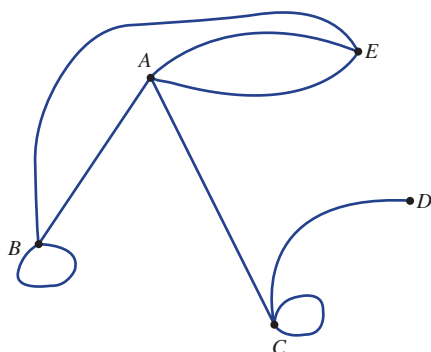
R  $\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}$

S  $\begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix}$

**3 a** None

**b** 2

**c**



**4 a** 8 **b** 5 **c** 5

**d**  $5 + 5 - 8 = 2$

**5** Yes

**6 a** 20 km

**b** 13 km

**7 a** It can be drawn so that no edges cross over each other.

**b**  $v = 7, f = 6, e = 11, 7 + 6 - 11 = 2$

**c i** C

**ii** (open) trail or semi-Eulerian trail

**d i**  $C - A - B - D - E - F - G - C$

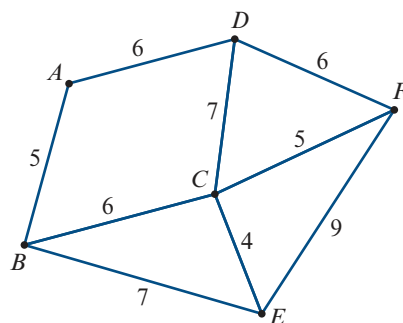
**ii** Hamiltonian cycle

**8 a** 7 km

**b** 3 km

**c** B and D

**d i**



**ii**

|   | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| A | 0 | 1 | 0 | 1 | 0 | 0 |
| B | 1 | 0 | 1 | 0 | 1 | 0 |
| C | 0 | 1 | 0 | 1 | 1 | 1 |
| D | 1 | 0 | 1 | 0 | 0 | 1 |
| E | 0 | 1 | 1 | 0 | 0 | 1 |
| F | 0 | 0 | 1 | 1 | 1 | 0 |

**9 a** Multiple answers are possible. One is  
 $A - B - D - E - C - A$

**b** BC, DE, CE

**c**  $B - C - E - D - E - C - A - B - D$ . Must start and end at odd-degree vertex.

**d** 9:54 a.m.

**10 a** 11 km

**b** There are exactly two odd-degree vertices in the network.

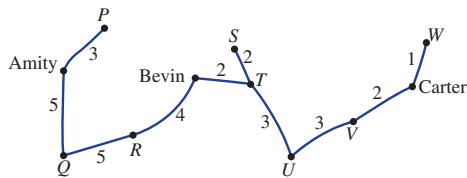
**c** Checkpoint V.

**d** Checkpoint U. If they do, they will have to travel one of the roads to Bevin a second time.

**e** Bevin - T - U - V - Carter or  
Bevin - T - U - Carter

f 21 km

g



## Section 12D

### Topic 4: Networks and decision mathematics 1

#### Multiple-choice questions

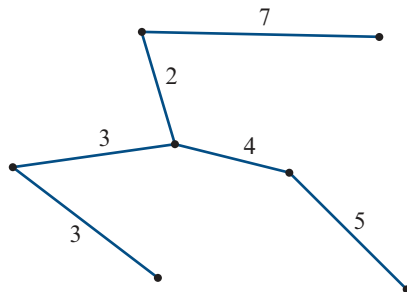
- 1 C      2 C      3 C      4 D  
5 C      6 D      7 A      8 B  
9 C      10 D

#### Short-response questions



2 A – B or C – B or A – C

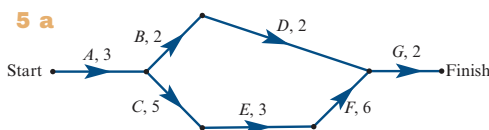
3 a



b 24

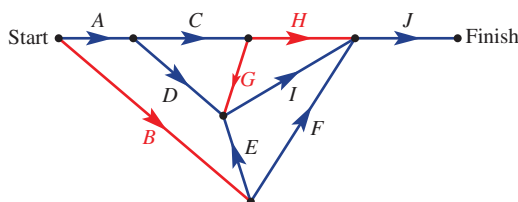
4 a 9      b 3      c F and G

5 a



b 19 hours

6



7 a EST for G is 5      b 24 hrs      c 7 hours

8 a 17

b 7

c 1

d A – C – G – J – L

9 a 9

b 7

c 1

d B – D – E – G

e 15

10 a A : LST = 1      D : duration = 4  
F : LST = 10      K : duration = 12

b B – C – E – G – J – K

11 a

| Activity | Immediate predecessors |
|----------|------------------------|
| A        | –                      |
| B        | A                      |
| C        | –                      |
| D        | –                      |
| E        | C                      |
| F        | B, E                   |
| G        | C                      |
| H        | D                      |
| I        | G, H                   |
| J        | F, I                   |

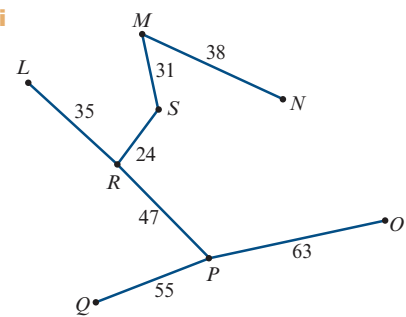
b 24 days

c C – E – F – J

12 a 112 km

b i minimum spanning tree

ii



iii 293 km

## Section 12E

### Topic 5: Networks and decision mathematics 2

#### Multiple-choice questions

- 1 D      2 C      3 A      4 D  
5 C      6 D      7 B      8 B  
9 A      10 C      11 D

#### Short-response questions

1 a 1

b Sharon

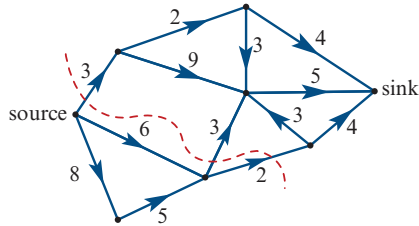
c Leah – Brazil, Sharon – Portugal, Kris – Tibet, Sue – Zimbabwe, Kathy – Fiji

2 a Cut 1 does not isolate the source from the sink.

b 26

c 22

- 3 a** Ahmet  
**b** Ahmet – Canapes, Beryl – Starter, Cynthia – Desert, Dario – Main  
**c i** 10 h  
**ii** 4 h  
**4 a i** 22  
**ii** 11



- 5 a**  
 Tyson — Flakey  
 Emma — Cherry Chomp  
 Gregory — Honey Crunch  
 Rose — Snacker  
**b** Tyson  
**c** Tyson – Snacker, Emma – Flakey, Gregory – Cherry Chomp, Rose – Honey Crunch

- 6 a**  $A - Z$ ,  $B - W$ ,  $C - X$ ,  $D - Y$ , or  $A - Z$ ,  $B - X$ ,  $C - W$ ,  $D - Y$   
**b** \$130  
**7 a** 11 megalitres per day  
**b** 33 megalitres  
**8** Min = 185, Max = 197  
**9** sink 1 = 8, sink 2 = 18  
**10** 700 kilolitres per minute for each outlet.

## Chapter 13

### Section 13A

#### Paper 1 revision questions

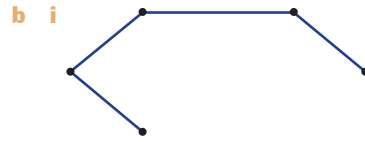
#### Multiple-choice questions

- 1** D    **2** C    **3** B    **4** C  
**5** D    **6** C    **7** D    **8** A  
**9** B    **10** D    **11** D    **12** A  
**13** D    **14** C    **15** A    **16** B  
**17** B    **18** D    **19** C    **20** A

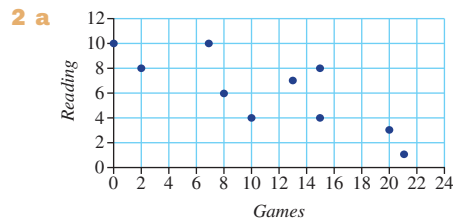
- 21** B    **22** C    **23** D    **24** D  
**25** B    **26** A    **27** C    **28** C  
**29** D    **30** B    **31** C    **32** D  
**33** B    **34** C    **35** C    **36** C  
**37** C    **38** D    **39** C    **40** A  
**41** A    **42** C    **43** C    **44** B  
**45** C    **46** D    **47** B    **48** D  
**49** D

#### Short-response questions

- 1 a i** It can be drawn with edges only intersecting at vertices.  
**ii** The number of vertices = 5.  
 The number of edges = 8.  
 The number of faces = 5.



- ii** The number of vertices = 5.  
 The number of edges = 4.  
 The number of faces = 1.



- b** There is a moderate, negative linear association between the number of hours the students spend playing computer games, and the time they spend reading.  
**3**  $bodyfat = 0.48 \times age - 6.40$   
**4** 2780 km  
**5** 13 689 km  
**6 a** 35    **b** 219    **c** 14 580    **d** 48  
**7 a**  $t_1 = 1, t_2 = 7, t_3 = 13$   
**b**  $t_1 = 2, t_2 = 10, t_3 = 50$   
**8 a**  $t_n = 220 - 10(n - 1)$   
**b**  $t_n = 3 \times 5^{n-1}$   
**9** The time series plot shows the following features. Trend — long term is positive because the amount spent on online sales generally increases as time increases. Seasonality — The data is seasonal peaking in Oct–Nov every year. Irregular fluctuation.  
**10** \$713 284  
**11 a i** 2.1 km  
**ii**  $P - Q - R - T - S - U$  or

$P-R-Q-S-T-U$  or

$P-R-Q-T-S-U$  or

$P-R-T-Q-S-U$

- b i**  $R-Q-P-R-T-Q-S-T-U-S$   
or  $R-Q-P-R-T-S-Q-T-U-S$

**ii** travel each road once

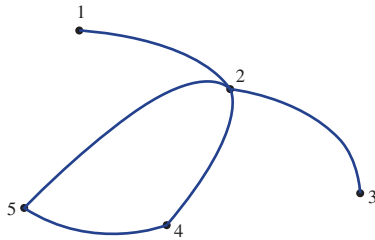
- 12 a** 9.0 litres/km

- b**  $m = 0.0218$ . On average, fuel consumption increases by 0.0218 litres/km for each 1 km increase in speed.

- 13 a** 0.4624

- b** 46.24% of the variation in sales is explained by the variation in advertising spend.

- 14 a**



- b** 10

- 15** \$182 660.82

- 16**  $y = -1.58x + 54.3$

- 17 a** Age group

| Enough? | < 30 years | ≥ 30 years |
|---------|------------|------------|
| Yes     | 65.0       | 43.3       |
| No      | 35.0       | 56.7       |
| Total   | 100        | 100        |

- c** There is an association between age group and agreement with the statement 'Is the government doing enough to reduce homelessness?'. A higher percentage of those aged under 30 (65.0%) agreed this with statement than those aged 30 or more (43.3%).

- 18 a** \$45 254.25, \$45 509.94, \$45 767.07

- b** Four months

- 19 a**  $A_0 = 17\ 650$ ,  $A_{n+1} = 1.0045 \times A_n$

- b** \$17 889.34 **c** \$19 658.18

- 20 a** Vertices  $D$  and  $E$  are odd.

- b**  $E$  and  $F$

- c**  $E-F-D-E-A-B-C-D$

- 21 a** 75 litres per minute

- b** 60 litres per minute

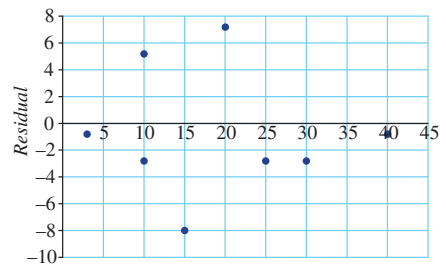
- 22**

|   | A | B | C | D |
|---|---|---|---|---|
| A | 0 | 1 | 1 | 0 |
| B | 1 | 0 | 1 | 1 |
| C | 1 | 1 | 0 | 0 |
| D | 0 | 1 | 0 | 0 |

- 23 a**

|      |      |     |      |     |      |      |      |
|------|------|-----|------|-----|------|------|------|
| -0.8 | -2.8 | 5.2 | -7.8 | 7.2 | -2.8 | -2.8 | -0.8 |
|------|------|-----|------|-----|------|------|------|

- b**



- 24**

| Q1   | Q2   | Q3   | Q4   |
|------|------|------|------|
| 1.01 | 1.15 | 1.32 | 0.52 |

- 25 a** 0.54%

- b**  $A_0 = 8200$ ,  $A_{n+1} = 1.0054 \times A_n$

- c** \$8747.43

- 26 a** \$74 777.29

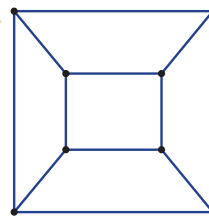
- b** \$9777.29

- 27** \$13 000

- 28 a**  $v = 4$ ,  $e = 6$

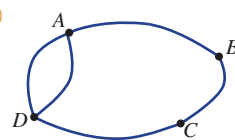
- b**  $v - e + f = 4 - 6 + 4 = 2$

- 29 a**



- b**  $v = 8$ ,  $e = 12$ ,  $f = 6$ ;  $8 - 12 + 6 = 2$

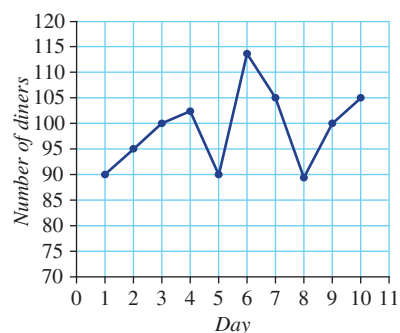
- 30**



- 31 a**

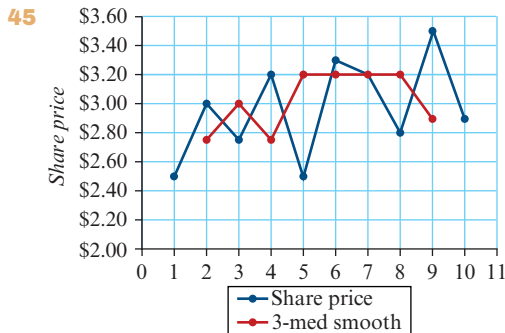
| Week | Wed | Thur | Fri | Sat | Sun |
|------|-----|------|-----|-----|-----|
| 1    | 90  | 96   | 100 | 103 | 90  |
| 2    | 114 | 106  | 88  | 100 | 105 |

- b**



- 32** The residual plot shows a curved structure, indication that the relationship between the length of a dugong and its age is not linear.

- 33 a**  $A_0 = 183\,000$ ,  $A_{n+1} = 1.00531 \times A_n - 2650$   
**b** \$174 519.06 **c** \$4769.06
- 34 a**  $272^\circ$  going west from MacKay or  $88^\circ$  degrees going east from Vancouver  
**b** MacKay is 18 hours ahead  
**c** 8:45 p.m.
- 35** ( $25^\circ\text{N}$ ,  $150^\circ\text{W}$ )
- 36**  $\text{relax} = -0.1 \times \text{work} + 8$
- 37 a** 173.9 **b** 188.7
- 38** 19682 km
- 39** 3:00 am Saturday
- 40** 2:00 am
- 41 a** 10.5%  
**b** \$365 240.92  
**c** \$15 840.92
- 42 a** \$436.05  
**b** 7.14%
- 43 a** \$13 603.83 **b** \$6311.05
- 44 a** \$10 687 **b** \$1687



The plot shows a slightly increasing trend with irregular fluctuations.

### Section 13B

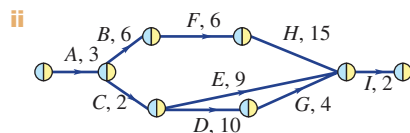
#### Paper 2 revision questions

#### Short-response questions

- 1 a i** 2 **ii** C

- b** A on breaststroke,  
 B on backstroke,  
 C on butterfly

- c i** 10

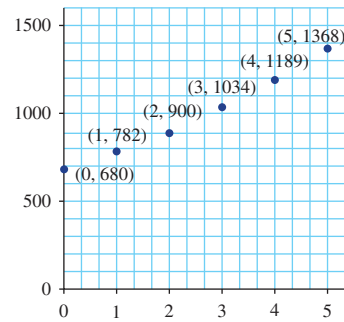


- iii** C, D, E, G

- iv** 21

- 2 a**  $C = 3000 + 1.2n$  **b** \$3300 **c** 1250 km

- 3**  $N_{n+1} = 1.15N_n$ ,  $N_1 = 782$



No. The population will not exceed 1500 during this period.

- 4** \$2655.99

- 5 a** 0.51% **b** 22 months

- 6 a** There is a strong, nonlinear relationship between efficiency and enthusiasm.  
**b** The relationship as shown in the scatterplot does not appear to be linear, therefore the correlation coefficient should not be calculated.

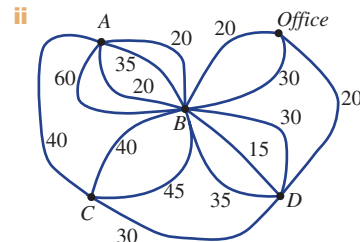
- 7** 1692 gm

- 8 a** 10:30 pm on Saturday **b** 9007 km

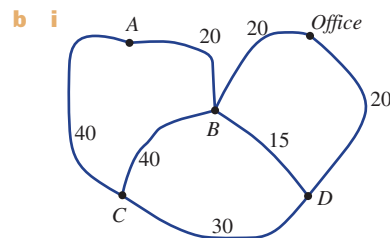
- c** 2:00 pm **d** 11 060 km

- 9 a** 3709 km **b** 71 km

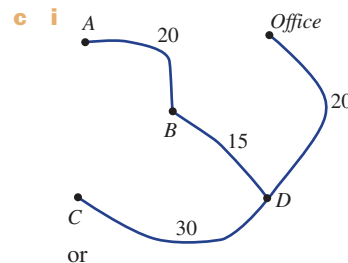
- 10 a i** 4

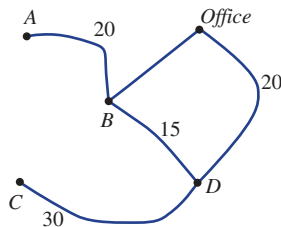


- iii** A, B, D

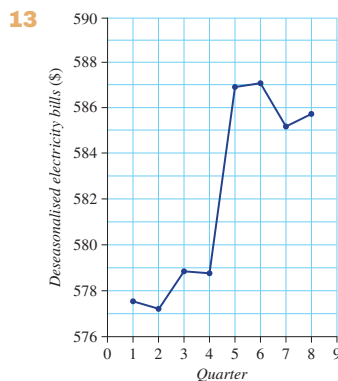


- ii** \$44 000

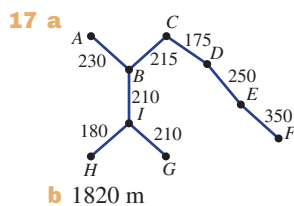




- ii** \$20 400  
**d i** 2.5 minutes **ii** 3.5 minutes  
**11 a** Earliest start time for  $E = 13$  days  
**b** Latest start time for  $D = 23$  **c** 31 days  
**d**  $A - B - E - F - G - H - I - J$   
 and  $A - B - C - I - J$  **e** 10 days  
**12 a** 135 **b** 128



- 14 a** 84.8%  
**b** 84.8% of the variation in fuel consumption is explained by the variation in speed.  
**15 a**  $V_n = 88\,000 - 7920n$  **b** \$40 480  
**c** \$39 600  
**16 a**  $t_n = 30\,800 + 3696n$  **b** \$41 888  
**c** 8 years



- b** 1820 m  
**18** 11:44 pm Thursday  
**19 a** There is a moderate, positive correlation between rainfall and tree growth. There is a moderate, negative correlation between temperature and tree growth.  
**b** 28.1% of the variation in tree growth is explained by the variation in rainfall, whereas 39.7% of the variation in tree growth is explained by the variation in temperature. Thus we can conclude that temperature is a better predictor of tree growth than rainfall.

- 20 a i** \$3595.62 **ii** 0.69  
**iii** Positive so overpaid by \$0.69, thus final payment should be \$0.69 less.  
**b** 184 months  
**21 a** \$560 000 **b** \$119 494.12  
**c** 77  
**22** \$2322  
**23 a** (10°N, 145°E) **b** (9°N, 145°E)  
**24 a** (26°S, 175°W) **b** (45°N, 46°W)  
**25 a** \$18 949.65 **b** \$609.65  
**26 a** \$10 **b** \$400  
**27** \$115 346.91  
**28 a** Loan 2: Less interest  
**b** Even with larger repayments of \$1900, loan 2 still pays less interest  
**29 a** 131.2 cents/litre  
**b** 141.1 cents/litre  
**c i** Victoria, slope = 2.1. On average, the price of fuel in Victoria is increasing by 2.1 cents/litre each year.  
**ii** NT, slope = 2.2. On average, the price of fuel in the NT is increasing by 2.2 cents/litre each year.  
**d i** 156.7  
**ii** 169.7  
**e** The difference is predicted to increase over time. The cost of petrol in the NT is already higher than the cost of petrol in Victoria, and the cost is increasing at a higher rate in the NT (on average 2.2 cents/litre each year) than it is increasing in Victoria (on average 2.1 cents/litre each year).  
**30**  $a = 12.5, b = 12.1$   
**31** 2031  
**32 a** Option A as the value is higher after 1 year  
**b** Option A is better as it has a higher value after 1 year and a higher interest rate so it will keep getting larger and larger. After three years, Option A is valued at \$87 842.08 and Option B is valued at \$86 294.16  
**33 a** 1557 km **b** (31°S, 130°E)  
**34 a** 2559 km **b** 2699 **c** 140 km  
**35 a** 5219 km **b** 1913 km  
**36 a** 127 km  
**b** 187 km  
**c** Path must end at C e.g.  
 $B - A - C - B - D - E - F - D - C - F - G - C$   
**37** \$481.77