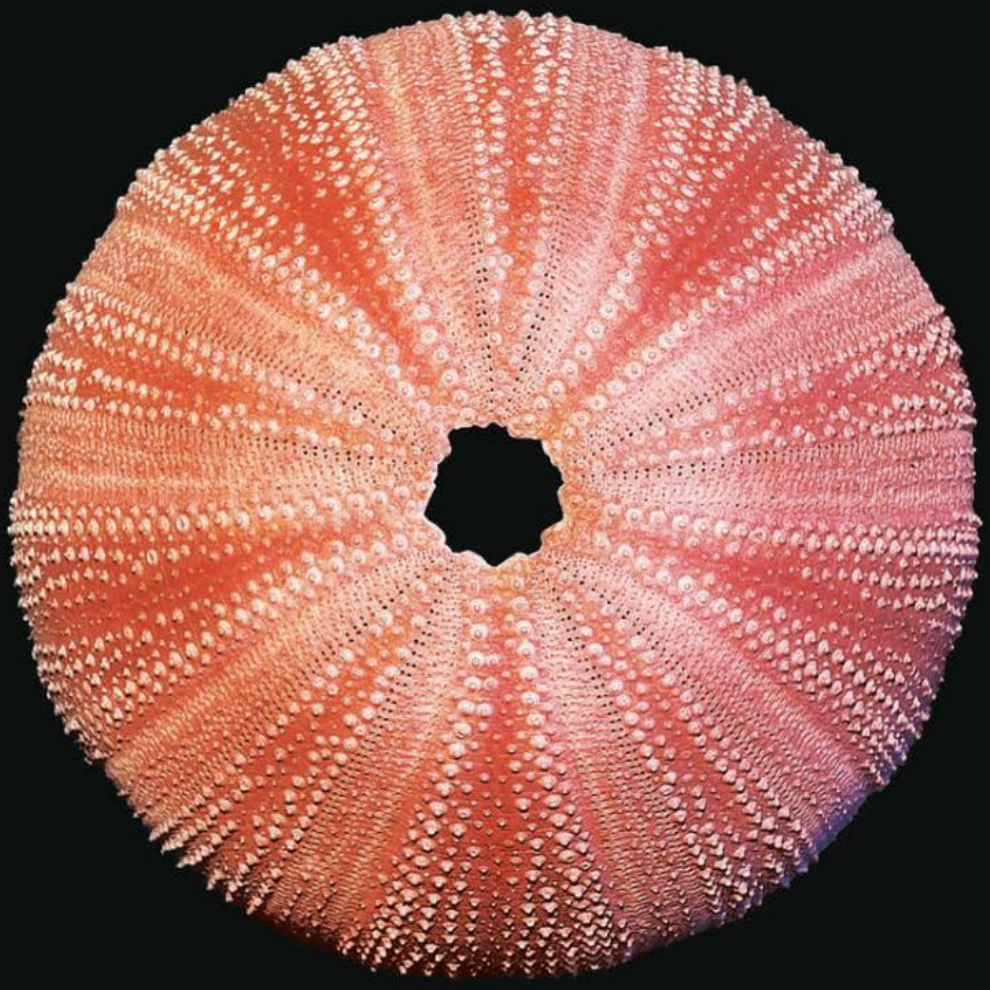
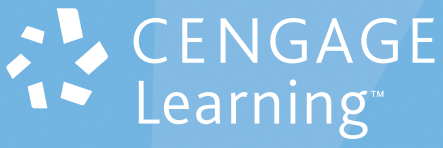
Natalie CARUSO
Karen McMULLEN
Tracey MACBETH-DUNN
Karen O'GORMAN
Sue GARNER
Greg NEAL



NELSON SENIOR MATHS

FOR THE AUSTRALIAN CURRICULUM | **GENERAL 12**



COPYRIGHT NOTICE

Copyright in this work is owned by Cengage Learning Australia ("the work"). A condition of purchase of this electronic version of the work is that you agree to respect the copyright in the work, abide by the Copyright Act 1968 and specifically agree not to transfer, sell, assign, misuse, copy or transmit an electronic or other version of the work to any third party.

Please note: This product is accompanied by a licence (single user, network or adoption) governing the terms and conditions of its use.

This is a legal agreement between the you, (the "Customer") and Cengage Learning Australia Pty Limited (ABN 14 058 280 149) (the "Licensor") which provides the terms and conditions of this non-exclusive licence and the limited warranty for the Product. Use of the Product indicates an acknowledgement that the Customer has read and agreed to be bound by the terms and conditions of this Agreement. If you do not agree to these terms and conditions, return the Product to the place of purchase within 15 days of the date of purchase (with proof of purchase) for a full refund

1. **Licence Grant**
   You do not receive title to the Product. Copyright in the Product (which includes all images, photographs, video, animations, audio, music and text incorporated in the Product, including all of the accompanying printed material) is owned by the Licensor and/or its suppliers and is protected by Australian copyright laws. The Licensor grants you a non-exclusive licence to use the Product subject to the restrictions and terms set out in this Agreement.

2. **A Licence allows you to:**
   Use the Product on your computer. The Customer represents that they shall in no way place the Product in the public domain or in any way compromise our copyright in the Material. You agree to take reasonable steps to protect our copyright.

3. **You may not:**
   Alter, modify, translate, reverse engineer, decompile, or adapt the software or create derivative works based on the Product. Make further copies by any means technological, electronic, digital whatsoever without the written permission of the Licensor. Rent or transfer all or any part of your rights under this Agreement. Remove or alter any copyright or other proprietary notice or label attached to the software.

4. **Termination**
   Any failure to comply with the terms and conditions of this agreement will result in the automatic termination of this licence. Upon termination of this licence for any reason, the Customer must destroy or return to the Licensor all copies of the software and accompanying documentation.

5. **Warranties**
   To the extent permitted by law, the Licensor's liability for any breach of the warranty or any term implied by law into this licence is limited to the lowest cost of replacing the goods, acquiring equivalent goods or having the goods repaired.

NATALIE CARUSO
KAREN MCMULLEN
TRACEY MACBETH-DUNN
KAREN O'GORMAN
SUE GARNER
GREG NEAL

NELSON SENIOR MATHS

FOR THE AUSTRALIAN CURRICULUM | **GENERAL 12**

NELSON
CENGAGE Learning

Nelson Senior Maths for the Australian Curriculum General 12
1st Edition
Natalie Caruso
Tracey MacBeth-Dunn
Greg Neal
Karen McMullen
Sue Garner
Karen O'Gorman

Publishing editor: Alan Stewart
Project editor: Alan Stewart
Editor: Anna Pang
Cover designer: Leigh Ashforth
Text designer: Nicole Melbourne
Art direction: Luana Keays and Danielle Maccarone
Cover image: Getty Images/Howard Mitchell
Permissions researcher: Miriam Allen
Production controller: Julie McArthur
Typeset by: Cenveo Publisher Services

Any URLs contained in this publication were checked for currency during the production process. Note, however, that the publisher cannot vouch for the ongoing currency of URLs.

ACARA Copyright Notice

All material identified by AC is material subject to copyright under the Copyright Act 1968 (Cth) and is owned by the Australian Curriculum, Assessment and Reporting Authority 2014.

For all Australian Curriculum material except elaborations: This is an extract from the Australian Curriculum.

Elaborations: This may be a modified extract from the Australian Curriculum and may include the work of other authors.

Disclaimer: ACARA neither endorses nor verifies the accuracy of the information provided and accepts no responsibility for incomplete or inaccurate information. In particular, ACARA does not endorse or verify that:

- The content descriptions are solely for a particular year and subject;
- All the content descriptions for that year and subject have been used; and
- The author's material aligns with the Australian Curriculum content descriptions for the relevant year and subject.

You can find the unaltered and most up to date version of this material at **http://www.australiancurriculum.edu.au**. This material is reproduced with the permission of ACARA.

© 2014 Cengage Learning Australia Pty Limited

Copyright Notice
This Work is copyright. No part of this Work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means without prior written permission of the Publisher. Except as permitted under the *Copyright Act 1968*, for example any fair dealing for the purposes of private study, research, criticism or review, subject to certain limitations. These limitations include: Restricting the copying to a maximum of one chapter or 10% of this book, whichever is greater; providing an appropriate notice and warning with the copies of the Work disseminated; taking all reasonable steps to limit access to these copies to people authorised to receive these copies; ensuring you hold the appropriate Licences issued by the Copyright Agency Limited ("CAL"), supply a remuneration notice to CAL and pay any required fees. For details of CAL licences and remuneration notices please contact CAL at Level 15, 233 Castlereagh Street, Sydney NSW 2000, Tel: (02) 9394 7600, Fax: (02) 9394 7601
Email: info@copyright.com.au
Website: www.copyright.com.au

For product information and technology assistance,
in Australia call **1300 790 853**;
in New Zealand call **0800 449 725**

For permission to use material from this text or product, please email **aust.permissions@cengage.com**

National Library of Australia Cataloguing-in-Publication Data
Caruso, Natalie, author.
Nelson Senior Maths for the Australian Curriculum General 12 /
Natalie Caruso, Tracey MacBeth-Dunn, Greg Neal,
Karen McMullen, Sue Garner, Karen O'Gorman.

9780170250283 (paperback)
Includes index.
For secondary school age.

Mathematics--Australia--Textbooks.

510

Cengage Learning Australia
Level 7, 80 Dorcas Street
South Melbourne, Victoria Australia 3205

Cengage Learning New Zealand
Unit 4B Rosedale Office Park
331 Rosedale Road, Albany, North Shore 0632, NZ

For learning solutions, visit **cengage.com.au**

Printed in China by China Translation & Printing Services.
1 2 3 4 5 6 7 18 17 16 15 14



ABOUT THIS SERIES

There are 8 books for this series currently. These cover the subjects Essential Mathematics, General Mathematics, Mathematical Methods and Specialist Mathematics.

These books have all been written specifically, from scratch, by a national author team for the Senior Australian Curriculum.

There are worked solutions available for purchase, which accompany each of the General, Methods and Specialist Student Books.

There are also ExamView question banks which can be purchased separately. These will contain a large number of multiple choice questions covering each course; General, Methods and Specialist. An ExamView test generator is also provided, so that teachers can construct their own tests for each topic.

Accompanying each printed textbook is a digital textbook called the NelsonNetBook and a NelsonNet website.

Go to www.nelsonnet.com.au to log in.

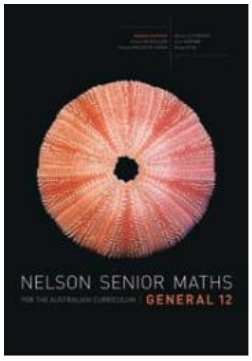
The resources are listed for each chapter. Simply click on the required resource.

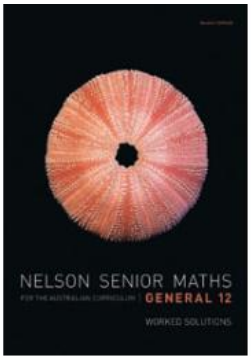
This is a web-based eBook which can be customised to suit your own learning needs.

The icons with the blue NelsonNet logo are 'hotspots'. Click on the icon and the resource will open.

The tools on the vertical toolbar allow you to personalise pages in a variety of ways, including voice recordings, drawings and links to favourite websites. You are also able to zoom in and out.

The tools on the horizontal toolbar allow you to navigate around your eBook and change settings.







CONTENTS

About this series iii
About this book vi
About the authors ix
Abbreviations and symbols x
1 Data analysis 2
1.01 Posing questions and collecting data 4
1.02 Summary statistics 10
1.03 Displaying and describing data 22
1.04 Looking for associations 36
1.05 Relationships between categorical variables 41
1.06 Relationships between numerical and categorical variables 50
1.07 Relationships between numerical variables 63
1.08 Interpretation of scatter plots 73
Summary 84
Chapter review 87
2 Linear models for numerical data 94
2.01 Fitting regression lines 96
2.02 The coefficient of determination 110
2.03 Making predictions 116
2.04 Residuals 122
Summary 137
Chapter review 139
3 Arithmetic and geometric sequences 144
3.01 Arithmetic sequences 146
3.02 Arithmetic series 156
3.03 Modelling an arithmetic sequence 161
3.04 Geometric sequences 167
3.05 Geometric series 175
3.06 Modelling a geometric sequence 181
3.07 Arithmetic and geometric modelling 184
Summary 191
Chapter review 192
Mixed revision for Chapters 1–3 195
4 Recurrence relations 200
4.01 Generating sequences using recurrence relations 202
4.02 Finding a recurrence relation for a given sequence 211
4.03 Recurrence relations for arithmetic sequences 214
4.04 Recurrence relations for geometric sequences 216
4.05 Graphical representation of first order recurrence relations 219
4.06 Long-term steady-state solution of a recurrence relation 232
4.07 Applications of recurrence relations 235
Summary 243
Chapter review 245
5 Undirected graphs 248
5.01 Basic concepts of a graph 250
5.02 Types of graphs 258
5.03 Applications of graphs 263
5.04 The adjacency matrix 268
5.05 Planar graphs and Euler's formula 274
5.06 Walks, paths and trails 280
5.07 Eulerian and semi-Eulerian graphs 283
5.08 Hamiltonian paths and cycles 288
5.09 Trees and spanning trees 292
5.10 Weighted graphs and minimum spanning trees 296
Summary 303
Chapter review 305

9780170250283

6	**Time series analysis**	312
6.01	Constructing and describing time series plots	314
6.02	Smoothing time series data using moving averages	323
6.03	Smoothing time series data using medians	331
6.04	Adjusting for seasonality	337
6.05	Fitting a least-squares regression line to time series data	342
	Summary	348
	Chapter review	350
Mixed revision for Chapters 4–6		355
7	**Loans and investments**	360
7.01	Simple interest review	362
7.02	Hire-purchase	366
7.03	Compound interest	371
7.04	The compound interest formula	376
7.05	Interest graphs	386
7.06	Effective interest rate	399
	Summary	403
	Chapter review	404
8	**Reducing balance loans and annuities**	408
8.01	Reducing balance loans	410
8.02	Using finance solvers for reducing balance loan calculations	416
8.03	Changing the terms of a loan	421
8.04	Annuities	428
8.05	Perpetuities	434
8.06	Compound interest investments with regular deposits	436
	Summary	442
	Chapter review	444
9	**Directed graphs and networks**	448
9.01	Directed graphs, reachability and dominance	450
9.02	Drawing network diagrams	459
9.03	Forward and backward scanning	468
9.04	Locating the critical path	475
9.05	Network flow problems	484
9.06	Bipartite graphs and the assignment problem	490
	Summary	501
	Chapter review	503
Mixed revision for Chapters 7–9		508
Answers		512
Glossary		569
Index		577

ABOUT THIS BOOK



A clear outline of chapter contents is provided.

Links to curriculum content descriptions are included.

A Prior learning worksheet is provided, enabling teachers to assess pre-requisite understanding.

Example 12

For the following sets of variables, state which is the explanatory variable and which is the response variable.

a Test scores (%) and time spent studying (hrs).

b Temperature (°C) and number of icy poles sold at the milk bar.

Solution

a Write the 2 sentences using variable names.	1 Test scores depend on the time spent studying. 2 Time spent studying depends on the test scores.
Decide which sentence makes sense and state RV and EV.	The first sentence makes more sense as the time spent studying may change the test scores. RV = Test scores EV = Time spent studying
Remember the RV is the variable stated first in the sentence which makes the most sense.	
b Write the 2 sentences using variable names.	1 Temperature depends on the number of icy poles sold. 2 Number of icy poles sold depends on the temperature.
Decide which sentence makes more sense and state RV and EV.	The second sentence makes more sense as the temperature may change the number of icy poles sold. RV = Number of icy poles sold EV = Temperature

Examples are clearly set out with reasoning (in xxxxx) and writing (in xxxxx) in separate columns.

Examples are sequenced in a logical order.

There are generally three, and no more than four, examples leading to an exercise.

Examples show solutions and steps which guide students through the use of the TI-Nspire and the CASIO ClassPad CAS calculators.

INVESTIGATION **Drawing your own network.**

Design your own network for a project. Some possible projects are

- building a tree house
- planting a vegetable garden
- baking a cake
- organising a children's party

Write your own activity table and estimated activity times. In your planning, you will need to decide on the sequence of activities, the activities that can be done at the same time as other activities, and the activities that depend on another activity being completed before they can start. Once this is done, the information can be added to the predecessor column in the activity table. You can now draw and label the network.

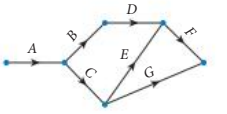
EXERCISE 9.02 Drawing network diagrams

Worksheet 9B

Concepts and techniques

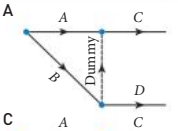
1 Which one of the following is **incorrect** about the directed graph on the right?

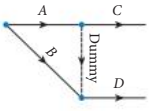
A *B* is preceded by *A*
B *C* is preceded by *F*
C *E* and *G* are preceded by *C*
D *F* is preceded by *D* and *E*
E *D* is preceded by *B*.

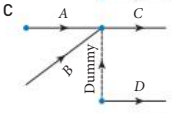


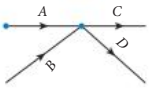
2 The directed graph for the network in the table below is

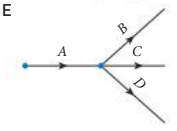
Activity	Immediate predecessor
A	None
B	none
C	*A*
D	*A*, *B*

A


B


C


D


E


9780170250283

Exercise questions are clearly divided into two categories: concepts and techniques and reasoning and communication.

Where appropriate, worksheets are provided for additional practice and consolidation of key concepts.

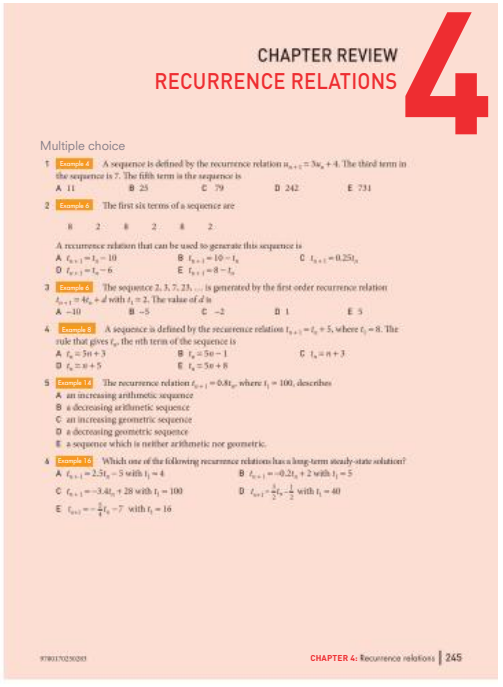
Each chapter contains at least one investigation, providing students with the opportunity to apply their understanding to a practical application.



CHAPTER SUMMARY
RECURRENCE RELATIONS
4

- A difference equation, also called a recurrence relation, is a rule that specifies a particular term in a sequence using the previous term or terms.
- A first order recurrence relation is a rule which specifies how a particular term in a sequence can be found from the previous term in the same sequence.
- A recurrence relation alone is not sufficient for defining a sequence. It is also necessary for a term in the sequence to be given. This is usually the first term but, in some cases, another term in the sequence is given.
- The rule for a first order recurrence relation takes the form $t_{n+1} = rt_n + d$ with $t_1 = a$, where r, d and a are constants.
- The recurrence relation for an arithmetic sequence is $t_{n+1} = t_n + d$ with $t_1 = a$, where d is the common difference and a is the first term.
- The recurrence relation for a geometric sequence is $t_{n+1} = rt_n$ with $t_1 = a$, where r is the common ratio and a is the first term.
- A recurrence relation of the form $t_{n+1} = rt_n + d$ with $t_1 = a$, $r \neq 1$ and $d \neq 0$, defines a sequence that is neither arithmetic nor geometric.
- When graphed, the terms of sequences generated by arithmetic recurrence relations lie on a straight line with gradient d.
  - If $d > 0$, the terms of the sequence lie on a straight line with a positive gradient, so the term values increase as n increases.
  - If $d < 0$, the terms of the sequence lie on a straight line with a negative gradient, so the term values decrease as n increases.
- When graphed, the terms of sequences generated by geometric recurrence relations lie on a curve.
  - If $r > 1$, the term values increase as n increases.
  - If $0 < r < 1$, the term values decrease as n increases.
  - If $r < 0$, the term values alternate in sign.

9780170250283 CHAPTER 4: Recurrence relations | 243



CHAPTER REVIEW
RECURRENCE RELATIONS
4

Multiple choice

1 Example 4 A sequence is defined by the recurrence relation $u_{n+1} = 3u_n + 4$. The third term in the sequence is 7. The fifth term is the sequence is
A 11 B 25 C 79 D 242 E 731

2 Example 6 The first six terms of a sequence are
8 2 8 2 8 2
A recurrence relation that can be used to generate this sequence is
A $t_{n+1} = t_n - 10$ B $t_{n+1} = 10 - t_n$ C $t_{n+1} = 0.25t_n$
D $t_{n+1} = t_n - 6$ E $t_{n+1} = 8 - t_n$

3 Example 6 The sequence 2, 3, 7, 23, ... is generated by the first order recurrence relation $t_{n+1} = 4t_n + d$ with $t_1 = 2$. The value of d is
A −10 B −5 C −2 D 1 E 5

4 Example 8 A sequence is defined by the recurrence relation $t_{n+1} = t_n + 5$, where $t_1 = 8$. The rule that gives t_n, the nth term of the sequence is
A $t_n = 5n + 3$ B $t_n = 5n - 1$ C $t_n = n + 3$
D $t_n = n + 5$ E $t_n = 5n + 8$

5 Example 14 The recurrence relation $t_{n+1} = 0.8t_n$, where $t_1 = 100$, describes
A an increasing arithmetic sequence
B a decreasing arithmetic sequence
C an increasing geometric sequence
D a decreasing geometric sequence
E a sequence which is neither arithmetic nor geometric.

6 Example 16 Which one of the following recurrence relations has a long-term steady-state solution?
A $t_{n+1} = 2.5t_n - 5$ with $t_1 = 4$ B $t_{n+1} = -0.2t_n + 2$ with $t_1 = 5$
C $t_{n+1} = -3.4t_n + 28$ with $t_1 = 100$ D $t_{n+1} = \frac{2}{5}t_n - \frac{1}{2}$ with $t_1 = 40$
E $t_{n+1} = -\frac{3}{4}t_n - 7$ with $t_1 = 16$

9780170250283 CHAPTER 4: Recurrence relations | 245

Each chapter concludes with a detailed chapter summary and a comprehensive review of concepts.

At the end of every three chapters a comprehensive mixed revision set is provided.

ICONS IN THE TEXTBOOK

Example 1 — Link from question to worked example

Prior learning — Prior learning

Worksheets — Worksheet

Spreadsheet — Interactive spreadsheet

Practice quiz — Practice quiz

ABOUT THE AUTHORS

Lead author: **Natalie Caruso**

Natalie is a very experienced Senior Mathematics teacher and former Head of Department. She is currently teaching mathematics at Loreto Mandeville Hall Toorak in Victoria. Natalie has authored a significant number of Mathematics texts and Study Guides. She has enjoyed many years of experience working with CAS calculators and has presented regularly at MAV. Natalie has also been a VCE exam Vetter and marker.

Tracey MacBeth-Dunn

Tracey MacBeth-Dunn (BSc Dip Ed) has been the Head Teacher Mathematics at Denison College Bathurst High Campus since 2005. She has taught in a number of schools and has 19 years experience as a classroom Mathematics teacher. Tracey has taught all levels of Mathematics and has been a HSC marker. She has been working in the publishing industry as a reviewer and writer for the last 5 years.

Karen McMullen

Karen has taught at Killester College, Springvale for nine years since graduating from Monash University in 2005 and currently holds the position of Mathematics Learning Leader. During this time she has completed her Master's in School Leadership (Numeracy) and taught VCE maths for eight years.

Karen O'Gorman

Karen has been a teacher of Mathematics for 17 years, teaching mainly in the senior year levels. She has a Bachelor of Education in addition to a Masters in Educational Leadership and Administration. She is currently entering her third year of employment at Haileybury, which caters from ELC to Year 12 with parallel education from Grade 5 on. Karen has held many varied leadership positions in past years. Currently her main focus is being back in the classroom and writing for Cengage.

Sue Garner

Sue Garner is assistant Head of VCE Mathematics at Ballarat Grammar in Victoria, where she has taught for twenty years, specialising in both Mathematical Methods (CAS) and Specialist Mathematics. Sue has been involved in the assessment of senior examinations in both of these subject areas. Sue has taught in secondary schools in Melbourne, Ballarat and the United Kingdom. She has lectured at the University of Ballarat, La Trobe University and the University of Melbourne and has published extensively.

Greg Neal

Greg has been teaching mathematics for 34 years. He is currently the Mathematics Learning Area Leader at Ballarat High School. Greg has been an assessor for VCAA Mathematical Methods for 10 years. He is a published author of a range of texts for Maths Methods and is extremely proficient with CAS technology. He has been involved with Cengage for a number of years as an author of support material for the Nelson Maths Junior Series.

The **Digital Resources Team** includes Melanie Koetsveld, Kate Sharp and Roger Walter.

ABBREVIATIONS AND SYMBOLS

=	is equal to
≠	is not equal to
≈	is approximately equal to
$\sqrt{\ }$	square root
<	is less than
>	is greater than
≤	is less than or equal to
≥	is greater than or equal to
±	plus or minus
∴	therefore
1 : 650 000	at a scale of 1 to 650 000
5 : 8	ratio of 5 to 8
%	percentage
/	per
LHS	left-hand side
RHS	right-hand side
Q_3	3rd quartile
$\bar{x}$	mean of x values
s_x	standard deviation
π	pi (approximately 3.14159…)
|||	is similar to
∠	angle
△	triangle
36° 27′ 15″	36 degrees 27 minutes and 15 seconds
053°T	a true bearing
N 53° E	a compass bearing

9780170250283

GREEK ALPHABET

Α,	α	alpha	Ι,	ι	iota	Ρ,	ρ	rho
Β,	β	beta	Κ,	κ	kappa	Σ,	σ	sigma
Γ,	γ	gamma	Λ,	λ	lambda	Τ,	τ	tau
Δ,	δ	delta	Μ,	μ	mu	Υ,	υ	upsilon
Ε,	ε	epsilon	Ν,	ν	nu	Φ,	ϕ	phi
Ζ,	ζ	zeta	Ξ,	ξ	xi	Χ,	χ	chi
Η,	η	eta	Ο,	ο	omicron	Ψ,	ψ	psi
Θ,	θ	theta	Π,	π	pi	Ω,	ω	omega





1

TERMINOLOGY

association/relationship
boxplot
categorical data
dependent variable
explanatory variable
discrete data
dot plot
histogram
independent variable
interquartile range
mean
median
mode
outlier
Pearson's correlation coefficient
population
range
response variable
sample
scatter plot
shape of data
standard deviation
stem-and-leaf plot
two-way table
variable

Prior learning

BIVARIATE DATA ANALYSIS

DATA ANALYSIS

1.01 Posing questions and collecting data
1.02 Summary statistics
1.03 Displaying and describing data
1.04 Looking for associations
1.05 Relationships between categorical variables
1.06 Relationships between numerical and categorical variables
1.07 Relationships between numerical variables
1.08 Interpretation of scatter plots
Chapter summary
Chapter review

THE STATISTICAL INVESTIGATION PROCESS

- review the statistical investigation process; for example, identifying a problem and posing a statistical question, collecting or obtaining data, analysing the data, interpreting and communicating the results. (ACMGM048)

IDENTIFYING AND DESCRIBING ASSOCIATIONS BETWEEN TWO CATEGORICAL VARIABLES

- construct two-way frequency tables and determine the associated row and column sums and percentages (ACMGM049)
- use an appropriately percentaged two-way frequency table to identify patterns that suggest the presence of an association (ACMGM050)
- describe an association in terms of differences observed in percentages across categories in a systematic and concise manner and interpret this in the context of the data. (ACMGM051)

IDENTIFYING AND DESCRIBING ASSOCIATIONS BETWEEN TWO NUMERICAL VARIABLES

- construct a scatter plot to identify patterns in the data suggesting the presence of an association (ACMGM052)
- describe an association between two numerical variables in terms of direction (positive/negative), form (linear/non-linear) and strength (strong/moderate/weak) (ACMGM053)
- calculate and interpret the correlation coefficient (r) to quantify the strength of a linear association. (ACMGM054)

FITTING A LINEAR MODEL TO NUMERICAL DATA

- identify the response variable and the explanatory variable (ACMGM055)
- use a scatter plot to identify the nature of the relationship between variables (ACMGM056)

ASSOCIATION AND CAUSATION

- recognise that an observed association between two variables does not necessarily mean that there is a causal relationship between them (ACMGM064)
- identify possible non-causal explanations for an association including coincidence and confounding due to common response to another variable, and communicate these explanations in a systematic and concise manner. (ACMGM065)

THE DATA INVESTIGATION PROCESS

- implement the statistical investigation process to answer questions that involve identifying, analysing and describing associations between two categorical variables or between two numerical variables; for example, is there an association between attitude to capital punishment (agree with, no opinion, disagree with) and sex (male, female)? Is there an association between height and foot length? (ACMGM066) AC

1.01 POSING QUESTIONS AND COLLECTING DATA

The statistical investigation process involves identifying an area to study and posing questions that easily allow for data collection, data analysis, interpretation and communicating results. Before collecting data, we must consider what type of data is required and how it is to be collected.

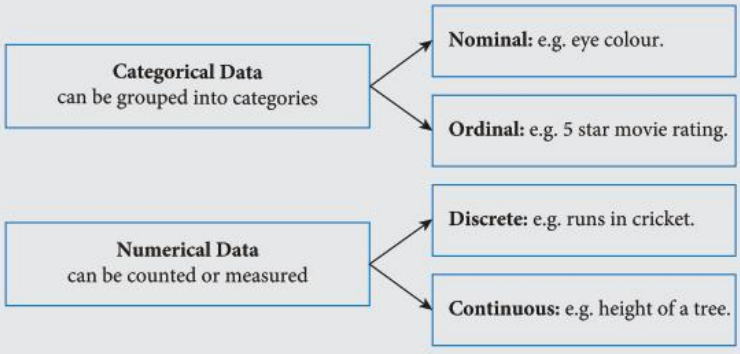
Types of data

A **variable** is a characteristic being observed, counted or measured.

Categorical data provides information about variables that can be grouped into categories. Categorical data can be further classified as nominal (having no order or relationship), or ordinal (having order).

Numerical data provides information about variables that can be counted or measured. Numerical data can be further classified as **discrete** (able to be counted), or **continuous** (able be measured).

IMPORTANT



Example 1

Classify the type of data that each of the following variables represent.

a Number of supporters at a soccer game.
b Types of breakfast cereal.
c Body temperature.
d Finishing position in a horse race.

Solution

a	Number of supporters is numerical data that can be counted: it can only use whole numbers.	Numerical and discrete.
b	Types of breakfast cereal can be put into groups and the order is not important.	Categorical and nominal.
c	Body temperature is numerical data that is measured.	Numerical and continuous.
d	Finishing position involves numbers, but you cannot count or measure these numbers. You can group position into categories such as 1st place, 2nd place, etc., and the order is important.	Categorical and ordinal.

Preparing to collect data

Data can be obtained by means such as making observations, taking measurements or asking questions. A questionnaire is a convenient way of obtaining data. The questions used should be unbiased, easy to understand and written in such a way that only a small range of appropriate responses are possible to allow for easy analysis. It is important to think about the type of data you are collecting, and therefore anticipating the possible responses beforehand.

Example 2

Explain whether the following questions would be appropriate to use on a survey. If not, state how you could make them appropriate.

a What colour eyes do you have?

b It is important to exercise at least three times a week. How often did you exercise this week?

c How many chocolate bars did you consume today?



iStockphoto/bravo1954

Solution

a This question could lead to too many responses, as people's idea of eye colour may vary.

Not appropriate, as you may get many different results.
Provide responses, e.g.

☐ Blue ☐ Green
☐ Brown ☐ Hazel
☐ Another colour

b This question is biased as it provides information that could influence a person's answer.

Not appropriate, as the first sentence may influence a person into providing an answer that is not entirely accurate. Remove the first sentence.

c This question should provide a reasonably small range of answers.

Appropriate due to small range of answers.

Collecting data

Once the data to collect has been identified and questionnaires have been prepared, we can then think about how we will collect our data.

A **population** is the entire group of interest in a study. When information is collected from every member of a group, this is taking a **census**. It is often impractical to collect information from every member of a population, so usually a smaller group or **sample** is selected from the population. It is important that the sample be representative of the population. Two factors affecting this are the size of the sample and the selection process.

Sampling

A sample is said to be biased if one or more groups from the population are over-represented. Samples must be of an appropriate size: too big and it becomes impractical, but too small and you risk your sample becoming biased.

To minimise the chance of bias, a **random sample** is often used. In a random sample, every member of the population has an equal chance of being selected, therefore minimising this risk. The most convenient way to select a random sample is by generating random numbers on a calculator.

Example 3

A sample of 25 needs to be collected from a population of 300 people. Use your calculator to select a random sample of 25.

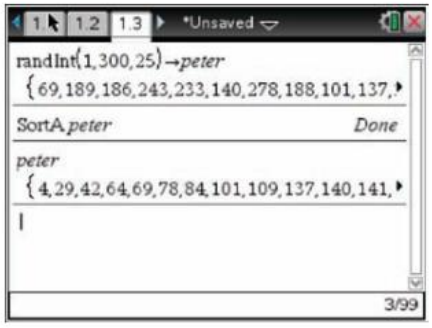
Solution

TI-Nspire CAS

Open a New Document with a Calculator page.
Set seed (only needs to be done once).
Make sure that each student enters a different value or they will all get the same numbers.
Press menu 5: Probability, 4:Random, 6: Seed.



Generate random integers.
menu 5: Probability, 4: Random, 2: Integer *randint(starting number, end number, amount of numbers to generate)* enter.
If you name the list using ctrl var followed by the name of the list (in this case '*peter*'), then you can sort the numbers by typing SortA then the name you used. Press enter. Type the list name again and the list will reappear sorted in ascending order.
Depending on the amount of numbers you may need to arrow across to see them all.

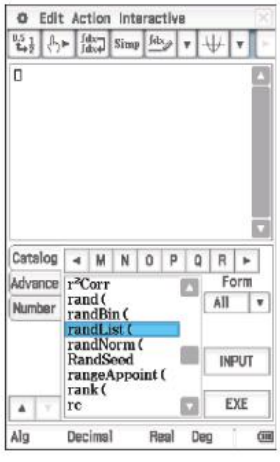


Data should be collected from person 69, 189, etc.

If a random number is repeated, you will need to generate more numbers.

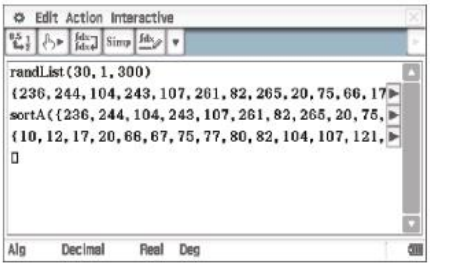
ClassPad

Use the Main $\sqrt{\alpha}$ application.
Press Keyboard, then tap ▼. From this Catalog menu, tap **R** and select **randList(** from the list of functions.



Complete the command on the screen so that you have *randList(amount of numbers to generate, starting number, end number)* where each value is separated by [,]. Press [EXE].

To sort the numbers, access the [Catalog] again, tap **S** and select **sortA** (then copy and paste the list of numbers and press [EXE].
The list will now be sorted in ascending order.
Tap [▶] to view the remaining numbers.



EXERCISE 1.01 Posing questions and collecting data

Concepts and techniques

1 Example 1 **a** Explain the difference between the two types of categorical data and give an example of each.
b Explain the difference between the two types of numerical data and give an example of each.

2 State which type of data each of the following variables represents.
a A person's birthplace.
b The number of televisions in a house.
c A child's height at 5 years of age.
d Favourite novel.
e Movie star rating.
f The amount of time spent doing homework nightly.
g Volume of water in a dam.
h Letter grades for an English essay.
i Speed of a car to the nearest km/h.

3 What type of data are the heights of players in a softball team?
A Categorical nominal
B Discrete numerical
C Continuous numerical
D Random sample
E Categorical ordinal

9780170250283

4 Example 2 Explain whether the following questions would be appropriate to use on a survey. If not, state how you could make them appropriate.
   a What is your highest run score for a single cricket innings?
   b How many times a week do you go to the supermarket?
   c What is your favourite movie?
   d Who do you think has been the best tennis player in the last 10 years?
   e The recommended amount of water to be consumed in one day is 8 glasses. How many glasses of water do you drink in a day?

5 Match each of the survey items **a** to **e** with an appropriate set of responses from **A** to **C**.
   a How many cars are there in your household?
   b Will you go on a holiday this year?
   c Is the MCG bigger than the SCG?
   d Women should have equality in the workplace.
   e The carbon tax will help decrease the damage to the environment.

   A ☐ Yes ☐ No ☐ Don't know
   B ☐ 0 ☐ 1 ☐ 2 ☐ 3 or more
   C ☐ Strongly disagree ☐ Disagree ☐ Neither agree nor disagree ☐ Agree ☐ Strongly agree

6 What is the best way of collecting data to find the three best-selling brands of chocolate?
   A Census
   B Survey
   C Experiment
   D Observation
   E Guess

7 Example 3 Use your calculator to generate the random sample required for the following populations.
   a A sample of 10 from a population of 60.
   b A sample of 15 from a population of 100.
   c A sample of 20 from a population of 120.
   d A sample of 50 from a population of 250.

8 From a population of 500 people, Sue has decided to collect her required data from a sample of 35 people. Identify a random sample that Sue could use.

Reasoning and communication

9 A group of researchers have decided to investigate what variables may influence a person's yearly income. They have decided to focus on age (in years) of the person, education level, hours worked per week and gender.
   a State what type of data each of the five variables represents.
   b Explain why the researchers would not take a census.
   c Explain why and how the researchers would select a random sample.
   d Using your calculator, identify a random sample of 100 that the researchers could use from the population of 10 000 workers.

1.02 SUMMARY STATISTICS

Once data has been collected, it can be analysed. Summary statistics describe the centre and the spread of the data.

IMPORTANT

Calculating measures of centre

Mean – is the arithmetic average of the data values. The mean is not an appropriate measure of centre if there are outliers in the data set.

The method used for calculating the mean depends on whether the data is a list of values, organised in a frequency table or grouped into class intervals in a frequency table.

List of values

$$\bar{x} = \frac{\text{sum of all the values}}{\text{number of values}} = \frac{\Sigma x}{n}$$

Discrete data in a frequency table

$$\bar{x} = \frac{\text{sum of (each value} \times \text{its corresponding frequency)}}{\text{sum of frequencies}} = \frac{\Sigma xf}{\Sigma f}$$

Grouped data in a frequency table

$$\bar{x} = \frac{\text{sum of (the midpoint of each group} \times \text{the frequency of that group)}}{\text{sum of frequencies}} = \frac{\Sigma x_m f}{\Sigma f}$$

where x_m is the midpoint of each class interval

Median – is the middle score when the data values are ordered from smallest to largest. The position of the median is the $\frac{n+1}{2}$th data value, where n is the total number of data values.
The median is not affected by outliers.

The median is the most appropriate measure of centre when the data is skewed or has outliers present.

Mode – the most common data value, or the data value with the highest frequency.
(The mode is seldom used in statistics as a reliable measure of centre).

9780170250283

Example 4

A test resulted in these scores: 5 8 9 5 7 3 6 8 6 5 4 2 9 5 6 7 4

Find the mean, median and mode. Answer correct to one decimal place where necessary.

Solution

Use formula for mean $\bar{x} = \frac{\sum x}{n}$
Give the answer correct to one decimal place.

$$\bar{x} = \frac{5+8+9+5+7+3+6+8+6+5+4+2+9+5+6+7+4}{17}$$
$$= \frac{99}{17}$$
$$= 5.8235...$$
$$\approx 5.8$$

For the median, first place the data in order.

2 3 4 4 5 5 5 5 6 6 6 7 7 8 8 9 9

Locate the position of the median.

The median position is the $\frac{17+1}{2}$th = 9th data value.

2 3 4 4 5 5 5 5 6 6 6 7 7 8 8 9 9
Therefore the median = 6.

When you have an even number of data points the median will be between two values. Add these two values then divide by two in order to find the middle.

The mode is the most frequent number.

Mode = 5

IMPORTANT

Calculating measures of spread

A measure of spread indicates how the data is spread out.

Range = **highest data value – lowest data value**. The range measures the spread of all the data and therefore is not an appropriate measure of spread when outliers are present.

Interquartile range – **IQR** = $Q_3 - Q_1$ where Q_1, the lower quartile, is the median of the lower half of data and Q_3, the upper quartile, is the median of the upper half of data. The IQR is not affected by outliers. The IQR is the most appropriate measure of spread when the data is skewed or has outliers present.

If there is an odd number of data values, the median is not included in either the upper group or the lower group when finding Q_1 and Q_3.

Standard deviation – standard deviation is calculated using $s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$.
The formula is seldom used to calculate the standard deviation; using a CAS calculator is the most efficient method.

Example 5

For the following data calculate the

a range

b IQR

c standard deviation, correct to 1 decimal place.

Stem	Leaf
4	4 7
5	2 6 8
6	1 3 5 5

Key: 4|4 = 44

Solution

a The range is the difference between the maximum and minimum data values.

Range = highest data value – lowest data value

$= 65 - 44$

$= 21$

b To find Q_1 and Q_3 we need to first divide the data set into two groups by finding the median.

Median = $\frac{9+1}{2}$th = 5th value.

Stem	Leaf
4	4 7
5	2 6 [8]
6	1 3 5 5

Median = 58

Locate Q_1 and Q_3 by finding the medians of each of the lower and upper groups.

Each group has four values, so the position of the quartiles is the $\frac{4+1}{2}$th = 2.5th value of each group (between the 2nd and 3rd value of each group).

Lower group: 44 47 52 56

$$Q_1 = \frac{47+52}{2}$$

$= 49.5$

Upper group: 61 63 65 65

$$Q_3 = \frac{63+65}{2}$$

$= 64$

Calculate the IQR using the formula.

$IQR = Q_3 - Q_1$

$= 64 - 49.5$

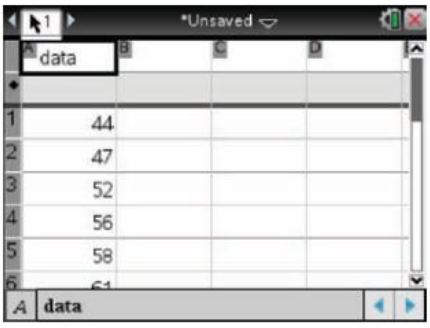
$= 14.5$

c Input the data into a calculator and calculate the one variable statistics.

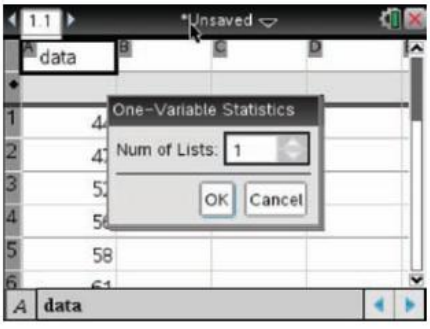
9780170250283

TI-Nspire CAS

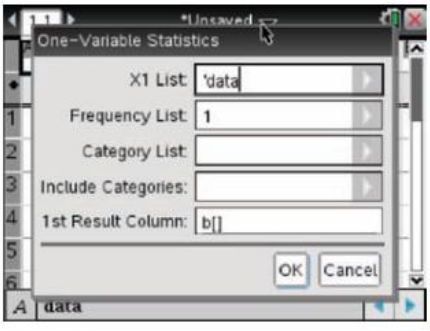
Open a New Document with a Lists & Spreadsheet page.
Name column A, data, then add the data into this list.



Press menu 4: Statistics 1: Stat Calculations 1: One-Variable Statistics.
In the pop up screen that appears, set the number of lists to 1 and press enter.

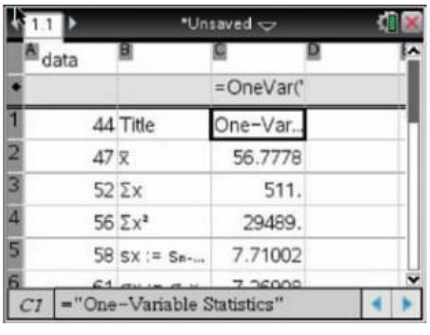


Select 'data for the X1 List, press ▶ then ▼ until 'data is selected then press enter. Leave the frequency list set as 1 then press tab down to OK and press enter.



The standard deviation is s_x.

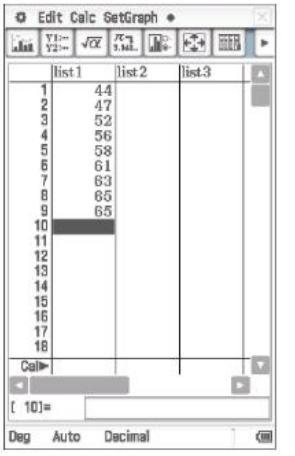
The One-Variable statistics also include the mean, median, quartiles, minimum and maximum values.
NOTE: The mode cannot be found using the TI-Nspire CAS.



$s = 7.7$

ClassPad

Use the **Statistics** application.
Enter the data in list 1.

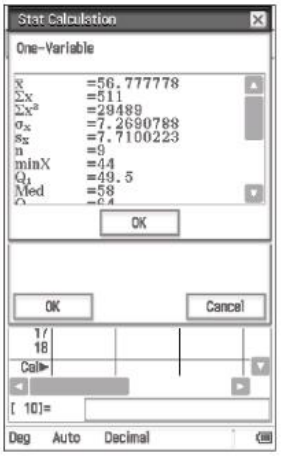


Tap **Calc** then **One-Variable.**
In the pop up screen that appears select list1 for XList then tap OK.



The standard deviation is s_x.

The One-Variable statistics also include the mean, median, mode, quartiles, minimum and maximum values.



Write the standard deviation, correct to 1 decimal place.

$s = 7.7$

9780170250283

Although all the summary statistics can be calculated by hand, using a CAS is a far more effective method, particularly when large sets of data are involved. Although a CAS does not calculate the IQR and range, it does calculate the values required for their formulas.

Example 6

Find the mean, median, range, interquartile range and standard deviation of the following data set.
Answer correct to 2 decimal places where appropriate.

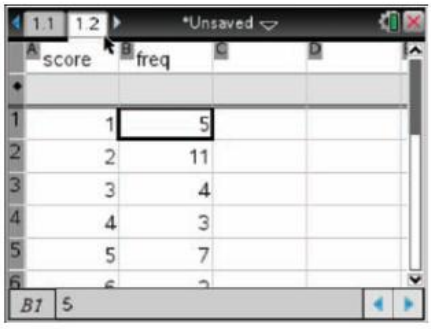
Score	Frequency
1	5
2	11
3	4
4	3
5	7
6	2

Solution

TI-Nspire CAS

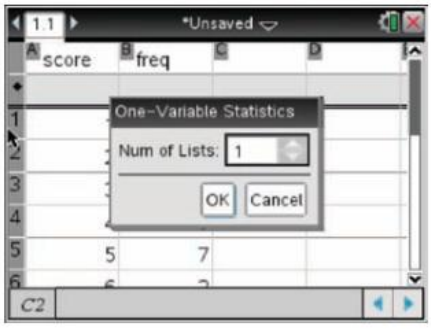
Open a New Document with a Lists & Spreadsheet page. Name column A 'score' and column B 'freq', then enter the scores into the list called 'score' and the corresponding frequencies into the list called 'freq'.

If the frequency table has class intervals, then the midpoint of each interval is entered into column A.

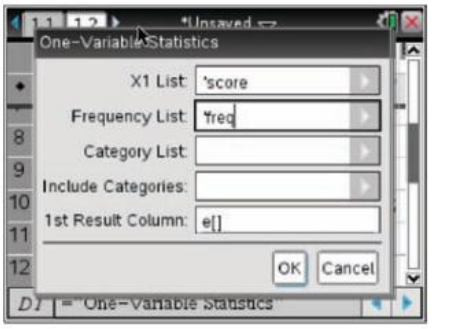


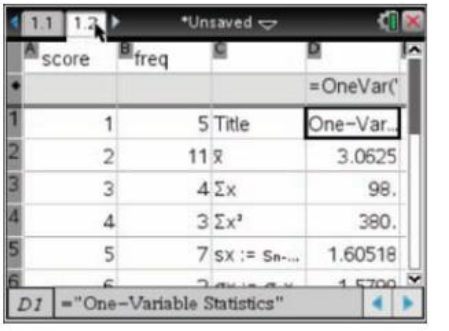
Press menu 4:Statistics 1:Stat Calculations 1:One-Variable Statistics.
In the pop up screen that appears, set the number of lists to 1 and press enter.

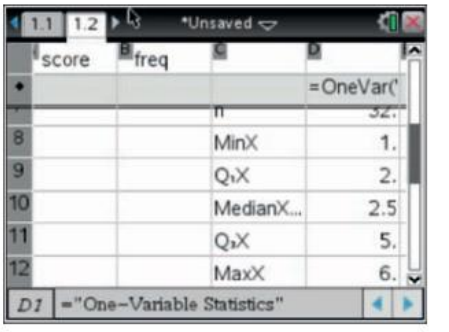
Although you have 2 columns in the table, only one list contains the values of the variable.



Select variable name 'score' for X1 List and select 'freq' for Frequency List, then press tab, down to OK and press enter.



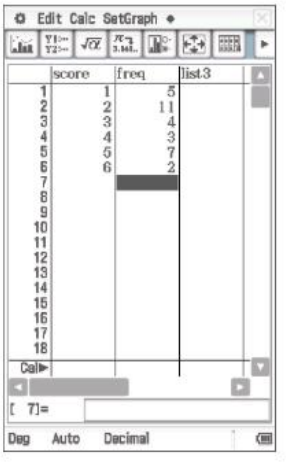




ClassPad

Use the Statistics application. Rename list1 as 'score' by first tapping the list1 cell. To access the letter keyboard, press Keyboard and tap abc.

Follow a similar process to rename the list2 column as 'freq' and enter values.



If the frequency table has class intervals, then the midpoint of each interval is entered into list 1.

Tap **Calc** then **One-Variable.**

9780170250283

In the pop up window that appears, set the **XList** to main\score and **Freq** to main\freq then tap OK.







The mean is $\bar{x}$.
The median is MedianX or Med.
The standard deviation is s_x. Round this answer to two decimal places.

$\bar{x} = 3.06$
Median = 2.5
$s = 1.61$

The range is the difference between the maximum and minimum data values.	Minimum = 1 Maximum = 6 Range = 6 – 1 = 5
IQR = $Q_3 - Q_1$	$Q_1 = 2$ $Q_3 = 5$ IQR = 5 – 2 = 3

Calculating the summary statistics is the first step but we must also interpret their meaning. Therefore, it is important to understand what each statistic tells you about the data.

IMPORTANT

Mean – gives the average of the data being investigated.

Median – determines the middle value of the data set when in order. This means that 50% of the data lies below the median and 50% lies above it.

Mode – describes the most frequent or popular data value.

Range – determines the difference between the highest data value and the lowest data value, therefore describing the total spread of scores.

Interquartile range –describes the spread of the middle 50% of the data.

Standard deviation – describes how data deviates from the mean; therefore a large standard deviation means the data is spread out from the mean and a small standard deviation means the data is clustered around the mean.

IMPORTANT

When interpreting results, it is crucial that you describe your results in context – always explain what the numbers mean in relation to what you have surveyed.

Example 7

Interpret the results from Example **6** if data set is the number of people living in houses on one particular street.

Solution

Interpret the measures of centre and spread in relation to the number of people living in houses on one particular street.	On this particular street, the average number of people in a household is 3.06. It varies somewhat; there is a difference of 5 people between the minimum and maximum household occupancy, although there is only a difference of 3 people in the middle 50% of houses.

EXERCISE 1.02 Summary statistics

Concepts and techniques

1 Example 4 For each of the following data sets, find the mean, median and mode.

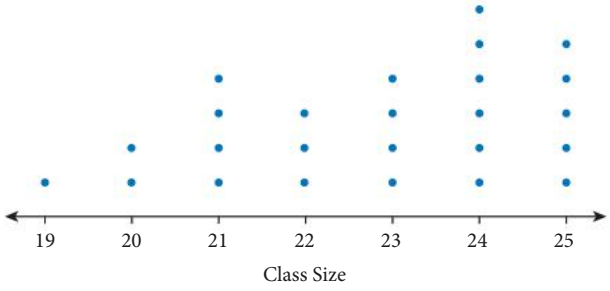
a Cost of two concert tickets.

$300 $400 $600 $440 $300 $700 $250 $580 $260

b Masses of large rocks to the nearest kg.

165 146 178 190 158 147 201 194

c Class sizes for Grade 1 at a local primary school.



d Times, in seconds, of competitors in a go-kart sprint race.

Stem	Leaf
9	1 7 9
10	4 5 6 6 8
11	0 2 3 4 4 5
12	1 2 3 3 3 9
13	2 3 7 7 7
14	6 9
15	0 1

Key: 9|1 = 91

2 Example 5 For each of the data sets in question **1**, find the range, interquartile range and standard deviation.

3 Example 6 Find the mean, median, mode, range, interquartile range and standard deviation of each of the following data sets. Answer correct to 2 decimal places where appropriate.

a 26 27 23 24 27 25 23
19 25 21 22 20 27 21

b 26 5 9 2 6 6 9 7 5
4 23 7 8 6 6 7 4 2

c

Score	Frequency
3	2
4	3
5	5
6	6
7	7
8	6
9	3
10	3
11	3
12	2

d

Score	22	23	24	25	26	27	28	29	30	31	32
Frequency	3	7	4	11	13	10	6	5	3	2	2

e

Score	Frequency
5–9	2
10–14	6
15–19	9
20–24	8
25–29	4

f

Score	Frequency
0–99	3
100–199	7
200–299	13
300–399	10
400–499	4
500–599	2

4 Example 7 Interpret the statistics found in questions **1** and **2** for each of the data sets.

Use the following information to answer questions **5** and **6**.

The test scores (in percentages) of a sample of Biology students are shown in the stem-and-leaf plot.

Stem	Leaf	
5	0 1 2 3 5	**Key: 5|1 = 51**
6	3 4 7 8	
7	1 5 6	
8	6 7	
9	1	

9780170250283

5 The average test score for the Biology students is closest to:
A 13.7% **B** 43% **C** 67% **D** 67.2% **E** 67.3%

6 The most appropriate measure of centre and spread for the Biology scores is:
A mean and range **B** median and range **C** mean and IQR
D median and IQR **E** mean and standard deviation

7 Joshua swam 1 km each morning for 10 days to prepare for a swimming carnival. His times (in minutes) were:
28 24 22 24 25 24 26 26 24 27
a What is his median swim time?
b What is his mean swim time?
c What is the interquartile range of times?
d What is the standard deviation of his times (correct to 1 decimal place)?
e What are the most appropriate measures of centre and spread for his times? Justify your answer.

8 The following data represents the heights (in cm) of students in a Year 12 class.
167 183 175 156 169 159 160 175 180 168 158 159
175 179 163 169 178 172 161 170 157 165 161
a Calculate the measures of centre and spread.
b Explain what each statistic tells you about the data set.

9 The make and prices of cars sold at a second-hand dealership in one week are shown below.
– Honda: Two at $13 000 and two at $14 000
– Holden: Three at $11 000 and two at $13 000
– Ford: One at $40 000 and two at $14 500
– Nissan: Three at $11 000
a How many cars were sold over the week?
b What was the most popular make of car sold during the week?
c Calculate the mean and median prices of the cars sold, correct to the nearest dollar.
d Explain why there is a difference between the mean and median and calculate this difference.
e Hence explain which measure of centre best describes the selling price of the cars for the week.

10 The following data was gathered on income per week in a small country town. Values are in dollars.

423	1234	152	876	912	334	5034	456	877	1205	768	723	612
489	531	332	429	516	199	367	428	687	699	785	321	850
579	666	239	751	648	466	883	991	1024	609	721	598	812
346	753	641	923	468	346	579	843	762	222	1000		

a Put the data into a table with appropriate class intervals.
b Calculate the measures of centre.
c Calculate the measures of spread.
d Interpret the results in relation to income per week.
Be sure to choose the measures of centre and spread that most accurately describe the data.

Class intervals should be chosen so that there are 5–10 classes (excluding outliers).

Reasoning and communication

11 For the data gathered below, calculate appropriate statistics and interpret the results.

Number of pets in a household

4	6	2	1	0	3	1	2	2	3	0	1	2
1	1	1	2	0	2	2	0	2	2	3	0	1
2	3	0	0	0	1	2	1	1	2	2	2	3
4	2	3	3	1	2	3	2	2	1	2	1	1
1	0	19	1	0	3	2	2	2	3	0	0	1
3	1	2	0	0	1	1	1	2	2	2	0	2
0	1	3	0	1	4	1	0	2	2	2	1	3
1	2	0	0	1	1	2	2	2				

12 Arrange the data below into an appropriate frequency table, calculate the appropriate statistics and interpret the results.

Time taken to complete a marathon (in minutes).

210	435	315	233	336	285	145	167	202	346	389	421	275
312	202	397	336	333	198	256	278	384	362	384	346	267
298	307	311	423	366	322	286	250	294	244	312	378	267
289	318	427	276	299	342	278	224	317	242	218		

13 The hair colour of 75 people was recorded and is displayed in the following table.

Hair colour	No. of people
Brown	16
Blonde	25
Black	
Red	
Other	14
Total	75

The researchers found that there were an equal number of people with red and black hair.

a Using the above information, complete the frequency table.

b What is the modal hair colour?

c Why is the mode the only valid measure of centre for this set of data?

1.03 DISPLAYING AND DESCRIBING DATA

There are a number of different displays that can be used to present data. When choosing a display for your data, you must carefully consider the type of data you have collected and the display that best shows the features that you wish to communicate.

Frequency tables

Frequency tables can be used to display both categorical and numerical data. The data values are listed in one column and the corresponding frequencies are displayed in a frequency column.

Ungrouped data

Number of mobile phones in a house	Frequency
0	1
1	3
2	9
3	13
4	12
5	10
6	8

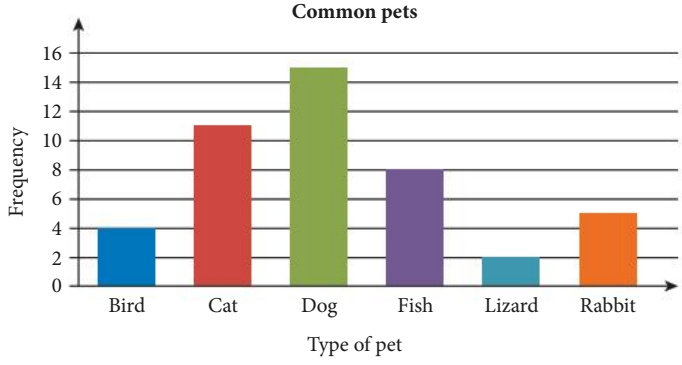
Grouped data

Time taken to get to school (mins)	Frequency
0–<5	4
5–<10	8
10–<15	16
15–<20	21
20–<25	19
25–<30	13
30–<35	7

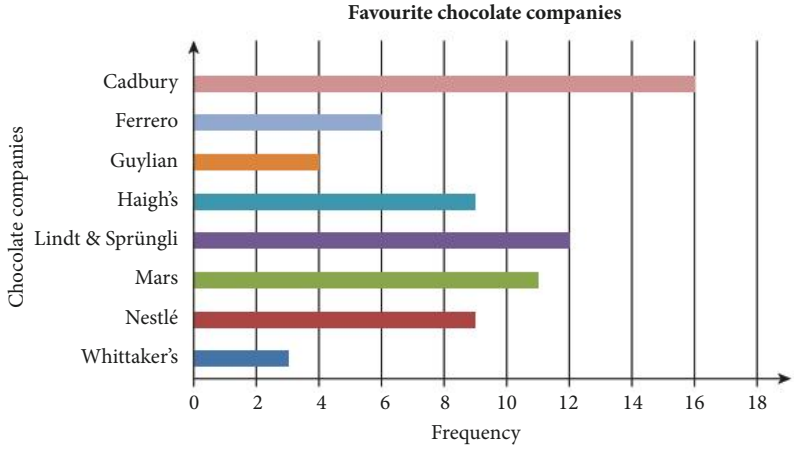
Column and bar graphs

Column and bar graphs are best used to display categorical data. Categories are represented on the horizontal axis with their corresponding frequency on the vertical axis for a column graph, and vice versa for a bar graph. Each category has a column or bar that represents its frequency. A column graph is usually preferred over a bar graph, but when category labels are long or there are many, a bar graph is easier to read.

Column graph

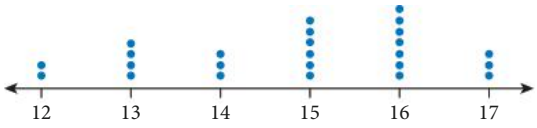


Bar graph



Dot plots

Dot plots are best used to display discrete numerical data where the data values are not too spread out, but can also be used to display categorical data with a small number of categories.



Ages of students at a school function

Stem-and-leaf plots

Stem-and-leaf plots are used to display small sets of numerical data where the data values lie within a small range. In this type of plot, each data value is observable. In an ordered stem-and-leaf plot, the data is ordered, usually from smallest to largest.

Stem	Leaf
0	0 2 9
1	1 6 7 7
2	3 3 7 8 8
3	0 1 3 4 5 5 6
4	0 2 4 6 8 8 9 9 9
5	1 1 5 6 7

Key: 1|1 = 11

Example 8

Construct a stem-and-leaf plot for the following data gathered on the lengths of newborn babies, in centimetres.

50.2	51.3	46.8	47.9	50.9	51.3	52.8	50.6	49.7	49.2	50.0
48.6	52.3	51.9	53.8	50.8	49.3	51.0	52.3	51.3	48.9	49.9
50.1	54.2	47.6	49.2	53.2	48.9	51.2	50.2	50.4		

Solution

Construct a plot that allows for the minimum and maximum values and fill in the plot unordered to ensure that you don't miss any data.

Stem	Leaf
46	8
47	9 6
48	6 9 9
49	7 2 3 9 2
50	2 9 6 0 8 1 2 4
51	3 3 9 0 3 2
52	8 3 3
53	8 2
54	2

9780170250283

Reorder the stem-and-leaf plot and ensure that a key is included.

The benefit of a stem-and-leaf plot is that the data is in order from smallest to largest.

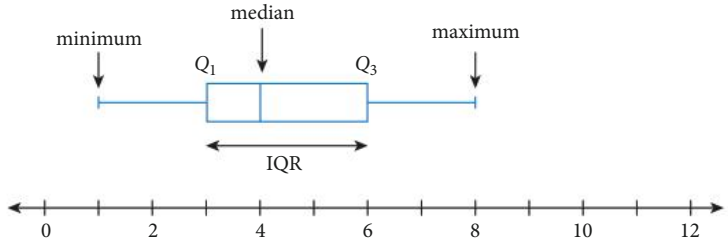
Stem	Leaf
46	8
47	6 9
48	6 9 9
49	2 2 3 7 9
50	0 1 2 2 4 6 8 9
51	0 2 3 3 3 9
52	3 3 8
53	2 8
54	2

Key: 51|6 = 51.6 cm

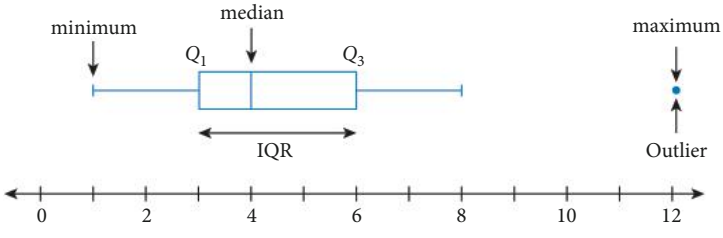
Boxplots

Boxplots are used to display numerical data and provide a visual summary of the data.

A five number summary is needed to construct a boxplot; minimum value, Q_1, median, Q_3, and maximum value.



Boxplots are a useful display when outliers are present within the data. An outlier is represented by a point. A whisker is only extended to the highest or lowest data value that is not an outlier.



Identification of outliers

An **outlier** is a data value which lies outside the interval:

$Q_1 - 1.5 \times \text{IQR} \leq x \leq Q_3 + 1.5 \times \text{IQR}$

Constructing a boxplot using a CAS will also identify outliers but the formula must be used to verify outliers.

Example 9

The following information represents the weekly wages ($) of university students who work part-time.

304	285	331	270	306	294	299	302	295	299
294	315	291	299	314	270	299	284	270	302

a Calculate a five number summary.

b Calculate $Q_1 - 1.5 \times \text{IQR}$ and $Q_3 + 1.5 \times \text{IQR}$

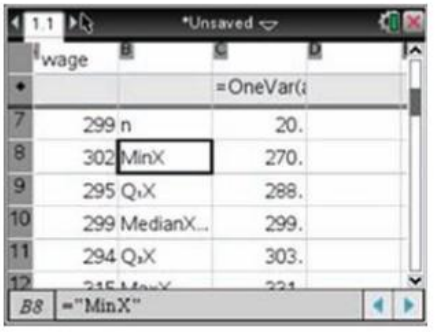
c Are there any outliers in this data set?

d Construct a boxplot that displays any outliers.

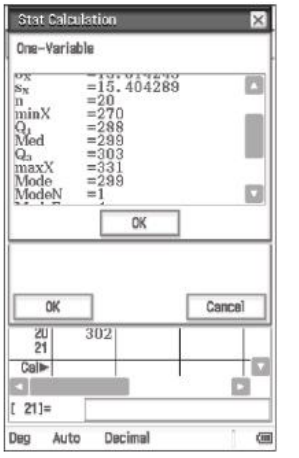
Solution

a Use technology to find the one variable statistics for the data set.

TI-Nspire CAS



ClassPad



Write down the five number summary.

Minimum = 270
Q_1 = 288
Median = 299
Q_3 = 303
Maximum = 331

b Calculate the interquartile range first.

$IQR = Q_3 - Q_1$
$= 303 - 288$
$= 15$

Find $Q_1 - 1.5 \times IQR$ with $Q_3 = 303$ and IQR = 15.

$Q_1 - 1.5 \times IQR = 288 - 1.5 \times 15$
$= 265.5$

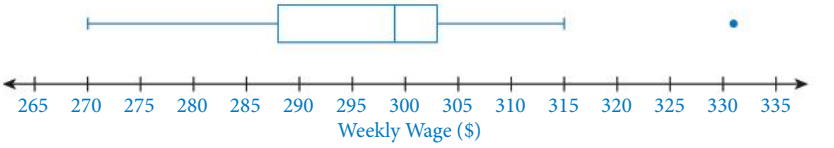
Find $Q_3 + 1.5 \times IQR$ with $Q_1 = 288$ and IQR = 15.

$Q_3 + 1.5 \times IQR = 303 + 1.5 \times 15$
$= 325.5$

c An outlier is any value below 265.5 or above 325.5.

331 is an outlier as it is larger than 325.5.

d Select an appropriate scale and construct the boxplot.
Since 331 is an outlier, the right whisker extends to 315 which is the next highest data value.



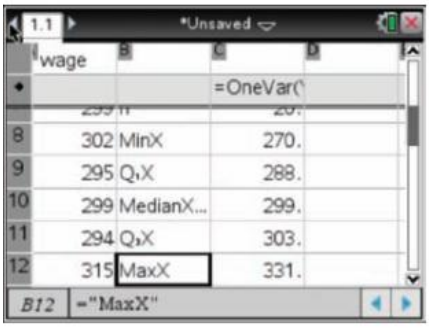
9780170250283

Using a CAS calculator to construct a boxplot

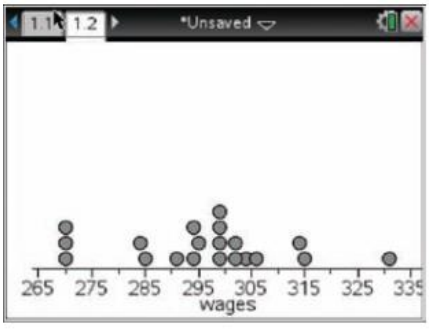
A CAS calculator can be used to construct a boxplot for the data in Example **9**.

TI-Nspire CAS

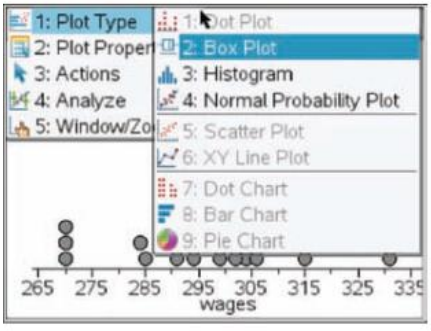
Using a New Document, enter data into the Lists & Spreadsheet page.



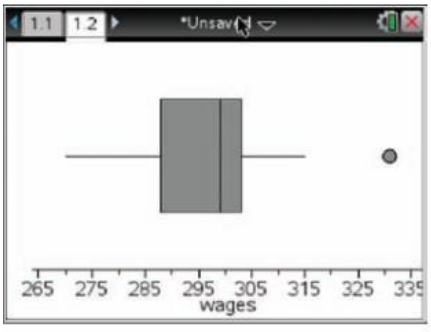
Add a Data & Statistics page, then press tab and select wage as the x variable.
A dot plot will appear.



To change it to a boxplot, press menu 1:Plot Type, 2:Boxplot.

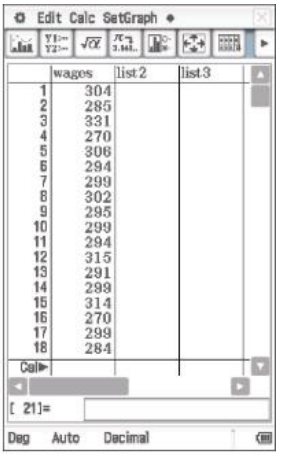


As you move the cursor over the graph the important values will appear.

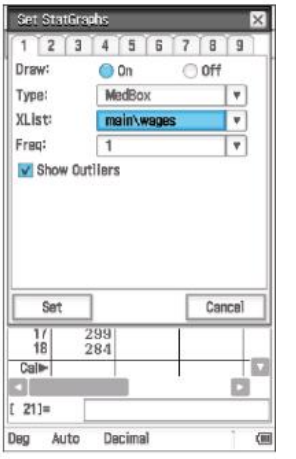


ClassPad

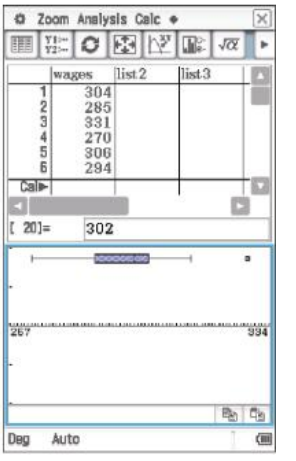
Using the Statistics application, enter the data.



Tap **SetGraph**, making sure **StatGraph1** only is ticked and **Stat WindowAuto** under the menu is on.
Tap **Setting.**
For **Type** select MedBox
For **Xlist** select main\wages
Make sure that the box for Show Outliers is checked.

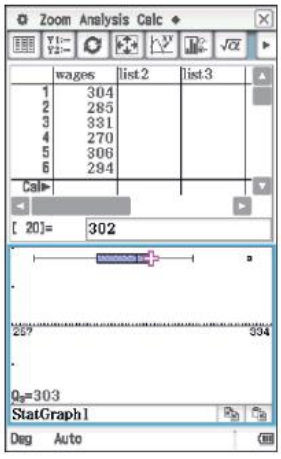


Tap **Set** then tap .



9780170250283

Tap on the graph screen and then tap **Analysis** followed by **Trace**. Using the right and left arrow keys enables you to read important values from the graph.

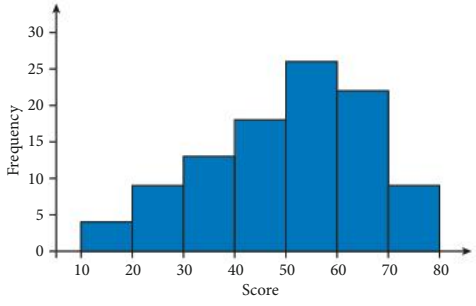


Histograms

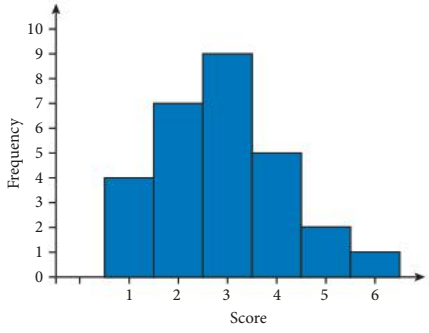
Histograms are typically used to display continuous numerical data grouped into class intervals.

They differ from column graphs in that there are no spaces between the columns (because the data is usually continuous), but there is a half space at the beginning of the graph. Columns in a histogram displaying grouped data sit in between the scale marks on the horizontal axis, rather than centred on them as in a column graph. Histograms can also be used to display discrete numerical data, but they are best reserved for continuous data.

Continuous data



Discrete data



Example 10

Use the following data to construct a histogram with appropriate class intervals.

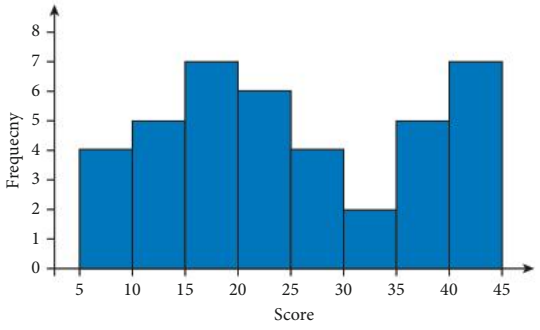
16 26 21 23 31 44 27 29 42 42 18 38 32
8 11 6 22 19 24 29 43 17 38 39 41 35
7 24 19 22 12 16 44 10 36 43 10 17 11 5

Solution

Create a frequency table with appropriate class intervals. A range of 40 would be best represented by class intervals of 5.

Score	Frequency
5–<10	4
10–<15	5
15–<20	7
20–<25	6
25–<30	4
30–<35	2
35–<40	5
40–<45	7

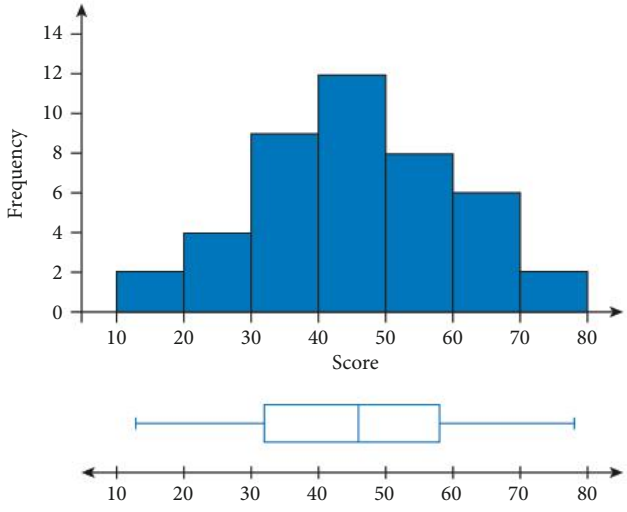
Draw the histogram, remembering to label axes.



Using a visual representation of data to describe the shape of a distribution

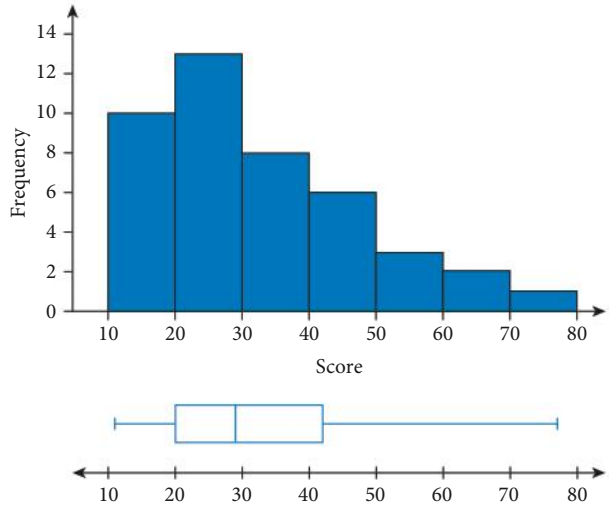
A graphical representation of data shows whether the data is symmetrical or skewed.

Symmetrical

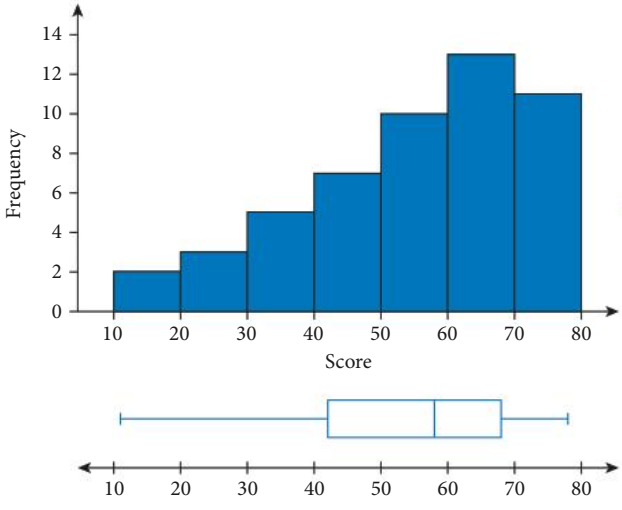


9780170250283

Positively skewed



Negatively skewed



If the IQR is roughly half the range, this indicates that there **may** not be any outliers and that the distribution is roughly symmetric. The likelihood of this increases if the mean and median are similar in value. Note that this is not a guarantee, just an indication.

Example 11

The following data was collected on test scores (out of 100) in a biology class.

83	59	74	48	68	92	81	90	87	15	83	62	75
49	69	94	84	63	76	51	72	95	85	64	78	56
72	97	33	64	78	56	85	98	34	67	79	59	86
99	37	68	79	89	86	73	42	23	43	29	82	91
88	21	91	89									

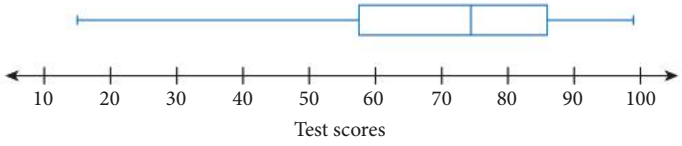
a Use the data to construct an appropriate display.

b Describe the **shape of the data**.

c If a pass is awarded to the top 75% of students, what is the lowest pass score?

Solution

a A histogram or boxplot would be appropriate, but as part **c** asks for the top 75%, a boxplot would make this task easier. Construct a boxplot.



b The tail of the boxplot is to the left. — The data is negatively skewed.

c If a pass is awarded to the top 75% of students, then to pass you must get a test score above the lower quartile. — The lower quartile is 57.5. The lowest score above this is 59, which is the lowest pass score.

EXERCISE 1.03 Displaying and describing data

Concepts and techniques

1 Example 8 This unordered stem-and-leaf plot represents the number of points scored per match by the Sharks in a football season.

Stem	Leaf
4	5 3 9
5	7 2 0 8
6	4 7 8 5 1 2
7	2 9 3 0
8	9 4 2

Key: 4|5 = 45

a Display the data using an ordered stem-and-leaf plot.

b How many matches were played in the season?

c What was the Shark's highest score for a match?

d For what percentage of matches did the Sharks score below 56 points?

9780170250283

2 The heights (in centimetres) of new applicants to a modelling agency are:

168	182	187	185	178	176	174	163	164
165	189	185	165	162	173	164	173	178
172	186	189	175	176	180	170	168	165

a Finish constructing the stem-and-leaf plot on the right.

b How many applicants were taller than 175cm?

c What percentage (correct to 1 decimal place) of the applicants were shorter than 170cm?

Stem	Leaf
16	
16*	
17	
17*	
18	
18*	

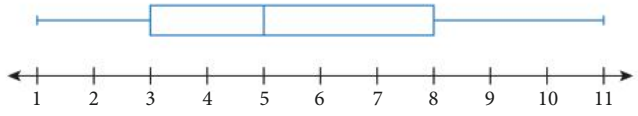
Key: 16|2 = 162

3 For the following boxplots, state

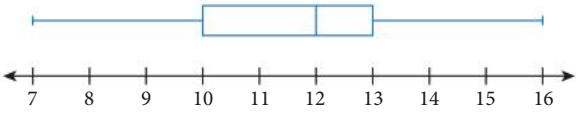
i the five-number summary

ii the values between which the middle 50% of the data lies between.

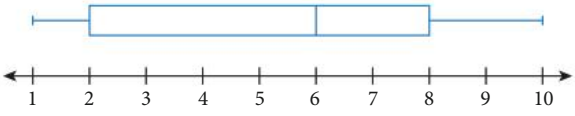
a



b



c

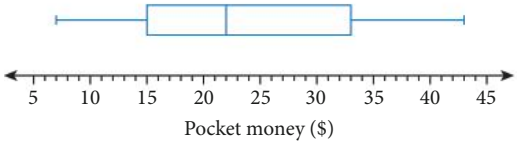


4 Example 9 Construct a boxplot for each of the following data sets, calculating and marking in any outliers.

a	45	50	44	54	37	44	15	50	41	52
	38	26	37	42	48	39	46	49	43	64
b	104	88	110	99	40	156	96	97	86	
	105	79	77	89	93	95	110	99	115	
c	15	22	25	25	25	27	17	27		
	23	20	24	26	29	27	21	25		

The following information relates to questions **5** and **6**.

This boxplot represents the amount of pocket money in dollars earned by a sample of 48 children.



5 What percentage of children earned less than $15?

A 10% **B** 15% **C** 25% **D** 50% **E** 75%

6 How many children earned $22 or more of pocket money?

A 22 **B** 24 **C** 25 **D** 42 **E** 50

7 Example 10 Construct a histogram for each of the following data sets.

a The number of people recorded in a 7-Eleven over a 48-hour period.

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

b The ages of people recorded at a skate rink.

22	17	26	21	18	42	49	8	10	11	13	53	
67	51	45	36	39	12	8	5	6	19	22	17	
22	29	13	14	15	14	13	77	52	53	14	19	
69	10	11	12	13	11	12	11	46	42	42	3	36

c The heights of Year 12 students (in cm) at a medium-sized school.

162	185	155	163	167	172	176	179	162	159	146	163
179	172	171	170	168	169	177	172	186	172	171	168
166	165	164	173	172	176	172	172	168	161	160	173
160	159	172	180	183	179	172	188	162	157	145	169
168	167	163	172	173	176	178	177	162	161	175	175
169	168	173	172	173	191	172	176	183	172	165	164
169	177	176	162	176	167	169	175				

8 Data was gathered on how much people spend on their phone bills every month.

99	45	20	30	39	49	50	59	69	129	19	29	79
39	49	49	49	25	30	79	99	59	69	49	39	15
0	49	79	69	29	39	49	59	50	30	80	49	79
59	19	49	60	100	49	79	30	49	29	89.		

a Draw a histogram with appropriate class intervals.

b What percentage of people spend $60 or more per month on their phone?

c What percentage of people spend between $20 and $60 per month on their phone?

d What is the price range of the middle 50% (the interquartile range) of money that people spend per month on their phones?

9 Example 11 Four different university classes completed the same test. The test scores (in percentages) are shown below. For each class:

i use the data to construct an appropriate display

ii describe the shape of the data.

iii If a credit is awarded to the top 25% of students (of each class), what is the highest non-credit score?

a

27	85	61	23	27	98	42	68	67	53	81	36
57	68	49	23	75	66	53	47	52	64	46	58
79	21	66	73	51	58	42	63	50	50	73	46
38	63	53	79								

b

47	89	78	62	88	51	61	99	91	87	82	83
64	85	89	87	82	76	72	76	73	71	54	91
93	94	82	86	87	84	86	72	77	73	84	86
81	90	92	57								

c

42	94	23	52	47	42	49	31	25	59	51	27
31	33	12	35	49	48	42	31	39	52	51	39
38	42	49	47	22	45	28	36	37	51	32	37
38	41	51	37								

9780170250283

d

88	72	99	91	82	83	87	70	100	98	94	92
81	88	89	78	72	82	72	73	95	76	74	82
81	93	72	98	88	82	84	76	77	91	93	95
72	84	88	97								

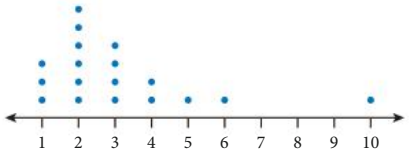
10 Which of these displays could be described as positively skewed with a possible outlier?

A

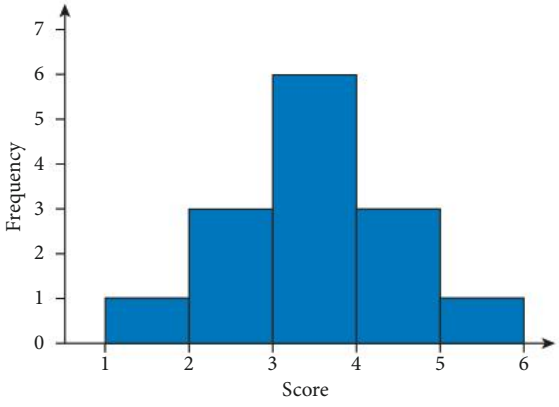
Stem	Leaf
7	1 5 6 8 9 9
8	0 0 2 3 4 6 8 9
9	0 2 5 7 9
10	2 3 5 6 7 7 8

Key: 7|9 = 79

B



C

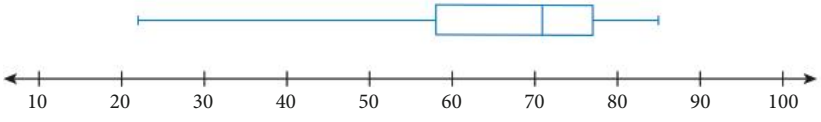


D

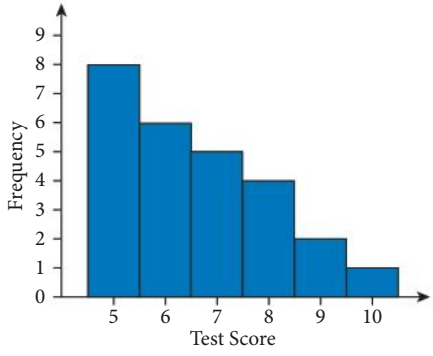
Stem	Leaf
12	3
13	
14	
15	4 7 9 9 9
16	0 2 3 7 8 8
17	0 0 2 4 4 5 5 7 7 7 9
18	3 9
19	0 1 2

Key: 12|3 = 123

E



11 The following histogram displays the results of a Year 12 History Quiz.



Score	Frequency
5	
6	
7	
8	
9	
10	
Total	

a How many students did the History quiz?

b Describe the shape of the data.

c Using the histogram complete the following frequency table.

d What percentage of students achieved a score greater than 8 for the History quiz? Give your answer correct to 1 decimal place.

Reasoning and communication

12 The heights of 30 students (to the nearest cm) were measured and recorded as follows.

116 125 132 119 154 143 127 136 131 126
131 138 121 139 141 135 130 128 122 132
125 127 140 127 135 147 132 143 134 131

Height	Frequency
115–119	
120–124	
125–129	
130–134	
135–139	
140–144	
145–149	
150–154	
Total	30

- **a** Classify the type of data being investigated.
- **b** Complete the following frequency table.
- **c** Hence construct a histogram.
- **d** Sort the original data from the shortest to the tallest student.
- **e** Construct a stem-and-leaf plot to represent the data.
- **f** Describe the shape of the data.
- **g** What percentage of the students are taller than 135cm?
- **h** Construct a boxplot of the data, calculating and marking in any outliers.
- **i** What range of heights are the shortest 25% of students?
- **j** Calculate the measures of centre and spread.
- **k** Describe the data using the information gathered from statistical measures and displays.

13 The number of stray dogs collected by the RSPCA in several suburbs are shown in the table.

No. of dogs	5	6	7	8	9	10
No. of suburbs	14	15	17	26	28	31

- **a** How many suburbs were stray dogs collected from?
- **b** Construct a histogram to represent the data.
- **c** Describe the shape of the data.
- **d** What is the mean number of dogs collected per suburb?
- **e** What is the standard deviation of the number of dogs? State your answer correct to 2 decimal places.
- **f** Are the answers to part **d** and **e** appropriate summary statistics for this data? Justify your answer.

14 Data was gathered on the number of items sold during May at an eBay store.

12 16 17 5 19 14 52 26 31 19 18 16 15
13 9 11 16 14 21 28 14 19 17 18 22 20
10 18 19 9 21

- **a** Calculate the measures of centre and spread.
- **b** Draw a histogram with appropriate class intervals.
- **c** Are there any outliers? Explain.
- **d** Draw a boxplot to represent the data.
- **e** Describe the shape of the data.
- **f** What percentage of daily sales are over 20 items?
- **g** How many items were sold in total on the five most successful days?
- **h** Use your answers to **a** – **e** to describe the data.

1.04 LOOKING FOR ASSOCIATIONS

Rather than just analysing data from a single variable (univariate data, e.g. height, gender, etc.), often we want to compare data from two variables and look for relationships between them (bivariate data, e.g., height vs gender).

When comparing two variables, it is first important to identify each variable as either the **explanatory variable** (or **independent variable**) or the **response variable** (or **dependent variable**).

IMPORTANT

An explanatory (independent) variable (EV) is not affected or changed by the other variable.

A response (dependent) variable (RV) is affected or changed by the other variable.

When deciding on which variable is the explanatory variable and which is the response variable, the following sentence may help.

Replace RV and EV with the variable names, trying each variable in both positions. Select the sentence that makes the most sense, and hence determine the explanatory and response variables.

IMPORTANT

The Response Variable (RV) depends on the Explanatory Variable (EV).

Example 12

For the following sets of variables, state which is the explanatory variable and which is the response variable.

a Test scores (%) and time spent studying (hrs).

b Temperature (°C) and number of icy poles sold at the milk bar.

Solution

a Write the 2 sentences using variable names.

1 Test scores depend on the time spent studying.

2 Time spent studying depends on the test scores.

Decide which sentence makes sense and state RV and EV.

Remember the RV is the variable stated first in the sentence which makes the most sense.

The first sentence makes more sense as the time spent studying may change the test scores.

RV = Test scores

EV = Time spent studying

b Write the 2 sentences using variable names.

1 Temperature depends on the number of icy poles sold.

2 Number of icy poles sold depends on the temperature.

Decide which sentence makes more sense and state RV and EV.

The second sentence makes more sense as the temperature may change the number of icy poles sold.

RV = Number of icy poles sold

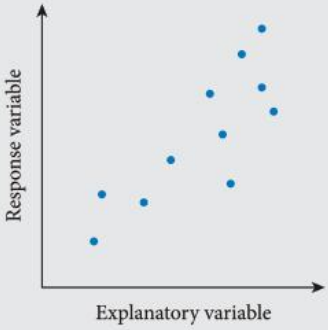
EV = Temperature

IMPORTANT

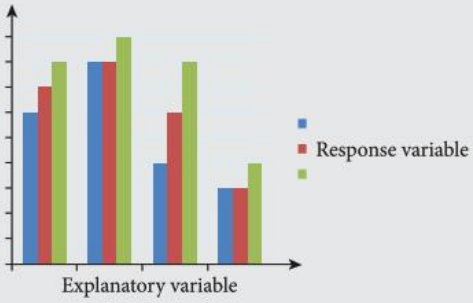
The RV and EV must be identified before drawing a graph.

When comparing numerical variables, the explanatory (or independent) variable belongs on the x-axis (or independent axis). The response (or dependent) variable belongs on the y-axis (or dependent axis).

When comparing categorical variables, the explanatory (or independent) variable is best placed on the x-axis (or independent axis). The response (or dependent) variable then forms part of the legend.



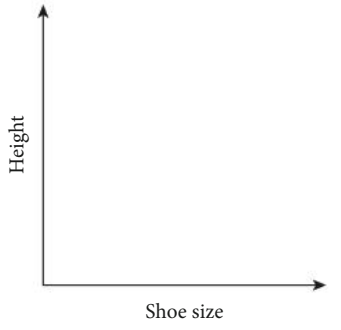
Side-by-side column graphs may have the response variable on the x-axis if it makes the comparison of the variables easier.



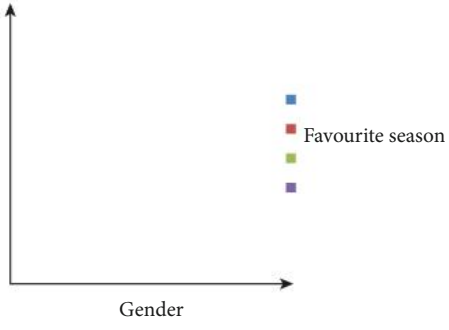
Example 13

For the following sets of axes, determine if the variables have been placed on the correct axis, and redraw the axes if necessary.

a



b



Solution

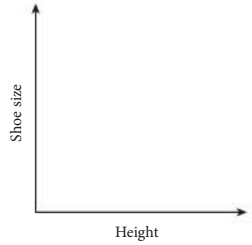
a The variables are numerical, so the explanatory variable belongs on the x-axis and the response variable belongs on the y-axis. Write out the two sentences and decide which makes more sense.

Height depends on shoe size.
Shoe size depends on height.
The second sentence makes more sense as a change in height is likely to create a change in shoe size.
RV = shoe size
EV = Height

9780170250283

Height belongs on the x-axis and shoe size belongs on the y-axis, which is not what is shown here.

Redraw the axis so that it displays the relationship correctly.



b The variables are categorical, so the explanatory variable belongs on the x-axis and the response variable belongs in the legend. Write out the two sentences and decide which makes more sense.

Favourite season depends on gender.
Gender depends on favourite season.
The first sentence makes more sense as gender might affect favourite season.
RV = favourite season
EV = gender

Gender belongs on the x-axis and favourite season is the legend, which is what is shown here.

EXERCISE 1.04 Looking for associations

Concepts and techniques

1 Example 12 For the following sets of variables, state which is the explanatory variable and which is the response variable.

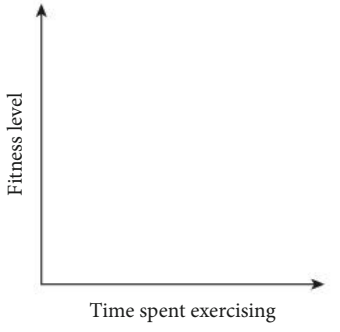
- **a** Temperature (°C) and amount of rainfall (mm)
- **b** Arm length and height
- **c** Time (mins) and distance covered (m)
- **d** Days spent dieting and weight loss (kg)
- **e** Hours of work and wage ($)
- **f** Yield of vegetables (kg) and size of a vegetable patch (m^2)
- **g** Age (years) and height (cm)
- **h** Weekly wage ($) and weekly budget ($)
- **i** Caffeine consumption and heart rate
- **j** Number of skiers and amount of snow

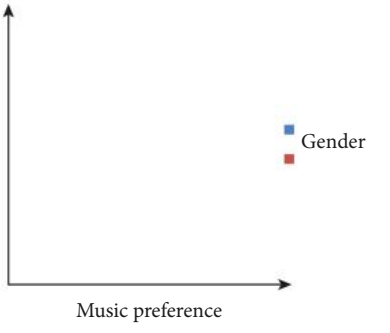
2 Explain why, when time is a variable, it is nearly always the explanatory variable.

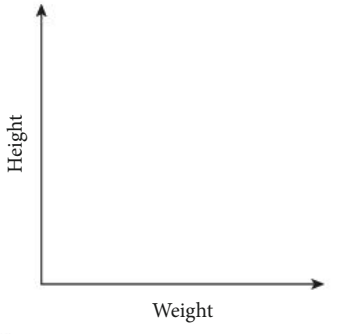
3 For the following sets of variables, determine if there is likely to be a relationship, and if so, identify the explanatory and response variables.
   a Basketball ability and height (cm)
   b Gender and hair colour
   c Size of house and time spent cleaning per week
   d Number of people in a house and number of rooms in a house
   e Age (years) and eye colour
   f Smoking status and health

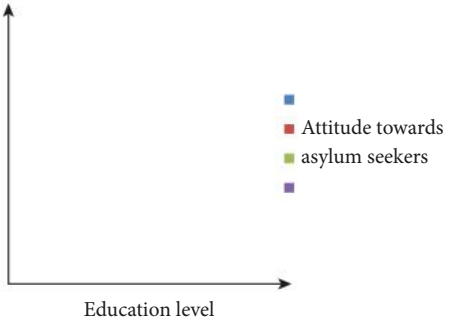
4 In a study on houses, the explanatory variable was identified as the amount of money spent on renovations. The response variable is most likely to have been:
   A the amount of renovations that were done
   B the time spent renovating the house
   C the amount of money the house owners have
   D the amount of money the house sells for
   E the number of builders working on the renovations

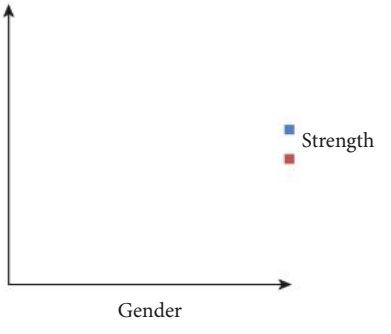
5 Example 13 For the following sets of axes, determine whether the variables have been placed on the correct axis.

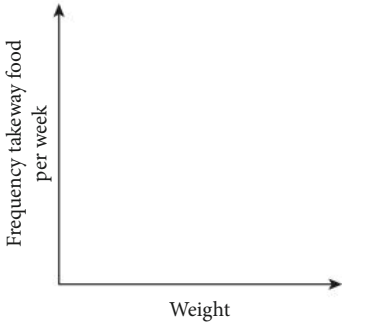
a


b


c


d


e


f


9780170250283

6 Draw and label a set of axes for the following pairs of variables that are likely to have a relationship.
   a Marital status and film preference
   b Number of pets and time spent caring for pets
   c Hair length and gender
   d Opinion on a current issue (for, against) and party preference (Labor, Liberal)
   e Height (cm) and number of children
   f How long it takes to read a book and how many pages a book has.

7 A study was done that looked at the relationship between the number of soft drinks consumed per week and the waist circumference of students. Which of the following statements is correct?
   A The number of soft drinks consumed per week is the explanatory variable and therefore belongs on the *y*-axis.
   B The waist circumference is the explanatory variable and therefore belongs on the *x*-axis.
   C The waist circumference is the response variable and therefore belongs on the *x*-axis.
   D The number of soft drinks consumed per week is the response variable and therefore belongs on the *y*-axis.
   E The number of soft drinks consumed per week is the explanatory variable and therefore belongs on the *x*-axis.

1.05 RELATIONSHIPS BETWEEN CATEGORICAL VARIABLES

When information about a group can be presented as two different frequency tables, a **two-way table** can be used instead. This is particularly valuable when comparing categorical variables. The purpose of investigating two variables is generally to determine if there is a **relationship** or **association** between them. When using a two-way table to investigate an association between two variables, the explanatory variable has frequencies shown in the columns and the response variable has frequencies shown in rows.

	Explanatory variable		
Response variable	**Characteristic 1**	**Characteristic 2**	
Characteristic 1			Row sum
Characteristic 2			Row sum
Characteristic 3			Row sum
	Column sum	Column sum	Total

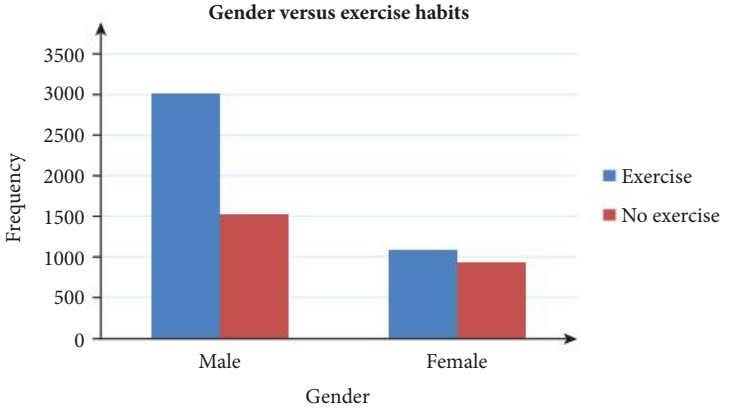
Once a two-way table has been constructed, the data contained within can be displayed as a side-by-side column graph or a segmented column graph. Below is an example of a two-way table and the matching graphical displays showing the exercise habits of males and females.

It is important to note that whilst we only ever compare two variables at a time, each variable can have multiple characteristics. In this table, the explanatory variable has two characteristics and the response variable has three characteristics.

Two-way table

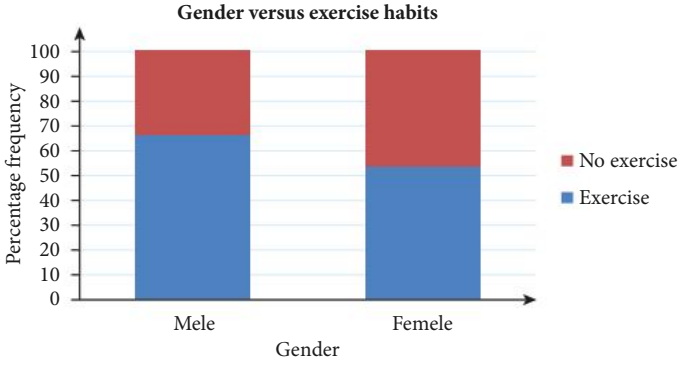
	Gender		
Exercise habits	**Male**	**Female**	**Total**
Exercise	3028	1084	4112
No exercise	1532	946	2478
Total	4560	2030	6590

Side-by-side column graph



Remember that the explanatory variable is expressed on the *x*-axis when comparing categorical variables. This means that the columns of a segmented column graph should match the columns of the two-way table.

Segmented column graph



The total columns in a two-way table are essential when analysing the data, particularly when percentages are required. Total columns should never be displayed in graphs.

Information can be extracted from a two-way table about both variables. For example, from the table above, we can see that:

- 4560 males were surveyed compared to 2030 females. Of these 6590 people, 4122 exercise compared to 2478 who do not
- $\frac{3028}{4560} \times 100\% \approx 66\%$ of males exercise and $\frac{1084}{2030} \times 100\% \approx 53\%$ of females exercise
- alternatively, of those who exercise, $\frac{3028}{4112} \times 100\% \approx 74\%$ are males, and of those who do not exercise, $\frac{1532}{2478} \times 100\% \approx 62\%$ are males.

9780170250283

Example 14

A market researcher interviewed 50 people about their preferences of ice-cream flavours. Of the 30 women, 20 said they preferred chocolate, while 14 of the men preferred other flavours. Construct a two-way table showing this information.

Solution

Create a table using gender as the explanatory variable (columns) and ice-cream flavour as the response variable (rows).

	Gender	
Ice-cream flavour	**Women**	**Men**
Chocolate		
Other		

Fill in information from the question.

- 50 people
- 30 women
- 20 women preferred chocolate
- 14 men preferred other flavours

	Gender		
Ice-cream flavour	**Women**	**Men**	
Chocolate	20		
Other		14	
	30		50

Complete the table using column and row totals.

$20 + 10 = 30$
$30 + 20 = 50$
$10 + 14 = 24$
$26 + 24 = 50$ and
$20 + 6 = 26$

	Gender		
Ice-cream flavour	**Women**	**Men**	
Chocolate	20	6	26
Other	10	14	24
	30	20	50

Comparisons could be made using the table from Example **14** (or a corresponding side-by-side column graph), but it could be misleading as more women were surveyed than men. Converting the data in the table to percentages minimises this effect and allows for easy comparisons. Generally speaking, percentages should be calculated in columns when the explanatory variable is shown in the columns of the table, unless you are asked otherwise.

Example 15

Convert the two-way table from Example **14** into a percentage two-way table.

Solution

The percentages should be calculated in columns, so redraw the two-way table using only the column totals.

	Gender	
Ice-cream flavour	**Women**	**Men**
Chocolate	20	6
Other	10	14
	30	20

Calculate the percentage of women who prefer chocolate.

$\frac{20}{30} \times 100 \approx 67\%$

Calculate all other required percentages.

Other women: $\frac{10}{30} \times 100 \approx 33\%$

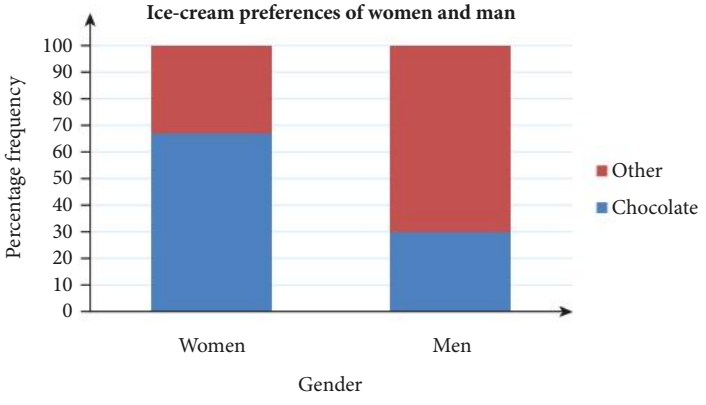
Men chocolate: $\frac{6}{20} \times 100 = 30\%$

Other men: $\frac{14}{20} \times 100 = 70\%$

Put the percentages into the two-way table.

	Gender	
Ice-cream flavour	**Women**	**Men**
Chocolate	67%	30%
Other	33%	70%
	100%	100%

From a percentage two-way table or a segmented column graph a comparison can be made to see if a relationship is present. From the segmented column graph below (using the data from Examples **14** and **15**), it appears that a higher percentage of women prefer chocolate ice-cream than men. More men prefer other flavours than women. This information suggests that the preferred flavour of ice-cream **may be associated** with gender.



IMPORTANT

It is very important that we never say one variable **causes** another variable to change, simply that there **may** be an association.

IMPORTANT

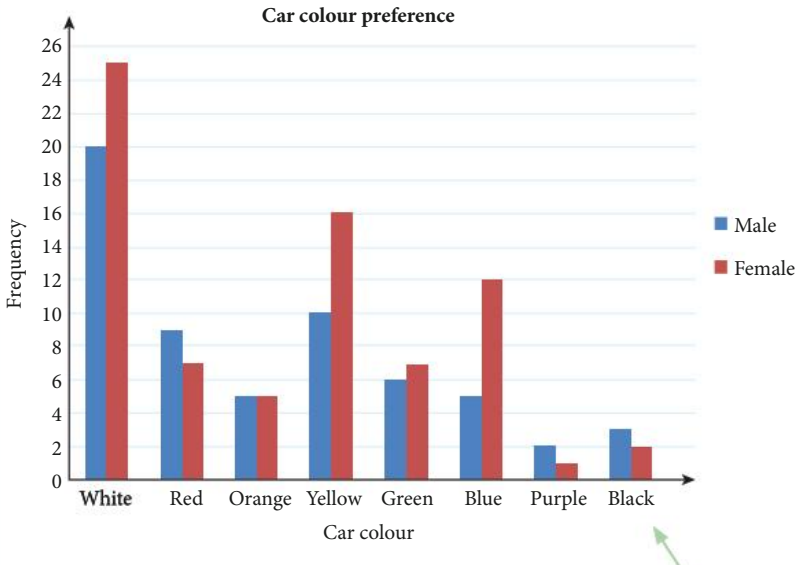
From a percentage two-way table or a segmented column graph it can be determined that:

- there may be an association as there are clear differences in the splits of each column, OR
- there is no association as the column splits are virtually the same.

9780170250283

Example 16

The graphical display below shows the results of a survey of preferred car colours of males and females.



a Construct a two-way table that will allow you to investigate if car colour depends on your gender.

b Construct a percentage two-way table and segmented column graph.

c Hence discuss if there is an association between car colour and gender.

Note that in this graph the response variable is on the *x*-axis, allowing easier comparison between male and female preferences for different colours.

Solution

a Create a table using gender as the explanatory variable (columns) and car colour as the response variable (rows).

	Gender	
Car colour	**Women**	**Men**
White		
Red		
Orange		
Yellow		
Green		
Blue		
Purple		
Black		

Complete the table by reading the frequency of each value from the side-by-side column graph, and then calculate the totals. For example, 25 females prefer white cars and 20 males prefer white cars.

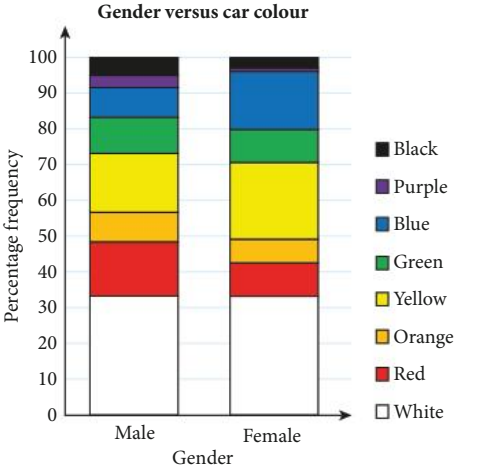
	Gender		
Car colour	Male	Female	Total
White	20	25	45
Red	9	7	16
Orange	5	5	10
Yellow	10	16	26
Green	6	7	13
Blue	5	12	17
Purple	2	1	3
Black	3	2	5
Total	60	75	135

b Calculate all the required percentages, remembering to use the column totals and not the row totals.

As the percentages have been rounded, they add to 99.9% not 100%. This will not have a major affect on the resulting graph.

	Gender	
Car colour	Male	Female
White	33.3%	33.3%
Red	15%	9.3%
Orange	8.3%	6.7%
Yellow	16.7%	21.3%
Green	10%	9.3%
Blue	8.3%	16
Purple	3.3%	1.3%
Black	5%	2.7%
	100%	100%

Construct the segmented column graph from the table.



c Determine if there is an association by looking for differences in the matching sections of the graph.

There are differences in the numbers of males and females who choose blue, yellow and red cars, which by itself would suggest that car colour may be associated with gender, although the five other colours were very similar for males and females. As the majority of the colours are similar, this suggests that there is no association between car colour and gender.

9780170250283

EXERCISE 1.05 Relationships between categorical variables

Two-way tables

Concepts and techniques

1 Why is it easier and more accurate to compare categorical variables when they are displayed as a segmented column graph rather than a side-by-side column graph?

2 Below is a partially completed two-way frequency table.

a State the explanatory and response variables.

b Copy and complete the two-way frequency table.

c How many people are not library members?

d How many women are represented in the table?

e Of those who are library members, what percentage are males?

f What percentage of females are not library members?

	Gender		
Library member	**Male**	**Female**	
Yes	20		52
No	8	15	

3 Example 14 80 people were asked about whether they use sunscreen at the beach. Of the 40 women, 30 said they do, with 35 of the men also stating that they use sunscreen at the beach. Construct a two-way frequency table using gender as the explanatory variable.

4 On Kirra beach, 50 people were asked whether they were staying locally or visiting and whether they intended to surf or swim. In total, there were 18 surfers and of the 31 people staying locally, 12 were surfers. Which of the following two-way frequency tables best represents this information?

A

	Locally	Visiting	
Surfer	12	19	31
Swimmer	6	13	19
	18	32	50

B

	Surfer	Swimmer	
Locally	12	6	18
Visiting	19	13	32
	31	18	50

C

	Locally	Visiting	
Surfer	12	6	18
Swimmer	29	23	42
	31	29	50

D

	Surfer	Swimmer	
Locally	12	19	31
Visiting	18	1	19
	30	20	50

E

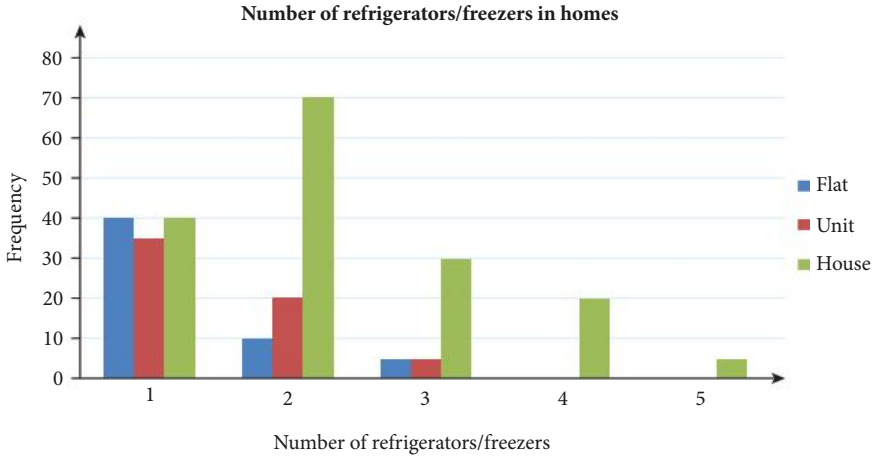
	Locally	Visiting	
Surfer	12	6	18
Swimmer	19	13	32
	31	19	50

5 Example 15 **a** Convert the two-way table from question **2** into a percentage two-way table (to the nearest percentage).
b Hence draw a segmented column graph.
c Determine if there is an association between the two variables.

6 **a** Convert the two-way table from question **3** into a percentage two-way table.
b Hence draw a segmented column graph.
c Determine if there is an association between the two variables.

Reasoning and communication

7 Example 16 The graph below shows the results of a survey about the type of residence and the number of refrigerators and freezers in it.



a Construct a two-way table that will allow you to investigate whether the number of refrigerators/freezers depends on the type of residence you live in.
b Construct a percentage two-way table and segmented column graph.
c Hence discuss if there is an association between the two variables being investigated.

8 Data was collected on whether people under or over 35 store their music digitally or on CD. Of the 75 people surveyed, 52 people preferred to store their music digitally, but only 14 of these were above the age of 35. The other 21 in this age group preferred CDs.
a Construct a two-way table to represent this data.
b Construct a percentage two-way table and segmented column graph.
c Construct a percentage two-way table using the row totals and then use these to construct another segmented column graph.
d What is the difference between the displays constructed in parts **b** and **c**?
e Use the different displays to find the percentage of:
i under 35s who store their music on CD
ii over 35s who store their music digitally
iii CD users who are over 35
iv digital users who are over 35.
f Explain why there is a difference between parts **ii** and **iv**.

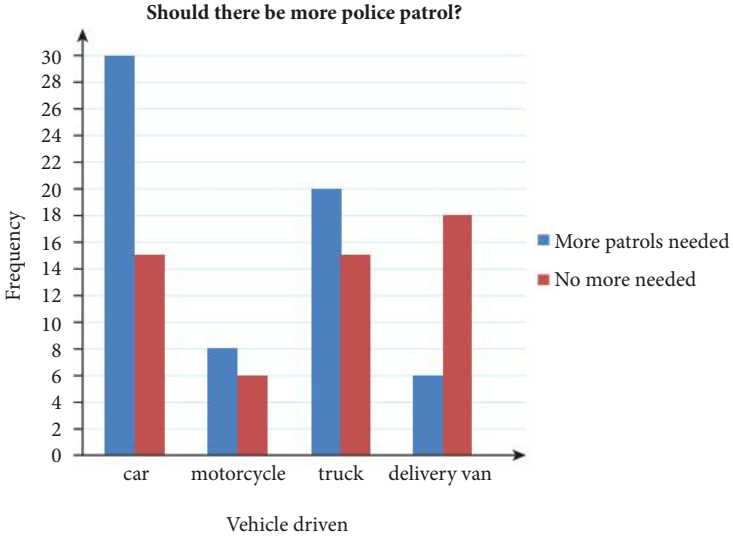
9780170250283

9 Data was collected on people who applied for their manual or automatic licences in the city versus the country. 36 of the 49 people from the country opted to apply for their manual licence. 52 of the 95 people surveyed chose to apply for their automatic licence.

Construct a two-way frequency table and use it to find the percentage of:

a people from the country who applied for their automatic licence
b people from the city who applied for their manual licence
c people who applied for their automatic licence who live in the city
d people who applied for their manual licence who live in the country.

10 The graph below shows the number of people driving different types of vehicles and their opinion on whether there should be more police patrols or not.



a Construct a two-way table that will allow you to investigate if a person's opinion on police patrols depends on the type of vehicle they drive.
b Construct a percentage two-way table and segmented column graph.
c Hence discuss if there is an association between the two variables being investigated.

11 A study of the deaths of 10 000 men found that 2500 of them were smokers. The number of men who died from lung cancer was 715, and 630 of these were smokers.

a State the two variables being investigated and determine which one is the explanatory variable.
b Construct a two-way table for the information.
c Convert the table to a percentage two-way frequency table with all percentages correct to the nearest whole number.
d Hence draw a segmented column graph.
e Using the answers from part **c** and **d**, write a report comparing the two variables being investigated.

1.06 RELATIONSHIPS BETWEEN NUMERICAL AND CATEGORICAL VARIABLES

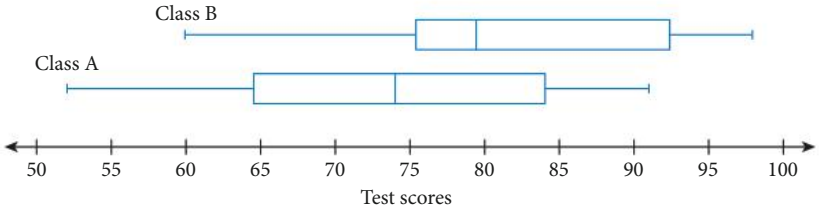
When comparing a numerical variable and a categorical variable we can use either a back-to-back stem-and-leaf plot or parallel boxplots. Parallel boxplots are more appropriate when the data set is large or when the categorical variable has more than two characteristics.

Back-to-back stem-and-leaf plot

Key: 5|6 = 56%

Class A	Stem	Class B
4 3 2	5	
8 5 4 3 1	6	0 0 2 4
9 6 4 4 3 1 0	7	3 5 6 7 8 8 8 9
8 8 7 5 3 2 0	8	0 1 4 6 9
1 0	9	2 3 5 6 7 7 8

Parallel boxplots



The purpose of constructing the graphical displays is so that comparisons can be made between the 2 categories in relation to their shape and summary statistics.

IMPORTANT

A report that compares data should not just state values, it should discuss what is higher, lower, similar or different between the two data sets.

IMPORTANT

Comparing the mean and standard deviation

- In particular situations, a higher mean is the better result (test scores) but in other circumstances, a lower mean is the desired result (time to complete a race).
- A higher standard deviation means that the data is more variable. Alternatively, a lower standard deviation means that the data is more consistent.

9780170250283

Example 17

Liz and George deliver pamphlets to letterboxes. The number of pamphlets delivered per hour over 12 hours is shown below.

Liz:	24	25	26	27	28	28	31	32	32	32	35	35
George:	15	18	21	24	25	29	31	31	32	38	38	45

a Display the data in a back-to-back stem-and-leaf plot.
b Comment on the shape of the data for each person.
c Calculate the mean and standard deviation for the number of pamphlets delivered by each person and comment on these results.
d Who would you say is the better delivery person and why?

Solution

a As the data range is reasonably small, split the stems so that the shape of the data will be easier to interpret. Remember to order the values of the leaves.

George	Stem	Liz
8 5	1	
4 1	2	4
9 5	2	5 6 7 8 8
2 1 1	3	1 2 2 2
8 8	3	5 5
	4	
5	4	

Key: 4|5 = 45

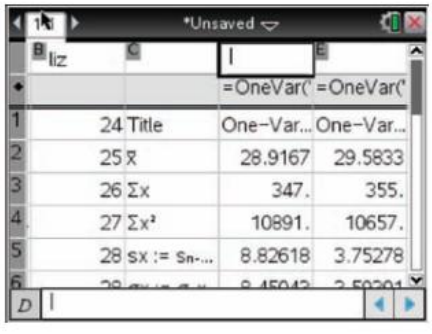
b Describe the shape of the data for each person.

Liz's data is approximately symmetrical whereas George's looks more negatively skewed.

c Calculate the one variable statistics using a CAS calculator.

TI-Nspire CAS

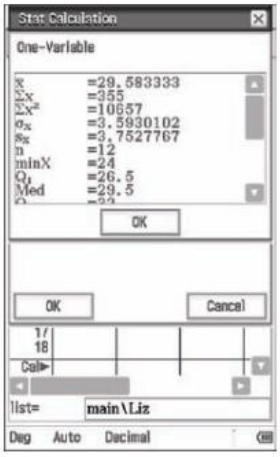
The one variable statistics can be calculated for both data sets simultaneously by setting the number of lists to 2 in the pop up box that appears.



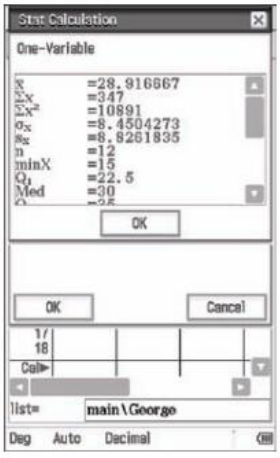
ClassPad

Enter the data into two lists and calculate the one variable statistics for each list.

Liz's results:



George's results:



Write the statistics for Liz and George.

Liz: $\overline{x} = 29.6$, $s_x = 3.8$
George: $\overline{x} = 28.9$, $s_x = 8.8$

The mean is the average of the data.

On average, Liz delivers slightly more pamphlets than George.

Standard deviation is a measure of spread of the data.

Liz has a smaller standard deviation, which indicates that there is less spread in her delivery rate. Therefore, she is more consistent than George.

d Use the results to decide who is the better delivery person.

Liz is the better delivery person as she has a slightly higher average (mean) delivery rate and she is more consistent (lower standard deviation).

9780170250283

Example 18

The test results of two Year 12 classes in Mathematical Methods are shown below.

Class A:	58	46	53	52	67	36	61	49	47	
	59	66	53	94	69	46	44	57		
Class B:	60	50	70	69	86	43	60	60	44	56
	49	50	56	56	42	65	47	67	25	46

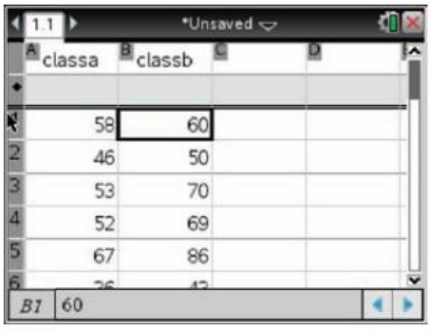
Construct parallel boxplots for the data.

Solution

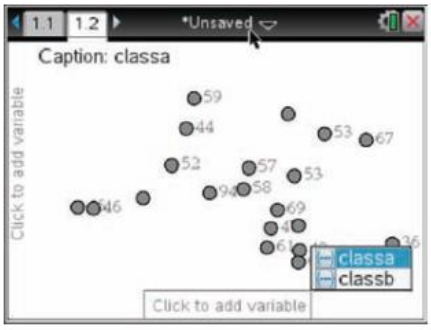
Use a CAS calculator to create the boxplots.

TI-Nspire CAS

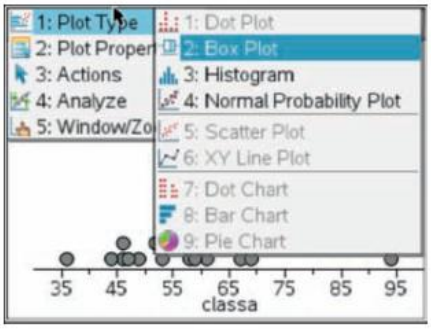
Open a New Document with a Lists & Spreadsheet page.
Type each set of data into a column.

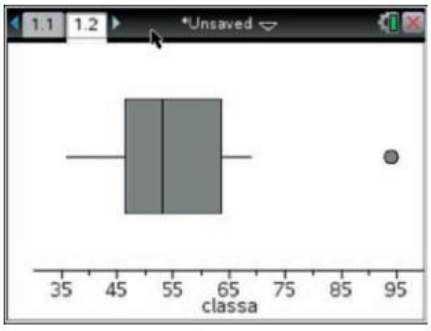


Add a Data & Statistics page.
Click in the Click to add variable space at the bottom of the page.
Select classa and a dot plot will appear.

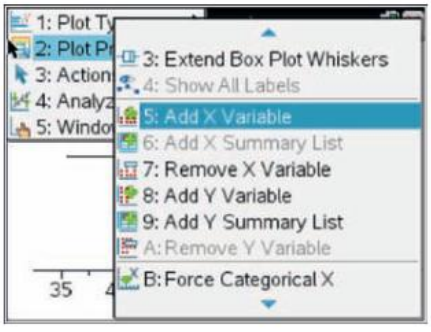


Press menu
1: Plot Type
2: Boxplot

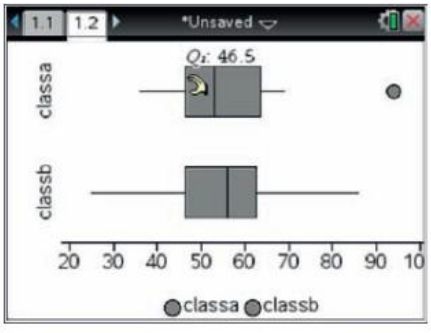




Press
2: Plot Properties
5: Add X Variable
Select classb.



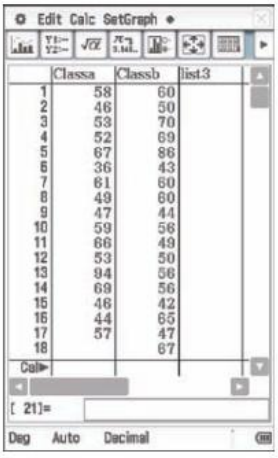
Use the touch pad to move around the screen and the five number summary values will appear.



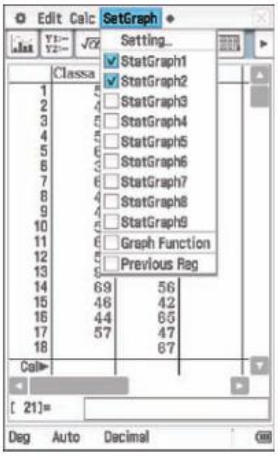
9780170250283

ClassPad

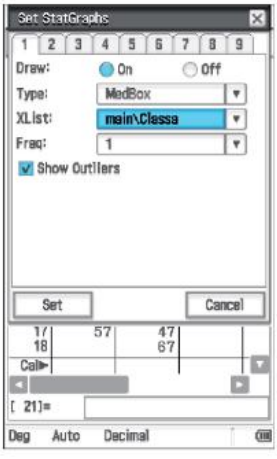
Use the Statistics application.
Type each set of data into a column.



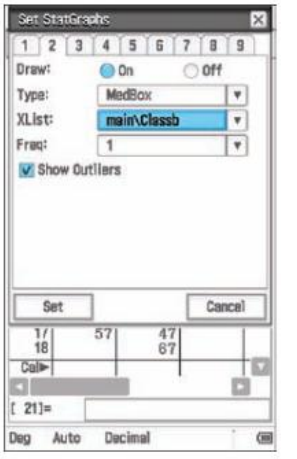
Tap **SetGraph**.
Make sure that StatGraph1 is checked and StatGraph2 is checked.
Tap **Setting**.



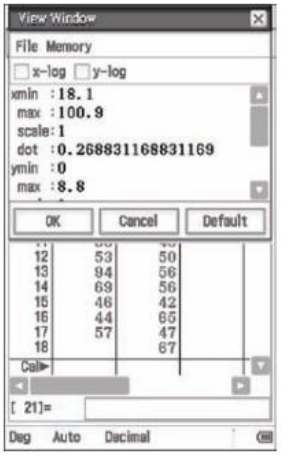
Complete the window as shown, then tap the tab numbered 2.



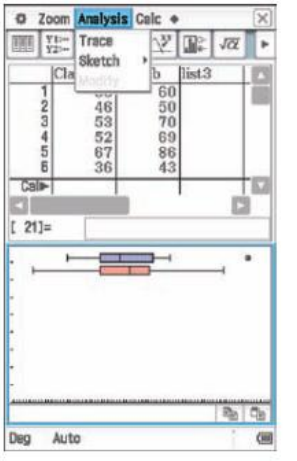
Complete the window as shown, then tap Set.



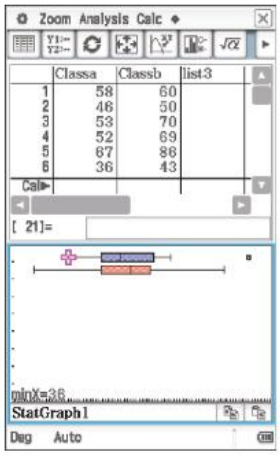
Tap ⚙ then **View Window** and set **xmin** and **max** below it and ***y*min** and **max** below that as shown on the screen on the right. Tap **OK**.



Tap 📊 then Analysis and Trace, use the arrow keys to move from one plot to the other and read the five number summary from each boxplot, which appears at the bottom of the screen.



9780170250283



Parallel boxplots can have more than 2 boxplots relating to the same numerical variable. When analysing boxplots, it important to remember that:

- boxplots split data into quarters, which means that each section of the boxplot represents 25% of the data
- the rectangular box section of the plot represents the IQR or the middle 50% of the data.

Making comparisons from boxplots is usually done by:

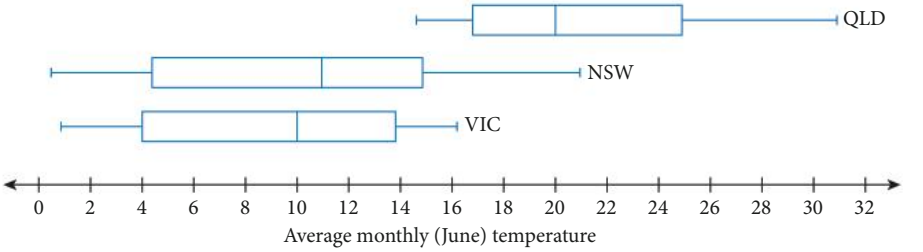
- comparing the centre or median of each data set
- comparing the spread of the data sets (range or interquartile range)
- comparing the shape (skewness)
- discussing the presence of outliers.

From parallel boxplots, it can be determined if there is an association between the categorical variable and the numerical variable. In general,

- if the boxplots are similar, then there is no association between the variables
- when there are significant differences, it can be stated that there **may** be an association between the two variables.

Example 19

The parallel boxplots below represent the average monthly temperature in June for Victoria, New South Wales and Queensland.



a Which state has the highest median average monthly temperature?
b Which state has the largest range of average monthly temperatures?
c Which state's data is best described as positively skewed?
d Which state had the lowest average monthly temperature?
e Write a brief summary comparing the average monthly temperature for each state.
f Is there an association between the states and their average monthly temperatures?

Solution

a	Look for the state whose median line is furthest along the scale.	Queensland
b	Look for the longest boxplot, including whiskers.	New South Wales
c	Look for the boxplot with its median to the left of the box and a right whisker being longer than the left one.	Queensland
d	Look for lowest left endpoint.	New South Wales
e	Compare the states, noting similarities and differences.	Victoria and New South Wales have similar average monthly temperatures, whereas Queensland has higher average monthly temperatures than the other two states.
f	Are there differences?	There may be an association between the state and average monthly temperature as Queensland's data is significantly different to the other states.

9780170250283

EXERCISE 1.06 Relationships between numerical and categorical variables

Comparing data

Concepts and techniques

1 Example 17 Two radar cameras that were positioned on different roads recorded car speeds (in km/h) as follows.

Camera 1: 78 63 75 69 71 83 80 67 74 72 73 74
90 83 65 73 69 89 76 102 83 78 69 71

Camera 2: 112 139 120 116 116 136 140 123 135 131 120 117
138 131 127 119 125 130 130 134 123 148 169 130

a Display this information in a back-to-back stem-and-leaf plot.
b Calculate the mean and standard deviation of the speed recorded by each camera, correct to 2 decimal places.
c Compare your results found in part **b**.

2 A class was given two different Science tests and their marks are displayed in the back-to-back stem-and-leaf plot.

a Find the range, interquartile range and median for each test.
b Calculate the mean and standard deviation for the mark achieved in each test.
c Which Science test was easier? Justify your answer.

Test A	Stem	Test B
2	0	3
5	1	
5	2	
	3	8
8	4	5
	5	2 4 7
5	6	2
7	7	
0	8	
2	9	4

Key: 2|5 = 25%

3 Twelve people took part in the QUIT smoking program. The number of cigarettes smoked per day was recorded before the start of the program and 6 weeks later.

Before: 21 10 36 16 32 42 9 14 21 18 34 45
6 weeks later: 6 24 31 21 16 19 16 18 28 32 8 13

a Construct a back-to-back stem-and-leaf plot for this data.
b What are the interquartile ranges of each data set?
c Comment on the shape of the Quit program data for the start of the program and 6 weeks later.
d Calculate the mean and standard deviation of the number cigarettes smoked per day for both time periods, correct to the nearest whole number.
e Compare the results you found in part **d**.
f Is the QUIT program generally working for this group of people? Justify your answer.

4 Example 18 **a** Construct parallel boxplots for the data in question **1**.
b For each camera, state the range of values that the lowest 25% of speeds fall within.
c Which camera recorded the highest median speed?
d Is there an association between the position of the cameras and the speed recorded? Justify your answer.

5 **a** Construct parallel boxplots for the data in question **2**.
b Which test had the largest range?
c Which test had the smallest interquartile range?
d Show why Test B has two outliers.
e Is there an association between the Science test that was taken and the resulting mark? Justify your answer.

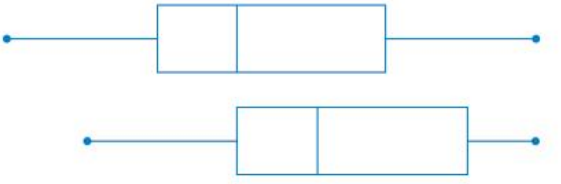
6 The ages of members (in years) in a yoga class at Husky Body Shop one morning are shown in the stem-and-leaf plot.

Women	Stem	Men
4	2	2 6
2	3	2 2 5 6 8
8 3	4	2 4 5 6
9 4 0 0	5	3
8 5 4 3 3	6	3 3 4

Key: 2|6 = 26

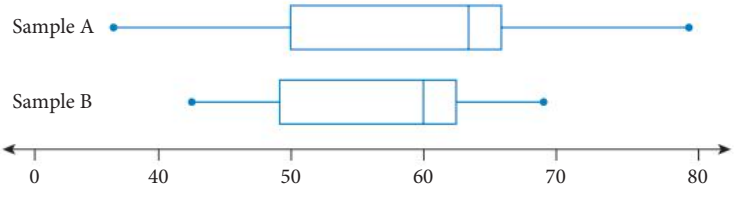
a Construct parallel boxplots for the above data.
b Find the median age of the men in the class.
c Compare the shapes of age distributions of the women and men.
d Which group was younger on average: women or men? Justify your answer.
e Is there an association between gender and the ages of members attending yoga classes? Justify your answer.

7 Which measure is the same for both of these boxplots?



A the mean **B** the range **C** the median
D the interquartile range **E** the mode

8 Two samples of 10 people each were given stopwatches and asked to stop the watches when they thought 60 seconds had elapsed. The results, in seconds, for the 'estimated minute' are shown in the parallel boxplots below.

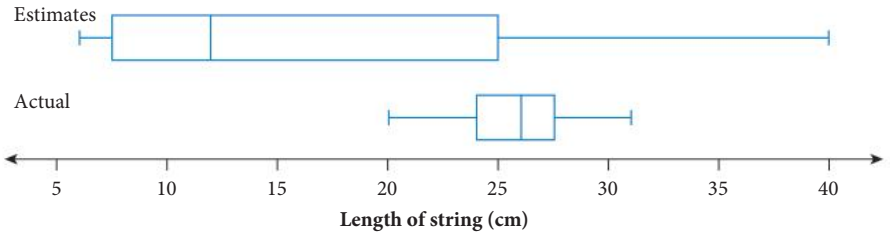


9780170250283

Which one of the following statements is **true**?

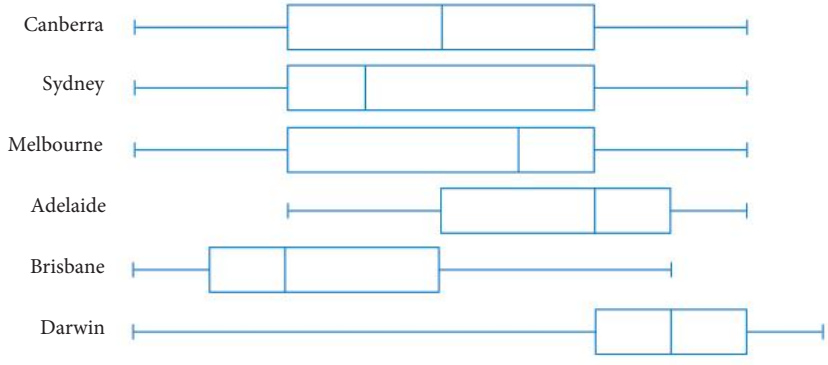
A The range for Sample B is larger than Sample A.
B Both data sets are positively skewed.
C The median for Sample A is higher than Sample B.
D Sample B recorded the smallest estimate of 60 seconds.
E The data sets have similar interquartile ranges.

9 Example 19 Yashneel decided to investigate his theory that 'People always underestimate the length of a piece of string'. He asked students to estimate the lengths of several pieces of string, measured the actual lengths and displayed the results in parallel boxplots.



a Write down the median of:
  i the estimated lengths ii the actual lengths.
b Compare the range of each data set.
c Compare the interquartile range of each data set.
d Classify the shape of each data set.
e Do you agree with Yashneel's theory? Justify your answer.

10 The box-and-whisker plots below show the distribution of petrol prices in six Australian capital cities.



a Which city has:
  i the greatest range of prices?
  ii the lowest median price?
  iii the smallest interquartile range of prices?
  iv a symmetrical distribution of prices?
b Is petrol cheaper in Sydney than in Melbourne? How can you tell?
c Is there an association between the city you are in and the cost of petrol? Justify your answer.

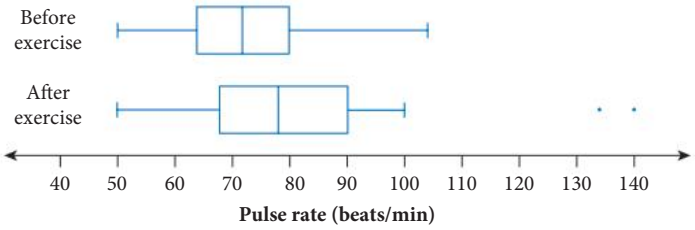
Reasoning and communication

11 The table below shows the average number of rainy days per month in Sydney and Melbourne.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
Melbourne	12	12	13	12	12	12	10	10	10	12	11	12
Sydney	8	7	9	12	14	14	15	16	15	14	12	11

a Construct a back-to-back stem-and-leaf plot with split stems.
b Discuss the shape of the data for each city.
c Calculate the five number summary for each city using the back-to-back stem-and-leaf plot.
d Hence construct a parallel boxplot.
e Write down one fact that is clearly evident in each graphical display.
f Is there an association between the city and the number of rainy days per month?

12 A group of 20 people had their pulse rates taken before and after an exercise class and the results are displayed below. Two outlier pulse rates (shown as '+') were excluded from the 'After exercise' data.



a What percentage of people had a pulse rate between 64 and 72 before the exercise class?
b How many people had a pulse rate between 64 and 72 before the exercise class?
c Which boxplot has a greater interquartile range?
d Between what two values do the middle 50% of pulse rates for after exercise lie between?
e What is the interquartile range of pulse rates after the class?
f Hence show why the two values were excluded from the 'after exercise' data.
g By how much did the median pulse rate increase after exercise?

13 Data was gathered on the test scores of four groups of students, as shown below.

No study: 46 35 48 32 34 19 51 28 22 19 87 44
42 11 62 17 23 55 27 47 52 43 38

Study: 69 43 93 63 85 72 29 64 81 68 73 43
68 56 87 46 72 72 29 65 89 42 63

Revision course: 53 46 88 63 75 46 53 38 49 51 48
52 39 37 42 24 31 51 61 63 42 42 42

Revision course + study: 77 95 49 82 81 93 59 95 94 56 76
86 88 97 67 98 79 67 88 42 82 81 82

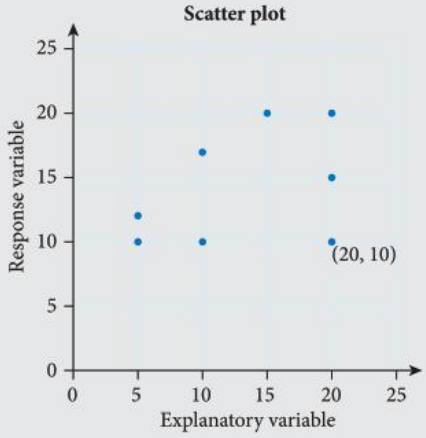
Construct a set of parallel boxplots to represent the data and comment on your results.

9780170250283

1.07 RELATIONSHIPS BETWEEN NUMERICAL VARIABLES

A **scatter plot** is a graphical display used for two numerical variables and is constructed by plotting points onto a Cartesian plane. The explanatory variable is plotted on the x-axis and the response variable on the y-axis.

IMPORTANT



Steps to constructing a scatter plot.

1 Data will be displayed in a table. Decide which variable is the explanatory variable (x) and which is the response variable (y).

Do not assume that the top row of the table is the explanatory variable. Use the sentence, RV depends on the EV, to decide.

2 Set up an appropriate scale for each axis and label them clearly with the variable names.

3 Plot each pair of data points from the table using a dot (•)

A scatter plot tells us about the association or **relationship** between the two variables being investigated.

IMPORTANT

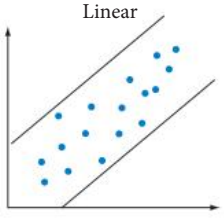
When interpreting a scatter plot, three areas need to be discussed.

1 Form
2 Direction
3 Strength

1 Form

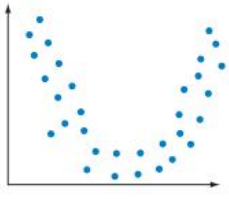
The form of a scatter plot shows either a linear, non-linear or no relationship.

Linear



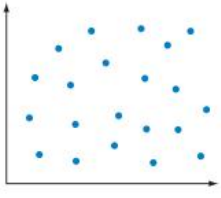
Data in general follows a straight line pattern.

Non-linear



Data does not occur in a straight line pattern and may have a curved pattern.

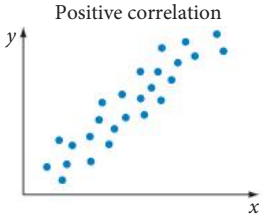
No relationship



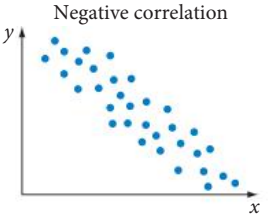
Data is randomly scattered and shows no pattern at all.

2 Direction

If the scatter plot demonstrates a **linear pattern**, we can then decide if the relationship is positive or negative.



The direction of a positive relationship rises from left to right. This indicates that the y (response) variable tends to increase as the x (explanatory) variable increases.



The direction of a negative relationship falls from left to right. This indicates that the y (response) variable tends to decrease as the x (explanatory) variable increases.

Example 20

For each pair of variables, state whether they have a positive relationship, a negative relationship or no relationship.

a Height and weight of a person.
b Salary level and heart rate of an employee.
c Price of a brand of car and the number of that brand sold.
d The amount of effort studying for an exam and the exam score.

Solution

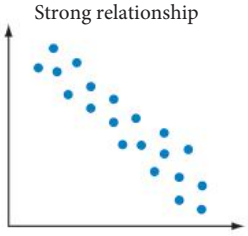
a	Decide if there is a relationship. If so, do both variables increase(positive), or does one increase while the other decreases(negative)?	Positive relationship, because weight usually increases as height increases.
b	Decide if there is a relationship. If so, do both variables increase(positive), or does one increase while the other decreases(negative)?	No relationship, because salary level has no effect on a person's heart rate.
c	Decide if there is a relationship. If so, do both variables increase(positive), or does one increase while the other decreases(negative)?	Negative relationship, because sales usually decrease as car prices increase.
d	Decide if there is a relationship. If so, do both variables increase(positive), or does one increase while the other decreases(negative)?	Positive relationship, because exam scores usually increase as study effort increases.

9780170250283

3 Strength

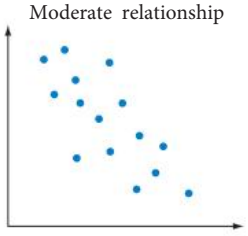
The strength of a relationship is categorised as strong, moderate or weak.

Strong relationship



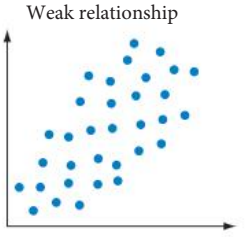
Data points are all reasonably close together.

Moderate relationship



Data points are more spread out than a strong relationship.

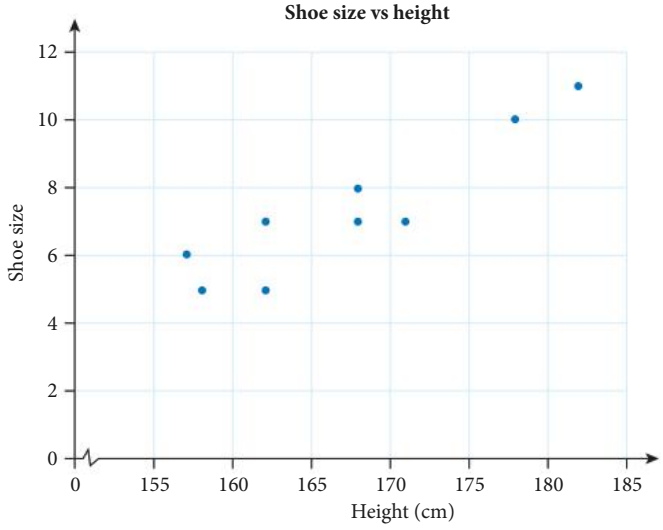
Weak relationship



Data parts are widely spread out.

Example 21

Describe the following scatter plot, referring to form, direction and strength.



Solution

The data appears to follow a straight line pattern, so it is linear.
As height increases, so too does shoe size, so it is a positive relationship.
The data points are reasonably close together, so the relationship can be classed as strong.

Linear, positive and strong.

Example 22

The number of hot drinks sold at a milk bar per day and the day's maximum temperature were recorded over a period of 2 weeks.

Temperature	35	33	27	22	18	29	38	36	24	25	29	34	21	19
No. of drinks sold	25	28	29	54	76	48	18	39	42	61	36	49	68	53

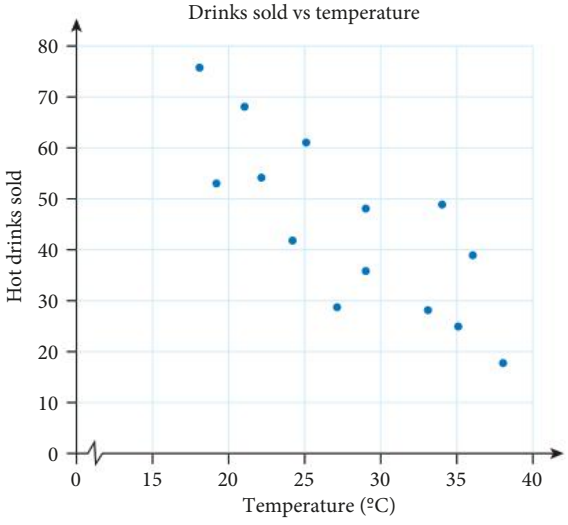
a Sketch a scatter plot for this information.
b Describe the relationship shown in the scatter plot.

Solution

a Decide which variable is the explanatory (x) and response variable (y).

1. Number of drinks sold depends on temperature.
2. Temperature depends on number of drinks sold.
Sentence 1 makes more sense, therefore temperature is the explanatory variable and drinks sold is the response variable.

Set up an appropriate scale for each axis and clearly label with the variable names and title. Plot the coordinates, remembering that temperature is x and drinks sold is y.



b The data appears to follow a straight line pattern, so it is linear.
As temperature increases, the number of hot drinks decreases, so it is a negative relationship.
The data points are somewhat close together, so the relationship can be classed as moderate.

The relationship between temperature and number of hot drinks sold can be described as linear, negative and moderate.

9780170250283

Scatter plots can also be drawn using CAS calculators.

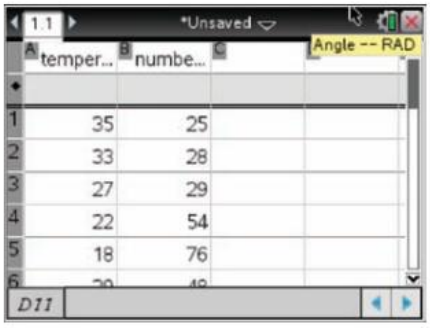
Constructing a scatter plot using a CAS calculator

A CAS calculator can be used to construct a scatter plot for the data in Example **22**.

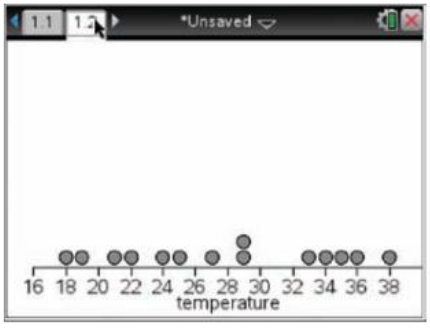
TI-Nspire CAS

Enter data into the Lists & Spreadsheet page, remembering to name each column. It is common practice to enter the values for the explanatory variable in the first column and the response variable in the second column.

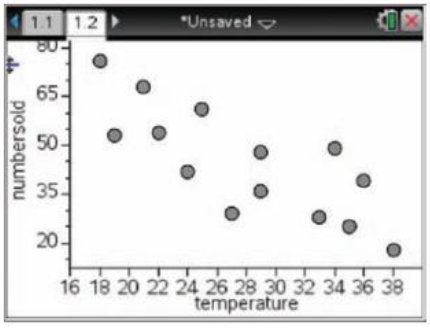
Remember, ctrl 7 will take you back to the top cell



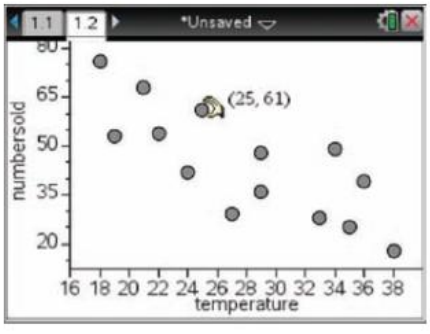
Press ctrl doc select 5: Add Data & Statistics. Scroll down to Click to add variable, then click and select explanatory variable name from the menu.



Scroll across to left side to Click to add variable, then click and select response variable name from the menu.

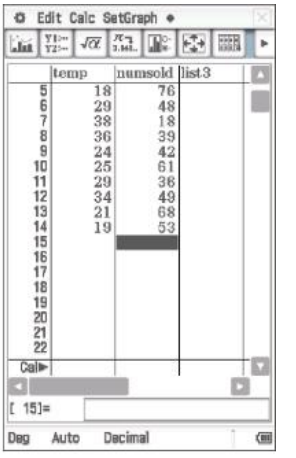


Move the hand around the screen to each dot and the coordinates of a point will appear.



ClassPad

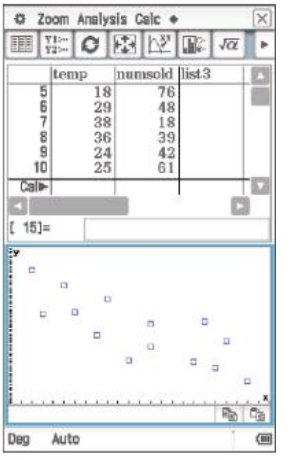
Using the Statistics application, rename columns and type in data. It is common practice to enter the values for the explanatory variable in the first column and the response variable in the second column.



Tap **SetGraph**, making sure only one StatGraph is ticked, then tap **Setting....** Complete the screen as shown and tab Set.

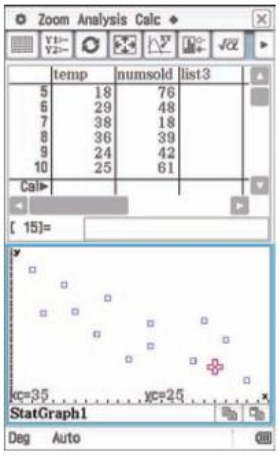


Tap  to insert the graph.



9780170250283

If you tap Analysis then Trace, you can move from point to point using the left and right arrow buttons and the coordinates will appear at the bottom of the screen.



EXERCISE 1.07 Relationships between numerical variables

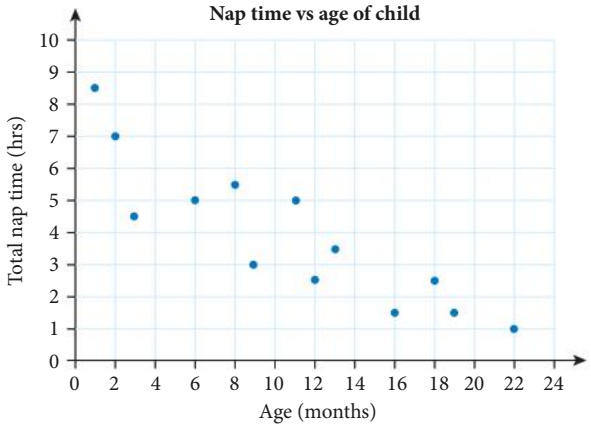
Concepts and techniques

1 Example 20 State whether the following pairs of variables have a positive, negative or no relationship.

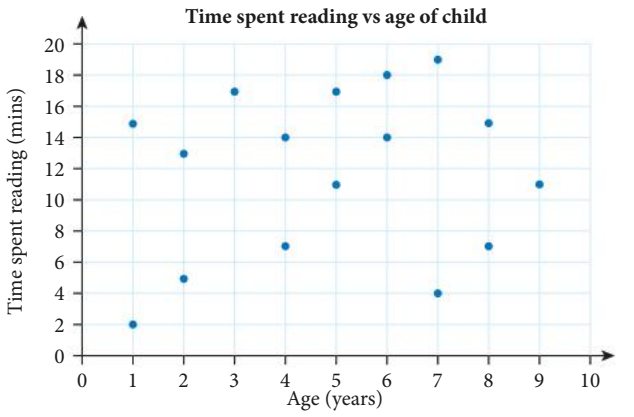
a Number of cigarettes smoked and probability of lung cancer.
b Number of police on roads and number of speeding cars.
c Number of chocolate bars eaten and how many pets a person owns.
d Calories consumed and weight gained.
e A person's alcohol consumption and reaction time.
f A person's height and the number of hours sleep they have per night.
g Temperature and number of people at a beach.
h Time spent travelling home and distance from home.

2 Example 21 Describe each of the following scatter plots with reference to their form, direction and strength.

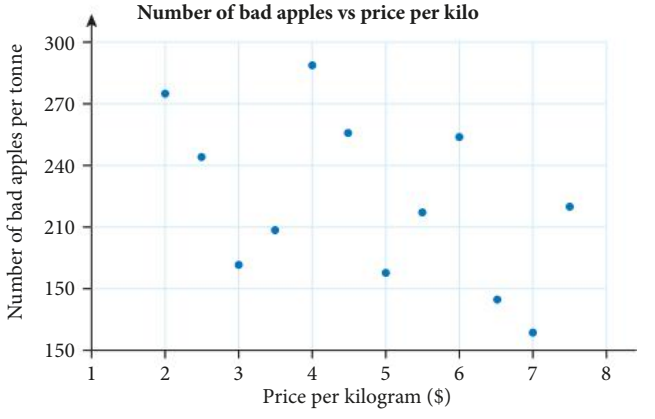
a



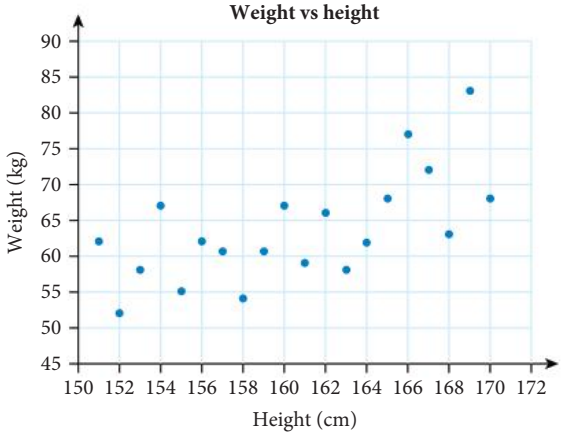
b



c

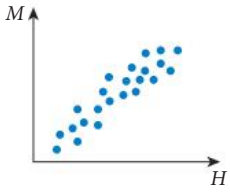


d



9780170250283

The following scatter plot relates to questions **3** and **4**.



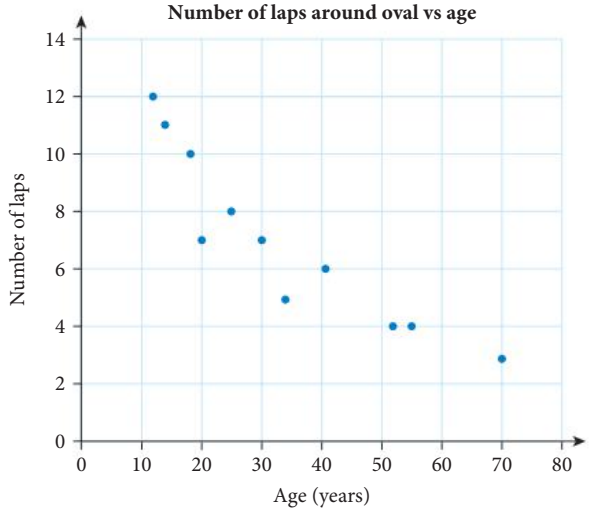
3 Which statement best describes the relationship between M and H?

A No relationship
B Non-linear relationship
C Linear, negative and weak
D Linear, positive and weak
E Linear, positive and strong

4 Which of the following sentences is **correct**?

A It can be concluded that M should decrease as H increases.
B There is little evidence to show that H should increase as M increases.
C It can be concluded that M should increase as H increases.
D There is some evidence to support that H will increase as M increases.
E There is some evidence to support that M will increase as H increases.

5 Eleven people of different ages ran around an oval for 30 minutes. The number of laps they ran was recorded and the results are displayed in the scatter plot below.



Which one of the following statements is **false**?

A The youngest person in the group ran the most laps.
B Age is the explanatory variable and number of laps is the response variable.
C The relationship between age and number of laps can be described as strong.
D The association between age and number of laps can be described as positive.
E The association between age and number of laps can be described as non-linear.

6 Example 22 For each of the following data sets:

i sketch a scatter plot showing the relationship between the two variables

ii describe the relationship shown in the scatter plot.

a The number of customers in a store over a two-week period was recorded and the results are shown below.

Time (days)	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Number of customers	76	92	63	79	81	64	102	71	86	119	86	78	111	92

b The hearing sensitivities (highest audible frequency measured in kilohertz) of different people from the same extended family were tested and the results are shown below.

Age	5	12	15	21	46	50	62	70	75
Frequency (kH)	30	25	23	22	19	18	16	17	15

c People at a milkbar were asked how many chocolate bars they ate on average a month.

Age (years)	8	16	23	87	46	38	24	76	11	33	19	26	31	59	65	50
Number of chocolate bars	4	13	15	4	13	11	11	6	12	9	16	8	12	6	3	8

d The ages of both parents for students in a Year 12 class were recorded and the results are shown below.

Mother's age	Father's age
50	52
53	55
40	42
39	42
44	47
36	39
53	56
56	59
58	58
48	49
50	49
40	41
45	50
47	46
44	50
49	53
45	49
41	44
52	58
54	62
37	39
43	43
45	45

9780170250283

7 Construct a scatter plot for each of the following data sets. For each scatter plot, comment on the association between variables.

a

x	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
y	5	2	9	11	7	9	16	8	16	22	17	15	24	19	11	28

b

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

c

x	123	162	196	108	177	153	119	133	142	181	102	165	192	127	113
y	38	16	21	23	8	27	32	19	5	11	36	21	13	8	18

d

x	42	76	62	37	59	71	51	63	48	75	39	55	67	31	35	44	61	77	35	45
y	27	74	46	37	55	64	57	51	43	59	42	61	55	42	52	34	66	87	32	51

1.08 INTERPRETATION OF SCATTER PLOTS

When the relationship between two variables is found to be linear, we can calculate **Pearson's correlation coefficient** (r). This is a value that can be used to accurately determine the direction and exact strength of the relationship.

IMPORTANT

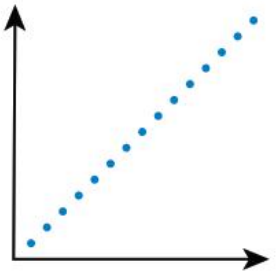
Interpreting the correlation coefficient

The correlation coefficient (r) is a number on a scale from -1 to 1 that measures the strength and direction of a linear relationship.

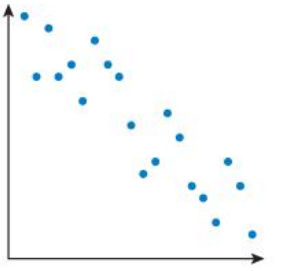
Value of r	What it means
1	Perfect positive relationship
$0.75 \leq r < 1$	Strong positive relationship
$0.5 \leq r < 0.75$	Moderate positive relationship
$0.25 \leq r < 0.5$	Weak positive relationship
$-0.25 < r < 0.25$	No relationship
$-0.5 < r \leq -0.25$	Weak negative relationship
$-0.75 < r \leq -0.5$	Moderate negative relationship
$-1 < r \leq 0.75$	Strong negative relationship
-1	Perfect negative relationship

The following are some examples of scatter plots with the corresponding correlation coefficients.

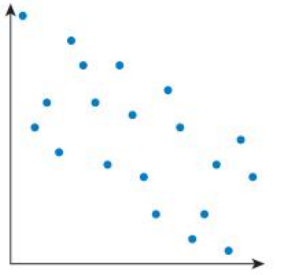
Perfect positive relationship
$r = 1$



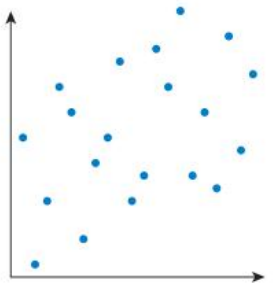
Strong negative relationship
$r = -0.9$



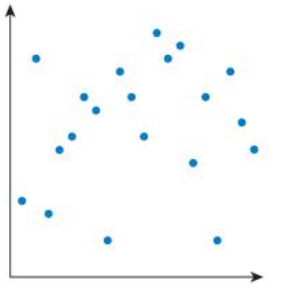
Moderate negative relationship
$r = -0.65$



Weak positive relationship
$r = 0.4$



No relationship
$r = 0.15$



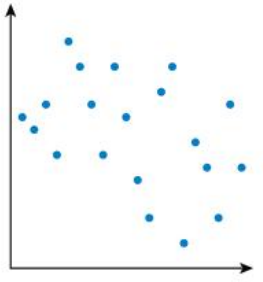
IMPORTANT

Using the correlation coefficient to describe the relationship between two variables is only reliable when the data is linear in form and there are no outliers present.

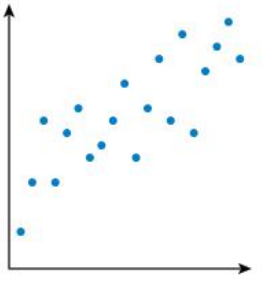
Example 23

Estimate the correlation coefficient (r) for the following graphs.

a



b



Solution

a The graph appears to have a weak, negative association. Estimate a 'higher' weak value that is negative. $r \approx -0.4$

b The graph appears to have a somewhat strong, positive association. Estimate a 'lower' strong value that is positive. $r \approx 0.8$

9780170250283

When interpreting a scatter plot, it is crucial that it is described in context of the relationship between the two variables it represents. It is not enough to simply say that the relationship is strong and positive, for example. Reference must be made to the individual variables demonstrating a relationship.

Some sample sentences that could be used to interpret the correlation coefficient are shown below with the words ***y* variable** and ***x* variable**. When writing these sentences, insert the actual variable names being investigated.

Linear, positive and strong	**Linear, positive and moderate**	**Linear, positive and weak**
It can be concluded that the ***y* variable** should increase as the ***x* variable** increases.	There is some evidence to suggest that the ***y* variable** should increase as the ***x* variable** increases.	There is limited evidence to suggest that the ***y* variable** should increase as the ***x* variable** increases.
Linear, negative and strong	**Linear, negative and moderate**	**Linear, positive and weak**
It can be concluded that the ***y* variable** should decrease as the ***x* variable** increases.	There is some evidence to suggest that the ***y* variable** should decrease as the ***x* variable** increases.	There is limited evidence to suggest that the ***y* variable** should decrease as the ***x* variable** increases.

IMPORTANT

If two variables are correlated, it does not necessarily mean that a change in one **causes** a change in the other. For example, a person's height and weight have a high correlation, which suggests that as height increases weight increases, but we must remember that a change in height does not necessarily always cause a change in weight. There may be external factors such as health, diet, amount of exercise, etc.

Example 24

Interpret the correlation values and write a brief sentence describing the relationship for the studies described below.

a A study is investigating whether the number of goals scored by a basketball player depends on their height. The correlation coefficient was found to be $r = 0.76$.

b A study is investigating whether a person's weekly savings depends on the amount of time they spend at shopping centres per week. The correlation coefficient was found to be $r = -0.68$.

c A study is investigating whether the number of packets of chips eaten per week depends on the height of a person. The correlation coefficient was found to be $r = 0.13$.

Solution

a	Interpret r.	The value $r = 0.76$ indicates a strong positive relationship.
	Write a sentence describing the relationship.	It can be concluded that the number of goals scored should increase as the height of a player increases.
b	Interpret r.	The value $r = -0.68$ indicates a moderate negative relationship.
	Write a sentence describing the relationship.	There is some evidence to suggest that the amount of savings should decrease as the time spent at shopping centres increases.
c	Interpret r.	The value $r = 0.13$ indicates that there is no relationship.
	Write a sentence describing the relationship.	There is no relationship between the number of packets of chips eaten per week and the height of a person.

In order to accurately calculate the correlation coefficient, there is a complex formula that can be used, but for the purposes of this course you are only required to calculate the correlation coefficient using a calculator. To do so, the data values need to be presented in a table or read carefully from a scatter plot.

Example 25

The number of umbrellas sold at a shopping centre and the amount of rain in the area were recorded over 8 weeks.

Umbrellas sold	10	35	12	26	27	19	17	18
Rainfall (mm)	60	110	56	58	105	75	48	90

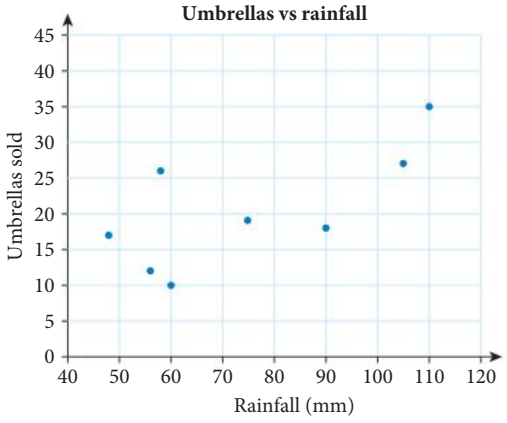
a Construct a scatter plot representing the information in the table.
b Calculate the correlation coefficient correct to four decimal places.
c Discuss if there is a relationship between the number of umbrellas sold and rainfall.

Solution

a	Decide which variable is the explanatory (x) and response variable (y).	1. Umbrellas sold depends on rainfall. 2. Rainfall depends on umbrellas sold. Sentence 1 makes more sense, therefore rainfall is the explanatory variable and umbrellas sold is the response variable.

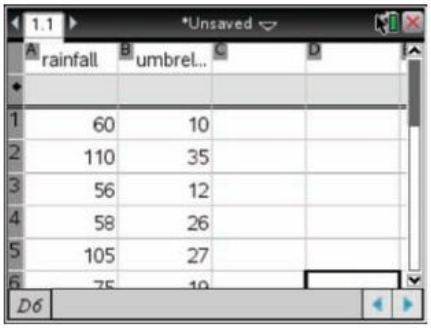
9780170250283

Set up an appropriate scale for each axis and label them clearly with the variable names and title. Plot the coordinates, remembering that rainfall is x and umbrellas is y.

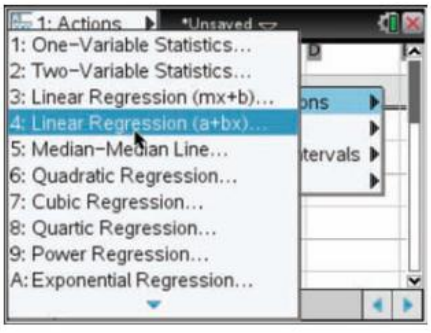


b **TI-Nspire CAS**

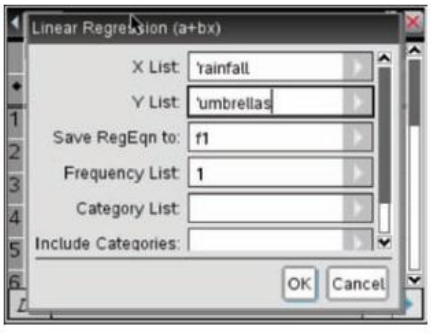
Enter the data into the Lists & Spreadsheet page, remembering to name each column.



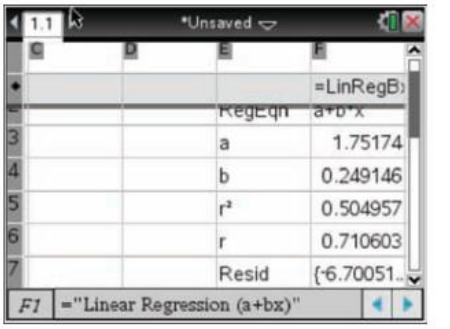
Press menu 4:Statistics, 1:Stat Calculations, 4:Linear Regression (a + bx).



Select the X List variable and the Y List variable, then click on OK.



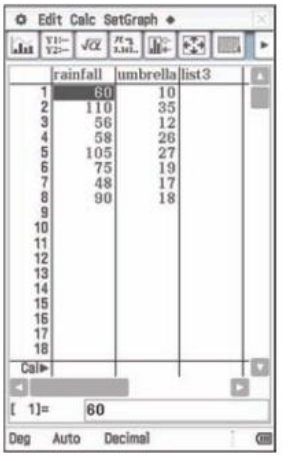
Using the screen that appears, scroll down to find the *r* value.



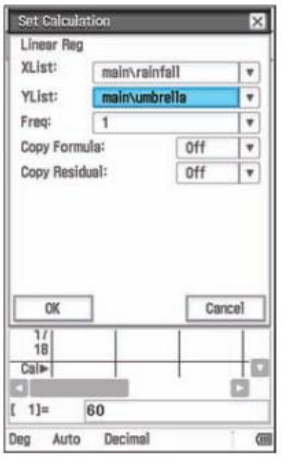
$r = 0.7106$

ClassPad

Using the Statistics application, rename two columns and type in the data.



Tap **Calc**, then select Regression, then **Linear Reg**.
Complete the screen as shown and tap OK.



9780170250283

A screen appears with the r value being the third value down.



$r = 0.7106$

c Use the r value to comment on the relationship between rainfall and umbrellas sold.

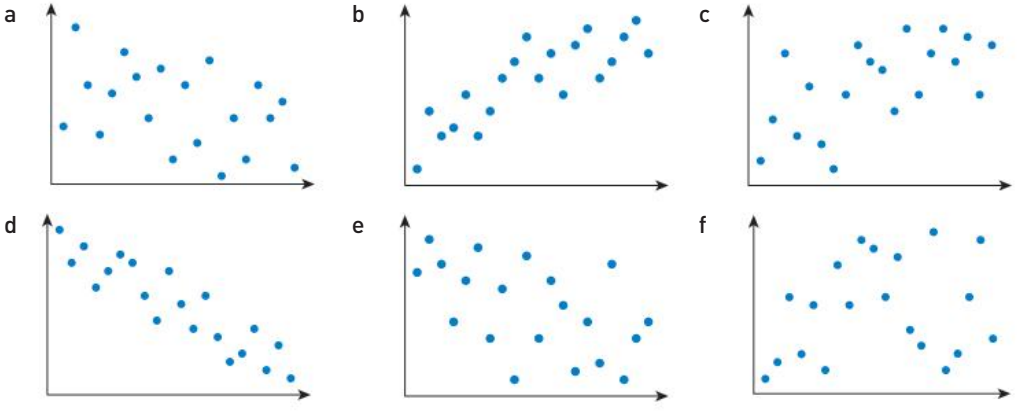
An r value of 0.7106 indicates that there is a moderate, positive linear relationship between rainfall and the numbers of umbrellas sold. There is some evidence to suggest that the number of umbrellas sold should increase as rainfall increases.

EXERCISE 1.08 Interpretation of scatter plots

Investigating relationships between two numerical variables

Concepts and techniques

1 Example 23 Estimate the correlation coefficient of the following graphs.



2 Describe the linear relationship that these r values represent.

a 0.7 b −0.2 c −0.62

d 0.33 e 0.95 f −0.8

3 Example 24 Interpret the correlation values and write a brief sentence describing the relationship for the studies described below.

a A study is investigating whether the amount of exercise a person does depends on their height. The correlation coefficient was found to be $r = 0.2946$.

b A study is investigating whether weight depends on the number of hours a person spends sitting down. The correlation coefficient was found to be $r = 0.5283$.

c A study is investigating whether the percentage of good peaches depends on their number of weeks in storage. The correlation coefficient was found to be $r = -0.9108$.

d A study is investigating whether the number of sales at a local store depends on the temperature. The correlation coefficient was found to be $r = 0.7241$.

e A study is investigating whether the number of pets owned depends on temperature. The correlation coefficient was found to be $r = -0.0628$.

f A study is investigating whether the number of home-cooked meals depends on income. The correlation coefficient was found to be $r = -0.4637$.

4 Example 25 The shoe sizes and heights of some students were measured and the results are shown below.

Shoe size	8	5	11	5	7	6	7	7	10
Height (cm)	168	158	182	162	162	157	168	171	178

a Draw a scatter plot representing the information in the table.

b Calculate the correlation coefficient correct to four decimal places.

c Discuss if there is a relationship between shoe size and height.

5 At a local cinema, people were asked their yearly income and the number of times they attended the cinema a year (on average).

Income ($000 per year)	63	125	32	19	136	49	102	67	82	25	91	42
Number of movies per year	4	12	1	8	16	13	13	6	7	6	12	10

a Draw a scatter plot representing the information in the table.

b Calculate the correlation coefficient correct to four decimal places.

c Discuss if there is a relationship between income and the number of movies watched per month.

6 The hot dog sales at a particular milk bar and the daily maximum temperature were recorded over a period of time.

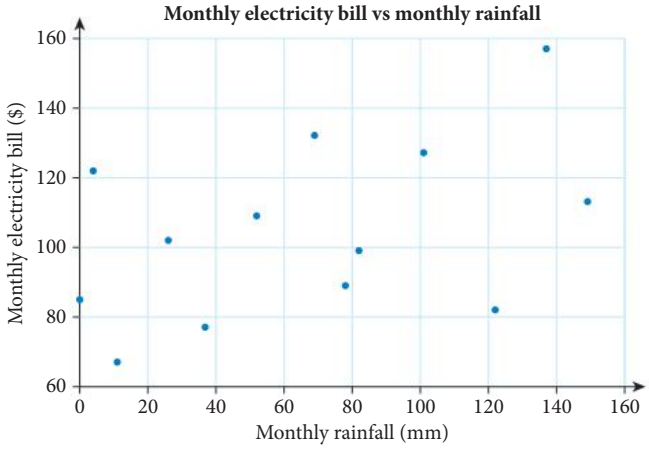
Temperature	5	19	38	24	31	16	21	36	41	28	32	12	35
Hot dog sales	96	62	5	22	46	32	37	11	16	42	24	88	67

a Calculate the correlation coefficient correct to four decimal places.

b Discuss if there is a relationship between temperature and hot dog sales.

9780170250283

7 Calculate the correlation coefficient (correct to four decimal places) for the scatter plot shown below and discuss if there is a relationship between monthly rainfall and monthly electricity bill.



Questions **8** and **9** refer to this table.

Hours per week spent doing paid work	26	37	49	19	0	11	40	38	29	32	78	39	62
Hours per week spent doing unpaid work	4	10	11	8	13	6	7	9	7	6	2	11	9

8 The correlation coefficient for this data is:

A −0.4174 **B** −0.3402 **C** −0.0421 **D** 0.0421 **E** 0.3402

9 Which of the following sentences best describes the relationship between the number of hours of paid work and the number of hours of unpaid work?

A There is limited evidence to sggest that the number of hours of unpaid work should increase as the number of hours of paid work increases.

B The number of hours of unpaid work should increase as the number of hours of paid work increases.

C There is some evidence to suggest that the number of hours of unpaid work should decrease as the number of hours of paid work increases.

D There is limited evidence that to suggest that the number of hours of unpaid work should decrease as the number of hours of paid work increases.

E There is no relationship between the number of hours of paid work and the number of hours of unpaid work.

Reasoning and communication

10 The annual rainfall and number of bushfires in a region were recorded over 8 years and the results are shown below.

Rainfall (cm)	16	14	19	28	22	24	23	20
No. of bushfires	20	32	26	10	18	11	9	21

a Sketch a scatter plot for this information.

b Interpret the scatter plot in terms of form, direction and strength.

c Calculate Pearson's correlation coefficient correct to 2 decimal places.

d Hence interpret the correlation value and write a brief sentence describing the relationship between the two variables.

11 There is a theory that growing old affects a woman's haemoglobin level. The ages of 20 women and their haemoglobin levels (in grams per decilitre) were recorded and are shown in the table.

a Sketch a scatter plot for this information.

b Interpret the scatter plot in terms of form, direction and strength.

c Calculate Pearson's correlation coefficient correct to 2 decimal places.

d Hence interpret the correlation value and write a brief sentence describing the relationship between the two variables.

e Justify whether you support the theory that growing old affects a woman's haemoglobin level.

Age (years)	Haemoglobin level (g/dL)
20	11.1
22	10.7
25	12.4
28	14.0
28	13.1
31	10.5
32	9.6
35	12.5
38	13.5
40	13.9
45	15.1
49	13.9
54	16.2
55	16.3
57	16.8
60	17.1
62	16.6
63	16.9
65	15.7
67	16.5

12 The population of birds in a particular area over ten years was recorded and the results are shown below.

Year	1	2	3	4	5	6	7	8	9	10
Population	984	699	512	403	347	320	310	304	299	296

a Draw a scatter plot for the data in this table and calculate the correlation coefficient.

b Describe the data using the *r* value.

c Explain why the *r* value is not useful in describing this data.

13 The following data was collected on the number of seconds for which people of different ages could hold their breaths.

Age	16	45	36	27	67	75	49	19	22	51	58	31
Time holding breath (seconds)	61	31	42	49	11	72	39	64	51	32	22	39

a Draw a scatter plot for the data in this table and calculate the correlation coefficient.

b Describe the data using the *r* value.

c Explain why the *r* value is not useful in describing this data.

d Recalculate the *r* value, excluding the outlier. What do you find?

INVESTIGATION How can you predict a person's height?

1 For all the members of your class (minimum 20 people), collect the following information and record it in a table.

- Middle finger length
- Height
- Arm length
- Circumference of head



Shutterstock.com/Pavel L Photo and Video

Student	Middle finger length	Height	Arm length	Head circumference
1				
2				

2 a Construct a scatter plot of finger length versus arm length, assuming that arm length depends on finger length.

b Visually interpret the relationship from the scatter plot, referring to form, direction and strength.

3 a Construct scatter plots for:

- height versus finger length
- height versus arm length
- height versus head circumference.

Assume that height is the dependent variable in each case.

b Calculate Pearson's correlation coefficient for each of the scatter plots constructed in question 3 part a.

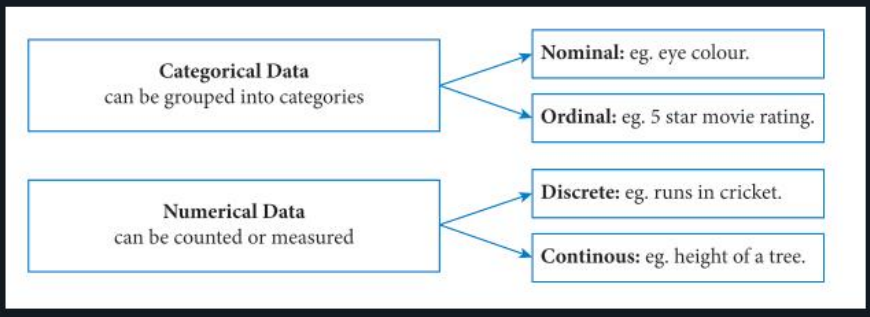
c Hence discuss the most efficient way to predict a person's height, using evidence to support your answer.

1

CHAPTER SUMMARY
DATA ANALYSIS

- The statistical investigation process involves identifying an area to study and posing a statistical question that allows for data collection, data analysis, interpretation and communicating results.

Types of data



- In a **random sample**, every member of the population has an equal chance of being selected, therefore minimising the chance of bias. The most convenient way to select a random sample is by generating random numbers on a calculator.
- Statistical measures of centre and spread help describe a data set.

Measures of centre are:

Mean	Median	Mode
$\bar{x} = \frac{\text{sum of all the values}}{\text{number of values}} = \frac{\sum x}{n}$	Middle number when in order and the position of the median is the $\frac{n+1}{2}$ th data value.	Most common or data value with the highest frequency.

The median is the most appropriate measure of centre when data is skewed or has outliers

Measures of spread are:

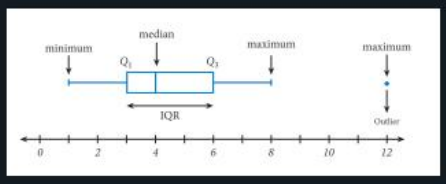
Range	Interquartile range	Standard deviation
highest data value − lowest data value	IQR = $Q_3 - Q_1$	$s_x = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$ OR using a CAS

The IQR is the most appropriate measure of spread when data is skewed or has outliers.

9780170250283

- Graphical displays such as frequency tables, column graphs, dot plots, stem-and-leaf plots, histograms and boxplots allow for visual interpretation of data and can often assist in the calculation of summary statistics. When constructing boxplots, an outlier is a data value that lies outside the interval

  $Q_1 - 1.5 \times \text{IQR} \leq x \leq Q_3 + 1.5 \times \text{IQR}$

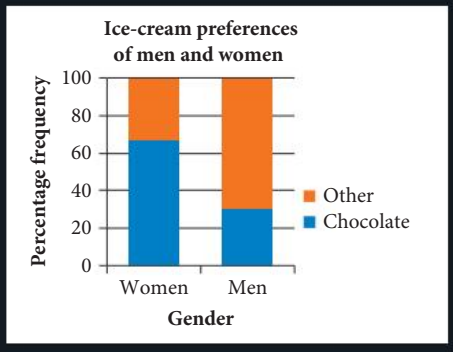


- The shape of the data represented on a graphical display can be described as symmetrical, positively skewed or negatively skewed.
- When looking for relationships between two variables, it is important to identify the explanatory variable (not affected by the other variable) and the response variable (is affected by the other variable).
- A two-way table is used to display two different categorical variables, each with two or more characteristics.

	Explanatory variable		
Response variable	**Characteristic 1**	**Characteristic 2**	
Characteristic 1			Row sum
Characteristic 2			Row sum
Characteristic 3			Row sum
	Column sum	Column sum	

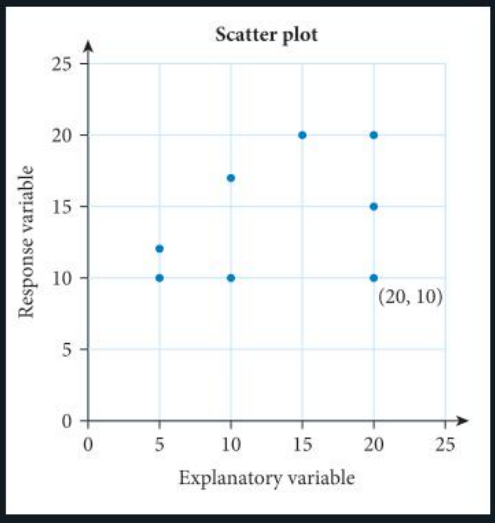
- A percentage two-way table allows for appropriate comparison of variables and can be used to construct a segmented column graph.

	Gender	
Ice-cream flavour	**Women**	**Men**
Chocolate	67%	30%
Other	33%	70%
	100%	100%



- When looking at the relationship between a numerical variable and a categorical variable, the appropriate graphical displays are back-to-back stem-and-leaf plots and parallel boxplots. Parallel boxplots are more appropriate when the data set is large. Comparisons can be made in relation to shape and summary statistics and a relationship between the two variables can be identified if it exists.

- A scatter plot is a graphical display used to investigate the relationship between two numerical variables.



- To describe the relationship for a scatter plot, discuss:

1 Form: linear, non-linear or no relationship
2 Direction: positive or negative
3 Strength: strong, moderate or weak

- Pearson's correlation coefficient (r) is a value that can be used to determine the direction and exact strength of a linear relationship. The value is found using a CAS calculator.
- The interpretation of a correlation coefficient is given in the table below.

Value of r	What it means
1	Perfect positive relationship
$0.75 \leq r < 1$	Strong positive relationship
$0.5 \leq r < 0.75$	Moderate positive relationship
$0.25 \leq r < 0.5$	Weak positive relationship
$-0.25 < r < 0.25$	No relationship
$-0.5 < r \leq -0.25$	Weak negative relationship
$-0.75 < r \leq -0.5$	Moderate negative relationship
$-1 < r \leq 0.75$	Strong negative relationship
-1	Perfect negative relationship

9780170250283

CHAPTER REVIEW
DATA ANALYSIS
1

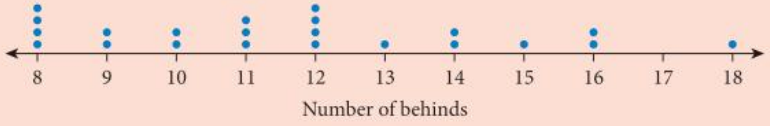
Multiple choice

1 A random sample is selected from a population to:
- A select data from the population to support your hypothesis
- B manipulate the data to falsify findings
- C use less data to save time
- D eliminate bias when selecting a sample because the population is large
- E eliminate patterns in the data

2 Peter recorded the number of fours he hit in a cricket season. The type of data collected is best described as;
- A Categorical, nominal data
- B Numerical, ordinal data
- C Categorical, ordinal data
- D Numerical, discrete data
- E Numerical, continuous data.

The following information relates to questions **3** and **4**.

The dot plot below displays the number of behinds kicked in each match by the Sydney Swans during the 2012 AFL season.

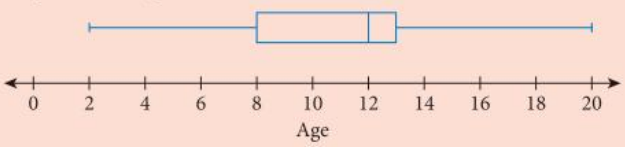


3 The measures of centre are:
- A mean = 11.5, median = 11.7, mode = 8
- B mean = 11.6, median = 11.5, mode = 8 and 12
- C mean = 11.7, median = 11.5, mode = 8 and 12
- D mean = 11.6, median = 11.5, no mode
- E mean = 11.7, median = 11.5, no mode

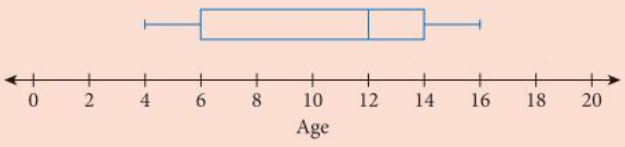
4 The measures of spread are:
- A range = 18, IQR = 5, standard deviation = 2.9
- B range = 10, IQR = 5, standard deviation = 2.8
- C range = 5, IQR = 10, standard deviation = 2.8
- D range = 18, IQR = 11.5, standard deviation = 2.9
- E range = 10, IQR = 5, standard deviation = 2.9

5 Select the boxplot below that represents the variable age that has a range of 12 and an interquartile range of 5.

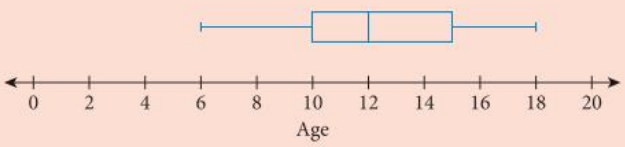
A



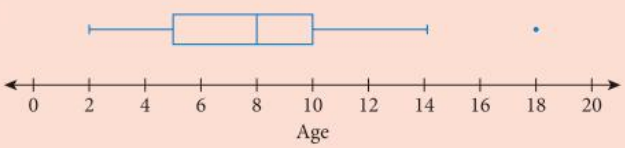
B



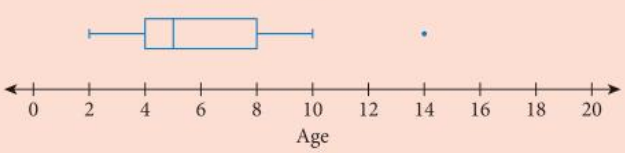
C



D



E



6 Which of the following sets of variables are classified **incorrectly** as explanatory (EV) and response (RV)?

A EV = age, RV = height
B EV = amount of exercise, RV = fitness level
C EV = amount of sugar consumed, RV = tooth decay
D EV = distance travelled, RV = time travelling
E EV = age of car, RV = selling price of car

7 The two-way table below shows the numbers of men and women who voted for the introduction of paid maternity leave.

	Gender		
Paid Maternity leave	**Men**	**Women**	
Yes	10	62	72
No	19	24	43
	29	86	115

Which one of the following statements is **true**?

A A lower percentage of women voted yes than men.
B A higher percentage of men voted no than women.
C 44% of people voted no.
D There is no relationship between gender and voting for paid maternity leave.
E Of those that voted yes, 72% were male.

9780170250283

8



The scatter plot above shows that:

A there is no relationship between a husband's height and a wife's height
B there is a strong negative relationship between a husband's height and a wife's height
C there is a non-linear relationship between a husband's height and a wife's height
D there is a moderate positive relationship between a husband's height and a wife's height
E there is a strong positive relationship between a husband's height and a wife's height.

9 The variables 'time spent exercising' and 'kilojoules burnt' are investigated. Which of the following graphical displays would be the most appropriate to use?

A scatter plot
B parallel boxplots
C segmented column graph
D back-to-back stem-and-leaf plot
E dot plot

Questions **10** and **11** refer to the following information.

The correlation coefficient for the time a salesman has been with a real estate company and the number of properties sold per week was found to be –0.28.

10 The relationship between the time a salesman has been with a real estate company and the number of properties sold per week can be described as:

A a weak positive relationship
B a weak negative relationship
C no relationship between the two variables
D a moderate negative relationship
E a strong positive relationship.

11 Which of the following sentences summarises this?

A It can be concluded that the number of properties sold per week should decrease as the time a salesman has been with a real estate company increases.

B There is no evidence to suggest that the number of properties sold per week is affected by the time that a salesman has been with a real estate company.

C There is limited evidence to show that the time a salesman has been with a real estate company should decrease as the number of properties sold per week increases.

D There is some evidence to show that the time a salesman has been with a real estate company should decrease as the number of properties sold per week increases.

E There is limited evidence to show that the number of properties sold per week should decrease as the time a salesman has been with a real estate company increases.

Short answer

12 Classify which type of data each of the following variables represent.

a Temperature of water in a swimming pool.

b Number of people who voted for The Greens in the last election.

c Response time when a patient's reflexes are tested.

d Breeds of dog.

13 For each data set below, calculate (correct to two decimal places):

i the mean, median and mode

ii the range, interquartile range and standard deviation

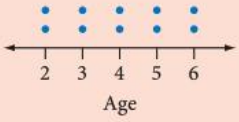
a 8 3 5 2 0 7 9 5 4

b 1 5 18 72 6 26 13 18 7 6 9 11 19
18 22 6 17 9 12 14 2 17 13 19 17

c

x	f
10	9
11	8
12	3
13	7

d



e

Stem	Leaf
3	1 2 4 4 5
4	0 3
5	3 5 5 5 7
6	8 9
7	3 3

Key: 7|3 = 7.3

9780170250283

14 Joshua swam 1 km each morning for 10 days to prepare for a swimming carnival. His times (in minutes) were:

28 24 22 24 25 24 26 26 24 27

a What is his median swim time?
b What is his mean swim time?
c What is the interquartile range of times?
d What is the sample standard deviation of his times (correct to one decimal place)?
e If Joshua asked you to tell him the most appropriate measures of location and spread for these times, which two would you choose? Justify your answer.

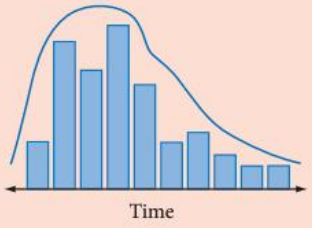
15 Describe the shape of each distribution shown below.

a Marks in a Japanese test (in percentages).

Stem	Leaf
3	1 2
4	3 5 9
5	0 2 6
6	4 5
7	3 5 6 7 7 8 9
8	0 2 4 6 6 8 8 9
9	1 2 4 8 9

Key: 7|3 = 7.3

b Amount of traffic on Sydney's roads on weekday mornings.



c



16 These are the ages of employees at the Berry Good Biscuit factory.

16 36 15 16 15 19 55 59 18 20

50 22 21 35 22 19 15 17 43 49

a Draw a stem-and-leaf plot for this data.
b Construct a boxplot to represent this information.
c Show if there are any outliers.
d Comment on the shape of the distribution.

17 The two-way table below describes the employment status of a group of men and women.

	Gender		
Employment status	**Male**	**Female**	
Casual work	438		980
Full-time work			
	1040		1851

a Complete the two-way table by filling in the missing information.
b Convert the table to a percentage two-way table.
c Hence construct a segmented column graph.
d Discuss if there is an association between gender and employment status.

18 For each of these data sets:
i draw a scatter plot
ii calculate Pearson's correlation coefficient (correct to four decimal places)
iii describe the relationship shown in terms of form, direction and strength
iv write a brief sentence to describe the relationship between the two variables.

a People were asked how many years they had studied for (including primary and high school as well as university and TAFE, etc.) and asked how many children they had.

Years of education	15	18	13	13	12	17	17	18	16	10	19	19	18	22	16	16
Number of children	2	4	6	3	1	2	1	2	0	7	1	3	4	2	1	2

b Students enrolled in a computer course were given an aptitude test before the course started and a final exam at the end of it. The marks of the two tests were as follows.

Aptitude test (%)	18	70	35	53	44	38	10	79	26	62
Final exam (%)	53	70	44	80	62	35	9	98	62	88

c Students enrolled in both English and Italian were asked their final exam mark at the end of the year.

English exam mark	23	78	46	89	16	77	63	52	97	46	56	57	36	72
Italian exam mark	16	92	63	52	25	72	49	15	92	89	52	43	42	81

9780170250283

Application

19 The back-to-back stem-and-leaf plot displays the lifetimes, in hours, of two brands of light globes.

Oso Bright	Stem	Brighta Longa
5 5 4 2	10	3 4 5
8 7 7 7 7 7 4 4 3	11	2 3 3 4 4 5 6
9 9 8 8 7 6 6 6 5 4 4 0	12	1 2 2 3 3 4 5 5 7 9 9 9
8 8 8 7 7 7 6 5 4 3 1	13	0 2 3 3 4 4 4 5 6 6 8 8 9
9 8 8 8 5 5 2 2	14	1 2 2 3 5 5 6 7 8 8
7 7 5 1	15	0 3 3 4 6

Key: 10|2 = 102 hours

a Find the mean, median and mode for each brand. Are they approximately equal?
b Calculate the standard deviation.
c Which brand has the most consistent lifetime?
d Construct parallel boxplots to represent this data.
e Between what lifetime hours does the middle 50% of the data lie for each brand of light globe?
f Discuss if there is an association between the brand of light globe and its lifetime.

20 Year 12 students were surveyed to find the average number of photographs they took a year (using their phone), and the amount of space free on their phone (in Gb).

a Calculate the measures of centre and spread for each data set (number of photos taken, amount of free space).
b Draw a boxplot for each data set, marking in any outliers.
c Comment on the centre and spread of both of these data sets.
d Draw a scatter plot of the data.
e Calculate Pearson's correlation coefficient, correct to four decimal places, and interpret the result.
f Summarise your findings, referring to both the individual data sets and the relationship between them.

Number of photographs taken using phone a year	Free space on phone (Gb)
1250	2
1000	11
500	16
750	8
200	22
1600	5
2200	2
100	14
1500	4
1400	10
2000	1
250	6
1200	15
1800	6
2000	12

Practice quiz





Shutterstock.com/zhu difeng

2

TERMINOLOGY

association/relationship
coefficient of determination
dependent variable
extrapolation
gradient
independent variable
interpolation
least squares regression line
numerical data
Pearson's correlation coefficient
predictions
residual
residual plot
scatter plot
y-intercept

Prior learning

BIVARIATE DATA ANALYSIS

LINEAR MODELS FOR NUMERICAL DATA

2.01 Fitting regression lines
2.02 The coefficient of determination
2.03 Making predictions
2.04 Residuals
Chapter summary
Chapter review

FITTING A LINEAR MODEL TO NUMERICAL DATA

- model a linear relationship by fitting a least squares line to the data (ACMGM057)
- use a residual plot to assess the appropriateness of fitting a linear model to data (ACMGM058)
- interpret the intercept and slope of the fitted line (ACMGM059)
- use the coefficient of determination to assess the strength of a linear association in terms of the explained variation (ACMGM060)
- use the equation of a fitted line to make predictions (ACMGM061)
- distinguish between interpolation and extrapolation when using the fitted line to make predictions, recognising the potential dangers of extrapolation (ACMGM062)
- write up the results of the above analysis in a systematic manner. (ACMGM063)

THE DATA INVESTIGATION PROCESS

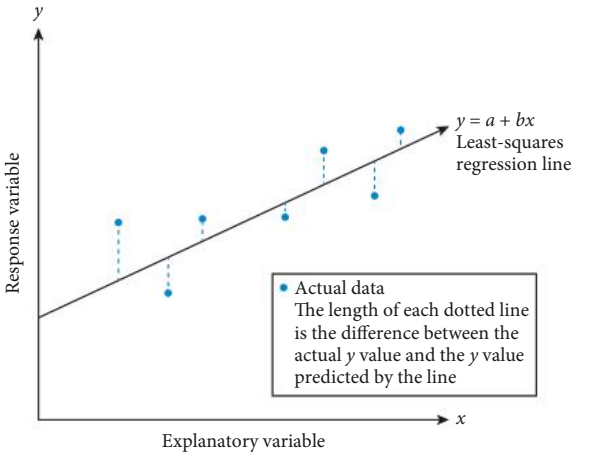
- implement the statistical investigation process to answer questions that involve identifying, analysing and describing associations between two categorical variables or two numerical variables; for example; is there an association between attitude to capital punishment (agree with, no opinion, disagree with) and sex (male, female)? is there an association between height and foot length? (ACMGM066) AC

2.01 FITTING REGRESSION LINES

In the previous chapter we studied how to interpret a scatter plot for a set of data in terms of form, direction and strength supported by the calculation of Pearson's correlation coefficient. Once it has been determined that two numerical variables have a relatively strong linear **association**, a straight line can be fitted into the data, creating a model that can be used to make predictions.

Diagram 1 shows a least squares regression line fitted to a data set.

Diagram 1



A least squares regression line is the linear fit that minimises the sum of the squares of the vertical distances (dotted lines in diagram 1) between the line and each point.

IMPORTANT

The general form of a **least squares regression line** is $y = a + bx$ where the a represents the **y-intercept** and b represents the **gradient**.

The coefficients a and b will often be decimal values that will require rounding to a specified number of decimal places.

The x-axis variable is the independent or explanatory variable and the y-axis variable is the dependent or response variable. When a least squares regression line equation is found, the equation should be written in terms of the variable names rather than using y and x.

The least squares regression line is generally considered to be the most accurate linear model. The calculations are quite tedious and as a result, for the purposes of this course, we are only required to use a CAS calculator to find its equation when given the bivariate data.

Example 1

The following table shows the results of an experiment that measures the temperature of a liquid as it cools down.

Time (minutes)	5	10	15	20	25	30
Temperature (°C)	87	78	69	56	53	41

This data is to be used to find a mathematical relationship that will enable the temperature at a certain time to be predicted.

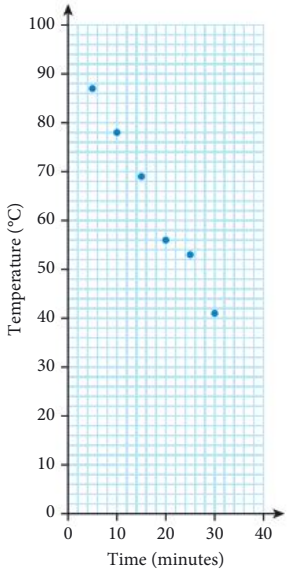
a Construct a scatter plot to represent this data.

b Calculate the correlation coefficient, correct to 2 decimal places, and interpret this value.

c Find the equation of the least squares regression line, writing the coefficients correct to 1 decimal place.

Solution

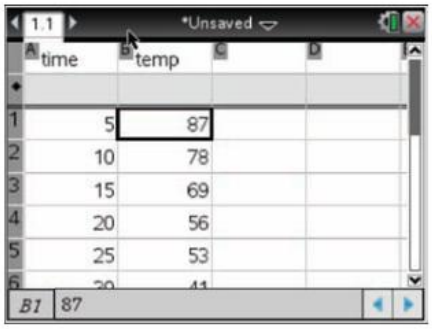
a Time is the explanatory variable and Temperature is the response variable. Construct axes with Time on the horizontal axis and Temperature on the vertical axis. Plot a point for each time and corresponding temperature.



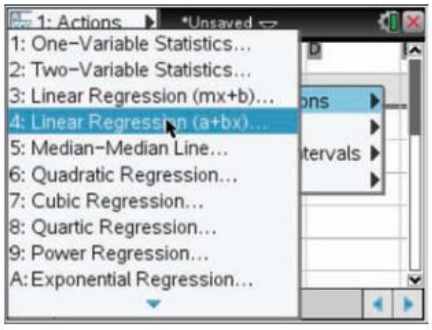
b Use a CAS calculator to calculate *r*.

TI-Nspire CAS

Enter the data into a Lists & Spreadsheet page, remembering to name each column.

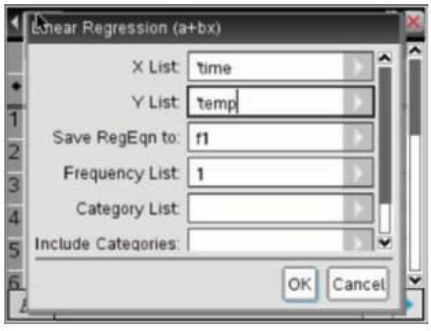


Press menu, then 4: Statistics, 1: Stat Calculations, 4: Linear Regression (a + bx).



Select 'time' for the X List variable and 'temp' for the Y List variable, then click on OK.

When you calculate *r* using a CAS calculator, the *a* and *b* values for the least squares regression line are also calculated.



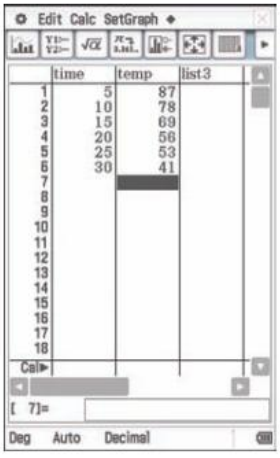
A screen appears, scroll down to find the *r* value.



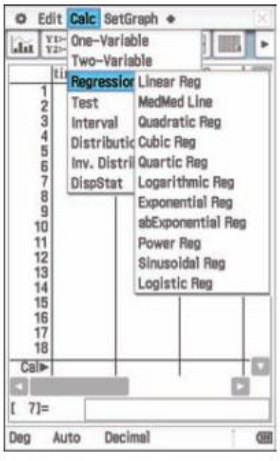
9780170250283

ClassPad

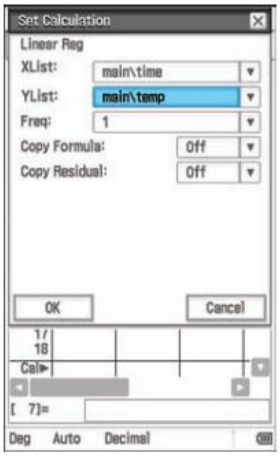
Using the **Statistics** application, rename two columns and type in the data.



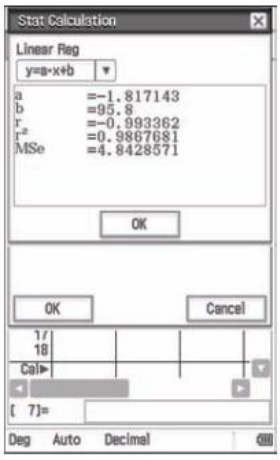
Tap **Calc**, then **Regression**, then select **Linear Reg**.



Complete the screen as shown and tap OK.



A screen appears with the r value being the 3rd value down.



r gives a measure of the strength of the linear relationship between the variables.

$r = -0.99$, which indicates that there exists a strong negative relationship between the variable temperature and time. It can be concluded that temperature decreases as time increases.

c **TI-Nspire CAS**

Read the a and b values from the screen and round to one decimal place.

$y = a + bx$
$a = 95.8$ and $b = -1.8$

Write the equation.

Replace y and x in the equation with the correct variable names.

$y = 95.8 - 1.8x$
$temperature = 95.8 - 1.8 \times time$

When the r value is negative, the gradient of the least squares regression line will also be negative and when the r value is positive, the gradient of the least squares regression line will be positive.

ClassPad

Read the a and b values from the screen and round to one decimal place.

$y = ax + b$
$a = -1.8$ and $b = 95.8$

Write the equation.

$y = -1.8x + 95.8$

Replace y and x in the equation with the correct variable names.
NOTE: the Classpad gives the least squares regression line in the form $y = ax + b$, so the gradient is a and the y-intercept is b.

$temperature = -1.8 \times time + 95.8$

Once a least squares regression line ($y = a + bx$) has been found, the y-intercept (a) and gradient (b) can be interpreted in terms of the variables being investigated.

9780170250283

For Example **1**:

The *y*-intercept was 95.8, which tells us that the temperature was 95.8°C when the time was 0 minutes, i.e., at the start.

The gradient was –1.8, which tells us that the temperature decreases by 1.8°C for every 1 minute increase in time.

IMPORTANT

To interpret the *y*-intercept, we generally state "The ***y* variable** is "*a*" **units** when the ***x* variable** is zero **units**.

To interpret the **gradient**, we generally state "The ***y* variable** increases/decreases by *b* **units** for every 1 **unit** increase in the ***x* variable.**

Use the word increases when gradient is positive and decreases when gradient is negative.

The words in bold are replaced with the appropriate variable names or units of measure for each variable.

In some cases the interpretation of the *y*-intercept will not be relevant to the situation. For example, a regression line relating the amount of hours studying for a test and the test result was found to be

$$\text{Test result} = -5.3 + 10.3 \times \text{hours of study}$$

The *y*-intercept is –5.3, which tells us that a test result will be –5.3 when the amount of study time is 0 hours. Clearly, a test result cannot be a negative value so the interpretation of the *y*-intercept is not sensible in this instance.

Example 2

Identify and interpret the gradient of the regression line with equation

$$\text{Hand span} = 2.1 + 0.2 \times \text{height}$$

where hand span and height are measured in cm.

Solution

The gradient is the coefficient of height in the equation.	The gradient is 0.2.
A gradient of 0.2 means that there is a 0.2 unit increase in the response variable for every 1 unit increase in the explanatory variable.	There is a 0.2 cm increase in hand span for every 1 cm increase in height.

The equation of the least squares regression line can also be found using the value of **Pearson's correlation coefficient** and the mean and standard deviation for each variable.

IMPORTANT

The equation of the least squares regression line is $y = a + bx$, where

the **gradient**, $b = r \times \frac{s_y}{s_x}$ and the y-intercept, $a = \bar{y} - b\bar{x}$

The first step in this method is to calculate the b value and then a. This is because b is required for the calculation of a.

$\bar{x}$ = the mean value of the x variable

$\bar{y}$ = the mean value of the y variable

r = Pearson's correlation coefficient

s_x = the standard deviation of the x variable

s_y = the standard deviation of the y variable

Example 3

The runs scored (y) and the balls faced (x) by a batsman were recorded and the values of the following statistics were determined.

$\bar{x} = 22.45$ $s_x = 15.81$ $\bar{y} = 14.91$ $s_y = 11.74$ $r = 0.89$

Calculate the least squares regression line that models this data, stating all coefficients correct to 2 decimal places.

Solution

Calculate the gradient (b) using the formula.

$$b = r \times \frac{s_y}{s_x}$$
$$= 0.89 \times \frac{11.74}{15.81}$$

Round correct to 2 decimal places.

$$\approx 0.66$$

Calculate the y-intercept (a) using the formula.

$$a = \bar{y} - b\bar{x}$$
$$= 14.91 - 0.66088 \times 22.45$$

Round correct to 2 decimal places.

$$\approx 0.09$$

It is important to use the unrounded answer for b when calculating the value of a.

Write the equation for the least squares regression line.

$a = 0.09$ and $b = 0.66$

$y = 0.09 + 0.66x$

Replace y and x in the equation with the correct variable names.

Runs scored = 0.09 + 0.66 × balls faced

Once a **scatter plot** has been constructed and the least squares regression line has been found it can be drawn onto the scatter plot. To draw a least squares regression line by hand, do the following:

Step1: Read the y-intercept from the regression line and plot a point at this value on the y-axis of the scatter plot.

Step 2: Select an x value that is represented on the horizontal scale of the scatter plot and substitute it into the regression line equation to find a corresponding value of y. Plot this point.

Step 3: Rule a line through the 2 points.

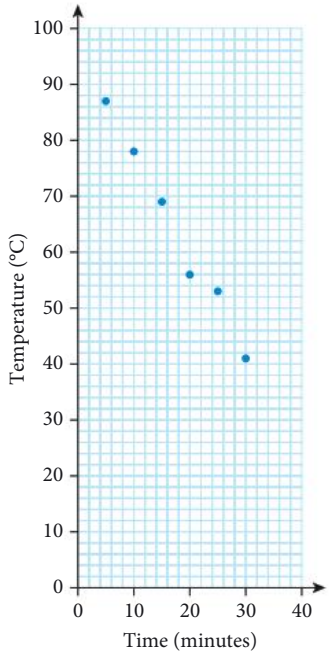
If the y-intercept cannot be plotted, skip Step 1 and repeat Step 2 to find a second set of coordinates.

Example 4

The regression line relating cooling time and liquid temperature was found to have the equation

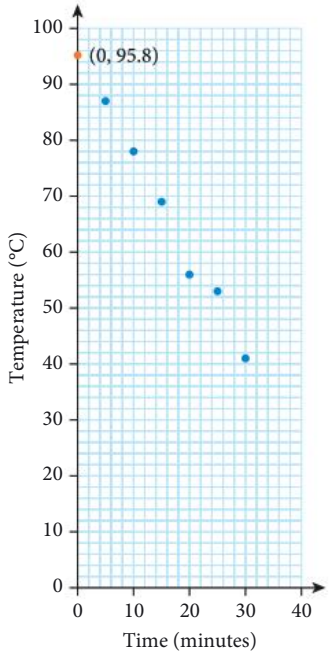
Temperature = 95.8 – 1.8 × time

Plot this regression line onto the scatter plot below.



Solution

Step 1: The y-intercept is 95.8, plot the point at (0, 95.8).



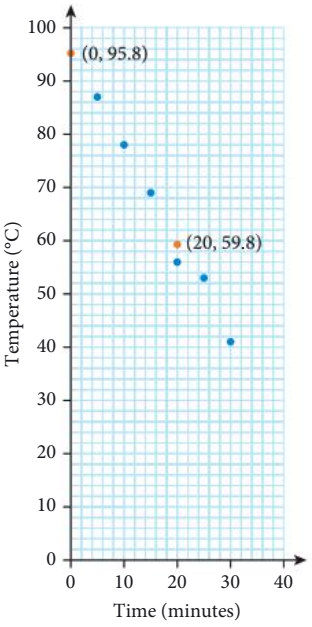
Step 2: Calculate a second set of coordinates.

When Time = 20

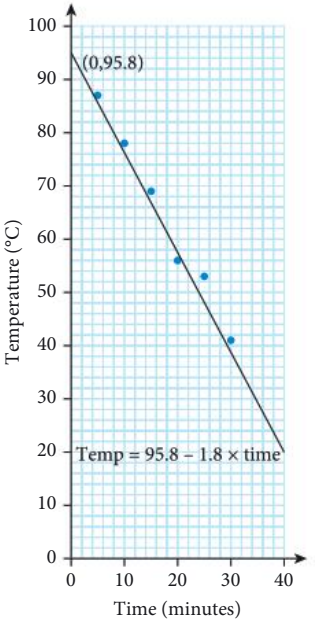
The value selected for time could have been any number between 0 and 30. Always try to select a value that is on a grid line and not too close to the vertical axis.

Temperature = 95.8 – 1.8 × 20
= 59.8
Therefore coordinates are (20, 59.8).

Plot the point (20, 59.8).



Step 3: Join the 2 points with a straight line.



9780170250283

Using a CAS calculator to add a linear regression line to a scatter plot.

Consider the data below.

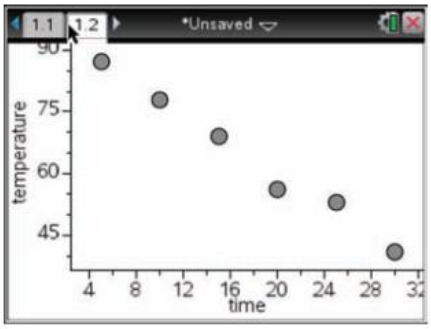
Time	5	10	15	20	25	30
Temp	87	78	69	56	53	41

The least squares regression line for this data has the equation

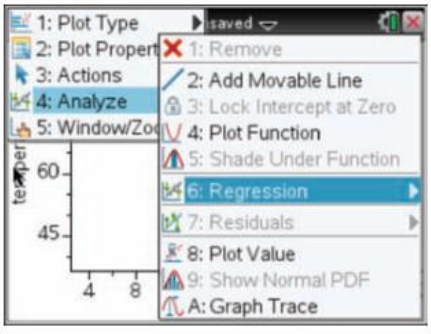
Temperature = 95.8 – 1.8 × time

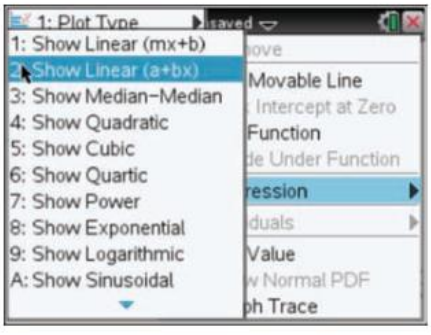
TI-Nspire CAS

Create a scatter plot.

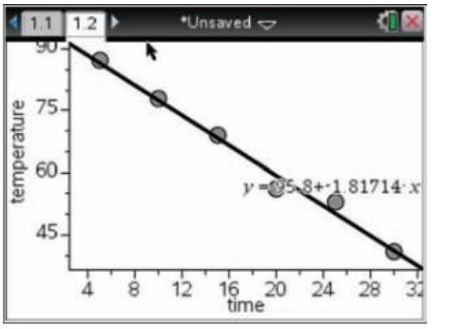


To add regression line, press menu 4: Analyze, 6: Regression, 2: Show linear(a + bx).





The regression line is now plotted on the scatter plot.

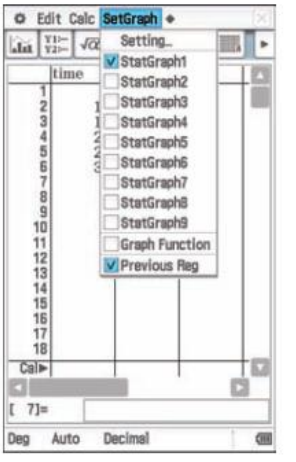


ClassPad

Using the **Statistics** application, enter the data into two lists.
Tap **Calc**, then select **Regression** then **Linear Reg** and complete the screen as shown.
Make sure that you select y1 in the Copy **Formula** menu, then tap OK.

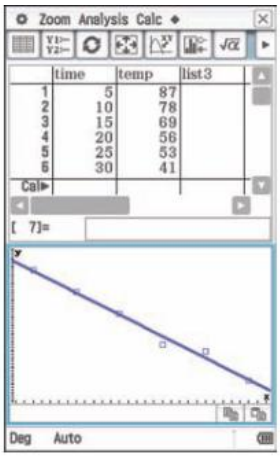


Create a scatter plot using **SetGraph**. The drop down menu should look like the one on the right.
Tap **Setting** and fill in the box to plot temp against time, then tap Set and to draw the graph.



9780170250283

The regression line will be plotted on the scatter plot. Remember to tap Resize for a full page.



EXERCISE 2.01 Fitting regression lines

Concepts and techniques

1 Example 1 Eleven students went on a hiking trip. Their heights and weights are shown in the table.

Height, H cm	150	151	151	152	153	153	156	157	158	160	161
Weight, W kg	54	55	58	61	60	62	61	64	63	60	65

a Construct a scatter plot to represent this data.
b Calculate the correlation coefficient, correct to 2 decimal places, and interpret this value.
c Find the equation of the least squares regression line, writing all coefficients correct to 1 decimal place.

2 The amounts of energy, E kilojoules, used per day by 9 men with weights W kg, are shown in the table.

Weight, kg	79	84	89	92	95	101	115	117	125
Energy, kJ	1620	1720	1820	1650	2030	1870	1980	2100	2340

This data is to be used to find a mathematical relationship that will enable the amount of energy used per day to be predicted from the weight.
a In this situation, the explanatory variable is _____
b Construct a scatter plot to represent this data.
c Calculate the correlation coefficient, correct to 2 decimal places, and interpret this value.
d Find the least squares regression line, writing all coefficients correct to 1 decimal place.
e Complete the following sentence: On average, for each kilogram of increase in weight, a person's energy usage increases by _____ .

3 Example 2 For the regression line found in question **1** part **c**, interpret the gradient.

4 For each of the following regression lines, interpret the gradient.
a Weight loss(kg) = 3 + 0.8 × amount of exercise(hrs)
b Height(cm) = −2.35 + 8.2 × age(years)
c Value of car($) = 32000−1235 × age of car(years)

5 Example 3 The annual rainfall (x) and number of bushfires (y) was recorded for a region and the values of the following statistics were determined.

$\bar{x} = 19.88$ $s_x = 6.31$ $\bar{y} = 18.38$ $s_y = 8.16$ $r = -0.88$

Calculate the least squares regression line that models this data, stating all coefficients correct to 2 decimal places.

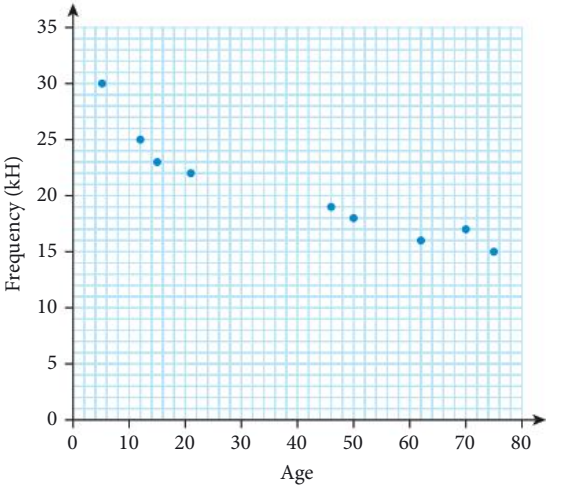
6 The number of umbrellas sold (y) and the amount of rain in millimetres (x) was recorded and the values of the following statistics were determined.

$\bar{x} = 76.5$ $s_x = 26.01$ $\bar{y} = 19.25$ $s_y = 13.90$ $r = 0.2$

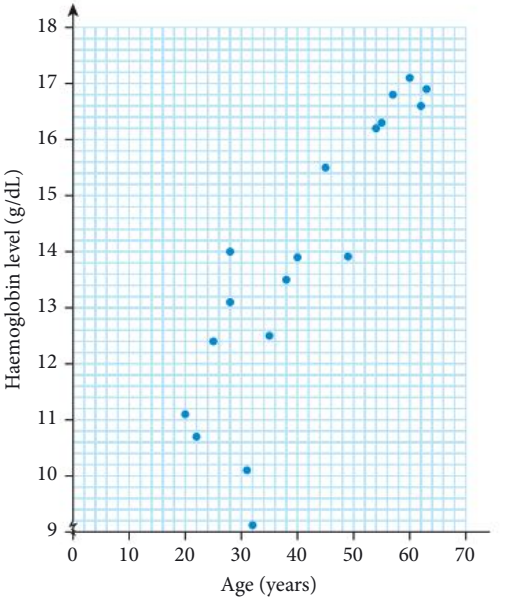
Calculate the least squares regression line that models this data, stating all coefficients correct to 2 decimal places.

7 Example 4 Sketch each of the following regression lines on the corresponding scatter plots.

a Frequency = 27.3 – 0.17 × age

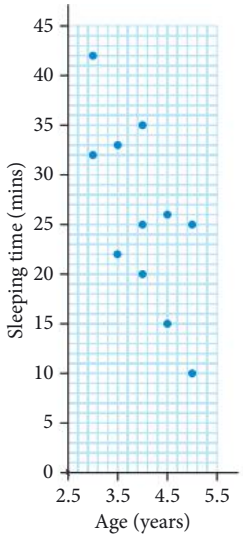


b Haemoglobin = 7.8 + 0.15 × age



9780170250283

c Sleeping time = 62.71 − 9.20 × age



Reasoning and communication

8 Happy Home Loans recorded the following data to see if there was a relationship between the interest rate and the number of home loan applications.

Interest rate (p.a.)	8.3	9.7	10.4	9.5	8.1	9.1	10.8	10.0
No. of applications	51	42	25	32	43	41	22	28

Where appropriate, state your answers correct to 2 decimal places.

a Calculate the mean and standard deviation for the variable interest rate.
b Calculate the mean and standard deviation for the variable number of applicants.
c State the explanatory variable and response variable.
d Hence construct a scatter plot.
e Calculate Pearson's correlation coefficient and interpret its meaning.
f Using your answers from parts **a**, **b** and **e**, find the equation of the regression line for the data, stating the coefficients correct to 1 decimal place.
g Interpret the y-intercept and gradient of the regression line found.
h Sketch the regression line on the scatter plot constructed in part **f**.

9 The numbers of enquiries to rent or buy a house were recorded over a 10-week period at Fastasales real estate agency. The data is shown in the table.

No. rental enquiries	16	34	70	58	52	25	38	46	62	49
No. buying enquiries	28	69	86	74	78	40	49	62	81	53

a Construct a scatter plot, assuming that the number of rental enquiries is the explanatory variable.
b Calculate Pearson's correlation coefficient and interpret its meaning.
c Find the equation of the least squares regression line for the data.
d Interpret the y-intercept and gradient of the regression line found.
e Sketch the regression line on the scatter plot constructed in part **a**.
f Based on all the previous answers, do you think the regression line is a suitable model for the data? Justify your answer.

10 The area of planting fields (m^2) and the number of pumpkins harvested was recorded for 15 different farms. The results are displayed in the table below.

A farmer is interested in finding a linear equation that allows the *number of pumpkins* harvested to be predicted from the *area* of a planting field.

a What will be the explanatory variable in the equation?

b Assuming there is a linear relationship, determine the least squares regression line that allows the *number of pumpkins* harvested to be predicted from the *area* of a planting field. State all coefficients correct to two decimal places.

c Using the equation found in part **b**, interpret the gradient of the least squares regression line in terms of the variables *area* and *number of pumpkins*.

No. of pumpkins	Area (m^2)
300	500
312	550
295	473
354	610
262	412
405	825
254	425
305	518
326	598
370	622
389	715
373	654
276	452
399	758
324	587

2.02 THE COEFFICIENT OF DETERMINATION

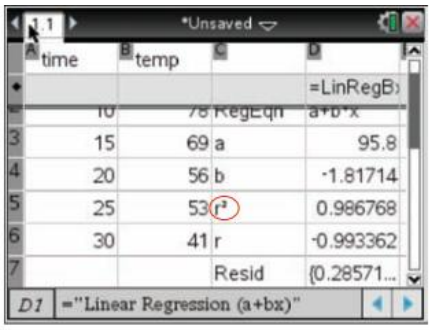
In the previous exercise we used least squares regression lines to model given sets of data. We should now ask ourselves "How well does our regression line actually represent our set of data?" To answer this question, the **coefficient of determination** or r^2 is used. It is a numerical value between 0 and 1 that is used to determine how useful a regression line is as a linear model for a given set of data. More specifically, when it is converted to a percentage it tells us the percentage of the variation in the dependent variable that is explained by the independent variable.

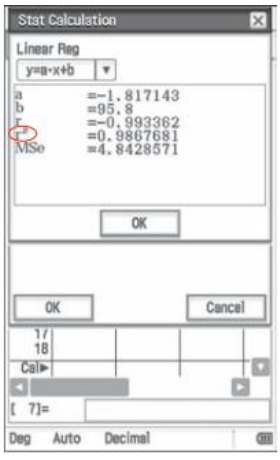
As the r^2 value ranges from 0 to 1 when converted to a percentage, it will range from 0% to 100%.

To calculate the coefficient of determination when a set of data is provided, a CAS calculator can be used. The r^2 value appears on the screen when a regression line is calculated (refer to Section 2.01, Example **1** for the steps).

Remember to determine the explanatory variable (x) and response variable (y) before performing any calculations.

9780170250283





IMPORTANT

When interpreting the coefficient of determination, a general sentence can be used; $r^2 \times 100\%$ of the variation in the **response** or **dependent variable** can be explained by the variation in the **explanatory** or **independent variable**.

Throughout the sentence, the **bold text** is replaced with the appropriate percentage and variable names of the set of data being investigated.

As the r^2 value gets higher the

1 relationship between the variables gets stronger

2 regression line becomes a more appropriate model for the data.

The r^2 value does not tell us about the direction (positive or negative) of a relationship as it is always between 0 and 1. It only informs us about how the variation in the two variables can be explained.

For example, if r^2 was 0.80, then we can state that 80% of the variation in the dependent variable can be explained by the variation in the independent variable or, alternatively, that 20% of the variation in the dependent variable is not explained by the variation in the independent variable.

Example 5

Data was collected to investigate the relationship between the minimum daily temperature and the maximum daily temperature and is displayed in the table below.

Minimum temperature (°C)	12	15	13.5	14.1	10.2	11.7	13.2	11.3
Maximum temperature (°C)	23.4	25.6	23.1	25.3	20	21.1	22.6	21

a Assuming that Minimum temperature is the explanatory variable, calculate the coefficient of determination, correct to 3 decimal places.

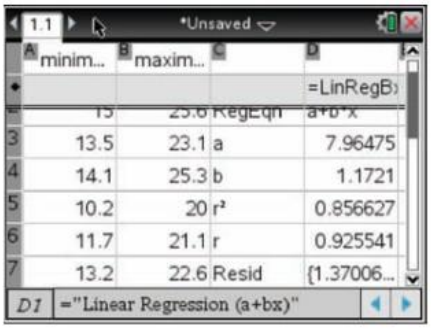
b Interpret the coefficient of determination.

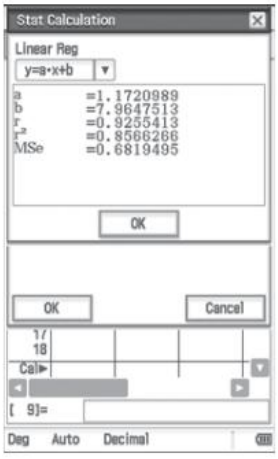
c State the regression line that models this data.

d Do you think that the model is appropriate? Justify your answer.

Solution

a Use a CAS calculator to find the linear regression line. The value of r^2 is also calculated.





$r^2 = 0.857$

b Calculate $r^2 \times 100\%$, then interpret the result using a general sentence.

85.7% of the variation in the maximum temperature can be explained by the variation in the minimum temperature.

c Read the a and b values for the linear regression line equation from the screen and round to two decimal places.
Write the equation using these constants.

$y = 7.96 + 1.17x$

Replace y and x in the equation with the correct variable names.

Maximum temperature = 7.96 + 1.17 × minimum temperature

d Determine the appropriateness of a linear model, using r^2 to support your decision.

Yes, this is an appropriate model due to the high r^2 value. Only 14.3% (100 – 85.7) of the variation between the variables is unexplained.

The coefficient of determination is denoted by r^2, which literally means $r \times r$. So the value of the coefficient of determination can be calculated by squaring the value of the correlation coefficient r.

9780170250283

For example,

- if $r = 0.54$, then $r^2 = (0.54)^2 = 0.2916$ OR
- if $r = -0.91$, then $r^2 = (-0.91)^2 = 0.8281$

Notice that when we square a positive or a negative r value we only get a positive r^2, which is why the coefficient of determination cannot be used to determine the direction of the relationship.

Conversely, if r^2 is given, then the r value can be calculated.

For example, if $r^2 = 0.49$, then $r = \pm\sqrt{0.49}$

$= \pm 0.7$

To determine if the r value is positive or negative, further information is required. This can be determined from

- a scatter plot, where you can see if the direction is positive or negative.
- the regression line equation – if the gradient is positive then r is positive and if the gradient is negative then r is negative.

Example 6

a The correlation coefficient for a set of bivariate data is 0.61. Find the coefficient of determination as a percentage.

b The least squares regression line that enables the percentage mark on a test to be determined from the number of hours spent studying is

$$\%\text{mark} = 15 + 9.5 \times \text{study hours}$$

The coefficient of determination for this data is 0.92. Find the value of Pearson's product-moment correlation coefficient.

Solution

a The coefficient of determination is the square of the correlation coefficient.

$r^2 = (0.61)^2$
$= 0.3721$

Multiply by 100 to convert to a percentage.

$r^2 = 0.3721 \times 100\%$
$= 37.21\%$

b Calculate the square root of the coefficient of determination. Remember that this value could be positive or negative.

$r = \pm\sqrt{0.92}$
$= \pm 0.96$

Determine the gradient of the regression line. It is the coefficient of *study hours* in the equation. The sign of the gradient enables us to determine if r is positive or negative.

The gradient of the least squares regression line is 9.5. As the regression line has a positive gradient, the r value is also positive.
$r = 0.96$

EXERCISE 2.02 The coefficient of determination

Interpreting regression lines and the coefficient of determination

Concepts and techniques

1 Example 5 The shoe sizes and heights of some students were measured and the results are shown below.

Shoe size	7	5.5	11	6	7	7.5	7	7	10	8	6.5
Height (cm)	168	158	182	162	162	157	168	171	178	171	168

a Calculate the coefficient of determination, correct to 3 decimal places.
b Interpret the coefficient of determination.
c Find the equation of the least squares regression line that models this data, stating all coefficients correct to 2 decimal places.
d Do you think the model is appropriate? Justify your answer.

2 The ages of both parents for students in a Year 12 class were recorded and the results are shown below.

a Calculate the coefficient of determination correct to 3 decimal places, assuming mother's age is the explanatory variable.
b Interpret the coefficient of determination.
c Find the equation of the least squares regression line that models this data, stating all coefficients correct to 2 decimal places.
d Do you think the model is appropriate? Justify your answer.

Mother's age	Father's age
51	52
53	55
40	41
39	42
44	47
37	38
53	56
57	60
58	58
48	49
52	49
40	41
45	51
47	46
44	50
46	53
45	48
43	44
52	59
55	62
37	39
43	43
44	45

9780170250283

3 The number of laps completed around an oval in 30 minutes and the age of the runner was recorded and is displayed in the table below.

Age	12	15	17	20	25	31	34	42	51	55	70
Number of Laps	12	10	10	8	9	7	5	6	5	4	3

a Calculate the coefficient of determination, correct to 3 decimal places.
b Interpret the coefficient of determination.
c Find the equation of the least squares regression line that models this data, stating all coefficients correct to 2 decimal places.
d Do you think the model is appropriate? Justify your answer.

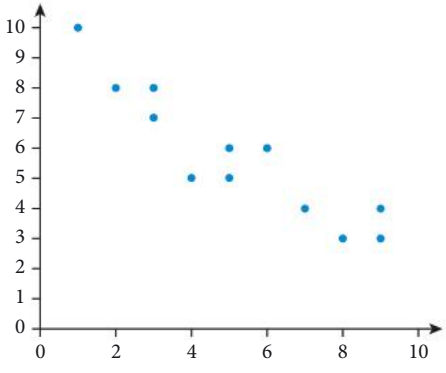
4 Example 6 Calculate the coefficient of determination for each of the following r values, giving your answer correct to 2 decimal places.

a $r = 0.45$ b $r = -0.756$ c $r = -0.3245$
d $r = 0.123$ e $r = -0.9999$

5 Example 6 a The least squares regression line that enables the distance travelled to be determined from the time spent travelling is

$$\text{Distance travelled} = 31 + 1.5 \times \text{time}$$

The coefficient of determination for this data is 0.88. Find the value of Pearson's correlation coefficient.
b There is a negative association between two variables and the value of the coefficient of determination is 0.46. Find the value of Pearson's product-moment correlation coefficient.
c The coefficient of determination for the data displayed in the scatter plot on the right is 0.72. Find the value of the correlation coefficient.



6 The correlation coefficient for a set of data was found to be $r = 0.922$. Written as a percentage, the coefficient of determination is closest to

A 0.85% B 0.92% C 83% D 85% E 92%

7 An investigation was conducted and the correlation coefficient between the amount earned by a group of 20 year old men and how many times a week they went out to a pub was found to be 0.66. Based on this information it can be concluded that

A 44% of the variation in the amount earned can be explained by the variation in the number of times at the pub
B 66% of the variation in the number of times at the pub can be explained by the variation in the amount earned
C 66% of the men go to the pub
D 44% of the variation in the number of times at the pub can be explained by the variation in the amount earned
E 66% of men who go to the pub more frequently earn more money.

Reasoning and communication

8 The number of hours doing homework each week and the number of hours spent gaming each week were recorded for a group of Year 12 students. The results are displayed in the table below. Where appropriate, give all answers correct to 2 decimal places.

Homework hours	10	6	12	14	15	12	11	7	17	15	13	10	8
Gaming hours	21	37	24	23	12	18	24	35	9	10	28	24	31

a Construct a scatter plot to represent the data, assuming that gaming hours is the explanatory variable.
b Calculate the value of Pearson's correlation coefficient and interpret it.
c Calculate the coefficient of determination and interpret it.
d Find the equation of the least squares regression line that models the data.
e Hence explain the gradient and y-intercept in context of the variables being investigated.
f Sketch the least squares regression onto the scatter plot constructed in part **a**.
g Do you think the model is appropriate? Justify your answer.

2.03 MAKING PREDICTIONS

Once the least squares regression line has been found and generally deemed an acceptable model, it can be used to **predict** values. The equation can be used to predict within the original data range as well as outside the data range. These values will differ from the actual data as the line is the linear model that fits best into the data.

When asked to predict for the y variable when given a value for the x variable, we use substitution. For example, the least squares regression line

$$\text{height(cm)} = 100 + 2.5 \times \text{age(years)}$$

can be used to predict the height of people of different ages.

What is the predicted height for a 8-year-old?

$$\begin{aligned}\text{height(cm)} &= 100 + 2.5 \times \text{age(years)} \\ &= 100 + 2.5 \times 8 \\ &= 120 \text{ cm}\end{aligned}$$

When stating the predicted value, make sure that you place the units of measure as well.

Example 7

The following regression line models data relating the maximum daily temperature (°C) to the minimum daily temperature(°C).

Maximum daily temperature = 7.25 + 1.25 × minimum daily temperature

Use the model to predict the maximum daily temperature when the
a minimum daily temperature is 5°C
b minimum daily temperature is 0°C
c minimum daily temperature is 18°C.

9780170250283

Solution

a Substitute 5 into the equation in place of minimum daily temperature.

Maximum daily temp = 7.25 + 1.25 × minimum daily temp
= 7.25 + 1.25 × 5
= 13.5°C

b Substitute 0 into the equation in place of minimum daily temperature.

Maximum daily temp = 7.25 + 1.25 × minimum daily temp
= 7.25 + 1.25 × 0
= 7.25°C

c Substitute 18 into the equation in place of minimum daily temperature.

Maximum daily temp = 7.25 + 1.25 × minimum daily temp
= 7.25 + 1.25 × 18
= 29.75°C

When asked to predict for the x variable when given a value for the y variable, we use substitution but then the equation needs to be transposed to read $x =$ _____ . For example, the least squares regression line

$$\text{height (cm)} = 100 + 2.5 \times \text{age (years)}$$

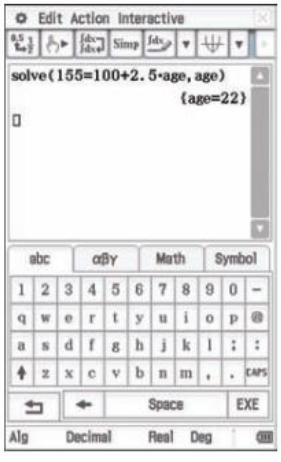
can be used to predict the age of people of different heights.

What is the predicted age for a person who is 155 cm tall?

$$\begin{aligned} \text{height (cm)} &= 100 + 2.5 \times \text{age (years)} \\ 155 &= 100 + 2.5 \times \text{age} \quad \text{(subtract 100 from both sides)} \\ 55 &= 2.5 \times \text{age} \quad \text{(divide both sides by 2.5)} \\ \text{age} &= 22 \text{ years old} \end{aligned}$$

It is often easier to use the solve function on a CAS calculator.





Example 8

The following regression line models data relating the maximum daily temperature (°C) to the minimum daily temperature(°C).

Maximum daily temperature = 7.25 + 1.25 × minimum daily temperature

Use the model to predict the minimum daily temperature when the

a maximum daily temperature is 18°C

b maximum daily temperature is 25°C.

Solution

a Substitute 18 into the equation in place of the maximum daily temperature and solve for the minimum daily temperature.	Maximum daily temp = 7.25 + 1.25 × minimum daily temp 18 = 7.25 + 1.25 × minimum daily temp 10.75 = 1.25 × minimum daily temp
Write the answer.	Minimum daily temp = 8.6°C
b Substitute 25 into the equation in place of the maximum daily temperature and solve for the minimum daily temperature.	Maximum daily temp = 7.25 + 1.25 × minimum daily temp 25 = 7.25 + 1.25 × minimum daily temp 17.75 = 1.25 × minimum daily temp
Write the answer.	Minimum daily temp = 14.2°C

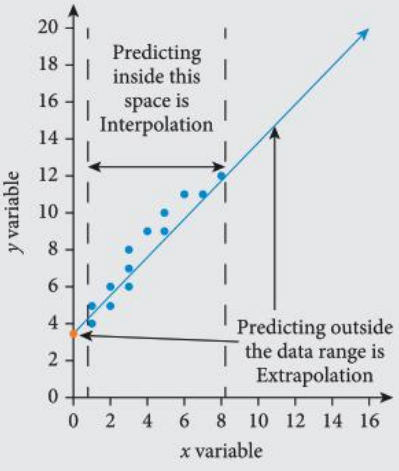
When we use the equation of the least squares regression to make predictions, this is referred to as either interpolating or extrapolating.

IMPORTANT

Interpolation – prediction made within the original range of the data given.

Extrapolation – prediction made outside the original range of the data given.

NOTE: Predictions classified as extrapolation tend to be less reliable than interpolation as the linear regression line models the data given and may not represent values outside the data range as effectively.



9780170250283

Example 9

Data was collected from people aged between 7 and 19 years of age and a linear regression line was found with the equation

$$\text{height (cm)} = 100 + 2.5 \times \text{age (years)}$$

a Predict the height of a 15-year-old. Is this prediction classified as interpolation or extrapolation?

b Predict the height of a 24-year-old. Is this prediction classified as interpolation or extrapolation?

c Which of the predictions in part **a** and **b** is the most reliable? Justify your answer.

Solution

a Substitute 15 into the equation in place of age.

$$\begin{aligned}\text{Height} &= 100 + 2.5 \times \text{age} \\ &= 100 + 2.5 \times 15 \\ &= 137.5 \text{ cm}\end{aligned}$$

Classify prediction as interpolation or extrapolation, according to whether the value used to make the prediction was within or outside the original data range.

As 15 years old is inside the original data range of 7 to 19 years, the prediction is classified as interpolation.

b Substitute 24 into the equation in place of age.

$$\begin{aligned}\text{Height} &= 100 + 2.5 \times \text{age} \\ &= 100 + 2.5 \times 24 \\ &= 160 \text{ cm}\end{aligned}$$

Classify prediction as interpolation or extrapolation according to whether the value used to make the prediction was within or outside the original data range.

As 24 years old is outside the original data range of 7 to 19 years, the prediction is classified as extrapolation.

c Decide on the most reliable prediction.

The prediction for the height of a 15-year-old is more reliable than that of a 24-year-old as it is classified as interpolation.

EXERCISE 2.03 Making predictions

Concepts and techniques

1 Example 7 The regression line with equation

$$\text{Number of laps} = 10.87 - 0.16 \times \text{age}$$

models the relationship between the number of laps run around an oval in 30 minutes and the age of the runner.

Use the model to predict the number of laps run in 30 minutes by a

a 12-year-old runner **b** 34-year-old runner **c** 55-year-old runner

2 The shoe sizes and heights of a group of students is modelled by a regression line with the equation

$$\text{Shoe size} = -10.3 + 0.14 \times \text{height}$$

Use the model to predict the shoe size for a student with a

a height of 97 cm **b** height of 121 cm **c** height of 165 cm

3 Example 8 The regression line with the equation given below models data relating the outside temperature (°C) to the amount of gas consumed (Kwh).

$$\text{Gas consumed} = 2212 - 125 \times \text{outside temperature}$$

Use the model to predict the outside temperature, to the nearest degree, when the

a gas consumed is 2000 Kwh **b** gas consumed is 1200 Kwh **c** gas consumed is 0 Kwh

4 The regression line with the equation given below models data relating the number of levels in a building to the height of the building (m).

$$\text{Height} = 5 + 0.57 \times \text{number of levels}$$

Use the model to predict the number of levels in a building, correct to the nearest whole number, when the

a height is 6.7 m **b** height is 12.1 m

5 Example 9 The weights (in grams) of boxes of chocolates containing between 15 and 50 chocolates were recorded. The regression model for the data was found to have the equation

$$\text{Weight} = 20 + 5 \times \text{number of chocolates}$$

a Predict the weight of a box with 25 chocolates. Is this prediction classified as interpolation or extrapolation?

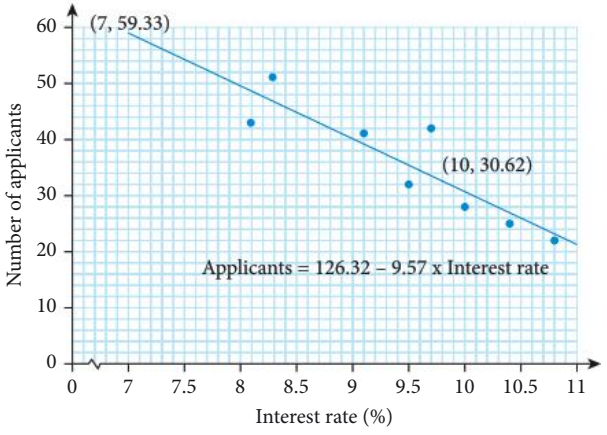
b Predict the weight of a box with 5 chocolates. Is this prediction classified as interpolation or extrapolation?

c Predict the weight of a box with 70 chocolates. Is this prediction classified as interpolation or extrapolation?

d Which of the above predictions is the most reliable? Justify your answer.

6 The scatter plot below represents the relationship between the interest rate and the number of home loan applicants. The least squares regression line was found to have the equation

$$\text{Number of applicants} = 126.32 - 9.57 \times \text{interest rate}$$



9780170250283

a Predict the number of applicants for an interest rate of 8.7%. Is this prediction classified as interpolation or extrapolation?
b Predict the number of applicants for an interest rate of 3.9%. Is this prediction classified as interpolation or extrapolation?
c Predict the interest rate when there are 40 applicants. Is this prediction classified as interpolation or extrapolation?
d Which of the above predictions are the most reliable? Justify your answer.

The following information relates to questions **7**, **8** and **9**.
A regression line modelling the height (cm) of 20 children aged between 5 and 15 years was found to have the equation

$$\text{Height} = 99.97 + 2.59 \times \text{age}$$

7 The predicted height for a 14-year-old is
A 136 cm **B** 136.23 cm **C** 137.5 cm **D** 146 cm **E** 146.34 cm

8 The predicted age of a child who is 120.69 cm is
A 5 **B** 6 **C** 7 **D** 8 **E** 9

9 Which one of the following statements is **false**?
A Predicting the height of a child aged 9 years is an example of interpolation.
B For every one year of increase in age, a child's height increases by 2.59 cm.
C Predicting the height of a 6-year-old will be more reliable than predicting the height of a 12-year-old.
D A child's height will increase as their age increases.
E Predicting the height of a 10-year-old is more likely to be reliable than predicting the height of an 18-year-old.

Reasoning and communication

10 The following data represents the diameter (cm) of a skateboard wheel and the rolling distance (m) of the skateboard.

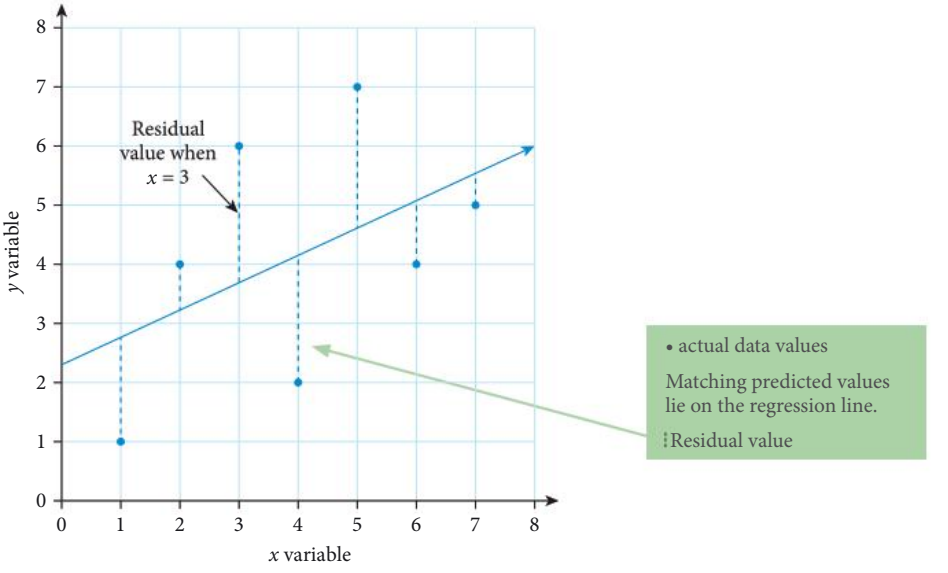
Rolling distance	17	20	22	24	23	25	27	30	32	35	36	40
Diameter	5	5.6	6.1	6.7	7	8.2	8.7	9.3	9.4	9.7	10	10.1

Where appropriate, state all answers correct to 2 decimal places.
a State the explanatory variable.
b Construct a scatter plot to represent the data.
c Calculate and interpret Pearson's correlation coefficient.
d Calculate and interpret the coefficient of determination.
e Find the equation of the least squares regression line.
f Hence interpret the gradient and y-intercept.
g Predict the rolling distance of a skateboard that has a wheel diameter of:
i 8.3 cm ii 10.7cm
h Which of the predictions in part **g** are the most reliable? Justify your answer.
i Predict the wheel diameters for skateboards that roll a distance of:
i 10 metres ii 21 metres
j Which of the predictions in part **i** are the most reliable? Justify your answer.

2.04 RESIDUALS

For any given x value, the **residual** value is the difference between the actual y variable value and the predicted value of the y variable using the equation of the least squares regression line.

Diagram 2



IMPORTANT

Residual = actual y value – predicted y value

The actual y value can either be stated, be found in the data table or read from a scatter plot.

The predicted y value needs to be calculated using the linear regression model for the data.

Some important features from Diagram 2 are:

1 actual data points that lie above the regression line will have a positive residual value

2 actual data points that lie below the regression line will have a negative residual value

3 actual data points that lie on the regression line with have a residual value of zero

4 the further the actual value is from the regression line, the larger the residual value.

9780170250283

Example 10

The regression line connecting height to age was found to have the equation

$$\text{height (cm)} = 100 + 2.5 \times \text{age (years)}$$

Calculate the residual value for

a John who is 12 years old and 142 cm tall

b Jane who is 10 years old and 118 cm tall.

Solution

a Calculate the predicted height for John by substituting 12 for age into the equation.

$$\begin{aligned}\text{height} &= 100 + 2.5 \times \text{age} \\ &= 100 + 2.5 \times 12 \\ &= 130 \text{ cm}\end{aligned}$$

Find the residual value using the formula.

$$\begin{aligned}\text{Residual} &= \text{actual } y \text{ value} - \text{predicted } y \text{ value} \\ &= 142 - 130 \\ &= 12 \text{ cm}\end{aligned}$$

The actual y value was stated in the question.

Remember to place units relating to the y variable in the final answer.

b Calculate the predicted height for Jane by substituting 10 for age into the equation.

$$\begin{aligned}\text{Height} &= 100 + 2.5 \times \text{age} \\ &= 100 + 2.5 \times 10 \\ &= 125 \text{ cm}\end{aligned}$$

Find the residual value using the formula.

$$\begin{aligned}\text{Residual} &= \text{actual } y \text{ value} - \text{predicted } y \text{ value} \\ &= 118 - 125 \\ &= -7 \text{ cm}\end{aligned}$$

When all the data is provided, a residual value for each of the actual data points can be calculated. When there are a large number of data points this would be quite tedious, therefore a CAS calculator can be used for this purpose. The CAS calculator automatically calculates each residual value when a regression line is found. These values can then be displayed in the spreadsheet page already created.

Example 11

Create a table of residual values for the following set of data, stating all answers correct to 1 decimal place.

x	1	2	3	4	5	6	7
y	1	4	6	2	7	4	5

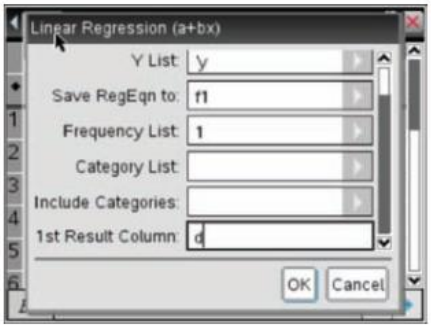
Solution

TI-Nspire CAS

Enter the data into a Lists & Spreadsheet page, remembering to name each column.

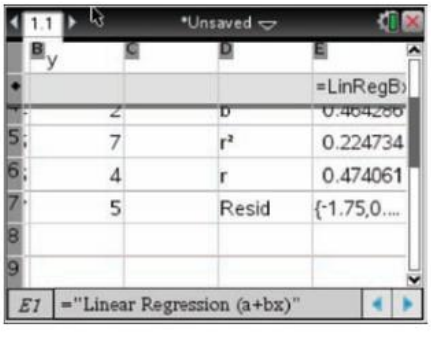


Press menu, then 4:Statistics, 1:Stat Calculations, 4:Linear Regression (a + bx).



Make sure that you select column D for answers so that column C can be used for the residual values.

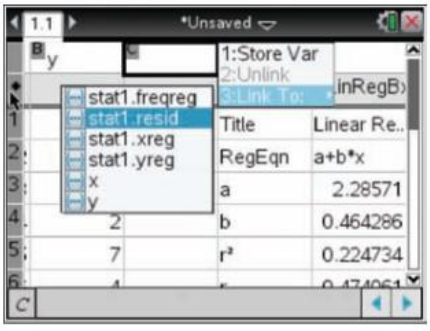
Scroll down the page to find the residuals. Notice that you can only see the first one until you move over the cell and the others appear.

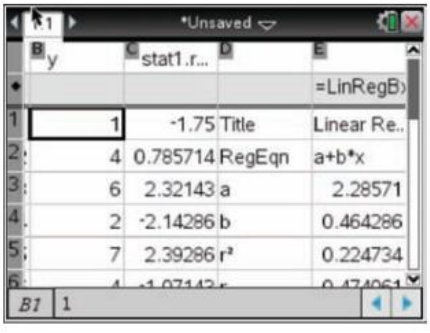


9780170250283

To see residuals more clearly add them to Column C of the spreadsheet.
Press var 3:Link to: stat1.resid

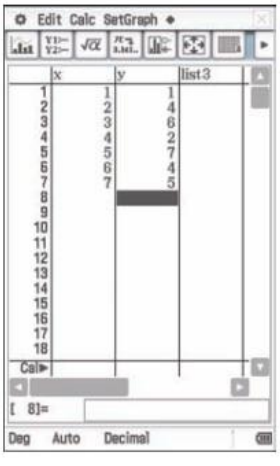
If you find another regression line without clearing your pages, then you would select stat2.resid.





ClassPad

Using the Statistics application, rename the columns and type in the data.

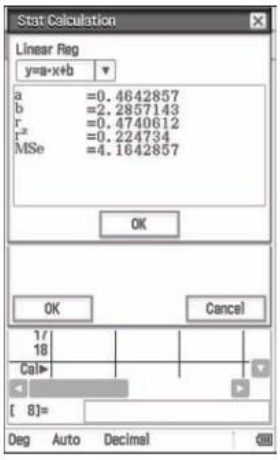


Tap **Calc** then **Regression** and select **Linear Reg**.
Complete the screen as shown.

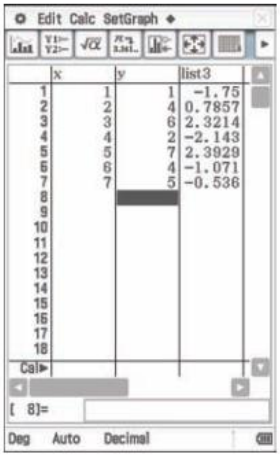
Make sure that you change Copy Residual from off to List 3. This means values will be entered into List 3.



Tap OK and the regression screen will be displayed.



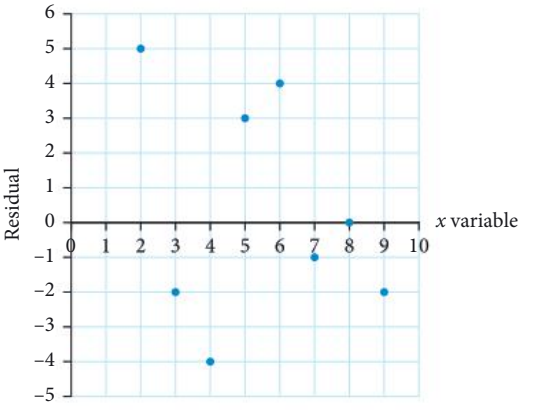
Tap OK and the residual values will appear in List 3.



Complete a table with the residual values added from the calculator screen.

x	1	2	3	4	5	6	7
y	1	4	6	2	7	4	5
residual	−1.8	0.8	2.3	−2.1	2.4	−1.1	−0.5

Once all of the residual values are found, a **residual plot** can be constructed. A residual plot is basically a scatter plot with the explanatory variable on the ***x***-axis and the residual values on the ***y***-axis. As residual values can be both positive and negative, the residual plot will consist of 2 quadrants.

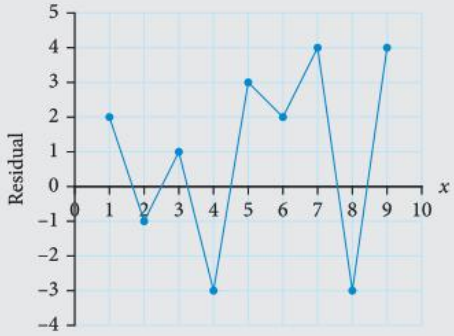


9780170250283

Previously, when determining whether the relationship between the variables being investigated is linear, we have used the form of the scatter plot, the correlation coefficient (r) and the coefficient of determination (r^2). A residual plot can be used as further evidence to verify that the relationship between the 2 variables being investigated is linear.

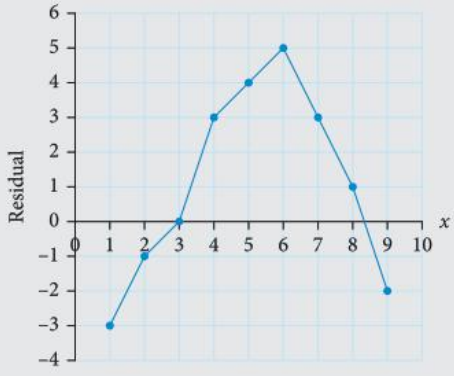
IMPORTANT

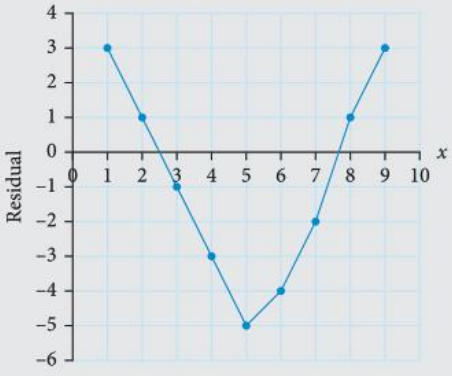
Data relating to 2 variables is considered to be 'probably' linear if the residual plot has values that are randomly scattered above and below the x-axis.



It is not common practice to join the dots on a residual plot; it has only been done here to show the general pattern or lack of pattern in the residuals.

Data relating to 2 variables is considered to be 'probably' non-linear if the residual plot forms a clear curved shape.





It should be noted that if the residual plot shows that the data being investigated is probably non-linear, then the least squares regression model is not appropriate.

Example 12

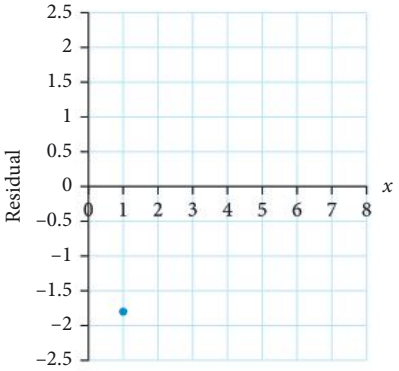
a Construct a residual plot for the data in Example **11**.
b Use the residual plot to decide if the data being investigated is linear or non-linear.

Solution

a Use data from Example **11**. Remember, we only use the x variable and the residual values.

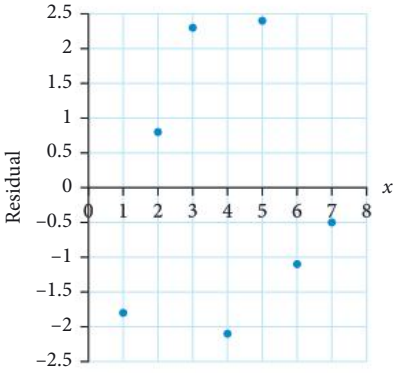
x	1	2	3	4	5	6	7
residual	−1.8	0.8	2.3	−2.1	2.4	−1.1	−0.5

Set up an appropriate scale and plot the first point.

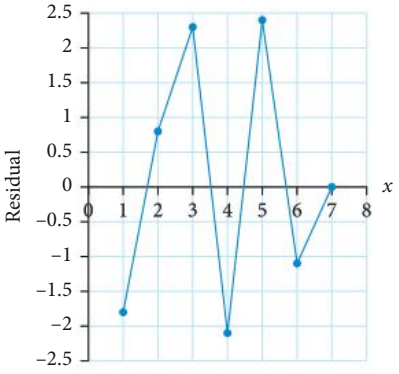


First point is (1 , −1.8)

Repeat for all other data points.

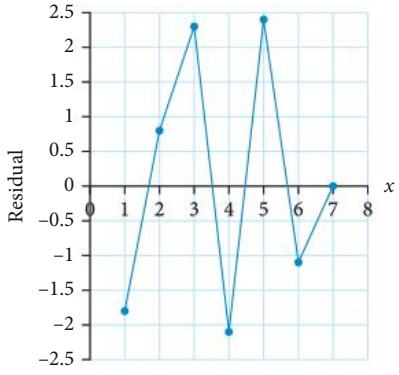


b Join the dots or visualise a residual plot with the dots joined.



9780170250283

Determine whether or not there is a clear pattern in the residual plot. The residual values appear to be randomly distributed around zero.



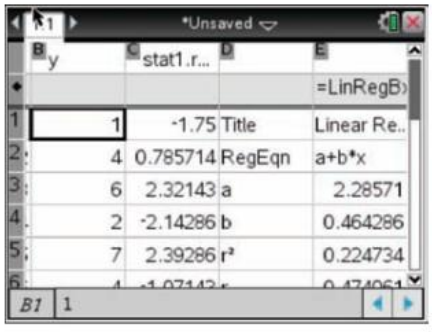
The relationship is probably linear as there is no clear pattern in the residual plot.

Using a CAS calculator to create a residual plot

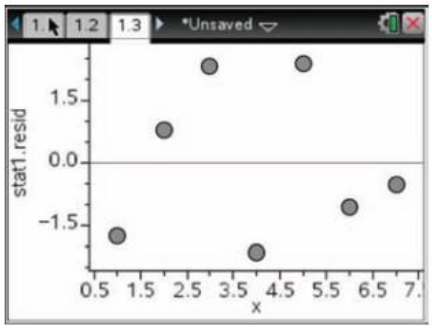
A CAS calculator can be used to construct the residual plot in Example **12**.

TI-Nspire CAS

First calculate the residuals.

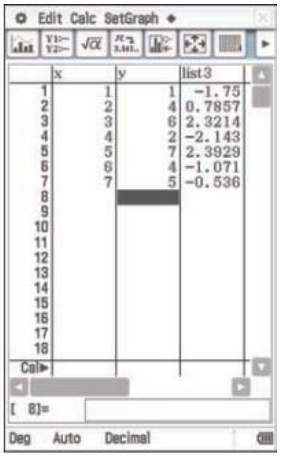


Press ctrl doc and select 5:Data and Statistics.
Move to the bottom of the screen and select x as the x variable.
Move to the left side of the screen and select stat1.resid as the y variable.
The resulting graph is the residual plot.

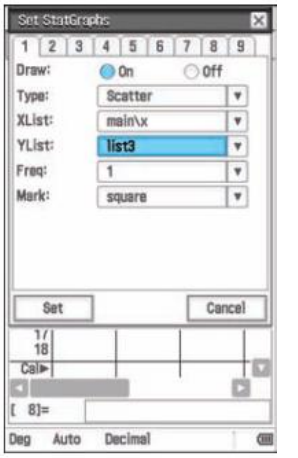


ClassPad

First calculate the residuals.

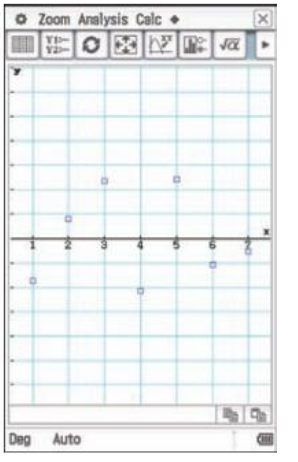


Tap **SetGraph** and ensure that only StatGraph1 is checked. Tap **Setting** and fill in the screen as shown. Tap Set.



Tap  to show the graph then Resize to get a full screen plot.

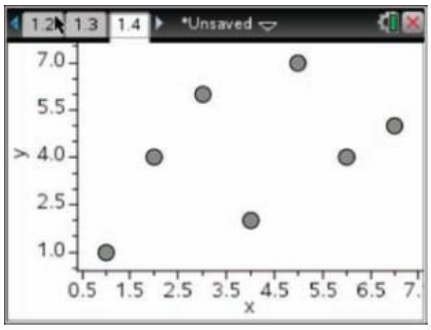
Tap 6 to change the scale if points are not clear



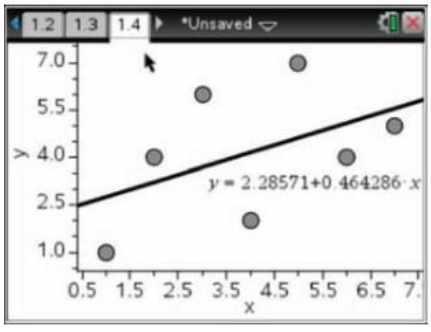
Using the TI-Nspire CAS, you can produce a scatter plot and residual plot on the one page. The steps are as follows.

9780170250283

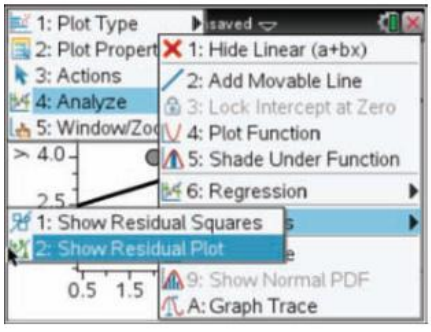
1 Create a scatter plot.



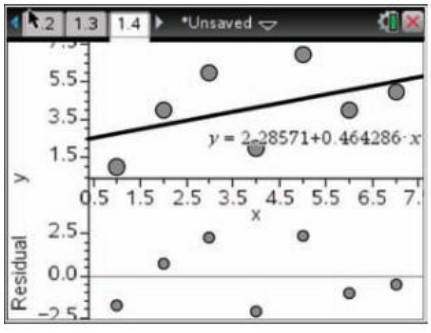
2 Add the regression line.
menu 4:Analyze, 6:Regression, 2: Show Linear(a + bx)



3 Add residual plot.
menu 4:Analyze, 7:Residual, 2:Show Residual Plot



The screen will appear with the scatter plot at the top of the screen and the residual plot directly below it.



Predictions and residuals

EXERCISE 2.04 Residuals

Concepts and techniques

1 Example 10 The regression line connecting the height of a building to the number of levels it has was found to have the equation

$$\text{Height} = 5 + 0.57 \times \text{number of levels}$$

Calculate the residual value for a:

a building with 12 levels that is 12.3 metres tall

b building with 5 levels that is 7.2 metres tall

2 The regression line connecting the amount of time spent on homework compared to gaming per week was found to have the equation

$$\text{Homework hours} = 19.08 - 0.33 \times \text{gaming hours}$$

Calculate the residual value for:

a Dylan who spent 10 hours doing homework and 16 hours gaming

b Hayley who spent 14.5 hours doing homework and 20 hours gaming

c Liam who spent 15.78 hours doing homework and 10 hours gaming

3 Example 11 Create a table of residual values for the following set of data, assuming that height is the independent variable. Give all answers correct to 1 decimal place.

Height (cm)	Femur length (cm)
178	50.2
173	48.4
165	45.1
164	44.6
168	45
165	42.6
155	39.9
155	38

4 Create a table of residual values for the following set of data, stating all answers correct to 1 decimal place.

Max temperature (°C)	16	18	19	20	22	23	25	27	29	31
No. of meat pies sold	54	56	52	49	43	44	39	25	20	15

5 Create a table of residual values for the following set of data, giving all answers correct to 1 decimal place.

Restaurant bill ($)	45	52.80	61.55	70.20	75.50	79	84.50	91
Tip ($)	1.20	1.60	1.80	2	2.40	3	3.50	4.20

6 Example 12 a Construct a residual plot for the data in question **3**.

b Using the residual plot, decide if the data being investigated is linear or non-linear.

7 a Construct a residual plot for the data in question **4**.

b Use the residual plot to decide if the data being investigated is linear or non-linear.

8 a Construct a residual plot for the data in question **5**.

b Use the residual plot to decide if the data being investigated is linear or non-linear.

9780170250283

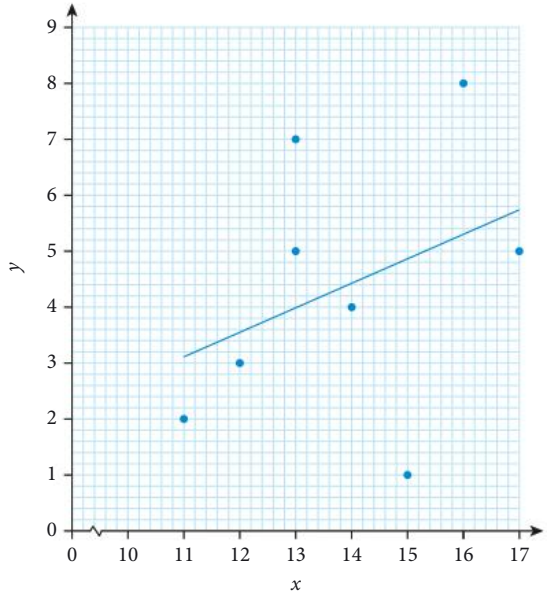
9 The following least squares regression line models the relationship between the velocity (m/sec) of an object and the time (secs) it has travelled.

$$\text{Velocity} = -1.81 + 1.93 \times \text{time}$$

The residual value of an object that travelled at a velocity of 16.5 m/sec at 10 seconds is:

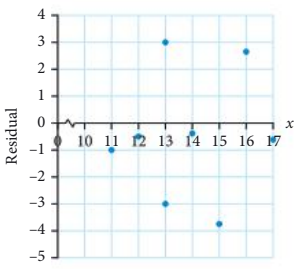
A −0.99 **B** 0.99 **C** 1 **D** 16.5 **E** 17.49

10 The scatter plot and the least squares regression line below shows the relationship between x and y.

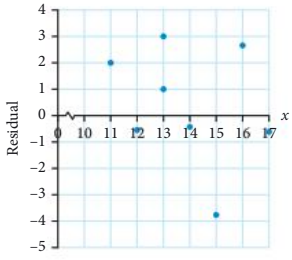


The residual plot for this data would look like:

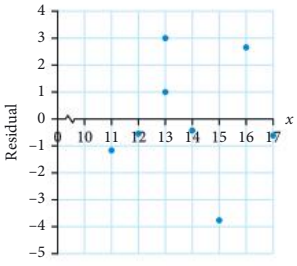
A



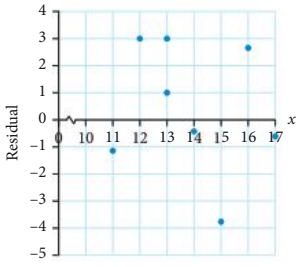
B



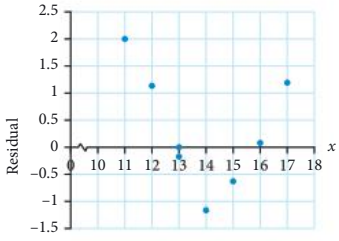
C



D



E



Reasoning and communication

11 Data was collected comparing the minutes of research that students did for an assignment and the final assignment mark out of 40. The data is displayed in the table below.

Research (mins)	27	28	30	42	42	58	60	60	61	89	90	88	87	110	115
Mark (out of 40)	8	21	10	12	18	27	20	23	27	22	30	32	33	33	35

a Explain why the amount of research completed is the explanatory variable.
b Construct a scatter plot to represent the data.
c Find the equation of the least squares regression line that models the data, stating all coefficients correct to 2 decimal places.
d Calculate r and r^2, correct to two decimal places, and use them to determine if the regression model in part **c** is appropriate.
e Use the regression model to predict the mark, correct to one decimal place, of a student who completed 58 minutes of research.
f Hence calculate the residual value for this student.
g Calculate a table of residual values, correct to 1 decimal place, and construct a residual plot.
h Does the residual plot support your answer to part **d**? Justify your answer.

INVESTIGATION Predicting hours of sleep per night for Year 12 students

The following numerical variables are displayed in the table below for 48 Year 12 students.

Sleep: The average hours of sleep on a school night

HW: The hours of homework completed per week

Sport: The hours of sport participated in per week

Pwork: The hours spent in paid work per week

Person	Sleep	HW	Sport	Pwork	Person	Sleep	HW	Sport	Pwork
1	9	3	8	13	14	5	10	20	31
2	8	22	21	4	15	7	15	1	15
3	10	17	3	0	16	10	6	8	15
4	7	1	19	14	17	9	0	7	4
5	8	2	5	7	18	9	10	15	2
6	9	4	9	8	19	16	33	0	0
7	5	4	10	3	20	13	5	20	20
8	8	12	14	16	21	7	25	10	15
9	8	2	0	15	22	5	8	0	15
10	8	7	3	0	23	9	16	16	3
11	9	6	6	9	24	7	8	0	20
12	8	3	2	12	25	9	3	8	13
13	7	9	25	4	25	8	22	21	4

Person	Sleep	HW	Sport	Pwork	Person	Sleep	HW	Sport	Pwork
27	10	17	3	0	38	7	5	2	10
28	7	2	1	10	39	6	6	2	0
29	8	3	6	0	40	4	2	1	3
30	7	0	3	21	41	10	11	40	0
31	8	8	9	8	42	8	7	3	5
32	6	4	1	12	43	7	7	40	40
33	5	10	4	15	44	6	0	3	0
34	9	11	11	15	45	9	1	8	24
35	7	1	20	6	46	9	12	6	10
36	8	10	4	12	47	8	2	5	5
37	8	10	4	10	48	6	5	1	10

The questions will lead you through investigating which variable is the most appropriate to use to predict the average number of hours of sleep on a school night. For the entire investigation, assume that hours of sleep is the response variable. Where appropriate, state your answers correct to 2 decimal places.

Use your CAS calculator to randomly select the 20 students you will use throughout the investigation. Highlight these students in the table.

1 a Sketch a scatter plot for Sleep against HW.

b Discuss the form, direction and strength of the data.

c Calculate the equation of the least squares regression line that models the data.

d Plot the regression line onto the scatter plot.

e Interpret the gradient of your equation.

f State the r and r^2 values and determine if the regression line is appropriate for making predictions.

g Construct a residual plot and discuss if it supports your findings from part **f**.

2 a Calculate the mean and standard deviation for the variable Sleep.

b Calculate the mean and standard deviation for the variable Sport.

c Calculate Pearson's correlation coefficient for the variables Sleep and Sport.

d Use you answers from part **a** to **c** to calculate the least squares regression line for Sleep Vs Sport.

e Using r , r^2 and a residual plot, discuss whether the regression line is an appropriate model for making predictions.

3 a Find a regression line between Sleep and Pwork.

b Use appropriate methods to determine if the regression line is appropriate for making predictions.

4 Use the correct regression line to make the following predictions.

a Predict the number of hours of sleep for a student who completes 14 hours of homework.

b Predict the number of hours of paid work for a student who sleeps for 7 hours.

5 Of the three regression lines calculated, choose the one that would be the best to use for predicting the number of hours of sleep. Justify your answer.



Shutterstock.com/Ikonoklast Fotografie

9780170250283

CHAPTER SUMMARY
LINEAR MODELS FOR NUMERICAL DATA
2

- Once it has been determined that two numerical variables have a linear relationship, a least squares regression line can be used to model the data.
- The general form of a least squares regression line is $y = a + bx$, where the a represents the y-intercept and b represents the gradient. When a least squares regression is found, the final equation should be written in terms of the variable names rather than using y and x.
- A least squares regression line is the linear fit that minimises the squares of the vertical distances between the line and each point.
- The least squares regression line ($y = a + bx$) can also be calculated using:

  $b = r \times \frac{s_y}{s_x}$ (gradient) and

  $a = \bar{y} - b\bar{x}$ (y-intercept)

  where

$\bar{x}$	= the mean of the x variable
$\bar{y}$	= the mean of the y variable
s_x	= the standard deviation of the x variable
s_y	= the standard deviation of the y variable
r	= Pearson's correlation coefficient for the 2 variables

- To interpret the y-intercept of a least squares regression line, we generally state "The ***y* variable** is "a" **units** when the ***x* variable** is zero **units**.
- To interpret the gradient of a least squares regression line, we generally state "The ***y* variable** increases/decreases by b **units** for every 1 **unit** increase in the ***x* variable.**
- The **coefficient of determination** or r^2 is a numerical value between 0 and 1 that is used to determine how useful a regression line is as a linear model for a given set of data. It is calculated using a CAS calculator.
- When interpreting the coefficient of determination, a general sentence can be used, **$r^2 \times 100\%$ of the variation in the response variable** can be explained by the variation in the **explanatory variable**.
- If $r = 0.54$, then $r^2 = (0.54)^2 = 0.2916$
- If $r^2 = 0.49$, then $r = \pm\sqrt{0.49} = \pm 0.7$

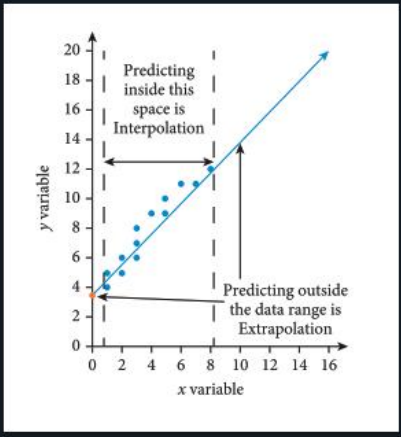
  To determine if the r value is positive or negative, further information is required. It will be given in one of the following forms.

  1. Scatter plot, where you can see if the direction is positive or negative.
  2. The regression line, if the gradient is positive, then r is positive and if the gradient is negative, then r is negative.
- Once the least squares regression line has been found and generally deemed an acceptable model, it can be used to **predict** values.

Height (cm) = 100 + 2.5 × age (years)
= 100 + 2.5 × 8
= 120 cm

Height (cm) = 100 + 2.5 × age (years)
155 = 100 + 2.5 × age
55 = 2.5 × age
age = 22 years old

- Interpolation – predictions made within the original range of the data given, therefore the predictions will be more reliable.
- Extrapolation – prediction made outside the original range of the data given, therefore the predictions may not be reliable.



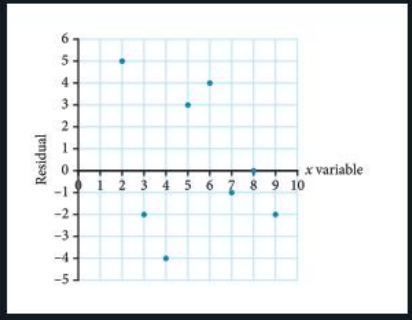
- A residual value is the difference between the actual y variable value and the predicted y variable value for a given x value.

Residual = actual y value – predicted y value

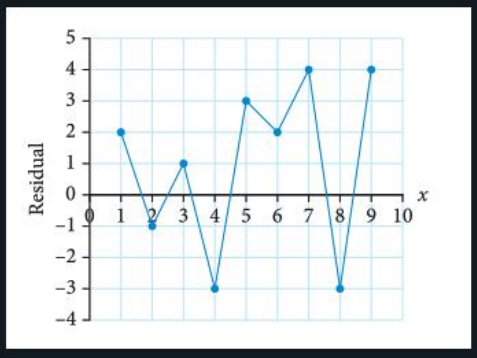
- Residual values for all the data can be quickly calculated using CAS and then displayed in a table. For example:

x	1	2	3	4
y	1	4	6	2
residual	–1.8	0.8	2.3	–2.1

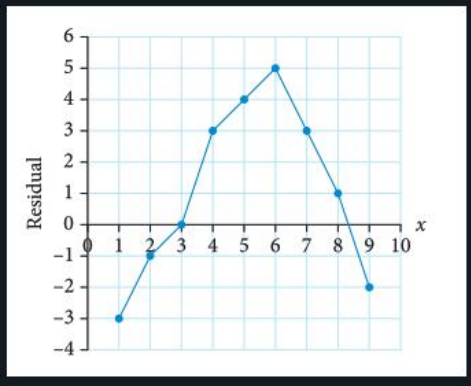
- Once all residual values have been found, they can be displayed in a residual plot. A residual plot is basically a scatter plot with the independent variable on the x-axis and the residual values on the y-axis.

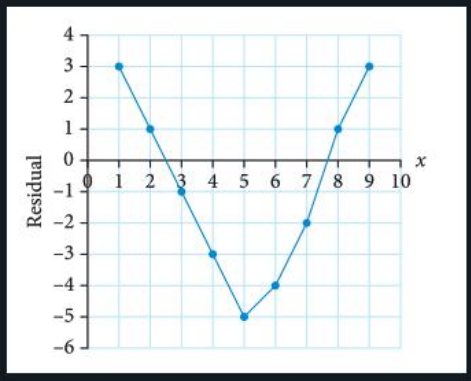


- A residual plot can be used as further evidence to verify that the relationship between the 2 variables being investigated is linear.
- Data relating to 2 variables is considered to be 'probably' linear if the residual plot has values that are randomly scattered above and below the x axis.



Data relating to 2 variables is considered to be 'probably' non-linear if the residual plot forms a clear curved shape.





9780170250283

CHAPTER REVIEW
LINEAR MODELS FOR NUMERICAL DATA
2

Multiple choice

The following information relates to questions **1**, **2** and **3**.

The effect of a dosage of medicine on a child's temperature over time was measured, with the following results.

Time (min)	2	4	6	8	10
Temp (°C)	39.5	39.1	38.9	38.2	37.7

1 The least squares regression line that models this data is
- A Temp = −0.23 + 40.03 × time
- B Temp = 40.03 − 0.23 × time
- C Time = 40.03 − 0.23 × temp
- D Temp = 172.2 − 4.31 × time
- E Time = 4.31 − 172.2 × temp

2 The value of Pearson's correlation coefficient is closest to:
A −0.98 B −0.92 C 0.92 D 0.98 E 1

3 The value of the coefficient of determination closest to:
A −0.99 B −0.65 C 0.95 D 0.97 E 0.99

The following information relates to questions **4** and **5**.

The number of mice in a colony as weeks pass can be modelled by a regression line with the equation

$$\text{Number of mice} = 12 + 10 \times \text{number of weeks}$$

4 Which one of the following statements is **true**?
- A The number of mice increases by 12 for every 1 week that passes.
- B The number of mice decreases by 12 for every 1 week that passes.
- C The number of mice increases by 10 for every 1 week that passes.
- D The number of mice decreases by 10 for every 1 week that passes.
- E The number of weeks increases by 10 for every extra mouse in the colony.

5 Which one of the following statements is **false**?
- A The number of mice in the colony is continually increasing.
- B The number of mice in the colony after 2 weeks is 32.
- C The initial population of mice in the colony was 12.
- D The number of mice in the colony after 10 weeks is 112.
- E It will take 4 weeks for the number of mice in the colony to reach 60.

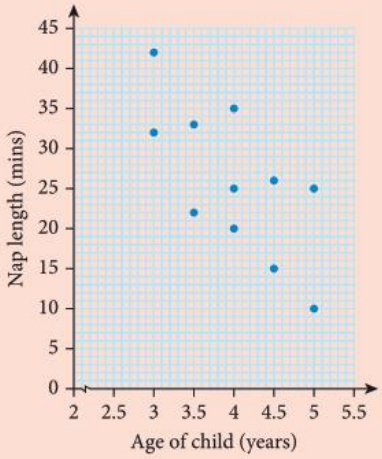
6 For a set of bivariate data involving x and y

$\bar{x} = 11.3$, $s_x = 7.2$, $\bar{y} = 12.5$, $s_y = 9.4$, $r = 0.86$

The gradient of the least squares regression line is closest to:

A 0.67 B 0.78 C 0.95 D 1 E 1.12

7 Given that the coefficient of determination is 0.496 for the scatter plot below, the r value is closest to:



A −0.70 B −0.25 C 0.25 D 0.70 E none of these

8 A regression line modelling the height (cm) of 30 adults aged between 21 and 60 years was found to have the equation

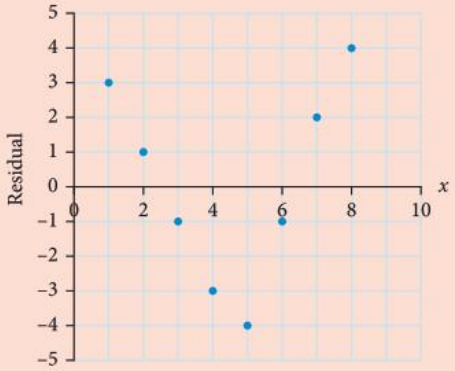
$$\text{Height} = 50.2 + 2.1 \times \text{age}.$$

Which one of the following statements is **false**?

A The predicted height of a 65-year-old is 186.7 cm.
B Predicting the height of an adult aged 25 years is an example of interpolation.
C Predicting the height of a 25-year old will be more reliable than predicting the height of a 12-year-old.
D For every one year of increase in age, an adult's height increases by 50.2 cm.
E Predicting the height of a 50-year-old is more likely to be reliable than predicting the height of a 10-year-old.

9780170250283

9 The following residual plot suggests that:



A The relationship between variables is probably non-linear.
B The relationship between variables is probably linear.
C The association between the two variables is strong.
D The association between the two variables is negative.
E There is no association between the two variables.

10 The least squares regression line with the equation below models the relationship between the speed of a push bike (km/h) and the strength of the head wind (km/h).

$$\text{Speed of bike} = 80 - 1.8 \times \text{wind speed}$$

The residual value of a bike travelling at a speed of 46 km/h into a 22 km/h headwind is:

A −5.6 B −2.8 C 2.8 D 3 E 5.6

Short answer

11 The weight of a mammal is measured as its length increases. The results are given in the table below.

Length (cm)	17	20	28	32	50	66	74	89	102
Weight (kg)	0.3	0.5	1.9	2.7	4.9	5.9	6.5	8.1	9.7

a Construct a scatter plot to represent this data.
b Calculate the correlation coefficient, correct to 2 decimal places, and interpret its value.
c Find the equation of the least squares regression line, writing all coefficients correct to 1 decimal place.

12 The number of cans of soft drink sold (y) and the daily temperature in °C (x) was recorded and the values of the following statistics were determined.

$\bar{x} = 27.2$ $s_x = 5.2$ $\bar{y} = 75.2$ $s_y = 10.1$ $r = 0.67$

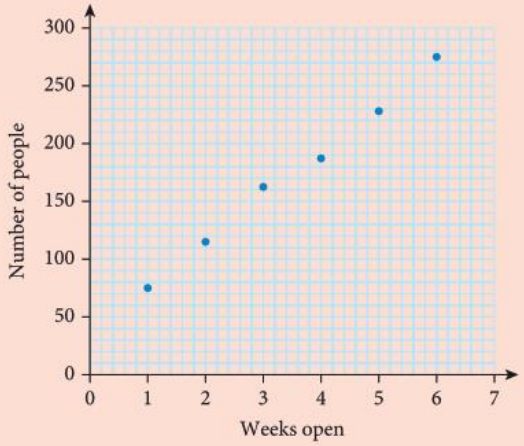
Find the equation of the least squares regression line that models this data, writing all coefficients correct to 2 decimal places.

13 For each of the following regression line equations, interpret the gradient.
a Heart rate (beats per minute) = 55 + 3.8 × amount of exercise (mins)
b Height (cm) = −5.6 + 2.1 × length of left arm (cm)
c Value of car ($) = 31 500 − 1112 × age of car (years)

14 The regression line equation connecting the number of people who attended a new restaurant over a number of weeks was found to be

$$\text{Number of people} = 38 + 38.74 \times \text{weeks}$$

Sketch the following regression line onto the scatter plot below.



15 Calculate and interpret the coefficient of determination for each of the following data sets.

a

x variable	0	2	5	6	8	10	11	17	18	19
y variable	5	11	26	38	41	47	51	67	69	74

b

x variable	1	2	3	4	5	6	7	8	9
y variable	58	52	44	43	36	29	24	21	17

16 Calculate the coefficient of determination for each of the following r values, giving your answers correct to 2 decimal places.

a $r = 0.62$ **b** $r = -0.123$ **c** $r = -0.527$ **d** $r = 0.44$

17 **a** The relationship between the variables *distance from the ground* and *time* is given by the equation

$$\text{Distance from ground} = 60 - 5.2 \times \text{time}$$

If the coefficient of determination is 0.65, find the value of Pearson's correlation coefficient.

b A positive linear relationship exists between two variables. Find the value of Pearson's correlation coefficient if the coefficient of determination is 0.87.

18 To obtain the regression line with the equation

$$\text{Horsepower (units)} = -10.5 + 0.09 \times \text{weight(kg)}$$

data was collected about cars weighing between 1300 kg and 2000 kg.

9780170250283

Give your answers correct to the nearest whole number.

a Predict the horsepower of a car that weighs 2100 kg. Is this prediction classified as interpolation or extrapolation?

b Predict the horsepower of a car that weighs 1550 kg. Is this prediction classified as interpolation or extrapolation?

c Predict the weight of a car that has a horsepower of 152 units. Is this prediction classified as interpolation or extrapolation?

d Which of the predictions in part **a**, **b** and **c** is the least reliable? Justify your answer.

19 The regression line connecting the length of a person's arm span to their age was found to have the equation

$$\text{Arm span length (cm)} = 76 + 2.4 \times \text{age (years)}.$$

Calculate the residual value for:

a Greg who is 12 years old and has a arm span of 115cm

b Trish who is 10 years old and has a arm span of 96.7cm

Application

20 The heights and weights of 12 students are shown in the table below.

Height (cm)	170	175	165	182	177	168	182	172	174	180	176	180
Weight (kg)	60	58	55	74	64	62	69	57	65	75	73	70

Where appropriate, state your answers correct to 2 decimal places.

a Construct a scatter plot for the data.

b Calculate and interpret Pearson's correlation coefficient.

c Find the equation of the least squares regression line that models the data.

d Use the coefficient of determination to determine if the regression line is an appropriate model.

e Use the regression model to predict

i the weight of a student who is 173cm tall

ii the weight of a student who is 151cm tall

iii the height of a student who weighs 91kg.

f Which of the predictions in part **e** are reliable? Justify your answer.

21 Twelve boxes containing 100 apples each were stored in a controlled atmosphere for different periods. After storage, the number of good apples in each box was counted.

Weeks in storage	13	3	4	5	3	7	16	10	9	10	2	12
Number of good apples	48	97	96	86	98	69	54	71	70	59	95	58

Where appropriate, state your answers correct to 1 decimal place.

a Which variable is the response variable?

b Construct a scatter plot to represent the data.

c Find the equation of the least squares regression line that represents the data.

d Construct a table of residuals for the data.

e Hence construct a residual plot.

f Using the residual plot and other statistics relating to the data, determine if the regression model found in part **c** is appropriate for making predictions.

Practice quiz





3

TERMINOLOGY

arithmetic sequence
arithmetic series
book value
common difference
common ratio
depreciation
geometric sequence
geometric series
infinite sum
limiting sum
*n*th term of a sequence
recurring decimal
sequence
series

Prior learning

GROWTH AND DECAY IN SEQUENCES

ARITHMETIC AND GEOMETRIC SEQUENCES

3.01 Arithmetic sequences
3.02 Arithmetic series
3.03 Modelling an arithmetic sequence
3.04 Geometric sequences
3.05 Geometric series
3.06 Modelling a geometric sequence
3.07 Arithmetic and geometric modelling
Chapter summary
Chapter review

THE ARITHMETIC SEQUENCE

- display the terms of an arithmetic sequence in both tabular and graphical form and demonstrate that arithmetic sequences can be used to model linear growth and decay in discrete situations (ACMGM068)
- deduce a rule for the nth term of a particular arithmetic sequence from the pattern of the terms in an arithmetic sequence, and use this rule to make predictions (ACMGM069)
- use arithmetic sequences to model and analyse practical situations involving linear growth or decay; for example, analysing a simple interest loan or investment, calculating a taxi fare based on the flag fall and the charge per kilometre, or calculating the value of an office photocopier at the end of each year using the straight-line method or the unit cost method of depreciation. (ACMGM070)

THE GEOMETRIC SEQUENCE

- display the terms of a geometric sequence in both tabular and graphical form and demonstrate that geometric sequences can be used to model exponential growth and decay in discrete situations (ACMGM072)
- deduce a rule for the *n*th term of a particular geometric sequence from the pattern of the terms in the sequence, and use this rule to make predictions (ACMGM073)
- use geometric sequences to model and analyse (numerically, or graphically only) practical problems involving geometric growth and decay; for example, analysing a compound interest loan or investment, the growth of a bacterial population that doubles in size each hour, the decreasing height of the bounce of a ball at each bounce; or calculating the value of office furniture at the end of each year using the declining (reducing) balance method to depreciate. (ACMGM074) AC

3.01 ARITHMETIC SEQUENCES

An **arithmetic sequence** is a **sequence** of numbers that is formed by adding or subtracting a constant number to each preceding term. This constant is called the **common difference**, denoted by the symbol d. For example, the arithmetic sequence 4, 7, 10, 13 has a common difference of 3. A way to find d, the common difference, is to either recognise it by inspection, or to subtract each term from the one after it.

IMPORTANT

The rule for finding d is $d = t_2 - t_1 = t_3 - t_2 = t_4 - t_3 = \cdots$

This gives the general common difference $d = t_n - t_{n-1}$

To describe an arithmetic sequence, the common difference (d) and the first term (a) must be stated. In the sequence above, $d = 3$ and $a = 4$.

Example 1

For the sequence 1, 8, 15, 22, 29, … find the first term (a) and the common difference (d).

Solution

State the first term in the sequence.

$a = 1$

Use $d = t_2 - t_1 = t_3 - t_2 = t_4 - t_3 = \cdots$ to show that there is a common difference.

$$\begin{aligned} d &= 8 - 1 \\ &= 15 - 8 \\ &= 22 - 15 \\ &= 29 - 22 \\ &= 7 \end{aligned}$$

Write the answer.

$a = 1$ and $d = 7$

9780170250283

Consider the sequence with first term 5 and common difference 3.

The 2nd term is $5 + 3 = 8$.

The 3rd term is $8 + 3 = 11$.

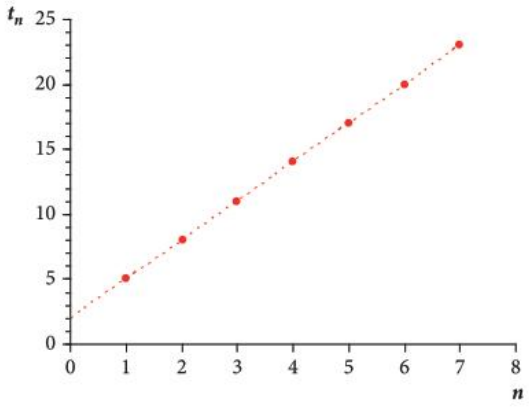
The 4th term is $11 + 3 = 14$.

The 5th term is $14 + 3 = 17$

You can continue and make a table:

n	1	2	3	4	5	6	7	...
t_n	5	8	11	14	17	20	23	...

You can plot a graph of the terms as shown below.



The graph is a straight line, so an arithmetic sequence is a linear function.

The slope of the line is 3. For every one along it goes up by 3. The slope of the line is the common difference. This is true for any arithmetic sequence.

IMPORTANT

The graph of an arithmetic sequence is a straight line with a slope equal to the common difference.

Notice that you can write the terms of the sequence above as follows:

$5 = 5 + 0$

$8 = 5 + 3$

$11 = 5 + 6 = 5 + 2 \times 3$

$14 = 5 + 9 = 5 + 3 \times 3$

$17 = 5 + 12 = 5 + 4 \times 3$

To get the second term, 3 is added to 5 once.

To get the third term, 3 is added to 5 twice.

To get the fourth term, 3 is added to 5 three times.

To get the fifth term, 3 is added to 5 four times.

You can do this for any number of terms and for any arithmetic sequence.

It is possible to work out a rule for the general term of an arithmetic sequence from the pattern of the terms. This is called the nth term of an arithmetic sequence. You can then use this rule to find any term in a sequence.

IMPORTANT

The **general term**, t_n, represents the ***n*th term of a sequence**

For an arithmetic sequence

$t_1 = a$

$t_2 = t_1 + d = a + d$

$t_3 = t_2 + d = (a + d) + d = a + 2d$

$t_4 = t_3 + d = (a + 2d) + d = a + 3d$

$t_5 = t_4 + d = (a + 3d) + d = a + 4d$

Following this pattern, the general term is

$t_n = a + (n - 1)d$

IMPORTANT

The rule for the ***n*th term**, t_n, of an **arithmetic sequence** is $t_n = a + (n - 1)d$

where a = the first term and d = common difference.

Example 2

If the first 6 terms of a sequence are 2, 7, 12, 17, 22, 27, state the value of the 10th term and the 20th term in the sequence.

Solution

Identify a (the first term) and d (the common difference).	$a = 2$ $d = 5$
Use the formula $t_n = a + (n - 1)d$ to find t_{10} and t_{20}.	$t_{10} = 2 + (10 - 1)5 = 47$ and $t_{20} = 2 + (20 - 1)5 = 97$
Write the answer.	$t_{10} = 47$ $t_{20} = 97$

9780170250283

Using a CAS calculator to generate the terms of an arithmetic sequence

A CAS calculator can be used to generate the terms of the sequence in Example **2**.

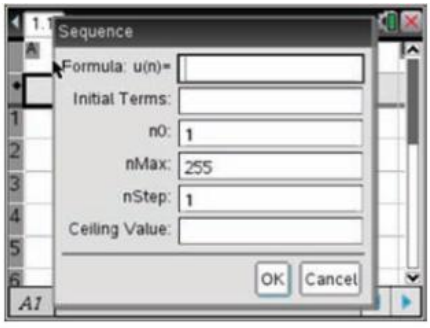
The sequence is defined with $a = 2$ and $d = 5$ so a rule for the nth term is

$t_n = a + (n-1)d$

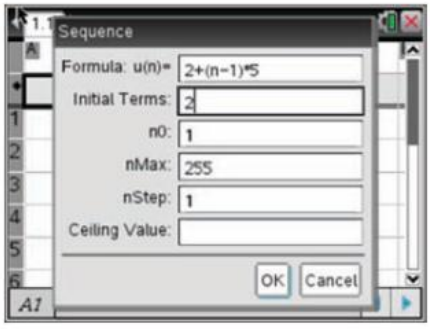
$= 2 + (n-1) \times 5$

TI-Nspire CAS

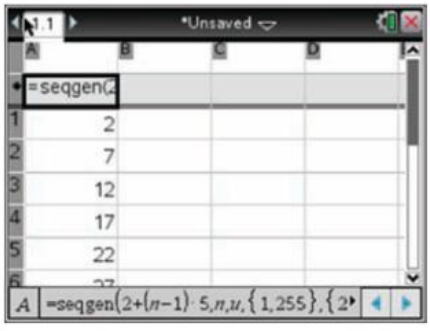
Open a New Document with a Lists & spreadsheet page.
Press menu then 3: Data, then 1: Generate Sequence.



The calculator uses u(n) to represent the nth term, which we call t_n
Enter the rule for t_n in the u(n) field.
Enter the value of a in the Initial Terms field.
Leave the values in the other fields unchanged.

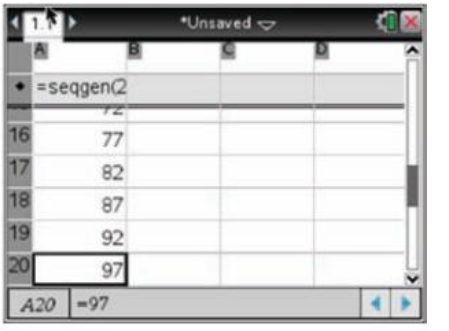


Click on OK.



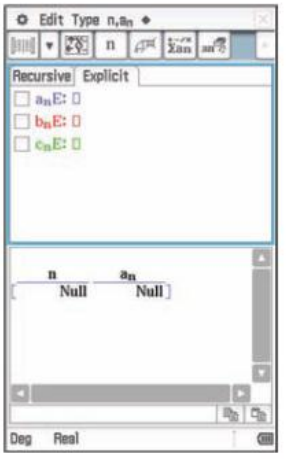
Scroll down by pressing ▼ to find the values of t_{10} and t_{20}.



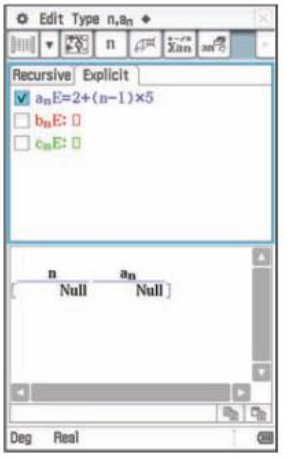


ClassPad

Use the **Sequence** application.
Tap on the **Explicit** tab.



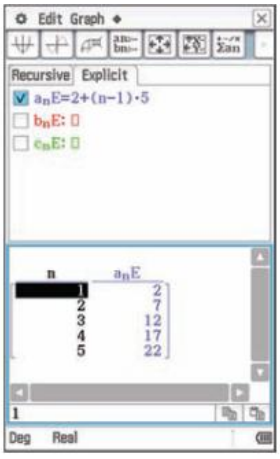
In the entry line next to a_nE: type the rule for the sequence: $2 + (n - 1) \times 5$
Tap [n] on the top toolbar to enter n in the rule.
Make sure that a_nE is checked.



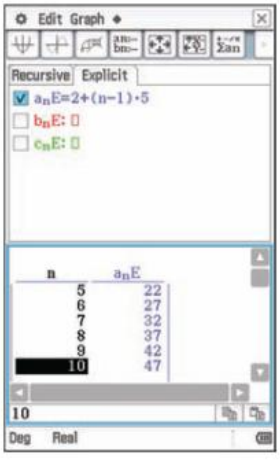
9780170250283

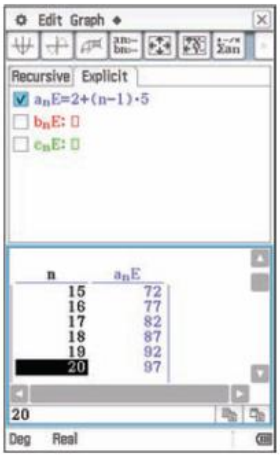
Tap ▦ to generate the terms of the sequence.

If the table does not start at n = 1 or does not contain enough terms, tap ▦ and change the start value to 1 and the end value to say 255. The first 255 terms of the sequence will be included in the table.



Scroll down by pressing ▼ to find the values of t_{10} and t_{20}.





The formula $t_n = a + (n-1)d$ can be used to find the term at which a sequence will reach a certain number.

Example 3

For the arithmetic sequence 54, 51, 48, …, find which term is equal to zero.

Solution

Identify a and d.	$a = 54$ $d = -3$
Describe the sequence using the formula $t_n = a + (n-1)d$	$t_n = a + (n-1)d$ $= 54 + [-3(n-1)]$
We are looking for which term is equal to 0. We don't know what n is equal to, but we know this unknown term (t_n) is equal to 0. Substitute this into the formula.	$0 = 54 + [-3(n-1)]$
Expand the brackets, rearrange and solve.	$0 = 54 - 3n + 3$ $0 = 57 - 3n$ $57 = 3n$ $n = 19$
Write the answer.	The 19th term is equal to zero.

The formula $t_n = a + (n-1)d$ can also be used to find a sequence if you are given 2 terms in an arithmetic sequence.

Example 4

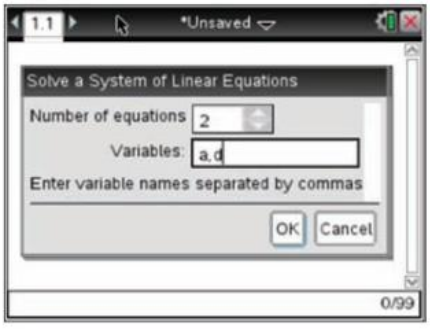
Find the arithmetic sequence in which the 3rd term is 2 and the 10th term is −12.

Solution

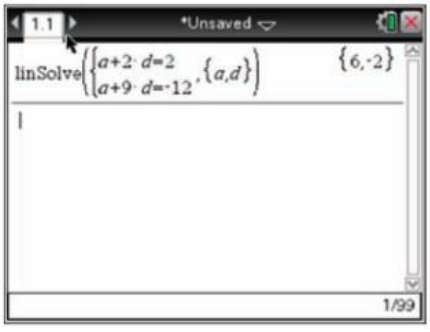
Use the formula $t_n = a + (n-1)d$ to set up an equation for t_3.	$t_3 = a + (3-1)d = 2$ $a + 2d = 2$
Use the formula $t_n = a + (n-1)d$ to set up an equation for t_{10}.	$t_{10} = a + (10-1)d = -12$ $a + 9d = -12$
Solve the simultaneous equations to find the values of a and d.	$a + 2d = 2 \rightarrow [1]$ $a + 9d = -12 \rightarrow [2]$
Use a CAS calculator to solve the simultaneous equations.	

TI-Nspire CAS

Open a New Document with a Calculator page.
Press menu then 3: Algebra, then 7: Solve System of Equations followed by 1: Solve System of Linear equations.
Complete the pop up screen as shown.

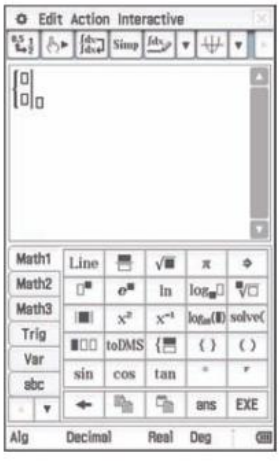


Type the two equations into the template and press enter.

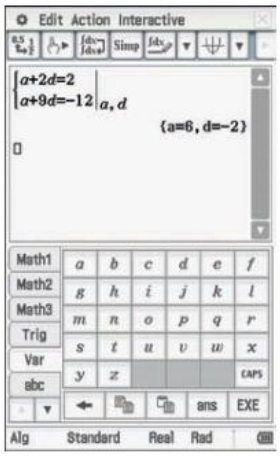


ClassPad

Use the Main $\sqrt{a}$ application.
Press Keyboard, then tap Math1 followed by {□}.



Type the two equations into the template and specify that the variables are a and d, as shown on the screen.
Press EXE.



Write the answers for a and d.

$a = 6$
$d = -2$

Write the answer.

The sequence is 6, 4, 2, 0, −2, −4, …

Arithmetic sequences

EXERCISE 3.01 Arithmetic sequences

Concepts and techniques

1 Generate the first 5 terms of the arithmetic sequence if

a a = 1, $d = 6$
b a = 10, $d = -6$
c a = 9, $d = 3$
d a = −2, $d = -4$

2 Example 1 For each of these sequences, find the first term (a) and the common difference (d).

a 11, 19, 27, 35, 43, …
b −10, −4, 2, 8, 14, …
c 31, 29, 27, 25, 23, …
d −9, −13, −17, …

3 Find the 10th term of the arithmetic sequence if

a a = −2, $d = 5$
b a = 1, $d = 6$
c a = 6, $d = 3.5$
d a = 2, $d = -5$

4 If the first 4 terms of a sequence are 10, 13, 16, 19, …, state

a the value of a
b the value of d
c the 10th term of the sequence
d the 20th term of the sequence

5 If the first 5 terms of a sequence are 1, −3, −7, −11, −15, …, state

a the value of a
b the value of d
c the 11th term of the sequence
d the 100th term of the sequence

6 Find a and d and an expression for t_n for the arithmetic sequence 3, 7, 11, 15, … Hence find the 25th term of the sequence.

7 **a** Find the common difference d for the sequence 1, 8, 15, 22, 29, …
b Hence find the 10th term of the sequence 1, 8, 15, 22, 29, …

9780170250283

8 Example 2 For each of these sequences, state the value of t_{10} and t_{20}.

a 50, 42, 36, 28, …
b −16, −25, −34, −43, …
c 4, 15, 26, 37, …
d −20, −13, −6, 1, 8, …

9 The common difference, d, of the arithmetic sequence 10, 6, 2, −2, … is

A −4 **B** −2 **C** 4 **D** 6 **E** 10

10 The 7th term of the arithmetic sequence 0, −2, −4, −6, … is

A −12 **B** −10 **C** −6 **D** −5 **E** 0

11 Show that the common difference for the sequence 21, 32, 43, 54, … is 11.

12 An arithmetic sequence has a first term of 18 and a common difference of −2.

a Find the first 6 terms.
b Draw a graph of the sequence.
c What is the shape of the graph?

13 An arithmetic sequence has a first term of 6 and a common difference of 4.

a Find the first 6 terms.
b Draw a graph of the sequence.
c What is the shape of the graph?

14 40, 37, 34, 31, . . . is an arithmetic sequence.

a Find the first 6 terms.
b Draw a graph of the sequence.
c What is the shape of the graph?

Reasoning and communication

15 Example 3 For the arithmetic sequence 28, 32, 36, 40, …, find which term is equal to 100.

16 **a** For the arithmetic sequence −189, −180, −171, −162, …, find which term is equal to 0.
b For the arithmetic sequence 43, 30, 17, 4, …, find which term is equal to −100.
c For the arithmetic sequence 218, 235, 252, 269, …, find which term is equal to 1000.
d For the arithmetic sequence −54, −35, −16, 3, …, find which term is equal to 250.

17 Example 4 Find the arithmetic sequence in which the 4th term is −3 and the 11th term is −40.

18 If an arithmetic sequence has $a = 6$ and $d = 2$, which term has the value of 34?

19 If an arithmetic sequence has $a = 2$ and $d = \frac{1}{4}$ which term has the value of 14?

20 Find the twentieth term of the arithmetic sequence in which $t_4 = 20$ and $t_{11} = 90$.

21 Find a value of w such that $6 + 2w$, $7 + 5w$, $5 + 12w$ will form successive terms of an arithmetic sequence.

22 If x, $2x + 3$, and $5x$ are the first three terms of an arithmetic sequence, calculate the value of x.

3.02 ARITHMETIC SERIES

A **series** is the sum of terms of a sequence. So for the arithmetic sequence 2, 5, 8, 11, 14, 17, … the equivalent **arithmetic series** would be $2 + 5 + 8 + 11 + 14 + 17 + \cdots$

We use the notation S_n to mean the sum of the sequence up to and including t_n.

INVESTIGATION Comparing arithmetic sequences

Five participants in a half marathon are to receive a share of the total prize pool of $100 000. Three proposals are produced:

Proposal 1: The prizes are to form an arithmetic sequence with the 5th prize being $5000.

Proposal 2: The prizes are to form an arithmetic sequence with the 1st prize being $80 000.

Proposal 3: The prizes are to form an arithmetic sequence with the 3rd and 4th prizes being $20 000 and $15 000 respectively.

Investigate which proposal is better for the:

a winner of the marathon.

b 5th place getter in the marathon.

Partial sum of an arithmetic sequence

The sum of the first n terms of an arithmetic series (nth partial sum) is given by the formula:

$$S_n = \frac{n}{2}(a+l)\text{, where } l = \text{last or } n\text{th term.}$$

We show this by the following.

Let the last or nth term be l.

$S_n = a + (a+d) + (a+2d) + \cdots + \ell$

also

$S_n = \ell + (\ell - d) + (\ell - 2d) + \cdots + a$

Solving simultaneously gives

$S_n = \frac{n}{2}(a+\ell)$

We can use this formula when we know the last term of the sequence under consideration.

Sum of an arithmetic sequence

$S_n = \frac{n}{2}[2a + (n-1)d]$

We can show this by the following.

Since $l = n$th term,

$l = a + (n-1)d$

and $S_n = \frac{n}{2}(a+\ell)$

9780170250283

$S_n = \frac{n}{2}[a + a + (n-1)d]$

This gives

$S_n = \frac{n}{2}[2a + (n-1)d]$

We use this formula when the nth term is unknown.

IMPORTANT

The **series, S_n, of an arithmetic sequence** is $S_n = \frac{n}{2}[2a + (n-1)d]$

where a = the first term, d = common difference,
n = number of terms.

Use $S_n = \frac{n}{2}[2a + (n-1)d]$ if we know a, d and n.

The **partial sum** of an arithmetic sequence is $S_n = \frac{n}{2}(a + \ell)$,
where l = last or nth term.

Use $S_n = \frac{n}{2}(a + \ell)$ if we know a and the last term.

Example 5

Consider the sequence 2, 7, 12, 17, 22, 27, … What is the sum of the first 10 terms?

Solution

Identify a and d.	$a = 2$ $d = 5$
Use the formula $S_n = \frac{n}{2}[2a + (n-1)d]$ where the sum to the 10th term is S_{10}.	$S_{10} = \frac{10}{2}[2 \times 2 + (10-1)5]$ $= 5[4 + 9 \times 5]$
Write the answer.	$S_{10} = 245$

Using a CAS calculator to find the sum of the first *n* terms of an arithmetic sequence

A CAS calculator can be used to find the sum of the first n terms of the sequence in Example **5**.

The sequence is defined with $a = 2$ and $d = 5$, so a rule for the sum for the first n terms is

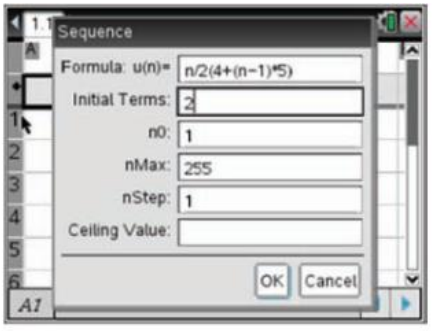
$S_n = \frac{n}{2}[2a + (n-1)d]$

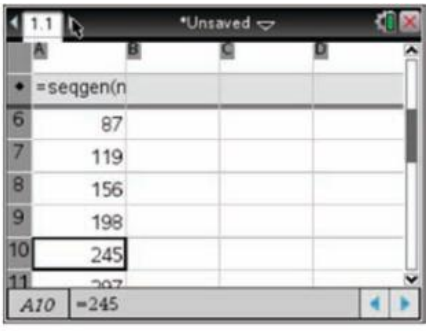
$= \frac{n}{2}[2 \times 2 + (n-1) \times 5]$

$= \frac{n}{2}[4 + (n-1) \times 5]$

$= \frac{n}{2}[4 + 5(n-1)]$

TI-Nspire CAS



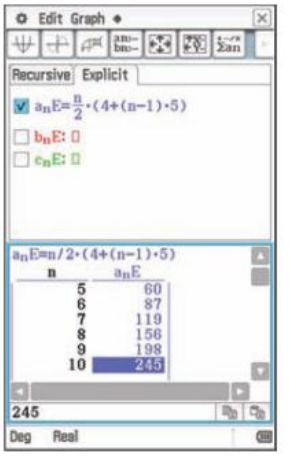


ClassPad

Use the **Sequence** application.

Tap on the **Explicit** tab. Enter the formula.

Tap to generate the terms of the sequence.



Using $S_n = \frac{n}{2}[4+5(n-1)]$ we get the 10th term equalling 245.



Shutterstock.com/PromesaArtStudio

9780170250283

You can also find the sum of an arithmetic sequence if you have the first and last terms of the sequence and can calculate the common difference.

Example 6

Find the sum of the arithmetic sequence 33, 40, 47, 54, …, 110.

Solution

State the values for a, d and l.

$a = 33$
$d = 7$
$l = 110$

We need to find n. Use the equation $t_n = a + (n-1)d$.

$t_n = l = 33 + (n-1)7 = 110$
$7(n-1) = 77$
$n - 1 = 11$
$n = 12$

There are now two methods we can use.

Using $S_n = \frac{n}{2}[2a + (n-1)d]$.

$S_{12} = \frac{12}{2}[2 \times 33 + (12-1)7]$
$= 6[66 + 77]$
$= 858$

Or we can use $S_n = \frac{n}{2}(a + l)$.

$S_{12} = \frac{12}{2}(33 + 110)$
$= 6(143)$
$= 858$

Write the answer.

$S_{12} = 858$

The formula can also be used to find general expressions for sequences where n is unknown.

Example 7

Find an expression for the sum of the first n counting numbers.

Solution

Write the first n counting numbers. This is an arithmetic sequence with $a = 1$ and $d = 1$.

$1, 2, 3, 4, \ldots, n$

Apply the rule for S_n.

$S_n = \frac{n}{2}(t_1 + t_n)$ or $S_n = \frac{n}{2}(a + l)$

Substitute for t_1 and t_n.

$S_n = \frac{n}{2}(1 + n)$

Write the answer.

$S_n = \frac{n(n+1)}{2}$

EXERCISE 3.02 Arithmetic Series

Concepts and techniques

1 Find the sum of the first 10 terms of an arithmetic sequence if

a $a = 1, d = 6$ **b** $a = 10, d = -6$

c $a = 9, d = 3$ **d** $a = -2, d = -4$

2 Find the sum of the arithmetic sequence to the stated number of terms.

a $a = -2, d = 5, S_7$ **b** $a = 1, d = 6, S_{12}$

c $a = 6, d = 3.5, S_6$ **d** $a = 2, d = -5, S_{20}$

3 Example 5 Find the sum of these arithmetic sequences to the stated number of terms.

a 6, 14, 22, 30, 38, …, S_{20} **b** −11, −14, −17, −20, −23, …, S_{25}

c 3, 9, 15, 21, 27, …, S_{100} **d** 19, 12, 5, −2, …, S_{10}

4 If the first 4 terms of a sequence are −2, −10, −18, −26, …, state what the sum of the sequence is to 9 terms.

5 Example 6 Find the sum of the following arithmetic sequences.

a 5, 17, 29, 41, …, 197 **b** −13, −17, −21, −25, …, −117

c 142, 133, 124, 115, …, 52 **d** −51, −45, −39, −33, …, 93

6 Evaluate $9 + 14 + 19 + \cdots + 224$

7 Example 7 The expression for the sum of the first n even numbers is

A n **B** $\frac{n}{2}[2a + (n-1)d]$ **C** $a + (n-1)d$ **D** $n(n+1)$ **E** $2n$

8 The expression for the sum of the first n odd numbers is

A n^2 **B** n **C** $2n$ **D** $n(n+1)$ **E** $\frac{n}{2}(a+l)$

9 Find an expression for the sum of the first n multiples of 5.

Reasoning and communication

10 For what value of n is the sum of n terms of $2 + 11 + 20 + \cdots$ equal to 618?

11 The 6th term of an arithmetic series is 23 and the sum of the first 10 terms is 210. Find the sum of 20 terms.

12 How many terms of the series $45 + 47 + 49 + \cdots$ give a sum of 1365?

13 For what value of n is the sum of the arithmetic series $5 + 9 + 13 + \cdots$ equal to 152?

14 How many terms of the series $80 + 73 + 66 + \cdots$ give a sum of 495?

15 The sum of the first 5 terms of an arithmetic series is 110 and the sum of the first 10 terms is 320. Find the first term and the common difference.

9780170250283

3.03 MODELLING AN ARITHMETIC SEQUENCE

Arithmetic sequences can be used to model and analyse practical situations involving linear growth or decay. Examples of this are analysing loans or investments, calculating taxi fares or calculating the value of an office photocopier at the end of each year. Graphs displaying arithmetic modelling will always use a straight line to indicate the progress of the value. With large numbers in these real-life situations, the use of technology, e.g. handheld calculators, computer spreadsheets, etc., helps enormously with these calculations.



Example 8

Sue borrows $5000 from an unauthorised broker and is unable to pay it back. After missing her first payment, she is charged another $3000 for every subsequent month she is unable to pay anything towards her debt. How much does Sue owe after 6 months if she continues to be unable to pay it back?

Solution

Use the sequence 5000, 8000, 11 000, … and $t_n = a + (n-1)d$ to find t_6.

$t_6 = 5000 + (6-1)3000$
$= 5000 + 15000$
$= 20000$

Write the answer.

Sue will owe $20 000.

Example 9

Alex plays an online game that rewards him every day he logs in. The first day he logs in he gets two tokens, the second day he gets four tokens, the third day he gets six tokens and so on. If he logs in to play the game every day in April, how many tokens will he be rewarded in total?

Solution

Identify a, d and n.

$a = 2$
$d = 2$
$n = 30$ (there are 30 days in April)

Use the formula $S_n = \frac{n}{2}[2a + (n-1)d]$ where the sum to the 30th term is S_{30}.

$S_{30} = \frac{30}{2}[2 \times 2 + (30-1)2]$
$= 15[4 + 29 \times 2]$

Write the answer.

$S_{30} = 930$

Most motor vehicles, computers, tools and equipment depreciate with age. Often when you go to a tax accountant and are claiming purchases as a tax deduction, the accountant will set up a depreciation schedule for you for future years. This is called the **book value** of the object.

There are three types of **depreciation**.

Straight-line depreciation: where the book value depreciates by a fixed dollar value each year.

Reducing balance depreciation: where the book value depreciates by a given percentage each year.

Unit cost method of depreciation: is dependent on the amount of use the asset has had.

An arithmetic sequence can be used to model the straight-line method or the unit cost method of depreciation.

Example 10

A $1400 chainsaw depreciates by $105 per year.

a What will be the chainsaw's book value after 4 years?

b When will the chainsaw be worth nothing?

Solution

a Identify a, d and n.

$a = 1400$, $d = -105$

$n = 5$ because it is 4 years after the 1st term (a).

Use the formula $t_n = a + (n-1)d$ to find t_5, where $a = 1400$, $d = -105$, $n = 5$.

$t_5 = 1400 + (5-1)(-105)$

$= 1400 - 420$

$= 980$

Write the answer.

The chainsaw's book value after 4 years will be $980.

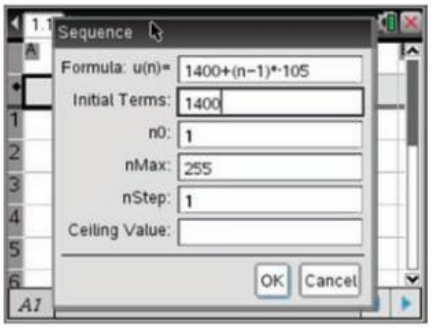
b Generate the terms of the sequence using a CAS calculator.

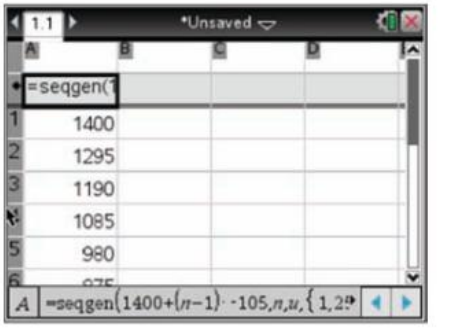
First find a rule for the nth term of the sequence.

$t_n = a + (n-1)d$

$= 1400 + (n-1) \times (-105)$

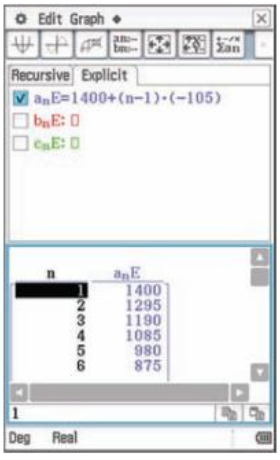
TI-Nspire CAS





9780170250283

ClassPad

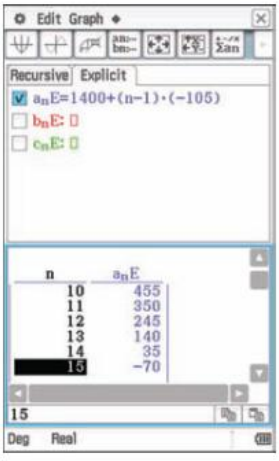


Look for the first term in the sequence that is less than zero.

TI-NSpire CAS



ClassPad



$t_{15} = -70$

Interpret this answer.

The chainsaw still has value at the end of the 13th year but this becomes zero in the 14th year.

Write the answer.

The chainsaw will be worth nothing after 14 years.

Example 11

A stack of cans on a display at a supermarket has 5 cans on the top row.
The next row down has 2 more cans and the next one has 2 more cans and so on.

a Calculate the number of cans in the 11th row down.

b If there are 320 cans in the display altogether, how many rows are there?

Solution

Identify this is an arithmetic sequence.

The first row (top row) has 5 cans, the 2nd row has 7 cans, the 3rd row has 9 cans and so on. This forms an arithmetic sequence with $a = 5$ and $d = 2$.

a Use the formula $t_n = a + (n-1)d$ to find the 11th term.

$t_{11} = 5 + (11 - 1)2 = 25$

Write the answer.

There are 25 cans in the 11th row.

b Use the formula $S_n = \frac{n}{2}[2a + (n-1)d]$ with the total of cans: $S_n = 320$.

$$S_n = \frac{n}{2}[2a + (n-1)d]$$

$$320 = \frac{n}{2}[2 \times 5 + (n-1) \times 2]$$

$$= \frac{n}{2}[10 + 2n - 2]$$

$$= \frac{n}{2}[2n + 8]$$

$$= n^2 + 4n$$

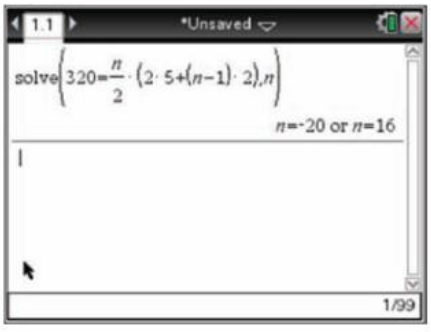
$$0 = n^2 + 4n - 320$$

$$= (n - 16)(n + 20)$$

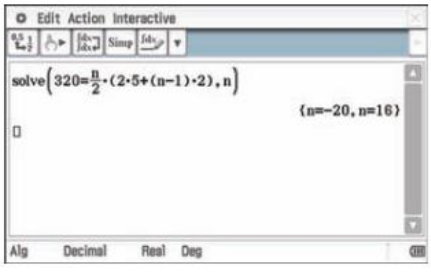
This gives the solutions $n = 16, n = -20$

The equation can be solved easily using a CAS calculator.

TI-NSpire CAS



ClassPad



Select the positive solution since n must be a positive integer.

$n = 16$

Write the answer.

There are 16 rows of cans.

9780170250283

EXERCISE 3.03 Modelling an arithmetic sequence

Reasoning and communication

1 Example 8 Mario puts $100 into a jar at the end of the first week of January. If he is able to put in another $20 every week, how much money will he have saved by the end of the year?

2 Elli goes for a hike and takes four litres of water with her. In her first kilometre she drinks 300 mL of water, but then manages to drink 200 mL for every subsequent kilometre that she hikes. If she hiked 13 km before reaching back home, how much water does she have left?

3 Example 9 Finn has roughly 13 000 songs on his mp3 player. In his first week at a new school, he added 1000 songs, and then adds an average of 100 songs every week thereafter. How many songs will Finn have on his mp3 player:
   a at the end of term (10 weeks)?
   b at the end of the school year (40 weeks)?

4 Amanda earned $20 000 in one year. At the beginning of the next year, she received a salary increase of $450. She now receives the same increase each year.
   a What will her salary be in the 10th year?
   b How much will Amanda earn altogether over the 10 years?

5 Teagan swims laps at a local pool. In her first week, she swims 10 laps. If she is able to increase the number of laps she swims by four every week, how many laps will she have swum in total if she swims every week for ten weeks?

6 Ahmed is selling chocolates for a fundraiser. In the first hour he sells two chocolates, in the second hour he sells five chocolates, in the third hour he sells eight chocolates, and so on.
   a How many chocolates will he sell in his eighth hour of selling?
   b How many chocolates will he have sold in total at the end of the eighth hour?
   c If it takes him twenty hours to sell all the chocolates, how many chocolates did he have to sell?

7 Example 10 A $3000 lounge suite depreciates by $225 per year.
   a What will be the lounge suite's book value after 15 years?
   b Graph the value of the lounge suite over 15 years.
   c When will the lounge suite be worth nothing?

8 Ryan is attempting to collect a set of 300 football cards. In the first month of collecting he collects 50 cards, and in each following month he collects 20 new cards.
   a How many cards has he collected in 6 months?
   b How long will it take him to collect all 300 cards?

9 Dane has won 1000 chocolate bars in a competition, and eats 50 the first few days that they arrive. After feeling sick he decides to give the rest away. On Monday he gives away 5 chocolate bars, on Tuesday he gives away 7 chocolate bars, and each day he is able to give away 2 more chocolate bars than the day before.

a How many chocolate bars will he have given away in two weeks time?

b How long will it be until he has no chocolate bars left?

10 Every week during a typing course, Tony improves his typing speed by 3 words per minute until he reaches 60 words per minute by the end of the course.

a If he can type 18 words per minute in the first week of the course, how many words per minute can he type by week 8?

b How many weeks does the course run for?

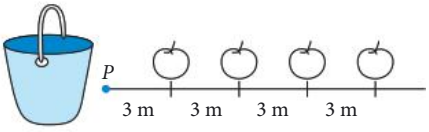
11 Example 11 A market gardener plants daffodil bulbs in rows, starting with a row of 45 bulbs. Each successive row has 5 more bulbs.

a Calculate the number of bulbs in the 34th row.

b Which row would be the first to have more than 100 bulbs in it?

c The market gardener plants 10 545 bulbs. How many rows are there?

12 In a game, a child starts at point P and runs and picks up an apple 3 m away. She then runs back to P and puts the apple in a bucket. The child then runs to get the next apple 6 m away, and runs back to P to place it in the bucket. This continues until she has all the apples in the bucket.



a How far is the kth apple from the bucket?

b How far does the child run to get to the kth apple?

c The child runs 270 m to fetch all the apples and return them to the bucket. How many apples are there?

13 For the arithmetic sequence 5, x, 11….

a find d

b find the sum of the sequence to 20 terms.

9780170250283

3.04 GEOMETRIC SEQUENCES

A **geometric sequence** or **geometric progression** is a sequence of numbers that is formed by multiplying the preceding term by a constant called the **common ratio**, denoted by the symbol r. For example, the geometric sequence 3, 6, 12, 24, 48, 96, … has a common ratio of 2. A way to find r, the common ratio, is to either recognise it by inspection, or to divide each term by the one before it.

IMPORTANT

The rule for finding r is $r = \frac{t_2}{t_1} = \frac{t_3}{t_2} = \frac{t_4}{t_3} = \cdots$.

This gives the common ratio $r = \frac{t_n}{t_{n-1}}$

To describe a geometric sequence, the common ratio (r) and the first term (a) must be stated. In the sequence above, $r = 2$ and $a = 3$.

Example 12

For the sequence $2, \frac{2}{3}, \frac{2}{9}, \frac{2}{27}, \frac{2}{81}, \ldots$ find the first term (a) and the common ratio (r).

Solution

State the first term in the sequence.

$a = 2$

Use $r = \frac{t_2}{t_1} = \frac{t_3}{t_2} = \frac{t_4}{t_3} = \cdots$ to show that there is a common ratio.

$$r = \frac{\frac{2}{3}}{2}$$
$$= \frac{\frac{2}{9}}{\frac{2}{3}}$$
$$= \frac{\frac{2}{27}}{\frac{2}{9}}$$
$$= \frac{1}{3}$$

Write the answer.

$a = 2$ and $r = \frac{1}{3}$

It is possible to work out a rule for a general term in a geometric sequence from the pattern of the terms. This is called the nth term of a geometric sequence. You can then use this rule to find any term in a sequence.

IMPORTANT

The **general term**, t_n, represents the ***n*th term** of a sequence.

For a geometric sequence

$t_1 = a$

$t_2 = t_1 r = ar$

$t_3 = t_2 r = ar^2$

$t_4 = t_3 r = ar^3$

$t_5 = t_4 r = ar^4$

Following this pattern, the general term is

$t_n = ar^{n-1}$

IMPORTANT

The rule for the nth term, t_n, of a geometric sequence is $t_n = ar^{n-1}$
where a = the first term and r = common ratio.

Consider the sequence with first term 120 and common ratio 0.5.

The 2nd term is $0.5 \times 120 = 60$.

The 3rd term is $0.5 \times 60 = 30$.

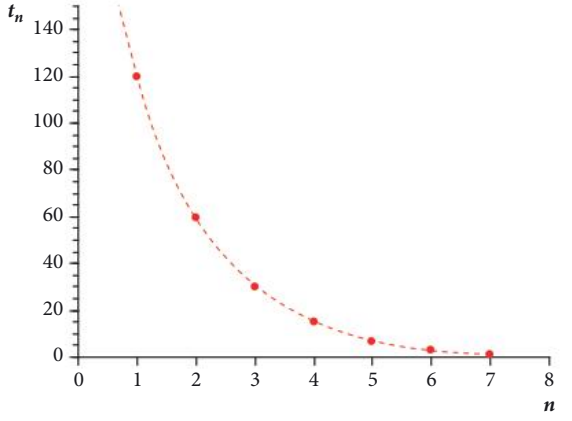
The 4th term is $0.5 \times 30 = 15$.

The 5th term is $0.5 \times 15 = 7.5$

You can continue and make a table:

n	1	2	3	4	5	6	7	...
t_n	120	60	30	15	7.5	3.75	1.875	...

You can plot a graph of the terms as shown below.



The graph is a curved line, so a geometric sequence is not a linear function.

9780170250283

Notice that you can write the terms of the sequence above as follows.

$120 = 120$

$60 = 120 \times 0.5$

$30 = 120 \times 0.5 \times 0.5 = 120 \times (0.5)^2$

$15 = 120 \times 0.5 \times 0.5 \times 0.5 = 120 \times (0.5)^3$

$7.5 = 120 \times 0.5 \times 0.5 \times 0.5 \times 0.5 = 120 \times (0.5)^4$

To get the second term, 120 is multiplied by 0.5 once.

To get the third term, 120 is multiplied by 0.5 twice.

To get the fourth term, 120 is multiplied by 0.5 three times.

To get the fifth term, 120 is multiplied by 0.5 four times.

You can do this for any number of terms and for any geometric sequence.

Example 13

If the first 6 terms of a geometric sequence are 2, 6, 18, 243, 729, 2187, …, state the value of the 8th term and the 15th term of the sequence.

Solution

Identify a and r where a = the first term and r = common ratio.	$a = 2$ $r = 3$
Use the formula $t_n = ar^{n-1}$ to find t_8 and t_{15}.	$t_8 = 2 \times 3^{8-1} = 2 \times 3^7 = 4374$ and $t_{15} = 2 \times 3^{15-1} = 2 \times 3^{14} = 9\,565\,938$
Write the answer.	$t_8 = 4374$ $t_{15} = 9\,565\,938$

Using a CAS calculator to generate the terms of a geometric sequence

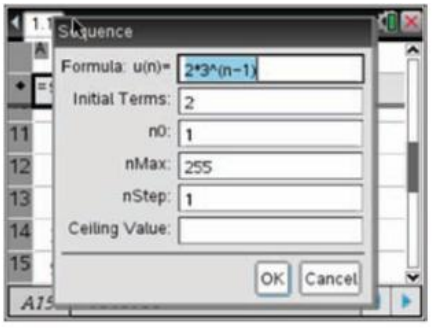
A CAS calculator can be used to generate the terms of the sequence in Example **13**.

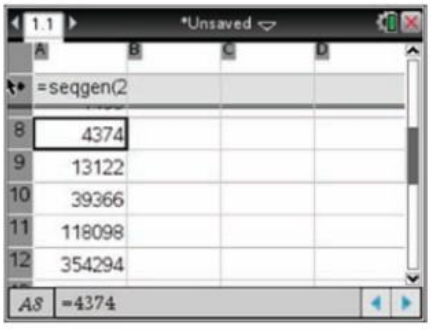
The sequence is defined with $a = 2$ and $r = 3$ so a rule for the nth term is

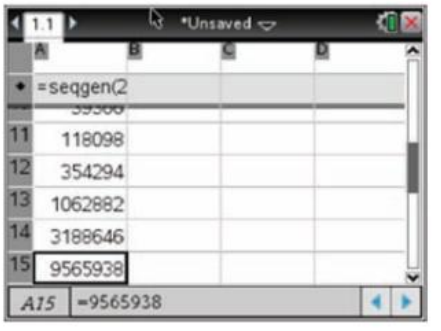
$t_n = ar^{n-1}$

$= 2 \times 3^{n-1}$

TI-Nspire CAS

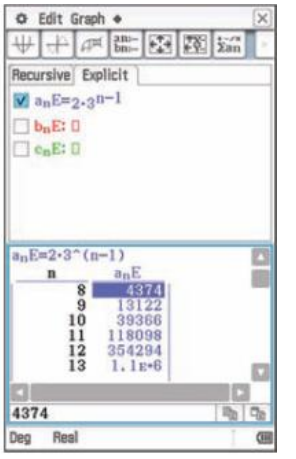


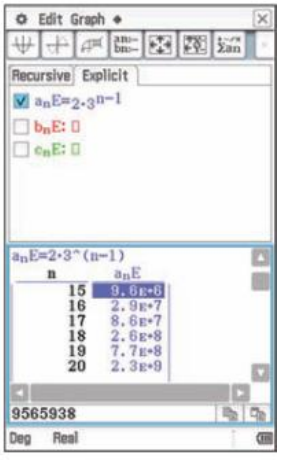




So the 8th term is 4374 and the 15th term is 9 565 938.

ClassPad





9780170250283

The formula $t_n = ar^{n-1}$ can be used to find the term at which a sequence will reach a certain number.

Example 14

For the geometric sequence 5, 10, 20, 40, …, find when the sequence first becomes larger than 1 000 000.

Solution

Identify a and r.

$a = 5$
$r = 2$

Describe the sequence using the formula $t_n = ar^{n-1}$

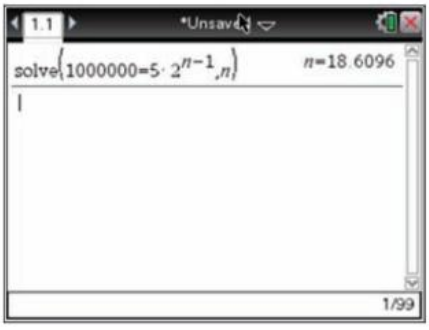
$t_n = ar^{n-1}$
$= 5 \times 2^{n-1}$

We are looking for the first term that is equal to or greater than 1 000 000 – we don't know what n is equal to, but we know this unknown term (t_n) is equal to 1 000 000. Substitute this into the formula.

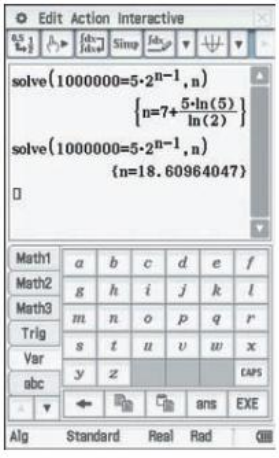
$1\,000\,000 = 5 \times 2^{n-1}$

Use a CAS calculator to solve for n.

TI-NSpire CAS



ClassPad



$n = 18.6$

It is important to recognise that because n is larger than 18, it means that after 18 terms, the sequence has not quite reached 1 000 000. Always round up in this circumstance.

Write your answer.

The sequence first becomes larger than 1 000 000 with the 19th term.

The formula $t_n = ar^{n-1}$ can be used to find a sequence if you are given 2 terms in a geometric sequence.

Example 15

Find the geometric sequence in which the 3rd term is 2 and the 10th term is 4374.

Solution

Use the formula $t_n = ar^{n-1}$ to set up an equation for t_3.

$$t_3 = ar^{3-1}$$
$$= ar^2$$
$$ar^2 = 2$$

Use the formula $t_n = ar^{n-1}$ to set up an equation for t_{10}.

$$t_{10} = ar^{10-1}$$
$$= ar^9$$
$$ar^9 = 4374$$

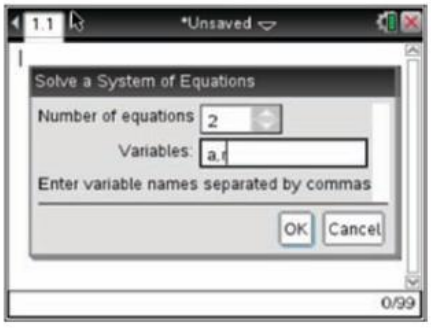
Write the simultaneous equations to be used to find the values of a and r.

$$ar^2 = 2 \quad [1]$$
$$ar^9 = 4374 \quad [2]$$

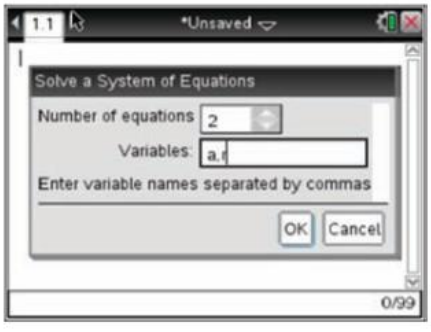
Use a CAS calculator to solve the simultaneous equations.

TI-Nspire CAS

Open a New Document with a Calculator page.
Press menu then 3: Algebra, then 7: Solve System of Equations followed by 1: Solve system of equations.
Complete the pop up screen as shown.



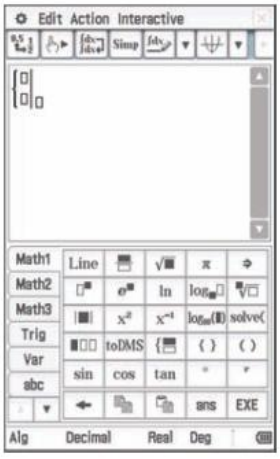
Type the two equations into the template and press enter.



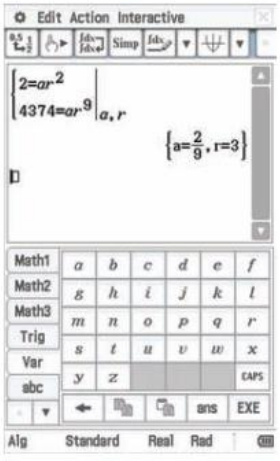
9780170250283

ClassPad

Use the Main application.
Press [Keyboard], then tap [Math1] followed by [ ].



Type the two equations into the template and specify that the variables are a and r as shown in the screen.
Press [EXE].



Write the values of a and r.

$a = \frac{2}{9}$

$r = 3$

Write the answer.

The sequence is $\frac{2}{9}, \frac{2}{3}$, 2, 6, 18, 54, 162, 486, 1458, 4374.

WS

Geometric sequences

EXERCISE 3.04 Geometric sequences

Concepts and techniques

1 Generate the first 5 terms of the geometric sequence if

a a = 1, $r = 6$ **b** a = 10, $r = -3$ **c** a = 9, $r = 3$ **d** a = −2, $r = -4$

2 Example 12 For each of these sequences, find the first term (a) and the common ratio (r)

a 7, 21, 63, 189, 567, … **b** 1000, 200, 40, 8, …

c −3, −12, −36, −108, … **d** −400, 200, −100, 50, …

3 Find the 5th term of the geometric sequence if

a $a = -2, r = 5$ **b** $a = 1, r = 6$ **c** $a = 6, r = \frac{1}{2}$ **d** $a = 2, r = -5$

4 If the first 4 terms of a geometric sequence are 5, 15, 45, 135, …, state

a the value of a **b** the value of r

c the 6th term of the sequence **d** the 10th term of the sequence

5 If the first 5 terms of a geometric sequence are $\frac{1}{4}, -\frac{3}{4}, \frac{9}{4}, -\frac{27}{4}, \frac{81}{4}, \ldots$, state

a the value of a **b** the value of r

c the 10th term of the sequence **d** the 20th term of the sequence

6 Find a and r, and an expression for t_n, for the geometric sequence 3, 9, 27, 81, 243, … Hence find the 20th term of the sequence.

7 **a** Find the common ratio r for the sequence 1, 0.2, 0.04, 0.008, 0.0016, …

b Hence find the 7th term of the sequence 1, 0.2, 0.04, 0.008, 0.0016, …

8 Example 13 For each of these sequences, state the value of t_{10} and t_{20}.

a 1, 4, 16, 64, 256, … **b** −48, 24, −12, 6, −3, …

c −10, −20, −40, −80, … **d** $12, 4, \frac{4}{3}, \frac{4}{9}, \ldots$

9 The 7th term of the geometric sequence 1, −2, 4, −8, 16, … is

A −64 **B** −8 **C** −2 **D** 1 **E** 64

10 The common ratio, r, of the geometric sequence 40, 10, 2.5, 0.625, 0.15625, … is

A −4 **B** 0.25 **C** 4 **D** 10 **E** 40

11 Show that the common ratio for the sequence $21, 6, \frac{12}{7}, \frac{24}{49}, \ldots$ is $\frac{2}{7}$.

Reasoning and communication

12 Example 14 For the geometric sequence 6, 30, 150, 750, … find when the sequence first becomes larger than 100 000.

13 **a** Find when the sequence 6, 18, 54, 162, … first becomes larger than 5 000 000.

b Find when the sequence 5 000, 1250, 312.5, … first becomes lower than 10.

c Find when the sequence −1, −2, −4, −8, … first becomes lower than −1000.

d Find when the sequence −75 000, −15 000, −3000, −600, … first becomes larger than −2.

14 Example 15 Find the geometric sequence in which the 10th term is −32 and the 13th term is −256.

15 If a geometric sequence has $a = 6$ and $r = 2$, which term has the value of 48?

16 If a geometric sequence has $a = 2$ and $r = \frac{1}{4}$, which term has the value of $\frac{1}{128}$?

17 Find the eleventh term of the geometric sequence in which $t_3 = \frac{1}{2}$ and $t_9 = \frac{729}{128}$.

18 Find a value(s) of m such that $\frac{1}{m}, \frac{m^2}{6}, \frac{m^3}{6}, \ldots$ will form successive terms of a geometric sequence.

3.05 GEOMETRIC SERIES

A **series** is the sum of terms of a sequence. So for the geometric sequence 2, 4, 8, 16, 32, 64, … the equivalent **geometric series** would be $2+4+8+16+32+64+\cdots$

We use the notation S_n to mean the sum of the sequence up to and including t_n.

INVESTIGATION Comparing arithmetic and geometric sequences

A poor girl saved a rich king from drowning one day. The king offered the girl the reward of sums of money in 30 daily payments. He gave the girl a choice of payments:

Choice 1: \$1 the first day, \$2 the second day, \$3 the third day and so on.

Choice 2: 1 cent the first day, 2 cents the second day, 4 cents the third day and so on, the payment doubling each day.

Investigate how much money the girl would receive for each choice.

Which plan would give the girl more money?

Sum of a geometric sequence

The sum of the first n terms of a geometric series is given by the formula:

$$S_n = \frac{a(r^n - 1)}{r - 1} \text{ for } |r| > 1 \quad \text{or} \quad S_n = \frac{a(1 - r^n)}{1 - r} \text{ for } |r| < 1$$

We can show this by the following.

The sum of a geometric series can be written as

$S_n = a + ar + ar^2 + \cdots + ar^{n-1}$ [1]

$rS_n = ar + ar^2 + ar^3 + \cdots + ar^n$ [2]

[1] – [2]:

$S_n(1 - r) = a - ar^n$

$= a(1 - r^n)$

$$S_n = \frac{a(1 - r^n)}{1 - r}$$

[2] – [1] gives the formula

$$S_n = \frac{a(r^n - 1)}{r - 1}$$

We use these formulas when the nth term is unknown.

IMPORTANT

The **series, S_n, of a geometric sequence** is $S_n = \dfrac{a(r^n - 1)}{r - 1}$
where a = the first term, r = common ratio, n = number of terms.

Example 16

Consider the series $3 + 12 + 48 + \cdots$ What is the sum of the first 10 terms?

Solution

Identify a, r, and n for the geometric sequence 3, 12, 48, …

$a = 3$
$r = 4$
$n = 10$

Since $r > 1$, we use the formula $S_n = \dfrac{a(r^n - 1)}{r - 1}$.

Apply the rule for S_n for 10 terms.

$$S_n = \frac{a(r^n - 1)}{r - 1}$$
$$S_{10} = \frac{3(4^{10} - 1)}{4 - 1}$$
$$= \frac{3(4^{10} - 1)}{3}$$
$$= 4^{10} - 1$$
$$= 1\,048\,575$$

Write the answer.

$S_{10} = 1\,048\,575$

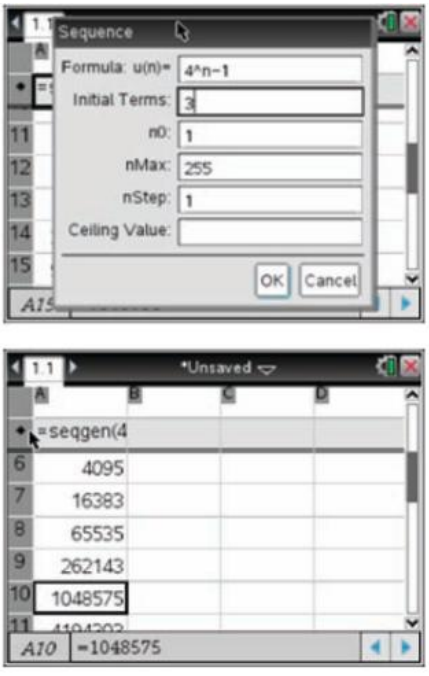
Using a CAS calculator to find the sum of the first *n* terms in a geometric sequence

A CAS calculator can be used to find the sum of the first n the terms of the sequence in Example **16**.

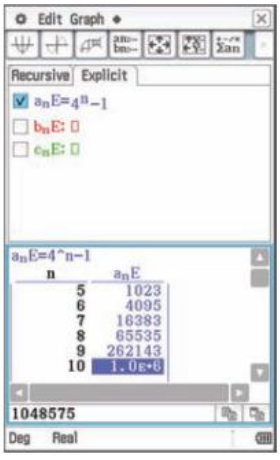
The sequence is defined with $a = 3$ and $r = 4$, so a rule for the sum for the first n terms is

$$S_n = \frac{a(r^n - 1)}{r - 1}$$
$$= \frac{3 \times (4^n - 1)}{4 - 1}$$
$$= \frac{3 \times (4^n - 1)}{3}$$
$$= 4^n - 1$$

TI-Nspire CAS



ClassPad



The sum of the first ten terms is 1 048 575.

Infinite sum of a geometric sequence

$S_\infty = \frac{a}{1-r}$, where $|r| < 1$ is the necessary condition for the **limiting sum** to exist.

We can show this by the following.

$S_n = \frac{a(1-r^n)}{1-r}$

For $|r| < 1$, as n increases, r^n decreases and approaches zero, e.g. when $r = \frac{1}{2}$:

$\left(\frac{1}{2}\right)^{10} = \frac{1}{1024}$ and $\left(\frac{1}{2}\right)^{20} = \frac{1}{1\,048\,576}$

$$S_\infty = \frac{a(1-0)}{1-r}$$
$$= \frac{a}{1-r}$$

We use this formula for an infinite number of terms where the sequence is converging to a limit.

IMPORTANT

The **sum of the infinite series, S_∞, of a geometric sequence is** $S_\infty = \frac{a}{1-r}$

where a = the first term, r = common ratio, infinite number of terms.

Example 17

Find the **infinite sum** of $2+1+\frac{1}{2}+\frac{1}{4}+\cdots$

Solution

Identify a and r where a = the first term and r = common ratio.

$a = 2$

$r = \frac{1}{2}$

Note that $r < 1$.

Use the formula $S_\infty = \frac{a}{1-r}$ for the infinite sum.

$$S_\infty = \frac{a}{1-r}$$
$$= \frac{2}{1-\frac{1}{2}}$$
$$= \frac{2}{\frac{1}{2}}$$
$$= 2 \times \frac{2}{1}$$
$$= 4$$

Write the answer.

$S_\infty = 4$

Example 18

The 3rd term of a geometric sequence is 54 and the 6th term is $11\frac{83}{125}$. Evaluate a, r and the infinite sum.

Solution

State the information given for the geometric sequence.

$t_3 = 54$

$t_6 = 11\frac{83}{125}$

Use the formula for the nth term $t_n = ar^{n-1}$ to write equations for t_3 and t_6.

$t_3 = ar^{3-1} = 54$

$t_6 = ar^{6-1} = 11\frac{83}{125}$

Write the simultaneous equations to be used to find the values of a and r.

$ar^2 = 54$ [1]

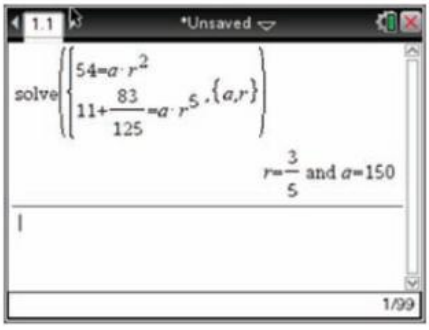
$ar^5 = 11\frac{83}{125}$ [2]

A CAS calculator can be used to solve for a and r.

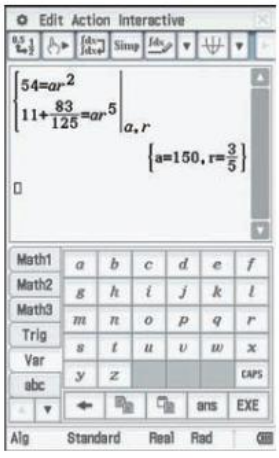
Note that $11\frac{83}{125}$ is typed as $11+\frac{83}{125}$

9780170250283

TI-NSpire CAS



ClassPad



Write the answers for a and r.

$a = 150$

$r = \frac{3}{5}$

As $r < 1$you can find the limiting sum of the geometric sequence using the formula $S_\infty = \frac{a}{1-r}$.

Substituting $a = 150$ and $r = \frac{3}{5}$ into the formula gives

$$S_\infty = \frac{150}{1-\frac{3}{5}}$$

$$= \frac{150}{\frac{2}{5}}$$

$$= 375$$

Write the answer.

$S_\infty = 375$

EXERCISE 3.05 Geometric series

Concepts and techniques

1 The expression for the partial sum to n terms of a geometric sequence is

A $\frac{a(r^n-1)}{r-1}$ B $\frac{n}{2}[2a+(n-1)d]$ C $a+(n-1)d$

D $\frac{a}{1-r}$ E ar^{n-1}

2 The expression for the infinite sum of a geometric sequence is

A $\frac{a(r^n-1)}{r-1}$ B $\frac{n}{2}[2a+(n-1)d]$ C $a+(n-1)d$

D $\frac{a}{1-r}$ E ar^{n-1}

WS Geometric series

WS Mixed sequences and series

3 The expression for the nth term of a geometric sequence is.

A $\frac{a(r^n-1)}{r-1}$ B $\frac{n}{2}[2a+(n-1)d]$ C $a+(n-1)d$

D $\frac{a}{1-r}$ E ar^{n-1}

4 Find the sum to 10 terms of a geometric sequence if:

a $a=1, r=6$ b $a=10, r=-6$ c $a=9, r=3$ d $a=-2, r=-4$

5 Example 16 Find the sum of 10 terms of these sequences.

a $6+24+96+\cdots$ b $3+15+75+\cdots$

6 Find the sum of 8 terms of these sequences.

a $-1+7-49+\cdots$ b $8+24+72+\cdots$

7 Find the sum of 15 terms of these sequences.

a $4+8+16+\cdots$ b $\frac{3}{4}-\frac{3}{8}+\frac{3}{16}-\cdots$

8 For the sequence $\frac{1}{2}$, 2, 8, 32, …

a the common ratio is

A $\frac{1}{4}$ B $\frac{1}{2}$ C 2 D 4 E 32

b the first term is

A $\frac{1}{4}$ B $\frac{1}{2}$ C 2 D 4 E 32

c the 6th term is

A 128 B 512 C 682.5 D 2048 E 2730.5

d the sum to 6 terms is

A 128 B 512 C 682.5 D 2048 E 2730.5

9 Find the sum to infinity for the geometric sequence if:

a $a=1, r=\frac{1}{2}$ b $a=10, r=\frac{1}{6}$

c $a=9, r=-\frac{1}{3}$ d $a=-2, r=-\frac{1}{2}$

10 Example 17 Find the sum to infinity for these geometric sequences.

a $6, 3, \frac{3}{2}, \frac{3}{4}, \ldots$ b 81, 27, 9, 3, …

c $-8, -2, -\frac{1}{2}, -\frac{1}{8}$ d $20, 4, \frac{4}{5}, \frac{4}{25}, \ldots$

11 Example 18 Find the sum to infinity for a geometric sequence when:

a the 3rd term is $\frac{5}{4}$ and the 5th term is $\frac{5}{256}$ and $r>0$

b the 4th term is 80 and the 6th term is 20 and $r>0$

c the 2nd term is 10 and the 5th term is $\frac{2}{25}$ d the 5th term is 1 and the 8th term is $\frac{1}{27}$

12 The geometric sequence $\frac{2}{3}, -\frac{2}{9}, \frac{2}{27}, -\frac{2}{81}, \ldots$ has

A $t_5=-\frac{2}{27}$ B $S_5=\frac{2}{243}$ C $S_5=\frac{4}{243}$

D $t_5=-\frac{2}{243}$ E $t_5=\frac{2}{243}$

Reasoning and communication

13 Find the value of n if the sum of n terms of the series $11+33+99+\cdots$ is equal to 108 251.

14 How many terms of the series $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots$ give a sum of $\frac{1023}{1024}$?

15 The common ratio of a geometric series is 4 and the sum of the first 5 terms is 3069. Find the first term.

16 Find the number of terms needed to be added for the sum to exceed 1 000 000 in the series $4+16+64+\cdots$

17 Evaluate the sum of these geometric sequences by first finding which term completes the sequence.

a $2+10+50+\cdots+6250$ **b** $18+9+4\frac{1}{2}+\cdots+\frac{9}{64}$ **c** $3+21+147+\cdots+7203$

d $\frac{3}{4}+2\frac{1}{4}+6\frac{3}{4}+\cdots+182\frac{1}{4}$ **e** $-3+6-12+\cdots+384$

3.06 MODELLING A GEOMETRIC SEQUENCE

You can use geometric sequences to model and analyse (numerically or graphically) practical situations involving geometric growth and decay; for example, analysing a compound interest loan or investment, the growth of a bacterial population that doubles in size each hour, the decreasing height of the bounce of a ball at each bounce; or calculating the value of office furniture at the end of each year using the declining (reducing) balance method to depreciate. With large numbers in these real-life situations, the use of technology (e.g. handheld calculators, computer spreadsheets, etc.) helps enormously with these calculations.

Example 19

A population of rabbits is seen to increase by 11% each month. If the population in the first month was 20 rabbits, how many rabbits will be in the population in 2 years?

Solution

State a, r and n.

$a = 20$
$r = 1 + 11\% = 1.11$
$n = 2 \times 12 = 24$

Describe the sequence using $t_n = ar^{n-1}$.

$t_n = ar^{n-1}$
$t_{24} = 20 \times 1.11^{24-1}$
$= 20 \times 1.11^{23}$

Solve the equation.

$= 220.525$

Write your answer, rounding to the nearest whole number.

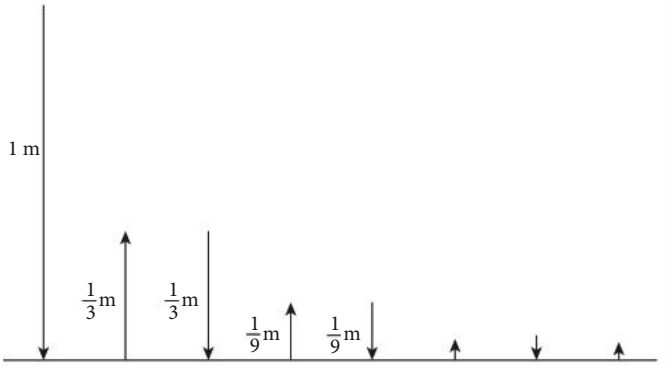
There will be approximately 221 rabbits after 2 years.

Example 20

A ball is dropped from a height of 1 metre and bounces up to $\frac{1}{3}$ of its height. It continues bouncing, rising $\frac{1}{3}$ of its height on each bounce until it eventually remains on the ground. What is the total distance through which it travels?

Solution

Use a diagram to visualise the bouncing ball.



Find a and r for the geometric sequence $\frac{1}{3}+\frac{1}{9}+\frac{1}{27}+\cdots$

$\frac{1}{3}+\frac{1}{9}+\frac{1}{27}+\cdots$ is a geometric series with $a=\frac{1}{3}$ and $r=\frac{1}{3}$

Use the formula for an infinite geometric series: $S_\infty=\frac{a}{1-r}$.

$$S_\infty=\frac{a}{1-r}$$
$$=\frac{\frac{1}{3}}{1-\frac{1}{3}}$$
$$=\frac{\frac{1}{3}}{\frac{2}{3}}$$
$$=\frac{1}{3}\times\frac{3}{2}$$
$$=\frac{1}{2}$$

Calculate the total distance travelled if the ball firstly drops 1 metre and every other bounce is an 'up and down' bounce, requiring doubling the distance.

$$\text{Total distance}=1+2\left(\frac{1}{2}\right)$$
$$=1+1$$
$$=2$$

Write the answer.

The ball travels 2 metres altogether.

9780170250283

A geometric sequence can be also used to express a **recurring decimal** as a fraction.

Example 21

Write $0.\overline{5}$ as a fraction.

Solution

Consider what $0.\overline{5}$ equals.

$$0.\overline{5} = 0.5555555\ldots$$
$$= \frac{5}{10} + \frac{5}{100} + \frac{5}{1000} + \cdots$$

Identify the geometric sequence.

This is a geometric series with $a = \frac{5}{10} = \frac{1}{2}$ and $r = \frac{1}{10}$.

Calculate the infinite sequence using the formula $S_\infty = \frac{a}{1-r}$ and $a = \frac{1}{2}$ and $r = \frac{1}{10}$.

$$S_\infty = \frac{a}{1-r}$$
$$= \frac{\frac{1}{2}}{1-\frac{1}{10}}$$
$$= \frac{\frac{1}{2}}{\frac{9}{10}}$$
$$= \frac{1}{2} \times \frac{10}{9}$$
$$= \frac{5}{9}$$

Write the answer.

$$0.\overline{5} = \frac{5}{9}$$

EXERCISE 3.06 Modelling a geometric sequence

Reasoning and communication

1 Example 19 Adele purchases a house, and it is valued at $650 000 in the first year. If it increases in value by 5% every year, how much will the house be worth in the 10th year of ownership? Answer correct to the nearest dollar.

2 A colony of bacteria contains 100 cells in its first twenty minutes. If the population doubles every 20 minutes, how many cells will the population contain in 4 hours?

3 A population of the endangered Eastern Barred Bandicoot consists of 162 individuals in its first year. If this population decreases by 16% each year, how many bandicoots will be left in its fifth year?

4 Example 20 A ball is dropped from a height of 2 m and bounces up to $\frac{3}{4}$ of its height. It continues bouncing, rising $\frac{3}{4}$ of its height on each bounce until it eventually remains on the ground. What is the total distance through which it travels?

5 A seed sapling grows to 80 cm high in its first year. If its growth rate halves each year, what is the maximum height that it will reach?

6 A novice weightlifter is able to dead lift 40 kg in his first month. In his second month he is able to increase this to 50 kg. Each month his improvement slows by 25%. What is the maximum weight that he will be able to lift?

7 A painting appreciates (increases its value) by 16% p.a. It is currently worth $20 000.
   a How much will it be worth in
      i 1 year? ii 2 years? iii 3 years?
   b How much will it be worth in 11 years?
   c How long will it take before it is worth $50 000?

8 Kim owes $1000 on her credit card. If she pays back 10% of the amount owing each month, she will never finish paying it off. Is this true or false?

9 Example 21 Use a geometric sequence to express these recurring decimals as fractions.
   a $0.\overline{7}$ b $0.\overline{2}$ c $0.\overline{16}$ d $0.\overline{234}$

10 When she started university, Chelsea bought herself a new car. At the start of her 10th year of ownership, she had it valued for resale at approximately $3860. She calculated that it had depreciated by an average of 14% every year. How much did Chelsea originally pay for the car? Answer correct to the nearest dollar.

11 Isabella started an online store and sold 50 items in her first year. She found that her sales increased by 20% each year. In which year did she first sell over 2000 items?

12 For the geometric sequence 4, x, 144, …
   a find r if $r > 0$ b find the sum of the sequence to 10 terms.

13 a For what values of p will the limiting sum exist for the series $1-2p+4p^2-\cdots$?
   b Find the limiting sum when $p=\frac{1}{5}$.
   c Evaluate p if the limiting sum of the series is $\frac{7}{8}$.

14 Show that in any geometric series the difference between the limiting sum and the sum of n terms is $\frac{ar^n}{1-r}$.

3.07 ARITHMETIC AND GEOMETRIC MODELLING

Arithmetic and geometric series are useful in solving real-life problems involving patterns. It is important to choose the correct type of series and know whether you are asked to find a term or a sum of terms.

To check which type of sequence is appropriate, look for keywords such as multiplied, divided, percentages (geometric sequences) and added and subtracted (arithmetic sequences). If a sequence is given, check each pair of terms for a constant difference or ratio.

IMPORTANT

Important formulas to remember.

	Arithmetic sequence	Geometric sequence
t_n	$t_n = a + (n-1)d$	$t_n = ar^{n-1}$
S_n	$S_n = \frac{n}{2}[2a + (n-1)d]$	$S_n = \frac{a(r^n - 1)}{r - 1}$ for $\lvert r\rvert > 1$ or $S_n = \frac{a(1 - r^n)}{1 - r}$ for $\lvert r\rvert < 1$
Special sums	**When the last term is known:** $S_n = \frac{n}{2}(a + l)$	**Infinite sum:** $S_\infty = \frac{a}{1 - r}$



Example 22

Mohamed is in charge of fundraising for his football club. In his first year, he raises $25 000 for the club. In his second year, he raises $20 000 and in his third year he raises $16 000.

a How much does he raise in his fifth year?

b Find the total maximum amount of money that Mohamed can raise if this pattern continues.

Solution

Determine if this is a geometric or arithmetic sequence.

$25\,000 - 20\,000 = 5000$

$20\,000 - 16\,000 = 4000$

This is not an arithmetic sequence.

$\frac{20\,000}{25\,000} = \frac{16\,000}{20\,000} = 0.8$

This is a geometric sequence.

Identify a and r.

$a = 25\,000$

$r = 0.8$

a Use the formula $t_n = ar^{n-1}$ to calculate t_5

$t_n = ar^{n-1}$

$t_5 = 25\,000 \times 0.8^4$

$= 10\,240$

Mohamed raises $10 240 in his fifth year.

b Total maximum implies an infinite sum. Use the formula $S_\infty = \frac{a}{1 - r}$ to find the total maximum amount of money that Mohamed can raise.

$S_\infty = \frac{a}{1 - r}$

$= \frac{25\,000}{1 - 0.8}$

$= 125\,000$

Mohamed can raise a total of $125 000 if this pattern continues.

Example 23

A large screen at a ski resort shows the current temperature updated every hour. One hour after opening the ski lifts, the temperature was displayed as 3°C, but it got steadily colder during the day and dropped by 0.6°C every hour. How many hours after opening the ski lifts is the temperature 0°C?



iStockphoto/harmatoslabu

Solution

Determine if this is a geometric or arithmetic sequence.	The temperature drops by a constant 0.6°C every hour, which means that this is an arithmetic sequence.
Identify a and d.	$a = 3$ $d = -0.6$
Describe the sequence using the formula $t_n = a + (n-1)d$	$t_n = a + (n-1)d$ $t_n = 3 + [-0.6(n-1)]$
We are looking for the term that is equal to 0. We don't know what n is equal to, but we know that this unknown term (t_n) is equal to 0. Substitute this into the formula.	$0 = 3 + [-0.6(n-1)]$
Expand the brackets, rearrange and solve.	$0 = 3 - 0.6n + 0.6$ $0 = 3.6 - 0.6n$ $3.6 = 0.6n$ $n = 6$
Write the answer.	The temperature is 0°C after 6 hours.

9780170250283

Example 24

A layer of tinting for a car window lets in 95% of light.

a What percentage of light is let in by

- **i** 2 layers
- **ii** 3 layers
- **iii** 10 layers of tinting?

b How many layers will let in 40% of light?

Solution

a Determine if this is a geometric or arithmetic sequence.

The first layer of tinting lets in 95% of light. Each subsequent layer lets in 95% of light. This forms a geometric sequence with $a = 0.95$ and $r = 0.95$.

i Use the formula $t_n = ar^{n-1}$ to calculate 2 layers, $n = 2$

$$\begin{aligned} t_2 &= 0.95 \times 0.95^{2-1} \\ &= 0.95 \times 0.95 \\ &= 0.9025 \end{aligned}$$

ii 3 layers, $n = 3$

$$\begin{aligned} t_3 &= 0.95 \times 0.95^{3-1} \\ &= 0.95 \times 0.95^2 \\ &= 0.857375 \end{aligned}$$

iii 10 layers, $n = 10$

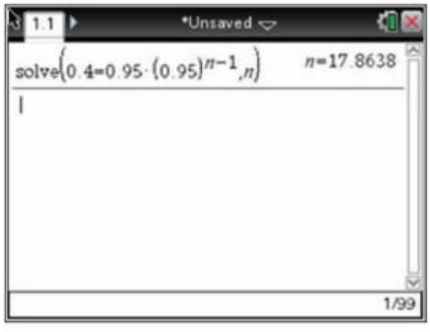
$$\begin{aligned} t_{10} &= 0.95 \times 0.95^{10-1} \\ &= 0.95 \times 0.95^9 \\ &= 0.598\,736\,939\,2 \end{aligned}$$

b We want to find n when the nth term is 40% or 0.4.

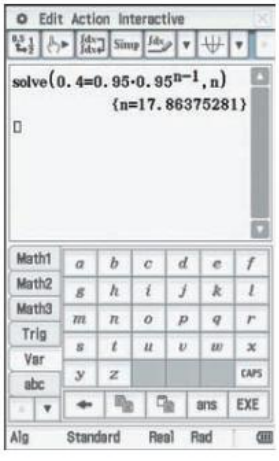
$$\begin{aligned} t_n &= ar^{n-1} \\ 0.4 &= 0.95(0.95)^{n-1} \end{aligned}$$

Use a CAS calculator to solve for n.

TI-NSpire CAS



ClassPad



Round up to the nearest whole number.

So around 18 layers of tinting will let in 40% of light.

Write the answer.

18 layers

Example 25

Madison's salary increases by $1500 each year. If after 10 years she has earned $417 500 in total, find her projected salary 20 years after starting.

Solution

Identify if this is a geometric or arithmetic sequence.

Her salary increases by a constant $1500 each year, so this is an arithmetic sequence.

We want to find t_{20}, but we are not given her initial salary (a), so we must find this first. Identify d, n and S_n

$d = 1500$
$n = 10$
$S_n = 417\,500$

Use the formula $S_n = \frac{n}{2}[2a + (n-1)d]$ to find a.

$$S_n = \frac{n}{2}[2a + (n-1)d]$$

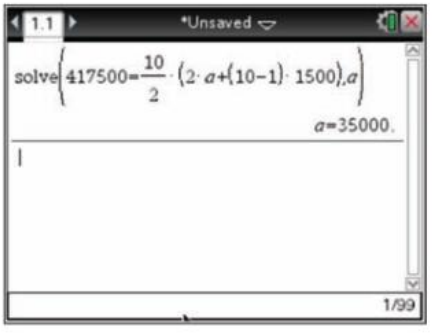
$$417\,500 = \frac{10}{2}[2a + (10-1)1500]$$

$417\,500 = 5[2a + 1500 \times 9]$
$417\,500 = 5[2a + 13\,500]$
$83\,500 = 2a + 13\,500$
$70\,000 = 2a$
$35\,000 = a$

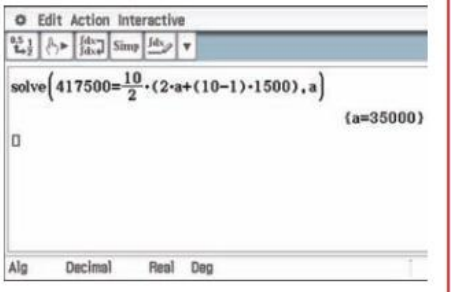
Madison's starting salary is $35 000.

A CAS calculator can be used to solve the equation.

TI-NSpire CAS



ClassPad



We can now use the formula $t_n = a + (n-1)d$ to find t_{20}.

$t_n = a + (n-1)d$
$t_{20} = 35\,000 + (20-1)1500$
$t_{20} = 35\,000 + 1500 \times 19$
$t_{20} = 63\,500$

Madison's projected salary after 20 years is $63 500.

9780170250283

EXERCISE 3.07 Arithmetic and geometric modelling

WS Modelling arithmetic and geometric sequences

Reasoning and communication

1 A geometric sequence is
- **A** the sum of all the multiples of 7 between 1 and 100
- **B** $8 + 14 + 20 + 26 + \cdots$
- **C** 0.3, 0.6, 0.9, …
- **D** $x, x + 3, x + 6, \ldots$
- **E** $x, x^2, x^3, \ldots$

2 Example 22 A new shop opens up and gets 215 customers in its first week. In its second and third week it gets 234 and 253 customers respectively. Assuming that the trend continues, find
- **a** the number of customers in the 10th week
- **b** the total number of customers the shop will receive for a full year.

3 Pine logs waiting in storage in a woodchip plant are stacked in a pile such that there are 25 in the top row, 26 in the next row down, 27 in the next row down and so on. If there are 120 logs in the bottom row
- **a** how many rows are in the pile?
- **b** how many logs are in the pile?

4 Daniel has a box containing 120 identical blocks. He constructs vertical towers using the blocks. The first tower is made using 3 blocks, the second using 5 blocks, the third using 7 blocks and so on. If he continues building the towers following this pattern, what is the maximum number of towers that he can make?

5 In a display, oranges are arranged on a sloping shelf to make a triangle. There are 50 oranges in the bottom row, 49 in the next row, 48 in the next row and so on. There is one orange at the top of the triangle. How many oranges are in the display?

6 Example 23 A pool contains 30 000 litres of water. A crack in the wall allows water to leak from the pool at a rate of 150 litres per minute. How much water is left in the pool after 2 hours?

7 Fiona is training to run a marathon. In the first week of training, she runs 5 km per day. In the second week she runs 8 km per day, in the third week she runs 11 km per day and so on.
- **a** How many kilometres per day does she run in the eighth week of training?
- **b** Fiona runs on four days each week. Assuming that this pattern of training continues, find the total distance she has run after ten weeks of training.

8 A tennis club embarked upon a fundraising campaign to raise money for new courts. In the first year of fundraising, they raised $20 000. Each year, after that they raised 10% more than in the previous year. Find the total amount of money that they had raised by the end of the fifth year.

9 Max, a part time cleaner, is employed on an initial salary of $32 000 per year under an agreement that states that his salary will increase by $3600 per annum. Find his total earnings after 10 years.

10 The value of an art collection appreciates at an average rate of 8% per year. The collection was initially purchased for $20 000. Find the value of the collection 12 years after its purchase. Answer correct to the nearest dollar.

11 Cars depreciate at varying rates, according to their make and how well they have been looked after by the owner. A car that was initially purchased for $39 000 depreciates at the rate of 16% per annum. Find its value 10 years after it was purchased. Answer correct to the nearest dollar.

12 Example 24 A layer of tinting on a car window lets in 90% of light.
- a What percentage of light is let in by 3 layers?
- b How many layers will let in less than 50% of light?

13 Maggie's porcelain doll collection increases by 3 dolls each year. She has 16 dolls at the end of the third year of collecting. How many dolls does she have at the end of the 10th year?

14 Membership fees at a gym rise by $15 each year. Sophie has been going to the gym for 5 years and her annual fee is currently $1020.
- a How much did Sophie pay in her first year of fees?
- b How much will Sophie have paid in total when she has been a member for 10 years?
- c When will Sophie pay $1200 in annual fees?

15 Sam is training for high jump. In her first week she increases her jump height by 10 cm. She improves every week, but her improvement drops by 50% each week. If her starting jumps were at 1.4 m, what is the maximum height she will be able to jump?

16
- a Calculate the sum of all the multiples of 7 between 1 and 100.
- b Calculate the sum of all numbers between 1 and 100 that are not multiples of 7.

17 Example 25 Lachlan's salary increases by 7% every year. If after 8 years he has earned $307 794 in total, find his projected salary 25 years after starting.

18 Over the past 7 years, Nick has collected a total of 189 books. If every year he collects 4 more books than he collected the previous year, find:
- a the number of books he collected in the first year
- b the number of books he collected in his 7th year
- c when his collection will first exceed 500 books.

19 Chantelle runs 5 km in 33 minutes. After a week of training she reduces this time to 30 minutes. If over the course of time she manages to cut a maximum of 15 minutes from her run time, find
- a the rate she cuts her time by each week
- b how much she cuts from her race time in her third week of training
- c her race time after 5 weeks of training
- d when her race time will drop below 20 minutes

20 The total weight of 100 packages is 10.4 tonnes. If the weights of consecutive packages differ by 2 kg, find the weights of the heaviest and lightest packages.

9780170250283

CHAPTER SUMMARY

ARITHMETIC AND GEOMETRIC SEQUENCES

3

- The **general term**, t_n, represents the ***n*th term** of a sequence
- The graph of an arithmetic sequence is a straight line with slope equal to the common difference.
- The rule for the ***n*th term, t_n, of an arithmetic sequence** is $t_n = a + (n-1)d$, where a = the first term and d = common difference.
- The rule for finding d is $d = t_2 - t_1 = t_3 - t_2 = t_4 - t_3 = \ldots$.

  This gives the general common difference $d = t_n - t_{n-1}$
- Partial sum of an arithmetic sequence

  The sum of the first n terms of an arithmetic series (nth partial sum) is given by the formula: $S_n = \frac{n}{2}(a+l)$ where l = last or nth term.
- The **series, S_n, of an arithmetic sequence** is $S_n = \frac{n}{2}[2a+(n-1)d]$,
  where a = the first term, d = common difference, n = number of terms.
- **Straight-line depreciation**: where the book value depreciates by a fixed dollar value each year.

  Reducing balance depreciation: the book value depreciates by a given percentage each year.

  Unit cost method of depreciation: is dependent on the amount of use the asset has had.
- The rule for the ***n*th term, t_n, of a geometric sequence** is $t_n = ar^{n-1}$, where a = the first term and r = common ratio.
- The graph of a geometric sequence is not a straight line.
- The rule for finding r is $r = \frac{t_2}{t_1} = \frac{t_3}{t_2} = \frac{t_4}{t_3} = \cdots$.

  This gives the general **common ratio** $r = \frac{t_n}{t_{n-1}}$
- The nth partial sum of a geometric sequence is given by

  $S_n = \frac{a(r^n - 1)}{r-1}$ for $|r| > 1$ or $S_n = \frac{a(1-r^n)}{1-r}$ for $|r| < 1$
- The **series, S_n, of a geometric sequence** is

  $S_n = \frac{a(r^n - 1)}{r-1}$ where a = the first term, r = common ratio, n = number of terms.
- **The sum of the infinite series, S_∞, of a geometric sequence is** $S_\infty = \frac{a}{1-r}$

  where a = the first term, r = common ratio, infinite number of terms.

3 CHAPTER REVIEW
ARITHMETIC AND GEOMETRIC SEQUENCES

Multiple choice

Use the following information to answer questions **1** and **2**.

As part of a savings plan, Tom saved $100 the first month and successively increased the amount that he saved each month by $10. That is, in the second month he saved $110, in the third month he saved $120, and so on.

1 Example 2 The amount that Tom will save in the 20th month is
A $280 B $290 C $300 D $3900 E $4200

2 Example 5 The total amount that Tom will save in twenty months is
A $280 B $290 C $300 D $3900 E $4200

3 Example 9 On the first day of a fundraising program, three girls sold 6 badges each. On the second day, each of those three girls sold double that amount. On the third day, each of those three girls sold double the amount they sold on the second day. The fundraising program continued in this pattern for seven days. The total number of badges sold in this fundraising activity was
A 126 B 1134 C 1152 D 2286 E 2304

4 Example 18 The second and third terms of a geometric sequence are 100 and 160 respectively.

The sum of the first ten terms of this sequence is closest to
A 4300 B 6870 C 11 000 D 11 290 E 11 350

5 Oscar planted a 20 cm high dwarf lemon tree in a pot. After it was planted, the height of the tree increased by 16 cm in the first month, 8 cm in the second month and 4 cm in the third month. Assuming that this pattern of growth continues, the maximum height that the tree will grow to is
A 32 cm B 36 cm C 40 cm D 48 cm E 52 cm

Short answer

6 Examples 1, 2 For the sequence 1, 5, 9, 13, 17, … find
a the first term (a) and the common difference (d)
b the 10th term of the sequence.

7 Example 3 For the arithmetic sequence −42, −35, −28, …, find which term is equal to zero.

8 Example 4 Find the arithmetic sequence in which the 4th term is 5 and the 9th term is 35.

9 Example 5 Find the sum of the first 20 terms for the sequence 3, 7, 11, 15, …

10 Example 6 Find the sum of the sequence 22, 29, 36, …,106.

11 Example 7 Find an expression for the sum of the first n odd numbers.

9780170250283

12 Example 8 Renee bought a pair of dogs, and at the end of the year decides to breed them. If the dogs have three puppies every year, find how many dogs Renee will own in 5 years.

13 Example 9 Stuart is trying to improve his fitness. In his first week he runs 10 km. He increases this to 12 km in the second week, 14 km in the third week and so on. If he continues this pattern, how many kilometres will he have run in total in 10 weeks?

14 Example 10 A car is bought for $10 500. If it depreciates by $750 each year,
a how much will it be worth in its 8th year?
b when will it be worth nothing?

15 Example 11 Elizabeth creates a cup pyramid, with 1 cup on the top row, 3 cups on the second row, 5 cups on the third row and so on.
a How many cups are in the 10th row?
b If there are 625 cups in the pyramid, how many rows are there?

16 Examples 12, 13 For the sequence 3, 6, 12, 24, … find:
a the first term (a) and the common ratio (r)
b the 10th term of the sequence

17 Example 14 For the sequence 4, 16, 64, … find when it first becomes larger than 500 000.

18 Example 15 Find the geometric sequence with $r > 0$ in which the 5th term is 250 and the 7th term is 156 250.

19 Example 16 Find the sum of the first 13 terms for the sequence $-6, -18, -54, \ldots$

20 Example 17 Find the infinite sum of the sequence 375, 75, 15, 3, …

21 Example 18 The 3rd term of a geometric sequence is 12 and the 6th term is $-\frac{3}{16}$. Find a, r and the infinite sum.

22 Example 19 A population of kangaroos is seen to increase by 17% each year. If the population in the first year was 215 kangaroos, how many kangaroos will be in the population in 10 years?

23 Example 20 A ball is dropped from a height of 1.7 metres and bounces up to $\frac{1}{4}$ of its height. It continues bouncing, rising $\frac{1}{4}$ of its height on each bounce until it eventually remains on the ground. What is the total distance through which it travels?

24 Example 21 Write the recurring decimal $0.\overline{4}$ as a fraction.

25 Example 22 A tree is planted in a backyard when it is 1.2 m tall. After it was planted, its height increased by 60 cm in the first month, 30 cm in the second month and 15 cm in the third month. Assuming that this pattern of growth continues, find the maximum height to which the tree will grow.

26 Example 23 Zoe has a large number of pamphlets to deliver. At the end of her first hour, she has 3000 pamphlets left. If she can deliver an average of 250 pamphlets an hour, how many hours will it take her to deliver the pamphlets?

27 Example 24 Water evaporates from a backyard pool at an average rate of 5% each week.

a What percentage of pool is left in the pond after

i 1 week? ii 2 weeks? iii 3 weeks?

b What percentage is left after 15 weeks?

c If there was no rain, approximately how long would it take for the pool to only have 25% of its water left?

28 Example 25 The number of customers in a store increase by 16% each week. If after 15 weeks, a total of 1085 customers have come to the store, find the projected number of customers in the 30th week.

Application

29 In the first month of land sales in a new housing estate, 100 blocks of land were sold.

In the second month of land sales, 105 blocks of land were sold.

In the third month of land sales, 110 blocks of land were sold.

The land sales continued in this pattern until all the blocks of land were sold.

The number of blocks of land sold each month forms the terms of an arithmetic sequence.
100, 105, 110, …

a Show that the common difference for this sequence is 5.

b How many blocks of land were sold in the sixth month of land sales?

c How many **more** blocks of land were sold in the tenth month than in the eighth month?

d How many blocks of land were sold, in total, in the first 12 months?

30 In the first year after land sales began, 15 building applications were approved by the council.

The number of building applications that the council approved each year after that formed the terms of a geometric sequence with a common ratio of 1.5.

a How many building applications were approved by the council in the third year after land sales began?

b In which year after land sales began did the number of building applications approved by the council each year first exceed 100?

c How many building applications were approved, in total, by the council, in the first five years after land sales began?

d In which year after land sales began did the total number of building applications approved by the council first exceed 1000?

Practice quiz

9780170250283

MIXED REVISION
CHAPTERS 1 • 2 • 3

1

Multiple choice

1 Which of the following statements is **not true**?
- A The mean is a measure of centre.
- B The mean is a type of average.
- C The mean is a measure of spread.
- D The mean is calculated by adding up all the values and dividing by the number of values.
- E The mean can be affected by outliers.

2 The following regression line was calculated from a set of collected data.

$$\text{homework hours} = 17.2 - 0.37 \times \text{gaming hours}.$$

Predict the number of hours spent gaming by a person who completed 10 hours of homework.

A 13.5 hours B 15.23 hours C 17.8 hours
D 19.46 hours E 21.34 hours

3 $1+3+5+7+\cdots+21$ equals

A 10 B 11 C 21
D 110 E 121

4 The sum of the infinite geometric sequence 48, –24, 12, –6, 3 . . . is equal to

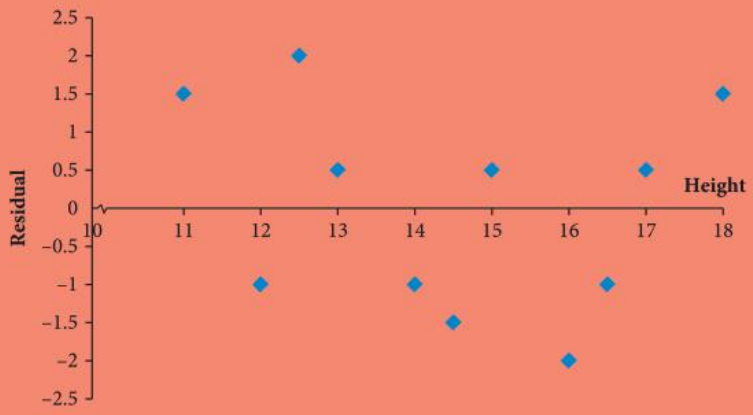
A 32 B 33 C 34
D 72 E 96

5 In a distribution that is negatively skewed:
- A the tail points to the left and scores are mostly high
- B the tail points to the left and scores are mostly low
- C the tail points to the right and scores are mostly low
- D the tail points to the right and scores are mostly high
- E the data is a mirror image through the median.

6 Calculate Pearson's correlation coefficient for the relationship with an r^2 value of 0.67 and a regression line of *distance travelled* = 30 + 1.9 × *time*.

A –0.82 B –0.45 C 0.45
D 0.82 E 1

7 The height and weight of 11 adults was recorded and the residual plot shown below was constructed.



The residual plot suggests that:

A there is a strong negative relationship between the weight and height of the 11 adults
B the original data was probably non-linear
C the original data was probably linear
D all the points on the scatter plot would lie above the regression line
E all the points on the scatter plot would lie below the regression line.

8 $1+\frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{512}=$

A $\frac{511}{256}$ B $\frac{1023}{512}$ C $\frac{2047}{1024}$

D $\frac{1023}{1024}$ E 460

9 A person's surname is best described as:

A numerical, discrete data
B numerical, continuous data
C categorical, nominal data
D categorical, ordinal data
E categorical, discrete data

9780170250283

Short answer

1 Consider the arithmetic series $9 + 13 + 17 + \ldots$. How many terms are required for this series to produce a sum of 1122?

The following information relates to questions **2** and **3**.

The following table shows the temperature of the water in a gas kettle and the time it has been on the gas.

Time (seconds)	10	25	37	40	50	70	90
Temperature (°C)	35	42	47	58	64	73	89

2 a Construct a scatter plot to represent this data.
b Calculate and interpret the correlation coefficient, correct to 2 decimal places.
c Find the least squares regression line, stating all coefficients correct to 1 decimal place.

3 a Complete the following residual table for the data, stating all answers correct to 1 decimal place.

Time (seconds)	10	25	37	40	50	70	90
Residual	1.5				3		

b Hence construct a residual plot
c From your residual plot, what can you say about the data?

4 a Calculate the mean, median and mode for the following data.

15 17 18 18 19 20 20 21 22 23 24

b An extra data point is added to the list, with a value of 35. Determine if the extra data point is an outlier or not.
c Calculate the mean, median and mode again and comment on the effect the outlier has had on each measure.

5 The heights, in centimetres, of a group of men and women are displayed in the back to back stem-and-leaf plot below.

Men	Stem	Leaf
9	15	2 4 5 5 6 7 8 9
8 7 5 6 2	16	0 2 2 3 4 5 5 5 6 6 7 8 8
9 9 8 8 7 5 5 4 4 3 2 1	17	2 3 4 9
8 5 3 0	18	3
5	19	

Key: 15/2 means 152cm

a Represent the data shown in parallel box plots.
b Which group had the higher median?
c Which group had the wider spread of heights?
d Comment on the skewness of each data set.
e Calculate the mean heights of men and women, correct to 1 decimal place.
f Which is the better measure of centre for these data sets? Why?

6 Find the 20th term of the series $3+10+17+\cdots$

7 Find which term of $3+7+11+\cdots$ is equal to 111.

Application

1 A study of causes of death in Australia found that, from 54 000 deaths, 9400 resulted from a heart attack. Of those who died from a heart attack, 3020 were obese. A total of 5840 of the 54 000 studied were considered to be obese.
- a Construct a two-way table for 'Cause of death' and 'Weight category'.
- b What percentage of people who died from a heart attack were obese? Answer correct to the nearest whole number.
- c Construct a percentage two-way table.
- d Display the information with a segmented column graph.
- e Discuss any findings from the graph.

2 The heights and weights of 12 students are shown in the table below.

Height (cm)	171	174	165	181	178	167	183	172	175	181	176	182
Weight (kg)	61	59	54	75	65	61	70	58	64	76	73	70

Where appropriate, state all answers correct to 2 decimal places.
- a Calculate the least-squares regression line that models the data.
- b Interpret the gradient and y-intercept of the regression line.
- c Calculate the coefficient of determination. Answer correct to 2 decimal places.
- d Use the coefficient of determination to justify if the regression line is an appropriate model.
- e Use the regression model to predict:
  - i the weight of a student who is 173cm tall
  - ii the weight of a student who is 151cm tall
  - iii the height of a student who weighs 91 kg.
- f Which of the predictions in part **e** are reliable? Justify your answer.

3 The nth term of a sequence is given by $t_n = 5n - 1$. Which term of the sequence is equal to 104?

4 A marble was rolled down a 3 metre plank and its speed was recorded at different distances along the plank. The results are shown in the following table.

Distance (m)	0	0.4	0.6	1	1.2	1.6	2	2.2	2.4	2.6	2.8	3
Speed (m/sec)	0	2	1.7	2.6	2.8	2.8	3.4	2.9	3.1	3.4	3.8	4.1

Where appropriate state all answers correct to 1 decimal places.
- a Which variable is the response variable?
- b Construct a scatter plot to represent the data.
- c Calculate the least squares regression line that represents the data.
- d Construct a table of residuals for the data.
- e Hence construct a residual plot.
- f Using the residual plot and other statistics relating to the data justify if the regression model found in part c is appropriate for making predictions.

5 The nth term of a sequence is given by $t_n = 103 - 3n$.
- a Find the first 3 terms in the sequence.
- b Find the value of n for the first negative term in the series.

9780170250283

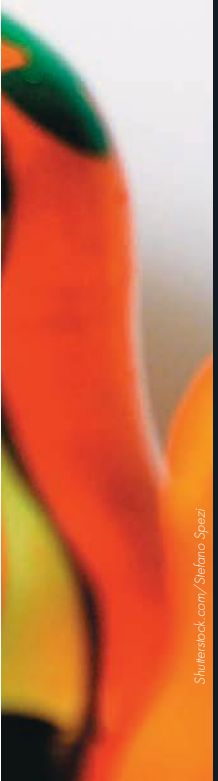
6 A survey was conducted to compare the ages of mothers and their daughters on the birth of their first child. The results are displayed below.

Mother's age	Daughter's age
19	25
19	33
23	28
21	30
24	21
25	29
23	26
23	22
22	34
18	17
26	29
26	35
26	26
16	24
20	39

a Calculate the mean age of the mothers on the birth of their first child. Answer correct to two decimal places.
b Calculate the mean age of the daughters on the birth of their first child. Answer correct to two decimal places.
c Compare your answers from part **a** and **b**.
d Construct a scatter plot to represent the data.
e Visually interpret the scatter plot, referring to form, direction and strength.
f Calculate the correlation coefficient for the data, correct to two decimal places.
g Hence interpret the correlation coefficient value and write a brief sentence describing the relationship between the ages of mothers and their daughters on the birth of their first child.

7 a Show that $(x+1)+(2x+4)+(3x+7)+\cdots$ is an arithmetic series.
b Find the sum of the first 50 terms of the series.





Shutterstock.com/Stefano Spezi

4

TERMINOLOGY

arithmetic sequence
common difference
common ratio
difference equation
first order difference equation
first order recurrence relation
first term
geometric sequence
long-term
recurrence relation
sequence
steady-state
term

Prior learning

GROWTH AND DECAY IN SEQUENCES

RECURRENCE RELATIONS

4.01 Generating sequences using recurrence relations
4.02 Finding a recurrence relation for a given sequence
4.03 Recurrence relations for arithmetic sequences
4.04 Recurrence relations for geometric sequences
4.05 Graphical representation of first order recurrence relations
4.06 Long-term steady-state solution of a recurrence relation
4.07 Applications of recurrence relations
Chapter summary
Chapter review

THE ARITHMTIC SEQUENCE

- use recursion to generate an arithmetic sequence (ACMGM067)

THE GEOMETRIC SEQUENCE

- use recursion to generate a geometric sequence (ACMGM071)

SEQUENCES GENERATED BY FIRST-ORDER LINEAR RECURRENCE RELATIONS

- use a general first-order linear recurrence relation to generate the terms of a sequence and to display it in both tabular and graphical form (ACMGM075)
- recognise that a sequence generated by a first-order linear recurrence relation can have a long term increasing, decreasing or steady-state solution (ACMGM076)
- use first-order linear recurrence relations to model and analyse (numerically or graphically only) practical problems; for example, investigating the growth of a trout population in a lake recorded at the end of each year and where limited recreational fishing is permitted, or the amount owing on a reducing balance loan after each payment is made. (ACMGM077) AC

4.01 GENERATING SEQUENCES USING RECURRENCE RELATIONS

A **recurrence relation**, also called a **difference equation**, is a rule that specifies a particular **term** in a sequence using the previous term or terms.

Consider the sequence, 3, 8, 13, 18, …

The first term in the sequence is 3 and each subsequent term in the sequence is 5 more than the previous term.

We can write this as $t_{n+1} = t_n + 5$, $t_1 = 3$

This is called a first order difference equation because it links consecutive terms in the sequence.

IMPORTANT

A first order difference equation or recurrence relation is a rule that specifies how a particular term in a sequence can be found from the previous term in the same sequence.

Example 1

Find the first six terms of the sequence defined by the difference equation

a $t_{n+1} = t_n - 4$, $t_1 = 12$

b $t_{n+1} = 3t_n + 2$, $t_1 = 5$

9780170250283

Solution

a The first term is t_1.

$t_1 = 12$

$t_{n+1} = t_n - 4$ means that each term in the sequence is 4 less than the previous term.
The second term is 4 less than the first term.

$t_2 = t_1 - 4$
$= 12 - 4$
$= 8$

The third term is 4 less than the second term.

$t_3 = t_2 - 4$
$= 8 - 4$
$= 4$

Continue this pattern to find all six terms.

The first six terms of the sequence are 12, 8, 4, 0, –4, –8.

TI-Nspire CAS

Open a New Document with a Calculator page.
Type 12 and press enter.
Type Ans – 4. Press ctrl (–) for ans.
Press enter.
Continue to press enter until you have the required number of terms.

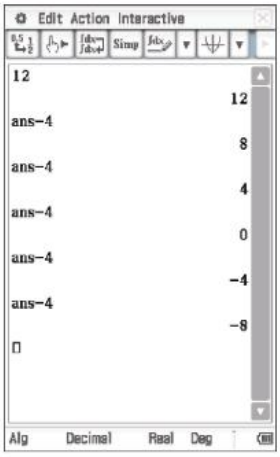




ClassPad

Use the Main $\sqrt{\alpha}$ application.
Type 12 and press EXE.
Press Keyboard and tap Math2 and ans.
Press (–) 4 then EXE.

Continue to press EXE until you have the required number of terms.



b The first term is t_1.

$t_1 = 5$

$t_{n+1} = 3t_n + 2$ means that each term in the sequence is 2 more than 3 times the previous term.

The second term is 2 more than 3 times the first term.

$t_2 = 3t_1 + 2$
$= 3 \times 5 + 2$
$= 17$

The third term is 2 more than 3 times the second term.

$t_3 = 3t_2 + 2$
$= 3 \times 17 + 2$
$= 53$

Continue using the rule until the first six terms are found.

$t_4 = 3t_3 + 2$
$= 3 \times 53 + 2$
$= 161$

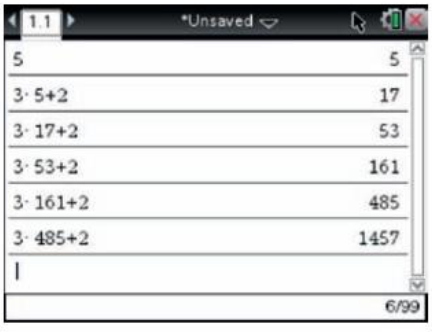
$t_5 = 3t_4 + 2$
$= 3 \times 161 + 2$
$= 485$

$t_6 = 3t_5 + 2$
$= 3 \times 485 + 2$
$= 1457$

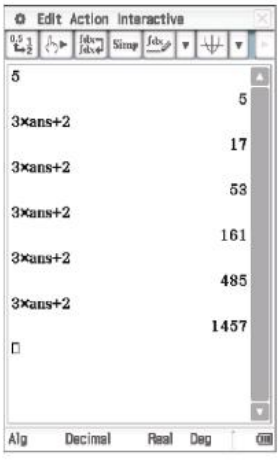
Write down the first six terms of the sequence.

The first six terms of the sequence are 5, 17, 53, 161, 485, 1457.

TI-Nspire CAS



ClassPad



9780170250283

In the difference equation $t_{n+1} = t_n + 5$, where $t_1 = 3$, t_1 represents the first term in the sequence, t_n represents a particular term in the sequence and t_{n+1} represents the term in the sequence that comes immediately after t_n.

The same recurrence relation could also be written as $t_n = t_{n-1} + 5$ where $t_1 = 3$, t_1 represents the first term in the sequence, t_{n-1} represents a particular term in the sequence and t_n represents the term in the sequence that comes immediately after t_{n-1}.

The first term in a sequence is usually denoted by t_1 but can also be denoted by t_0.

Example 2

Find the first three terms for the following recurrence relations.

a $t_n = 5t_{n-1}$, where $t_1 = 2$

b $t_n = 4 - 2t_{n-1}$, where $t_0 = -2$

Solution

a The first term is t_1.

$t_1 = 2$

$t_n = 5t_{n-1}$ means that each term in the sequence is 5 times the previous term.
The second term, t_2, is 5 times the first term.

$t_2 = 5t_1$
$= 5 \times 2$
$= 10$

The third term, t_3, is 5 times the second term.

$t_3 = 5t_2$
$= 5 \times 10$
$= 50$

Write down the first three terms.

The first three terms of the sequence are 2, 10, 50.

b The first term is t_0.

$t_0 = -2$

$t_n = 4 - 2t_{n-1}$ means that each term in the sequence is 4 minus twice the previous term.
The second term, t_1, is 4 minus 2 times the first term.

$t_1 = 4 - 2t_0$
$= 4 - 2 \times (-2)$
$= 8$

The third term, t_2, is 4 minus 2 times the second term.

$t_2 = 4 - 2t_1$
$= 4 - 2 \times 8$
$= -12$

Write down the first three terms.

The first three terms of the sequence are -2, 8, -12.

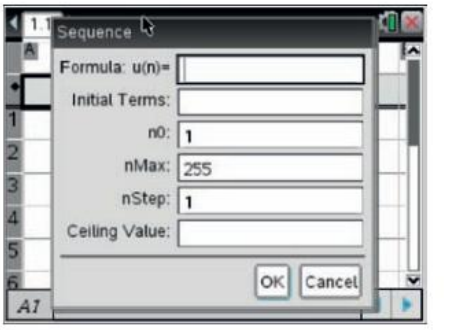


A CAS calculator can be used to generate terms of a sequence from a difference equation.

TI-Nspire CAS

Follow the steps below to find the first five terms of the sequence $t_{n+1} = 2t_n - 3$, where $t_1 = -5$.

Open a New Document with a Lists & Spreadsheet page.
Press menu
3: Data
1: Generate Sequence

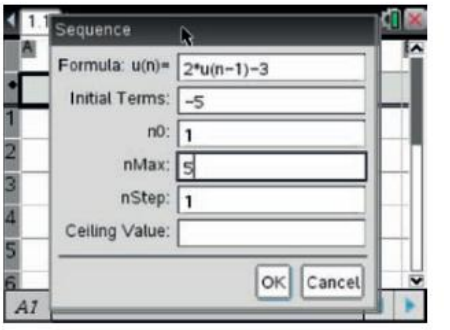


The formula entry line has u(n) =
Rewrite the difference equation in an equivalent form with t_n as the subject.

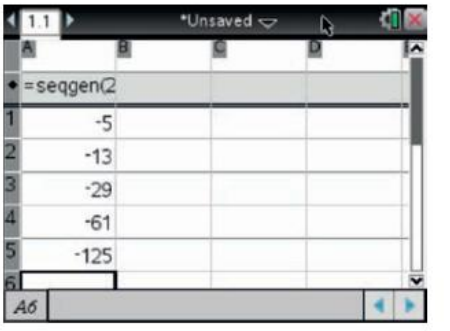
$t_{n+1} = 2t_n - 3,\ t_1 = -5$
$t_n = 2t_{n-1} - 3,\ t_1 = -5$

Enter the right-hand side of the difference equation in the formula entry line. Use u(n–1) to represent t_{n-1} in the rule.
Enter t_1 as the initial term.
Set n0 to 1.
Set nMax to 5.
Set nStep to 1 (this should be the default setting).

To find the first 10 terms of a sequence, for example, set n0 to 1 and set nMax to 10.



Press tab until OK is highlighted, then press enter.

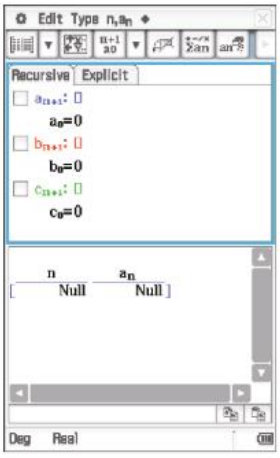


9780170250283

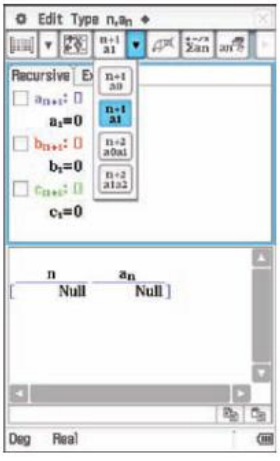
ClassPad

Follow the steps below to find the first five terms of the sequence $t_{n+1} = 2t_n - 3$, where $t_1 = -5$.

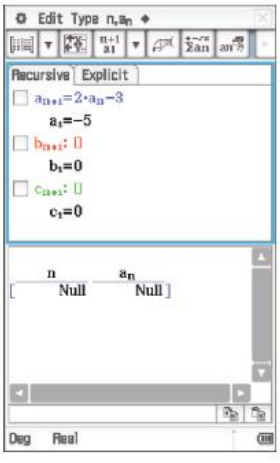
Use the **Sequence** application.
Tap **Recursive**.



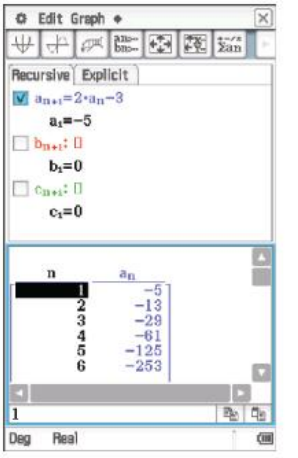
Tap the arrow in the middle and a drop down menu will appear with different templates for difference equations. Tap $\boxed{\frac{n+1}{a1}}$.



The difference equation
$t_{n+1} = 2t_n - 3$, $t_1 = -5$ is equivalent to
$a_{n+1} = 2a_n - 3$, $a_1 = -5$
Type the difference equation in the entry line after $a_{n+1} =$
You do not need to type a_n. Tap **n**, $\mathbf{a}_n$ in the panel at the top of the screen and tap $\mathbf{a}_n$ in the drop down menu.



Make sure that the difference equation is checked, then tap .
The column headed n gives the term numbers in the sequence and the column labelled a_n gives the terms of the sequence.



A difference equation alone is not sufficient to define a sequence. It is also necessary for a term in the sequence to be given. This is usually the first term but, in some cases, another term in the sequence is given. The difference equation can still be used to find other terms in the sequence as required.

Example 3

The difference equation $w_{n+1} = 3w_n - 4$ can be used to generate a sequence. If $w_3 = -1$, find w_5.

Solution

w_3 is known so w_4 can be found using the rule.

$$\begin{aligned} w_4 &= 3w_3 - 4 \\ &= 3 \times (-1) - 4 \\ &= -7 \end{aligned}$$

Find w_5 using w_4.

$$\begin{aligned} w_5 &= 3w_4 - 4 \\ &= 3 \times (-7) - 4 \\ &= -25 \end{aligned}$$



Shutterstock.com/Pedro Nogueira

9780170250283

Example 4

The third term in a sequence is 24. If a recurrence relation that generates the sequence is $u_{n+1} = 2u_n + 6$, find the first term in the sequence.

Solution

The third term in a sequence is 24, so $u_3 = 24$.
Write down the recurrence relation that connects u_3 and u_2.

$$u_{n+1} = 2u_n + 6$$
$$u_3 = 2u_2 + 6$$
$$24 = 2u_2 + 6$$

If u_{n+1} is u_3, then u_n is u_2

Solve to find u_2.

$$24 - 6 = 2u_2$$
$$18 = 2u_2$$
$$\frac{18}{2} = u_2$$
$$u_2 = 9$$

Write down the recurrence relation that connects u_2 and u_1.
Solve to find u_1.

$$u_{n+1} = 2u_n + 6$$
$$u_2 = 2u_1 + 6$$
$$9 = 2u_1 + 6$$
$$9 - 6 = 2u_1$$
$$3 = 2u_1$$
$$u_1 = \frac{3}{2}$$

Answer the question.

The first term in the sequence is $\frac{3}{2}$.

Alternatively, transpose the recurrence relation to make u_n the subject first.

$$u_{n+1} = 2u_n + 6$$
$$u_{n+1} - 6 = 2u_n$$
$$u_n = \frac{u_{n+1} - 6}{2}$$

Write down the recurrence relation that connects u_3 and u_2.
Substitute and simplify to find u_2.

$$u_n = \frac{u_{n+1} - 6}{2}$$
$$u_2 = \frac{u_3 - 6}{2}$$
$$= \frac{24 - 6}{2}$$
$$= 9$$

Write down the recurrence relation that connects u_2 and u_1.
Substitute and simplify to find u_1.

$$u_n = \frac{u_{n+1} - 6}{2}$$
$$u_1 = \frac{u_2 - 6}{2}$$
$$= \frac{9 - 6}{2}$$
$$= \frac{3}{2}$$

Answer the question.

The first term in the sequence is $\frac{3}{2}$.



EXERCISE 4.01 Generating sequences using recurrence relations

Concepts and techniques

1 Example 1 Find the first six terms of the sequences defined by each of the following difference equations.

a $t_{n+1} = t_n - 2,\ t_1 = 15$
b $t_{n+1} = 4t_n,\ t_1 = 2$
c $t_{n+1} = t_n + 7,\ t_1 = -8$
d $t_{n+1} = -0.5t_n,\ t_1 = 64$
e $u_{n+1} = u_n - 5,\ u_1 = 50$
f $w_{n+1} = 2w_n + 1,\ w_1 = 3$
g $K_{n+1} = 3 - 5K_n,\ K_1 = 1$
h $P_{n+1} = 0.5P_n + 2,\ P_1 = 40$

2 Example 2 Use each recurrence relation to generate the first three terms of the sequence it defines.

a $t_n = t_{n-1} + 3$, where $t_0 = 4$
b $t_n = 4t_{n-1}$, where $t_0 = -10$
c $t_n = 10 - 2t_{n-1}$, where $t_0 = 5$
d $P_n = 0.4P_n + 80$, where $P_0 = 100$
e $S_n = 0.5S_{n-1} + 8$, where $S_0 = 18$
f $u_n = 3u_{n-1} - 7$, where $u_0 = 4$

3 Example 3 The difference equation $u_{n+1} = 2u_n + 5$ can be used to generate a sequence. If $w_4 = 9$, find w_6.

4 Example 4 The third term in a sequence is 172. If a recurrence relation that generates the sequence is $u_{n+1} = 3u_n - 2$, find the first term in the sequence.

5 A sequence is defined by the recurrence relation $t_{n+1} = -0.5t_n + 4$. The fourth term in the sequence is −1.5. Find

a the fifth term
b the first term.

6 For the difference equation $d_{n+1} = 6d_n - 4$, if $d_2 = 5$, then d_4 will be equal to

A 6 **B** 12 **C** 26 **D** 152 **E** 5588

Reasoning and communication

7 Consider the second order difference equation $t_{n+2} = t_{n+1} + t_n$, where $n \geq 1$ with $t_1 = 3$ and $t_2 = 7$.

a Explain how the third term of the sequence can be found using the values of the first and second term of the sequence.
b Find the first five terms in the sequence.
c Which term in the sequence first exceeds 500?

8 Find the third and fourth terms of each sequence defined by the following recurrence relations.

a $t_{n+2} = 2t_{n+1} + t_n$, where $n \geq 1$ with $t_1 = 5$ and $t_2 = 8$.
b $t_{n+2} = 4t_{n+1} - 2t_n$, where $n \geq 1$ with $t_1 = 1$ and $t_2 = 3$.
c $t_{n+2} = t_{n+1} + 3t_n$, where $n \geq 1$ with $t_1 = 1$ and $t_2 = 2$.

4.02 FINDING A RECURRENCE RELATION FOR A GIVEN SEQUENCE

Sometimes it is necessary to determine a difference equation or a recurrence relation for a given sequence. In some cases, this can be done by inspection.

Example 5

Find a difference equation that generates the sequence 50, 42, 36, 28, …

Solution

Each term is 8 less than the previous term. Write down a rule that connects consecutive terms in the sequence.

$t_{n+1} = t_n - 8$

The first term needs to be stated in order to complete the difference equation.

$t_{n+1} = t_n - 8$, where $t_1 = 50$.

The rule for a recurrence relation cannot always be found by inspection. In such cases, algebraic techniques need to be used.

IMPORTANT

The rule for a first order recurrence relation takes the form $t_{n+1} = rt_n + d$ with $t_1 = a$, where r, d and a are constants.



Shutterstock.com/rulgsantos

Example 6

The sequence 8, 26, 80, … is generated from a first order recurrence relation. Find this recurrence relation.

Solution

Write down the general rule for a first order recurrence relation.

$t_{n+1} = rt_n + d$ with $t_1 = a$

Write each term in the sequence with its term number.

$t_1 = 8$, $t_2 = 26$, $t_3 = 80$

Since $t_1 = a$, state the value of a.

$t_1 = 8$
$a = 8$

The recurrence relation $t_{n+1} = rt_n + d$ has two unknowns, d and r. Use the three terms to write down two equations in terms of d and r.

$t_2 = rt_1 + d$
$26 = 8r + d$ [1]
$t_3 = rt_2 + d$
$80 = 26r + d$ [2]

Solve simultaneously to find the values of d and r.

[2] − [1]: $54 = 18r$

$\frac{54}{18} = r$

$r = 3$

Substitute $r = 3$ into [1].

$26 = 8 \times 3 + d$
$= 24 + d$
$d = 2$

Write the recurrence relation.

$t_{n+1} = 3t_n + 2$ with $t_1 = 8$



Shutterstock.com/Judy Dietz

9780170250283

EXERCISE 4.02 Finding a recurrence relation for a given sequence

Concepts and techniques

First order recurrence relations

1 Example 5 Find recurrence relations that generate each of the following sequences.

a 4, 8, 12, 16, …
b 3, 6, 12, 24, …
c 100, 50, 25, 12.5, …
d 49, 42, 35, 28, …

2 The sequence 10, 25, 70, … is generated from the recurrence relation $t_{n+1} = rt_n - 5$ with $t_1 = 10$.

a Find the value of r.
b Find the sixth term in the sequence.

3 The recurrence relation $t_n = 2t_{n-1} + d$, where $t_1 = 12$, defines the sequence 12, –6, –42, …

a Find the value of d.
b Find the seventh term of the sequence.

4 The first three terms of the sequence defined by the recurrence relation $t_{n+1} = rt_n + d$ with $t_1 = a$ are 28, 50 and 94. Find the values of r, d and a.

5 For the first order recurrence relation $t_{n+1} = rt_n + d$ with $t_1 = a$, $t_2 = 40$, $t_3 = 32$ and $t_4 = 28$:

a find the values of r and d
b find the value of a
c find the fifth term in the sequence.

6 Example 6 For each of the following, find a first order recurrence relation that defines the sequence.

a 2, 3, 5, …
b 8, –4, 0, …
c 20, 12, 8, …
d 40, –6, 5.5, …
e 1, 7, 31, …

7 The sequence 4, 9, 24, 69, … is generated from the recurrence relation $t_{n+1} = kt_n - 3$, where $t_1 = 4$. The value of k is

A 2 **B** 3 **C** 5 **D** 7 **E** 15

Reasoning and communication

8 The nth term of a sequence is given by the rule $t_n = 3^{n-1} + 5$.

a Find the first three terms in the sequence.
b Find the recurrence relation that defines the sequence.

9 The numbers 5, x and 49 are the first three terms of the sequence defined by the recurrence relation $t_{n+1} = rt_n + 1$, with $t_1 = 5$.

a Find the values of r and x, given that $x > 0$.
b Find the fourth term in the sequence.

10 Carmel breeds canaries. Each year the number of canaries that she owns increases by 15% and she sells 50 canaries. She started with 400 canaries in 2014.

a Write a recurrence relation to model this situation, with C_n being the number of canaries in year n.
b How many canaries will Carmel own in 2016?

4.03 RECURRENCE RELATIONS FOR ARITHMETIC SEQUENCES

An **arithmetic sequence** is one in which the difference between any two consecutive terms in the sequence is constant. The **first term** in the sequence is a and the **common difference** between the terms is d, where $d = t_2 - t_1 = t_3 - t_2 = t_4 - t_3 = \cdots = t_{n+1} - t_n$.

Recall that the general **first order recurrence relation** is $t_{n+1} = rt_n + d$, with $t_1 = a$. For an arithmetic sequence, $r = 1$.

IMPORTANT

The recurrence relation for an arithmetic sequence is $t_{n+1} = t_n + d$, with $t_1 = a$, where d is the common difference and a is the first term.

Example 7

Find a recurrence relation that generates the arithmetic sequence 30, 27, 24, …

Solution

Write down the first term.	$t_1 = 30$
Find the common difference. The terms in the sequence are decreasing by 3.	$d = -3$
Write the recurrence relation.	$t_{n+1} = t_n + d$ with $t_1 = a$ $t_{n+1} = t_n - 3$ with $t_1 = 30$

The general term or nth term of an arithmetic sequence is given by $t_n = a + (n-1)d$.

Example 8

A sequence is defined by the difference equation $t_{n+1} = t_n - 8$, where $t_1 = 4$. Find a rule for the nth term of the sequence.

Solution

Determine the values of a, the first term, and d, the common difference.	$t_1 = 4$ so $a = 4$. $t_{n+1} = t_n - 8$ so $d = -8$.
Substitute the values of a and d into the rule for the nth term of an arithmetic sequence. Expand and simplify.	$t_n = a + (n-1)d$ $= 4 - 8(n-1)$ $= 4 - 8n + 8$ $= 12 - 8n$

9780170250283

Example 9

The seventh term of an arithmetic sequence is 35 and the common difference is 4.

a Find the first term of the sequence.

b Write down the recurrence relation that describes this sequence.

Solution

a The sequence is arithmetic so $t_n = a + (n-1)d$.
Substitute $t_7 = 35$ and $d = 4$.
Solve for a.

$t_n = a + (n-1)d$
$35 = a + (7-1) \times 4$
$35 = a + 6 \times 4$
$35 = a + 24$
$a = 11$

b Write down the recurrence relation for an arithmetic sequence.
Substitute $a = 11$ and $d = 4$.

$t_{n+1} = t_n + d$ with $t_1 = a$.
$t_{n+1} = t_n + 4$ with $t_1 = 11$.



EXERCISE 4.03 Recurrence relations for arithmetic sequences

Concepts and techniques

1 Example 7 Write difference equations that define each of the following arithmetic sequences.

a $-4, -1, 2, 5, \ldots$
b $30, 21, 12, 3, \ldots$
c $1, 3.5, 6, 8.5, \ldots$
d $\frac{1}{4}, \frac{3}{4}, \frac{5}{4}, \frac{7}{4}, \ldots$
e $3.8, 2.7, 1.6, 0.5, \ldots$

2 Which of the following recurrence relations define arithmetic sequences?

a $t_{n+1} = t_n + 5$, where $t_1 = -2$
b $t_{n+1} = 2t_n - 1$, where $t_1 = 10$
c $t_{n+1} = t_n - \frac{1}{2}$, where $t_1 = \frac{2}{3}$
d $t_{n+1} = 8 - t_n$, where $t_1 = -3$
e $t_{n+1} = 0.5t_n$, where $t_1 = 20$
f $t_{n+1} = 3t_n + 3$, where $t_1 = 0$

3 Example 8 A sequence is defined by the recurrence relation $t_{n+1} = t_n + 40$, where $t_1 = 20$. Find a rule for the nth term of the sequence.

4 Find a rule for the nth term of the sequence defined by each of the following difference equations.

a $t_{n+1} = t_n - 3$, $t_1 = 5$
b $S_{n+1} = S_n + 4$, $S_1 = 3$
c $u_n = u_{n-1} - 5$, $u_1 = 15$
d $w_n = w_{n-1} + 10$, $w_1 = 8$

5 A sequence is described by the recurrence relation $t_{n+1} = rt_n + d$, where $t_1 = 1$.
This is an arithmetic sequence for

A $r = -1, d = 0$
B $r = 2, d = 0$
C $r = 1, d = 3$
D $r = -3, d = 1$
E $r = 2, d = -1$

6 Example 9 The fifth term in an arithmetic sequence is 25 and the common difference is −4.
   a Find the first term of the sequence.
   b Write down the recurrence relation that describes this sequence.
   c Write down the first three terms in the sequence.

7 The sixth term in an arithmetic sequence is 8 and the first term is 348.
   a Find the common difference.
   b Write down the recurrence relation that describes this sequence.
   c Write down the first three terms in the sequence.

8 An arithmetic sequence is defined by the rule $t_n = 3n - 2$, $n = 1, 2, 3, \ldots$ The difference equation that defines this sequence is
   A $t_{n+1} = 3t_n - 2$, where $t_1 = 1$
   B $t_{n+1} = 3t_n$, where $t_1 = 1$
   C $t_{n+1} = t_n - 2$, where $t_1 = 1$
   D $t_{n+1} = t_n + 3$, where $t_1 = 1$
   E $t_{n+1} = t_n + 1$, where $t_1 = 1$

9 The nth term in an arithmetic sequence is given by the rule $t_n = 4n - 2$.
   a Write down the first three terms in the sequence.
   b Express the sequence as a difference equation.

Reasoning and communication

10 The tenth term of an arithmetic sequence is 14 and the twenty-third term is 105. Find a recurrence relation that defines this sequence.

11 Daniel is participating in a walkathon to raise money for charity. He receives $5 for walking the first kilometre and $10 for walking each subsequent kilometre.
   a Write down a recurrence relation that gives the sequence representing the total amount Daniel receives for each kilometre he walks.
   b How much does Daniel raise for his charity if he walks 8 kilometres?

4.04 RECURRENCE RELATIONS FOR GEOMETRIC SEQUENCES

A **geometric sequence** is one in which the ratio of any two consecutive terms in the sequence is constant. The **first term** in the sequence is a and the **common ratio** between terms is r, where $r = \frac{t_2}{t_1} = \frac{t_3}{t_2} = \frac{t_4}{t_3} = \cdots = \frac{t_{n+1}}{t_n}$.

Recall that the general first order recurrence relation is $t_{n+1} = rt_n + d$, with $t_1 = a$. For a geometric sequence, $d = 0$.

IMPORTANT

The recurrence relation for a geometric sequence is $t_{n+1} = rt_n$, with $t_1 = a$, where r is the common ratio and a is the first term.

Example 10

Write the geometric sequence 72, –36, 18, –9, … as a recurrence relation.

Solution

Write down the first term. $t_1 = 72$

Find the common ratio. $r = \frac{-36}{72} = \frac{18}{-36} = \frac{-9}{18}$

$= -\frac{1}{2}$

Write the recurrence relation. $t_{n+1} = rt_n$ with $t_1 = a$

$t_{n+1} = -\frac{1}{2}t_n$ with $t_1 = 72$

The general term or nth term of an geometric sequence is $t_n = ar^{n-1}$.

Example 11

Express the geometric sequence defined by the rule $t_n = 4(0.25)^{n-1}$, $n = 1, 2, 3, \ldots$ as a recurrence relation.

Solution

Write down the rule for the nth term of a geometric sequence. $t_n = ar^{n-1}$.
$t_n = 4(0.25)^{n-1}$

Identify the values of a and r from the given rule. $a = 4$ and $r = 0.25$.

Write down the recurrence relation for a geometric sequence. $t_{n+1} = rt_n$ with $t_1 = a$

Substitute the values of a and r. $t_{n+1} = 0.25t_n$ with $t_1 = 4$



Shutterstock.com/Matthew Jacques

Example 12

The fourth term in a geometric sequence is 128 and the common ratio is 2.

a Find the first term of the sequence.

b Write down the difference equation that describes this sequence.

Solution

a The sequence is geometric, so $t_n = ar^{n-1}$

$t_n = ar^{n-1}$

Substitute $t_4 = 128$ and $r = 2$.

$128 = a \times 2^3$

$128 = 8a$

Solve for a.

$a = \frac{128}{8}$

$= 16$

b Write down the difference equation for a geometric sequence.

$t_{n+1} = rt_n$ with $t_1 = a$

Substitute $a = 16$ and $r = 2$.

$t_{n+1} = 2t_n$ with $t_1 = 16$

EXERCISE 4.04 Recurrence relations for geometric sequences

Concepts and techniques

WS

Recurrence relations for arithmetic and geometric sequences

1 Example 10 Write each of the following geometric sequences as recurrence relations.

a 2, 4, 8, …

b 3, 15, 75, …

c 5, –30, 180, …

d 48, –24, 12, …

e 2, 5.6, 15.68, …

f 81, –27, 9, …

2 For the sequence defined by $t_{n+1} = 1.2t_n$, $t_1 = 4.5$, find

a the fourth term in the sequence

b the rule for the nth term of the sequence.

3 State whether the sequences that obey the following recurrence relations are arithmetic, geometric or neither.

a $t_{n+1} = \frac{3}{2}t_n$, where $t_1 = 4$

b $t_{n+1} = t_n - 6$, where $t_1 = 22$

c $t_{n+1} = 5 - t_n$, where $t_1 = 2$

d $t_{n+1} = -0.2t_n + 1.6$, where $t_1 = 100$

e $t_{n+1} = t_n + 8$, where $t_1 = 0.5$

f $t_{n+1} = -1.8t_n$, where $t_1 = 50$

g $t_{n+1} = t_n$, where $t_1 = 1$

4 Example 11 Express the geometric sequence defined by the rule $t_n = 3(2)^{n-1}$, $n = 1, 2, 3, \ldots$ as a recurrence relation.

5 Express the geometric sequences defined by each of the following rules as recurrence relations.

a $t_n = 4(3)^{n-1}$, $n = 1, 2, 3, \ldots$

b $t_n = -2(2)^{n-1}$, $n = 1, 2, 3, \ldots$

c $t_n = 3(0.2)^{n-1}$, $n = 1, 2, 3, \ldots$

d $t_n = 8\left(-\frac{1}{2}\right)^{n-1}$, $n = 1, 2, 3, \ldots$

e $t_n = 5(1.3)^{n-1}$, $n = 1, 2, 3, \ldots$

9780170250283

6 Example 12 A geometric sequence has a common ratio of –3. The fifth term in the sequence is 405.
   a Find the first term in the sequence.
   b Write down the recurrence relation that describes this sequence.

7 The first term of a geometric sequence is 12 and the third term in the sequence is 108.
   a Given that $r < -1$, find the difference equation that defines this sequence.
   b If $r > 1$, find the recurrence relation that defines this sequence.

8 A sequence is defined by the recurrence relation $t_{n+1} = -0.2t_n$, where $t_1 = 20$. The rule that gives t_n, the nth term of the sequence is
   A $t_n = -0.2(20)^{n-1}$ B $t_n = -20(0.2)^{n-1}$ C $t_n = -20(-0.2)^{n-1}$
   D $t_n = 0.2(-20)^{n-1}$ E $t_n = 20(-0.2)^{n-1}$

Reasoning and communication

9 A rubber ball bounces to a height of 8 metres on its first bounce. The height of each subsequent bounce is $\frac{3}{4}$ of the height of the previous bounce.
   a Write a recurrence relation that defines the sequence of bounce heights.
   b On which bounce does the height first fall below 2 metres?

10 The value of an art collection appreciates at an average rate of 8% per year. The collection was first purchased for $20 000.
   a Write a recurrence relation that gives the sequence representing the values of the art collection each year.
   b Find a rule for the appreciation sequence.
   c Find the value of the art collection 10 years after it is purchased.

4.05 GRAPHICAL REPRESENTATION OF FIRST ORDER RECURRENCE RELATIONS

The recurrence relation $t_{n+1} = t_n + d$, with $t_1 = a$, represents an arithmetic sequence with a common difference of d between successive terms and a first term of a.

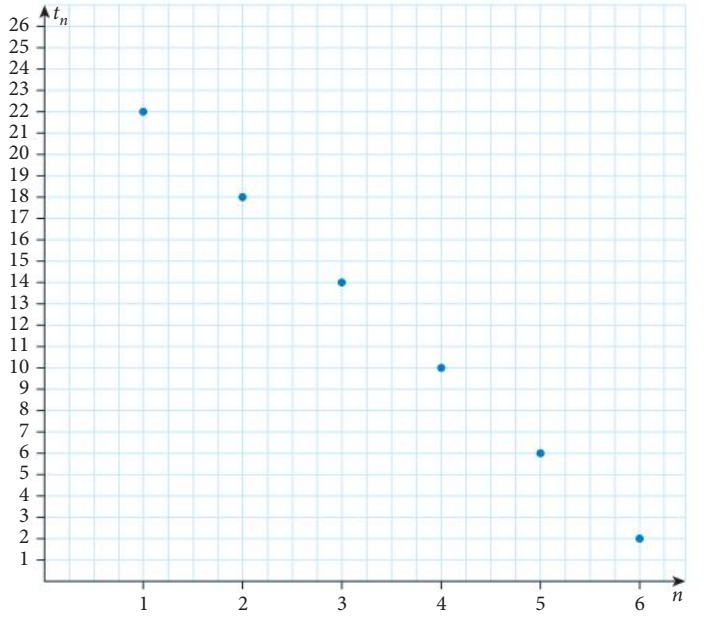
IMPORTANT

When graphed, the terms of sequences generated by arithmetic recurrence relations lie on a straight line with gradient d.

- If $d > 0$, the terms of the sequence lie on a straight line with a positive gradient, so the term values increase as n increases.
- If $d < 0$, the terms of the sequence lie on a straight line with negative gradient, so the term values decrease as n increases.

Example 13

The first six terms of a sequence are graphed below.



Find the recurrence relation that generates the terms of this sequence.

Solution

The points lie on a straight line and therefore can be generated by an arithmetic recurrence relation.	$t_{n+1} = t_n + d$ with $t_1 = a$
The first term in the sequence is 22.	$t_1 = 22$
As n increases by 1, the term value decreases by 4.	$d = -4$
Write the recurrence relation.	$t_{n+1} = t_n - 4$, with $t_1 = 22$

The recurrence relation $t_{n+1} = rt_n$, with $t_1 = a$, represents a geometric sequence with a common ratio of r between successive terms and a first term of a.

IMPORTANT

When graphed, the terms of sequences generated by geometric recurrence relations lie on a curve.

- If $r > 1$, the term values increase as n increases.
- If $0 < r < 1$, the term values decrease as n increases.
- If $r < 0$, the term values alternate in sign.

9780170250283

Example 14

a Plot a graph of the first five terms of the sequence defined by the recurrence relation $t_{n+1} = 2.5t_n$, with $t_1 = 2$.

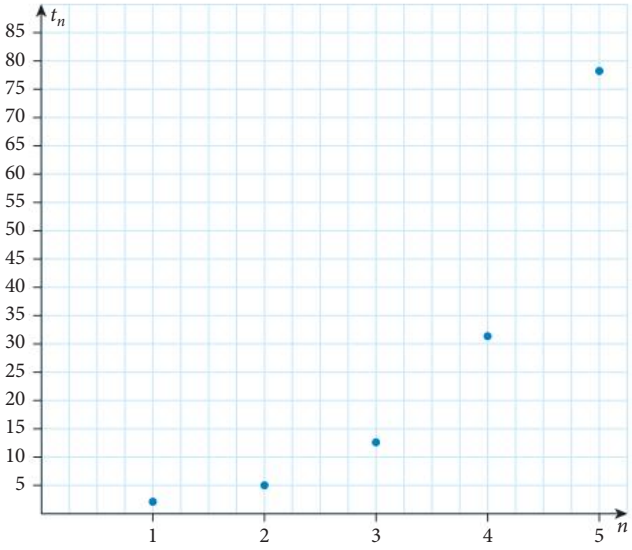
b Does the recurrence relation have a long-term increasing, decreasing or steady-state solution?

Solution

a Find the first five terms of the sequence.

$t_1 = 2$
$t_2 = 2.5 \times 2 = 5$
$t_3 = 2.5 \times 5 = 12.5$
$t_4 = 2.5 \times 12.5 = 31.25$
$t_5 = 2.5 \times 31.25 = 78.125$

Plot the points (1, 2), (2, 5), (3, 12.5), (4, 31.25), (5, 78.125) on a graph.
Label the vertical axis t_n and the horizontal axis n.



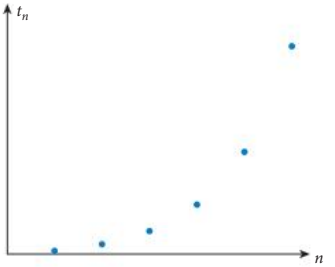
b As n increases, the terms in the sequence increase in value.

The recurrence relation has a long-term increasing solution.

The first order recurrence relation $t_{n+1} = rt_n + d$, where $t_1 = a$, with $r \neq 1$ and $d \neq 0$, represents a sequence that is neither arithmetic nor geometric.

When graphed, the terms in the sequence may increase indefinitely, decrease indefinitely, or approach a particular value in the **long-term** as illustrated in the examples below.

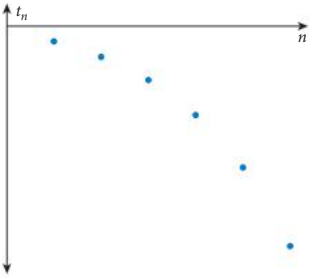
$t_{n+1} = 2t_n + 5, t_1 = 2$



As n increases, t_n increases.

The sequence is said to have a long-term **increasing** solution.

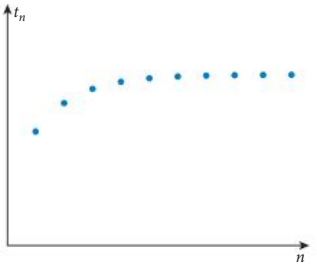
$t_{n+1} = 1.5t_n - 1, t_1 = -2$



As n increases, t_n decreases.

The sequence is said to have a long-term **decreasing** solution.

$t_{n+1} = 0.5t_n + 3, t_1 = 4$



As n increases, t_n approaches a particular value.

The sequence is said to have a long-term **steady-state** solution.

The graphs obtained when the terms of a first order recurrence relation are plotted are not limited to those illustrated in the examples above.

A CAS calculator may be used to plot the terms of a sequence generated by a recurrence relation.

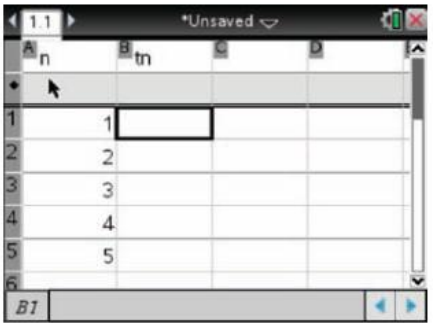
TI-Nspire CAS

Follow the steps below to plot the first five terms of the sequence $t_{n+1} = 2t_n - 3$, where $t_1 = -5$.

Open a New Document with a Lists & Spreadsheet page.
Name Column A, n and Column B, tn.

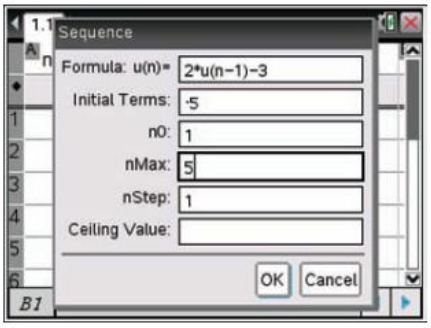


To graph the first five terms of a sequence enter the term numbers 1 to 5 in Column A.



9780170250283

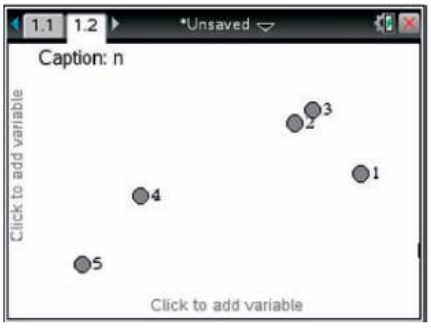
Move the cursor to the formula entry line for Column B.
Press menu
3: Data
1: Generate Sequence
Define the sequence as shown.



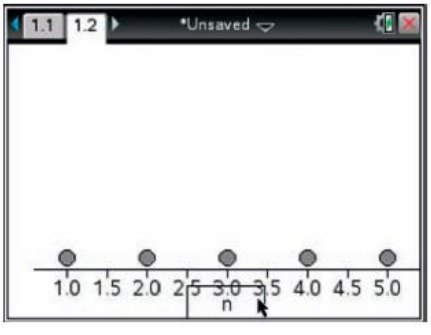
Select OK.



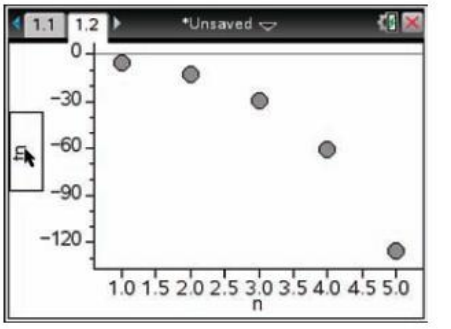
Add a Data & Statistics page.



Click to add variable at the bottom of the screen and select n.



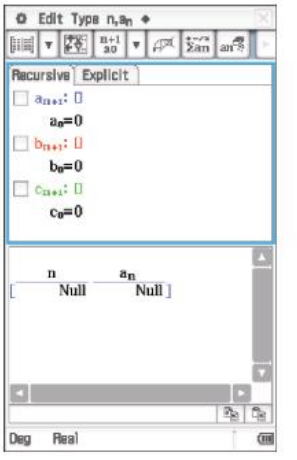
Click to add variable at the left of the screen and select tn.



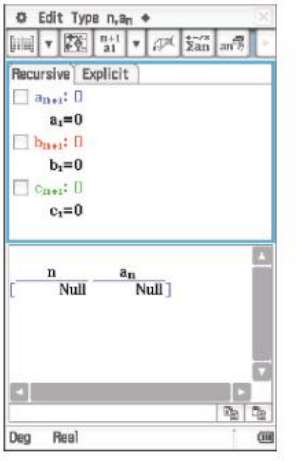
ClassPad

Follow the steps below to plot the first five terms of the sequence $t_{n+1} = 2t_n - 3$, where $t_1 = -5$.

Use the **Sequence** application.
Tap Recursive.

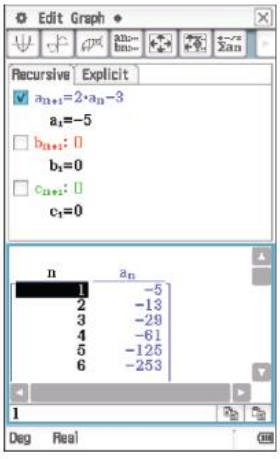


Tap the arrow in the middle of the top tool bar and a drop down menu will appear with different templates for recurrence relations.
Tap $\boxed{\frac{n+1}{a1}}$.

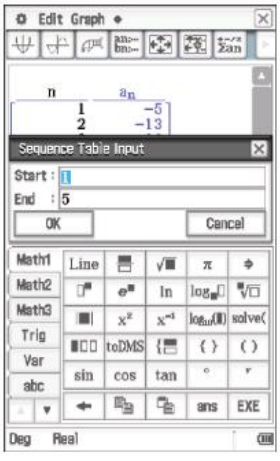


9780170250283

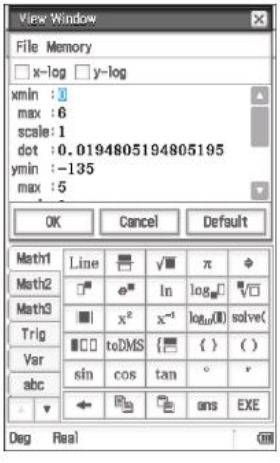
Type the difference equation in the entry line after $a_{n+1} =$.
Make sure that the difference equation is checked, then tap [▤].



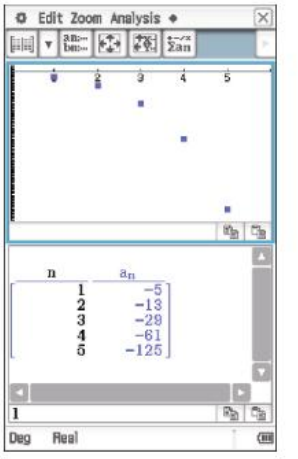
It is necessary to graph the first five terms of the sequence, so tap [⊞] and set **start** to 1 and **End** to 5. Tap [OK].



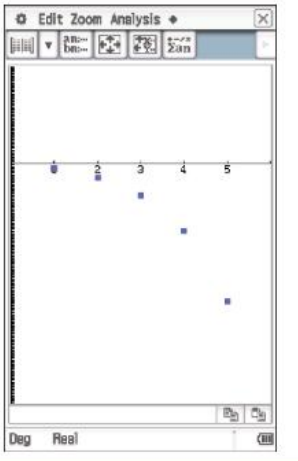
Tap [⊞] to set a suitable viewing window.
Tap [OK].



Tap to graph.

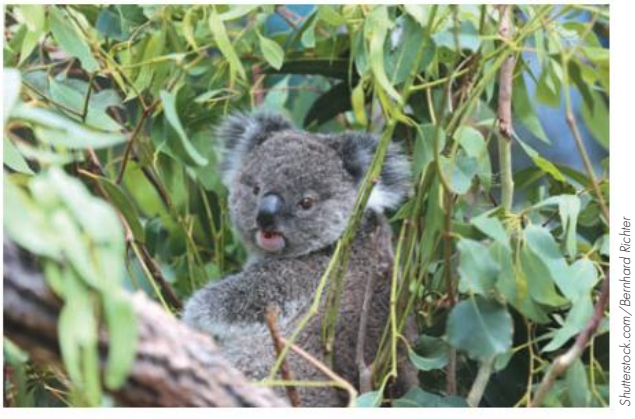


To view only the graph, tap the graph window, then tap Resize.



Example 15

A population of 600 koalas is relocated to a national park. The number of koalas living in a park is modelled by the recurrence relation $k_{n+1} = 1.1k_n - 50$, where $k_1 = 600$. Graph the koala numbers for ten years and comment on the size of the population over time.



Shutterstock.com/Bernhard Richter

9780170250283

Solution

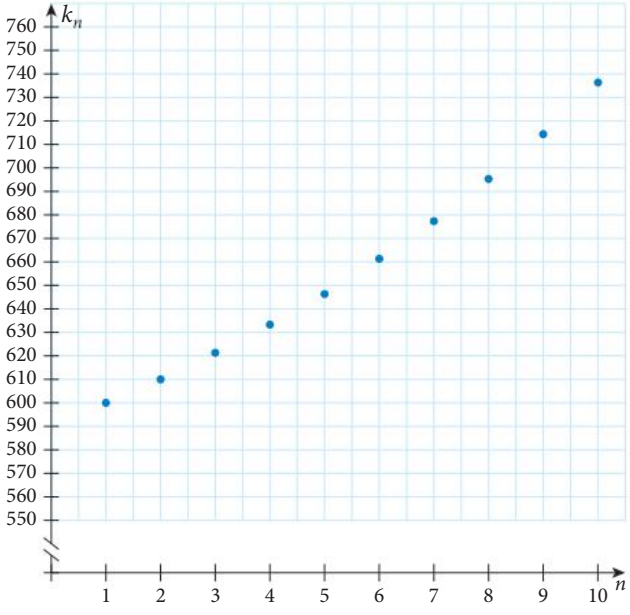
Find the first ten terms in the sequence.

The first ten terms of the sequence are most easily found using a CAS calculator. Refer to section 4.01 for instructions.

n	k_n
1	600
2	610
3	621
4	633
5	646
6	661
7	677
8	695
9	714
10	736

Koala numbers are given correct to the nearest whole number.

Plot a graph of k_n against n.



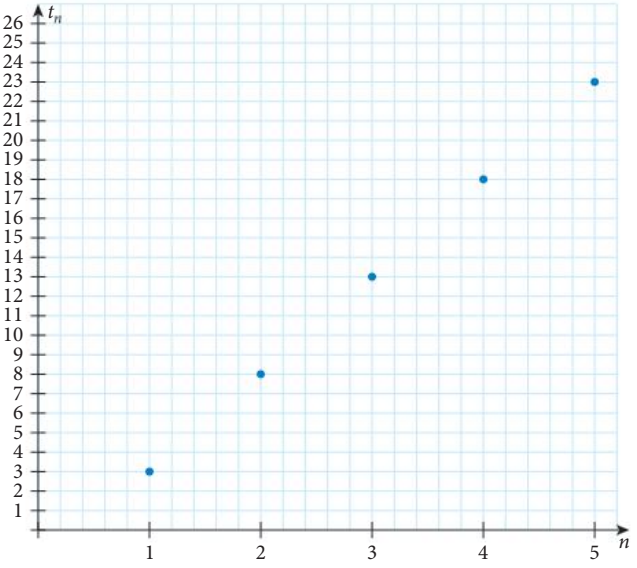
The terms of the sequence increase as n increases.

The size of the koala population increases indefinitely over time.

EXERCISE 4.05 Graphical representation of first order recurrence relations

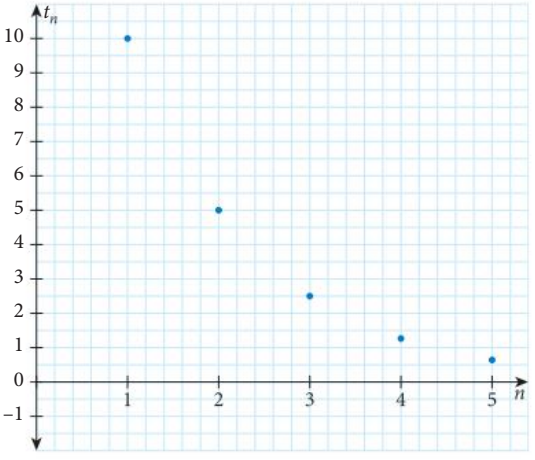
Concepts and techniques

1 Example 13 The first five terms of a sequence are graphed below.



Find the recurrence relation that generates the terms of the sequence.

2 The first five terms of a sequence are shown in the graph below.



Which one of the following difference equations describes the sequence?

A $t_{n+1} = t_n - 5, t_1 = 10$
B $t_{n+1} = t_n - 0.5, t_1 = 10$
C $t_{n+1} = -0.5t_n, t_1 = 10$
D $t_{n+1} = 0.5t_n, t_1 = 10$
E $t_{n+1} = 0.5t_n, t_1 = 20$

9780170250283

3 Example 14 **a** Plot a graph of the first six terms of the sequence defined by the recurrence relation $t_{n+1} = 0.2t_n + 8$, with $t_1 = 4$.

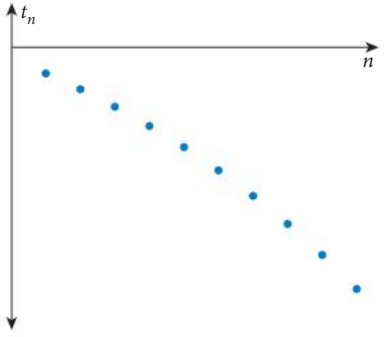
b Does the recurrence relation have a long-term increasing, decreasing or steady-state solution?

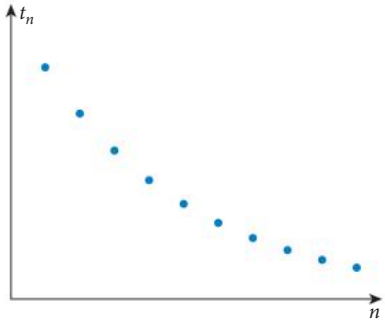
4 The fourth term in a geometric sequence is 54 and the common ratio is 3.

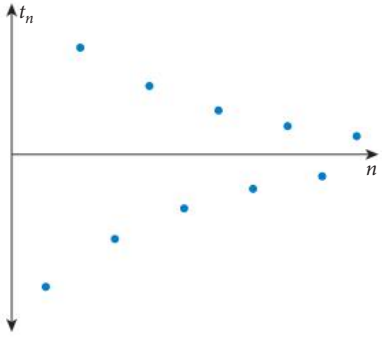
a Find the first term in the sequence.

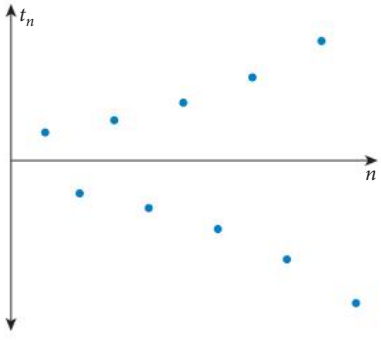
b Graph the first five terms.

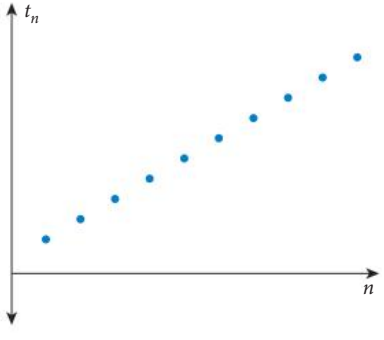
5 For the difference equation $t_{n+1} = -1.2t_n$, where $t_1 = k$, a graph that could be used to represent the terms of this sequence is

A 

B 

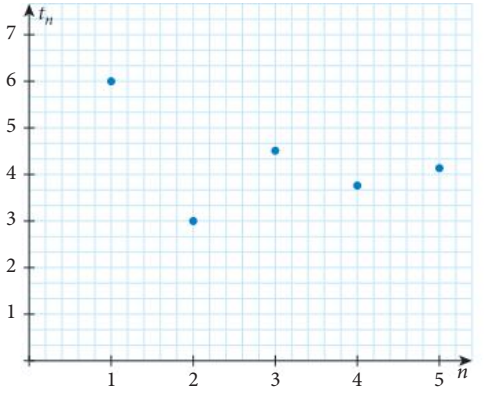
C 

D 

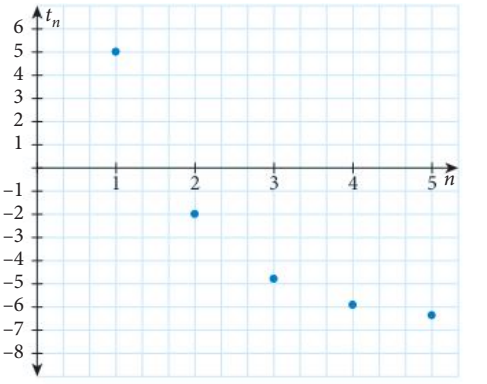
E 

6 Match each of the following graphs **a** to **f** with one of the recurrence relations **A** to **F**.

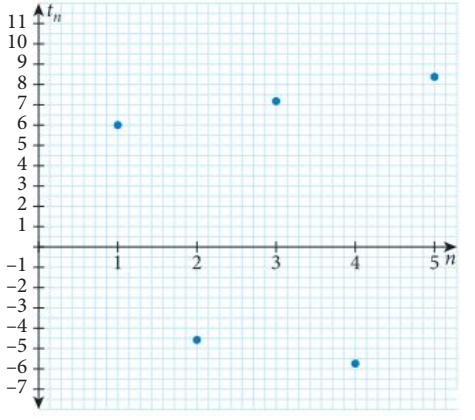
a



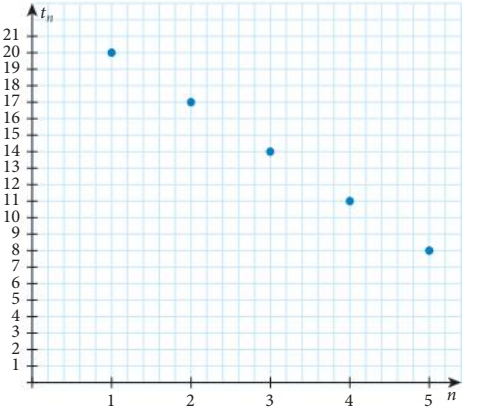
b



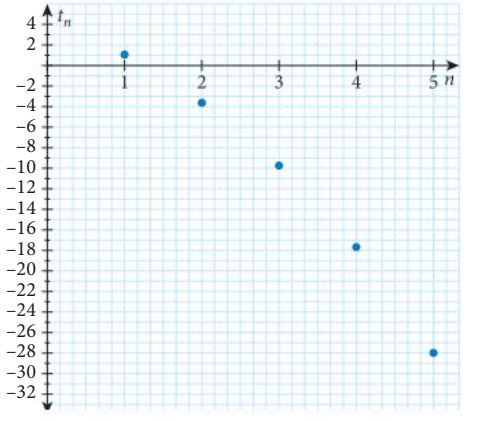
c



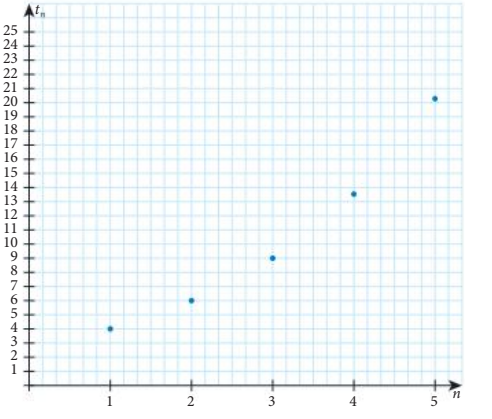
d



e



f



A $t_{n+1} = t_n - 3$, where $t_1 = 20$

B $t_{n+1} = 1.5t_n$, where $t_1 = 4$

C $t_{n+1} = -0.5t_n + 6$, where $t_1 = 6$

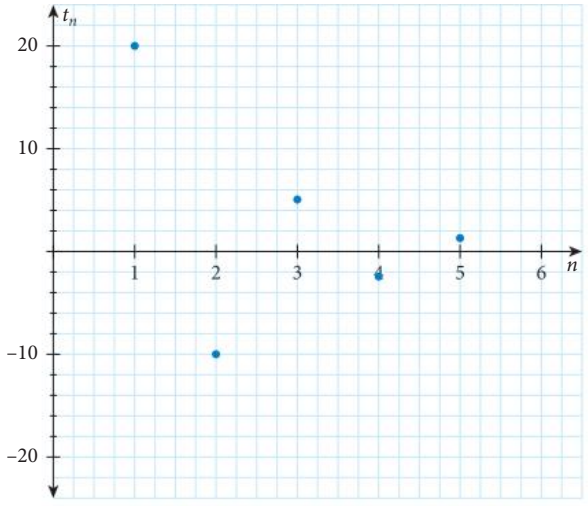
D $t_{n+1} = 1.3t_n - 5$, where $t_1 = 1$

E $t_{n+1} = 0.4t_n - 4$, where $t_1 = 5$

F $t_{n+1} = -1.1t_n + 2$, where $t_1 = 6$

9780170250283

7 The graph below shows the first five terms of a sequence.



A recurrence relation that generates the terms of this sequence is

A $t_{n+1} = t_n - 30$, where $t_1 = 20$

B $t_{n+1} = -2t_n$, where $t_1 = 20$

C $t_{n+1} = \frac{1}{2}t_n$, where $t_1 = 20$

D $t_{n+1} = -\frac{1}{2}t_n$, where $t_1 = 20$

E $t_{n+1} = 10 - t_n$, where $t_1 = 20$

8 The sequence defined by which of the following recurrence relations has a long-term decreasing solution?

A $t_{n+1} = t_n + 3$, where $t_1 = 10$

B $t_{n+1} = 3t_n$, where $t_1 = 10$

C $t_{n+1} = t_n - 4$, where $t_1 = 10$

D $t_{n+1} = 0.1t_n + 9$, where $t_1 = 10$

E $t_{n+1} = 5t_n + 4$, where $t_1 = 10$

Reasoning and communication

9 For each of the following, view a graph displaying the first ten terms of the sequence defined by the difference equation and hence identify the difference equations that generate sequences with a long-term steady-state solution.

a $t_{n+1} = 1.4t_n - 5$, where $t_1 = -5$

b $t_{n+1} = -0.4t_n + 9$, where $t_1 = 10$

c $t_{n+1} = 0.5t_n + 4$, where $t_1 = 4$

d $t_{n+1} = -\frac{3}{2}t_n + 6$, where $t_1 = 2$

e $t_{n+1} = \frac{3}{4}t_n - 3$, where $t_1 = 1$

f $t_{n+1} = 0.8t_n - 4$, where $t_1 = -2$

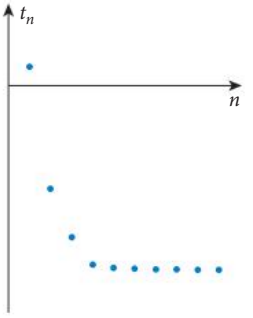
10 By referring to your answers for question **9**, what do you notice about the value of r in the first order difference equation that generates a sequence with a long-term steady-state solution?

11 Example 15 The number of tonnes of pollutant in a lake each year after a clean-up operation commences is modelled by the recurrence relation $t_{n+1} = 0.8t_n + 3$, where $t_1 = 25$. Graph the amount of pollutant in the lake for ten years and comment on the amount of pollutant in the lake over time.

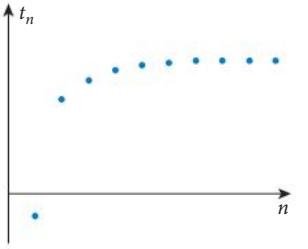
4.06 LONG-TERM STEADY-STATE SOLUTION OF A RECURRENCE RELATION

When $-1 < r < 1$, the term value in a sequence defined by a recurrence relation of the form $t_{n+1} = rt_n + d$, where $t_1 = a$, approaches a particular value as n becomes larger. This can be observed graphically. Some examples are given below.

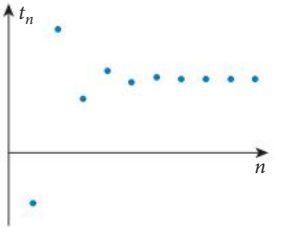
$t_{n+1} = 0.4t_n - 8$, where $t_1 = 2$.



$t_{n+1} = 0.5t_n + 3$, where $t_1 = -1$.



$t_{n+1} = -0.4t_n + 6$, where $t_1 = -3$.



The value that the terms of a sequence approach in the long run (as n becomes very large) is referred to as the **steady-state solution** of the sequence.

IMPORTANT

To find the steady-state solution of a sequence, we assume that the term value ceases to change in the long run. That is, in the long run, $t_{n+1} = t_n$.



Shutterstock.com/luminaimages

9780170250283

Example 16

A sequence is defined by the recurrence relation $t_{n+1} = 0.4t_n + 18$, where $t_1 = 4$. What value do the terms in this sequence approach in the long run?

Solution

Since $r = 0.4$, the sequence defined by the recurrence relation has a long-term steady-state solution.	$t_{n+1} = 0.4t_n + 18$
For a steady state, $t_{n+1} = t_n$.	Let $t_{n+1} = t_n = x$
Substitute x for t_{n+1} and t_n in the recurrence relation.	$x = 0.4x + 18$
Solve for x.	$x - 0.4x = 18$ $0.6x = 18$ $x = \frac{18}{0.6}$ $= 30$
Answer the question.	In the long run, the terms in the sequence approach the value of 30.



Example 17

The sequence defined by the recurrence relation $u_{n+1} = ku_n + 12$, where $u_1 = 10$, has a steady-state solution of 37.5. Find the value of k.

Solution

Write the recurrence relation.	$u_{n+1} = ku_n + 12$
In the long term, $u_{n+1} = u_n$. In this case, the long-term value is 37.5.	Let $u_{n+1} = u_n = 37.5$ $37.5 = 37.5k + 12$
Solve for k.	$37.5 - 12 = 37.5k$ $25.5 = 37.5k$ $k = \frac{25.5}{37.5}$ $= \frac{17}{25}$

EXERCISE 4.06 Long-term steady-state solution of a recurrence relation

Graphical representation of a sequence and long-term solution of a recurrence relation

Concepts and techniques

1 Example 16 What values do the terms in each of the following sequences take in the long run?

a $t_{n+1} = 0.6t_n + 24$, where $t_1 = 8$

b $t_{n+1} = 0.5t_n - 32$, where $t_1 = 1$

c $u_{n+1} = 0.3u_n + 28$, where $u_1 = 4$

d $w_{n+1} = -0.2w_n + 36$, where $w_1 = 20$

e $v_{n+1} = -0.8v_n - 54$, where $v_1 = -4$

2 A sequence is defined by the recurrence relation $t_{n+1} = 0.9t_n + 4$, where $t_1 = 5$.

a Find the value of t_3.

b What is the smallest value of n for which $t_n > 20$?

c In the long run, what value will the terms in the sequence take?

3 Which one of the following recurrence relations has a long-term steady-state solution?

A $t_{n+1} = 1.5t_n - 3$ with $t_1 = 0.4$

B $t_{n+1} = -2t_n + 1$ with $t_1 = 3$

C $t_{n+1} = -1.7t_n - 14$ with $t_1 = 10$

D $t_{n+1} = \frac{5}{4}t_n + \frac{1}{2}$ with $t_1 = 20$

E $t_{n+1} = \frac{3}{4}t_n - 9$ with $t_1 = 8$

4 Example 17 The sequence defined by the recurrence relation $v_{n+1} = kv_n + 20$, where $v_1 = 5$, approaches a long-term value of 25. Find the value of k.

Reasoning and communication

5 Two sequences are defined by the recurrence relations $t_{n+1} = 2t_n - 0.3$ with $t_1 = 2$, and $u_{n+1} = 0.6u_n + 5$ with $u_1 = 2$.

a Explain why only one of these recurrence relations has a long-term steady-state solution.

b Find the long-term value for this sequence.

6 **a** Find a rule, in terms of r and d, that gives the long-term value of the sequence defined by the recurrence relation $t_{n+1} = rt_n + d$.

b Use the rule found in **a** to find the long-term value of the sequence

i $t_{n+1} = 0.7t_n + 6$ **ii** $u_{n+1} = -0.5u_n - 8$

7 A sequence is defined by the recurrence relation $t_{n+1} = ct_n + 5c$, where c is a real constant. If the long-term steady-state solution of the sequence is 7.5, find the value of c.

8 The terms in the sequences defined by the recurrence relations $t_{n+1} = 0.7t_n + 27$ with $t_1 = 5$, and $w_{n+1} = 0.9w_n + k$ with $w_1 = 10$, approach the same value in the long term. Find the value of k.

9 A deer population in a wildlife reserve initially numbers 200 and is estimated to drop in size by 8% each year. Each year, an extra 20 deer are introduced to the reserve. What will be the long-term size of the deer population?

9780170250283

4.07 APPLICATIONS OF RECURRENCE RELATIONS

Recurrence relations can be used to model many real-life situations, including compound interest, loans and investments, depreciation, drug doses and population growth and decay.

When writing recurrence relations, it is important to identify or calculate the value of the first term and to determine the relationship between successive terms in the sequence.

Example 18

A business purchased a photocopier for $10 000 in 2014. It depreciated at a rate of 18% per annum.

a Write down a recurrence relation that gives the value of the photocopier at the start of each year.

b The photocopier is replaced when its value first falls below $3000. At the start of which year will the copier be replaced?

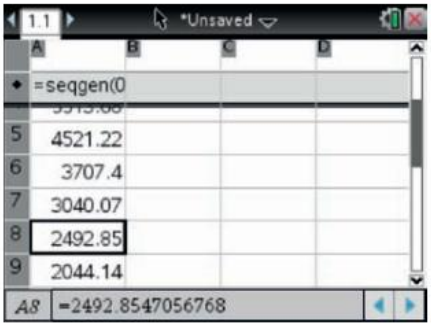
Solution

a If the photocopier depreciates by 18% per annum, this means that it loses 18% of its value from one year to the next. Thus, the value of the photocopier in any one year is 82% of its value the previous year.
Write this as a recurrence relation and include the initial value of the photocopier as the first term.

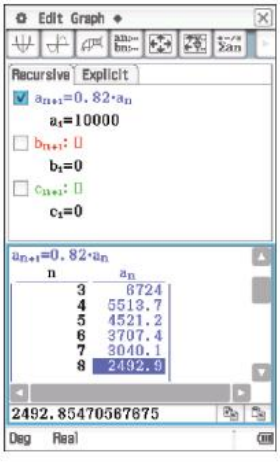
$V_{n+1} = 0.82V_n$ where $V_1 = 10\,000$.

b Use technology to generate the sequence of values from the recurrence relation.

TI-Nspire



ClassPad



Identify the term in the sequence that gives the first value that is less than \$3000. Establish which year this term corresponds to. The value of the photocopier at the start of 2014 is V_1 so V_8 is the value of the photocopier at the start of 2022.	$V_8 = 2492.85$
Answer the question.	The photocopier is replaced at the start of 2022.

Example 19

Samuel takes a walk, each day on his holiday. The distances he walks on the first three days are shown in the table below.

Day	1	2	3
Distance (km)	2	3	4.4

a Show that the sequence of distances walked is neither arithmetic nor geometric.

b A difference equation that models the distance walked on the nth day of Samuel's holiday is

$$D_{n+1} = rD_n + d, \text{ where } D_1 = k.$$

Find the values of r, d and k.

c What distance will Samuel walk, on the 5th day of his holiday? Answer in kilometres, correct to one decimal place.

Solution

a

An arithmetic sequence has a common difference between any two consecutive terms.
Establish that there is not a common difference.

$$D_2 - D_1 = 3 - 2$$
$$= 1$$
$$D_3 - D_2 = 4.4 - 3$$
$$= 1.4$$
$D_2 - D_1 \neq D_3 - D_2$ so the sequence of distances is not arithmetic.

A geometric sequence has a common ratio between any two consecutive terms.
Establish that there is not a common ratio.

$$\frac{D_2}{D_1} = \frac{3}{2}$$
$$= 1.5$$
$$\frac{D_3}{D_2} = \frac{4.4}{3}$$
$$= 1.466\ldots$$
$\frac{D_2}{D_1} \neq \frac{D_3}{D_2}$ so the sequence of distances is not geometric.

9780170250283

b $D_1 = k$, so k is the value of the first term.

$k = 2$

The difference equation $D_{n+1} = rD_n + d$ has two unknowns, d and r. Use the three terms to write down two equations in terms of d and r.

$D_2 = rD_1 + d$
$3 = 2r + d$ [1]
$D_3 = rD_2 + d$
$4.4 = 3r + d$ [2]

Solve simultaneously to find the values of d and r.

[2] – [1]: $1.4 = r$
Substitute $r = 1.4$ into [1].
$3 = 2 \times 1.4 + d$
$3 = 2.8 + d$
$d = 0.2$

c Write the difference equation that defines the distance Samuel walks each day.

$D_{n+1} = 1.4D_n + 0.2$

Start with $D_3 = 4.4$ and find D_4.

$D_4 = 1.4D_3 + 0.2$
$= 1.4 \times 4.4 + 0.2$
$= 6.36$

Use the value of D_4 to find D_5.

$D_5 = 1.4D_4 + 0.2$
$= 1.4 \times 4.96 + 0.2$
$= 9.104$

Answer the question, giving the distance Samuel walks on the fifth day, correct to one decimal place.

On the fifth day, Samuel walks 9.1 km.



Example 20

Marion is injected with 60 mL of a pain killing drug. Every four hours, the amount of drug in her bloodstream is reduced by 40%. To compensate for this loss, she is given an extra 15 mL of the drug every four hours.

a Find a recurrence relation for the amount of drug in Marion's bloodstream when she is given her nth dose.

b How much of the drug is present in her bloodstream 12 hours after the initial dose is administered?

c Eventually, the amount of pain killing drug in Marion's bloodstream stabilises. In the long run, how many mL of pain killer will be present in Marion's bloodstream after she receives her dose of pain killing drug?

Solution

a Write the general rule for a first order recurrence relation. Let A_n be the amount of pain killing drug in Marion's bloodstream when she is given her nth dose.

$A_{n+1} = rA_n + d$ where $A_1 = a$.

Each hour, the amount of pain killing drug in her bloodstream is reduced by 40% so 60% remains.

$r = 0.6$

Each hour, she receives an additional dose of 15 mL.

$d = 15$

Her initial dose is 60 mL.

$A_1 = 60$

Write the recurrence relation.

$A_{n+1} = 0.6A_n + 15$, where $A_1 = 60$

b Marion receives a dose of pain killing drug every four hours. After 12 hours, she will have received four doses, her initial dose of 60 mL and then three subsequent doses of 15 mL.

$A_{n+1} = 0.6A_n + 15$

Find A_4.

$$\begin{aligned} A_2 &= 0.6A_1 + 15 \\ &= 0.6 \times 60 + 15 \\ &= 51 \end{aligned}$$

$$\begin{aligned} A_3 &= 0.6A_2 + 15 \\ &= 0.6 \times 60 + 15 \\ &= 45.6 \end{aligned}$$

$$\begin{aligned} A_4 &= 0.6A_3 + 15 \\ &= 0.6 \times 45.6 + 15 \\ &= 42.36 \end{aligned}$$

Answer the question.

After 12 hours, Marion has 42.36 mL of pain killer in her bloodstream.

c In the long term, the recurrence relation has a steady-state solution as indicated by the statement 'Eventually, the amount of pain killing drug in Marion's bloodstream stabilises.'
For the steady-state solution, $A_{n+1} = A_n$.

Let $A_{n+1} = A_n = x$

$$\begin{aligned} A_{n+1} &= 0.6A_n + 15 \\ x &= 0.6x + 15 \\ x - 0.6x &= 15 \\ 0.4x &= 15 \\ x &= \frac{15}{0.4} \\ &= 37.5 \end{aligned}$$

Answer the question.

In the long run, Marion will have 37.5 mL of the drug in her bloodstream when she receives her dose of pain killing drug.

9780170250283

EXERCISE 4.07 Applications of recurrence relations

Concepts and techniques

Applications of recurrence relations

1 Write a recurrence relation to represent each of the following situations.
   a Elly currently receives an annual salary of $40 000. At the start of each year, she receives a 4% pay rise.
   b A frog jumps 48 cm in its first jump and the length of each subsequent jump is 80% of its previous jump length.
   c Carla parks her car in a car park. She is charged $12 for the first hour and an additional $8 for each subsequent hour.
   d A machine is purchased for $20 000. It depreciates by 15% each year.
   e A photocopier is depreciated by $0.05 for each copy it makes and it makes 200 000 copies per year. It was purchased for $68 000.
   f A rain water tank contains 5000 litres. Each day 12% of the volume of water in the tank is used and in the evening an extra 200 litres of water is added to the tank.

2 As part of a program to improve her fitness, Kelly walks 1 kilometre on the first day. She plans to increase the distance she walks by 0.2 kilometres each day.
   a Write down a recurrence relation that describes the distance that Kelly walks each day.
   b How far with Kelly walk on the 20th day?

3 Example 18 A company car is purchased for $80 000 at the start of 2012. Each year the value of the car depreciates by 14%.
   a Write down a recurrence relation that gives the value of the car at the start of each year.
   b The car is replaced when its value first falls below $50 000. At the start of which year will the car be replaced?

4 A crop of sweet corn is estimated to contain 45 000 kg of corn. Each day, 2500 kg of corn is picked from the plants and sent to market.

   The number of kilograms of corn remaining unpicked n days after picking commences, S_n, can be modelled by the recurrence relation $S_{n+1} = S_n + d$, where $S_0 = a$.
   a Write down the values of d and a.
   b How many kilograms of corn remain unpicked five days after picking commences?
   c How many days will it take to pick all of the sweet corn from the plants?

5 Example 19 A council employs contract gardeners to plant and maintain public gardens in a new housing estate. In the first month, the gardeners work for 800 hours, in the second month, they work for 680 hours and in the third month, they work for 584 hours. This decreasing pattern of work continues.
   a Show that the sequence of hours worked is neither arithmetic nor geometric.
   b Write a recurrence relation that models the number of hours the gardeners will work in the nth month.
   c How many hours will the gardeners work for in the sixth month? Answer correct to one decimal place.
   d How many more hours will the gardeners work for in the sixth month than the seventh month?

6 At the start of each year, the number of antelope living in an African wildlife reserve is given by the rule $A_{n+1} = 0.8A_n + 600$.

From this rule, it can be concluded that each year

A 80% of the antelope living in the reserve die and 600 new antelope are introduced
B 20% of the antelope living in the reserve die and 600 new antelope are introduced
C 600 antelope are lost from the reserve each year
D 80% of the 600 antelope living in the reserve each year continue to live in the reserve
E 20% of the 600 antelope living in the reserve each year continue to live in the reserve.

7 Warm Pools Solar is a company that produces solar heating systems for pools. The number of solar heating systems sold in the *n*th year of business is generated by the recurrence relation $S_n = 1.4S_{n-1} - 300$, where $n \leq 5$ and $S_3 = 3200$.

a How many solar pool heating systems were sold in the first year?
b How many solar pool heating systems were sold in
  i the fifth year? **ii** total over five years?
c What percentage of solar pool heating systems sold in the first four years were sold in the first year? Answer correct to two decimal places.

8 A particular species of flightless bird lives only on Waddle Island. An introduced species infects the birds with a disease. The number, B_n, of infected birds on the island in week n obeys the difference equation $B_{n+1} = 2B_n - 12$, for $n = 1, 2, 3, \ldots$, where $B_1 = a$.

a The number of infected birds on the island in week 4 is 44. How many birds were infected in the first week?
b How many birds are infected in week 7?
c Is the sequence generated by the rule for B_n arithmetic, geometric or neither? Justify your answer.

9 Park authorities need to control the size of a kangaroo population in a national park. Each year there is a 15% increase in the size of the population due to breeding and at the end of each year 300 kangaroos are relocated out of the park. The size of the kangaroo population at the start of 2012 was 2000.

a The number of kangaroos living in the park at the start of 2014 is
  A 0 **B** 600 **C** 1600 **D** 2000 **E** 2645
b A recurrence relation that could be used to model the size of the kangaroo population is
  A $k_{n+1} = 1.15k_n - 300$, where $k_1 = 2000$
  B $k_{n+1} = 1.15(k_n - 300)$, where $k_1 = 2000$
  C $k_{n+1} = 0.15k_n - 300$, where $k_1 = 2000$
  D $k_{n+1} = 0.15(k_n - 300)$, where $k_1 = 2000$
  E $k_{n+1} = 0.85k_n - 300$, where $k_1 = 2000$

10 At the start of 2014, Hugh's salary was \$60 000. Each year, his salary is increased by 4% of his previous year's salary and also by \$5000, according to an agreement he has with his employer.

a Use a recurrence relation to model this situation.
b What is Hugh's annual salary for the year 2017?
c In which year will Hugh's salary first exceed \$100 000?

11 Daniel wins $200 000 in the lottery. He invests this sum in an account that earns 0.5% interest per month, compounded monthly. Each month, he withdraws $800 from the account and reinvests the amount remaining.

a Find a recurrence relation that models this situation.

b How much is Daniel's investment worth after 1 year?

c After how many months will the amount of interest that Daniel has earned on his investment first exceed $3000?

12 Example 20 A vegetable farmer uses water from a storage tank to water his crops. The tank initially contains 50 000 litres. Each day, 10% of the water in the tank is used to water the crops and each evening an extra 3000 litres is pumped into the tank from a nearby river. This pattern continues each day.

The number of litres of water, L_n, in the tank of the nth morning is modelled by the difference equation

$$L_{n+1} = rL_n + d, \text{ where } L_1 = k.$$

a Write down the values of r, d and k.

b Determine how many litres of water will be in the tank on the fifth morning.

c On which morning will the volume in the tank first fall below 40 000 litres?

d How many litres of water will be contained in the tank in the long term?

13 A free range chicken farmer has 200 hens. Each month, 20% of the chickens are attacked and killed by foxes and the farmer buys an additional 32 hens to replace the lost stock.

a Find a recurrence relation that models this situation.

b If this pattern continues, how many hens will the farmer have in the long run?

14 A hedge is planted around the border of a country property. Each year, the hedge grows by an estimated 80 cm and at the end of the year it is trimmed to 60% of its total height. When the hedge is first planted, it is 50 cm tall.

a Find a recurrence relation that gives the height of the hedge at the start of the nth year.

b How tall is the hedge at the start of the third year after planting?

c Once the hedge reaches a certain height, this height is maintained. What is this height? Answer correct to the nearest centimetre.



Shutterstock.com/Imagenet

INVESTIGATION Investing money

At the start of the year, Madeleine invests $250 000 in an account that pays 6% per annum with interest compounded yearly.

a Write a recurrence relation of the form $v_{n+1} = rv_n$, where $v_1 = a$, that gives the value of Madeleine's investment at the start of each year.

b Use the recurrence relation to complete the following table.

Year	1	2	3	4	5
Value of investment					

c Plot a graph that shows the value of Madeleine's investment at the start of each of the first five years.

d Describe the long-term value of Madeleine's investment.

Instead of leaving her investment untouched, at the end of each year, after interest is paid into the account, Madeleine withdraws $12 000 from her investment.

e Write a recurrence relation of the form $v_{n+1} = rv_n + d$, where $v_1 = a$, that gives the value of Madeleine's investment at the start of each year.

f Use the recurrence relation to complete the following table.

Year	1	2	3	4	5
Value of investment					

g Plot a graph that shows the value of Madeleine's investment at the start of each of the first five years.

h Describe the long-term value of Madeleine's investment. Compare this with the long-term value of her investment when no withdrawals were made from the account.

Madeleine withdraws $20 000 at the end of each year, after interest is paid, instead of withdrawing $12 000.

i Write a recurrence relation of the form $v_{n+1} = rv_n + d$, where $v_1 = a$, that gives the value of Madeleine's investment at the start of each year.

j Use the recurrence relation to complete the following table.

Year	1	2	3	4	5
Value of investment					

k Plot a graph that shows the value of Madeleine's investment at the start of each of the first five years.

l What does the graph suggest about the value of Madeleine's investment over time?

m Use a suitable table of values to determine what happens to the value of Madeleine's investment over 24 or so years.

n How much can Madeleine withdraw from her account after interest is paid each year and keep the balance constant?

o Write a recurrence relation that models this situation.

9780170250283

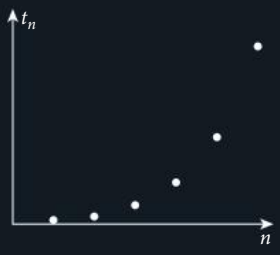
CHAPTER SUMMARY
RECURRENCE RELATIONS

4

- A recurrence relation, also called a difference equation, is a rule that specifies a particular term in a sequence using the previous term or terms.
- A first order recurrence relation is a rule which specifies how a particular term in a sequence can be found from the previous term in the same sequence.
- A recurrence relation alone is not sufficient for defining a sequence. It is also necessary for a term in the sequence to be given. This is usually the first term but, in some cases, another term in the sequence is given.
- The rule for a first order recurrence relation takes the form $t_{n+1} = rt_n + d$ with $t_1 = a$, where r, d and a are constants.
- The recurrence relation for an arithmetic sequence is $t_{n+1} = t_n + d$ with $t_1 = a$, where d is the common difference and a is the first term.
- The recurrence relation for a geometric sequence is $t_{n+1} = rt_n$ with $t_1 = a$, where r is the common ratio and a is the first term.
- A recurrence relation of the form $t_{n+1} = rt_n + d$ with $t_1 = a$, $r \neq 1$ and $d \neq 0$, defines a sequence that is neither arithmetic nor geometric.
- When graphed, the terms of sequences generated by arithmetic recurrence relations lie on a straight line with gradient d.
  - If $d > 0$, the terms of the sequence lie on a straight line with a positive gradient, so the term values increase as n increases.
  - If $d < 0$, the terms of the sequence lie on a straight line with a negative gradient, so the term values decrease as n increases.
- When graphed, the terms of sequences generated by geometric recurrence relations lie on a curve.
  - If $r > 1$, the term values increase as n increases.
  - If $0 < r < 1$, the term values decrease as n increases.
  - If $r < 0$, the term values alternate in sign.

- When graphed, the terms in the sequence generated by the first order recurrence relation $t_{n+1} = rt_n + d$, where $t_1 = a$, $r \neq 1$ and $d \neq 0$, may increase indefinitely, decrease indefinitely or approach a particular value in the long term.

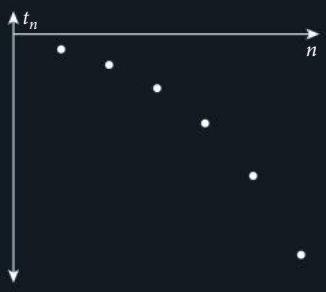
$t_{n+1} = 2t_n + 5$, $t_1 = 2$



As n increases, t_n increases.

The sequence is said to have a long-term increasing solution.

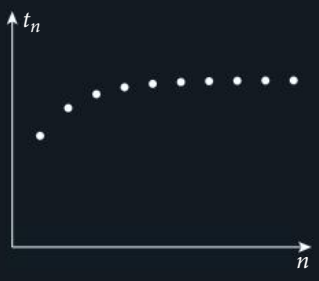
$t_{n+1} = 1.5t_n - 1$, $t_1 = -2$



As n increases, t_n decreases.

The sequence is said to have a long-term decreasing solution.

$t_{n+1} = 0.5t_n + 3$, $t_1 = 4$



As n increases, t_n approaches a particular value.

The sequence is said to have a long-term steady-state solution.

- To find the steady-state solution of a sequence, we assume that the term value ceases to change in the long run. That is, in the long run, $t_{n+1} = t_n$.

9780170250283

CHAPTER REVIEW
RECURRENCE RELATIONS
4

Multiple choice

1 Example 4 A sequence is defined by the recurrence relation $u_{n+1} = 3u_n + 4$. The third term in the sequence is 7. The fifth term is the sequence is

A 11 B 25 C 79 D 242 E 731

2 Example 6 The first six terms of a sequence are

8 2 8 2 8 2

A recurrence relation that can be used to generate this sequence is

A $t_{n+1} = t_n - 10$ B $t_{n+1} = 10 - t_n$ C $t_{n+1} = 0.25t_n$
D $t_{n+1} = t_n - 6$ E $t_{n+1} = 8 - t_n$

3 Example 6 The sequence 2, 3, 7, 23, … is generated by the first order recurrence relation $t_{n+1} = 4t_n + d$ with $t_1 = 2$. The value of d is

A -10 B -5 C -2 D 1 E 5

4 Example 8 A sequence is defined by the recurrence relation $t_{n+1} = t_n + 5$, where $t_1 = 8$. The rule that gives t_n, the nth term of the sequence is

A $t_n = 5n + 3$ B $t_n = 5n - 1$ C $t_n = n + 3$
D $t_n = n + 5$ E $t_n = 5n + 8$

5 Example 14 The recurrence relation $t_{n+1} = 0.8t_n$, where $t_1 = 100$, describes

A an increasing arithmetic sequence
B a decreasing arithmetic sequence
C an increasing geometric sequence
D a decreasing geometric sequence
E a sequence which is neither arithmetic nor geometric.

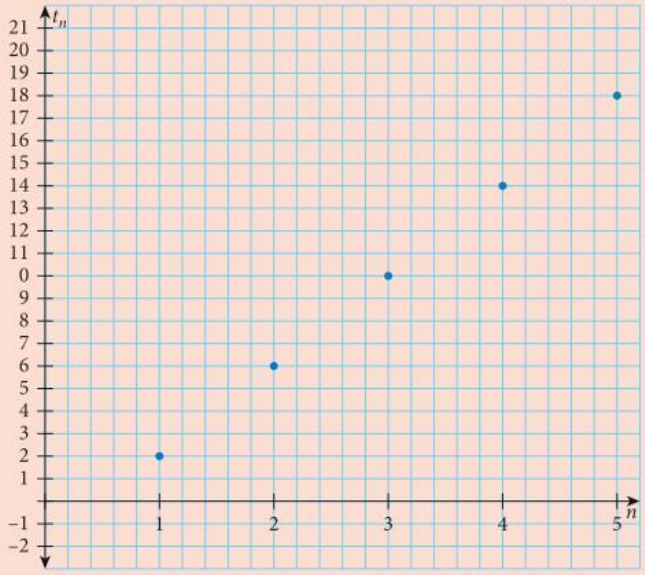
6 Example 16 Which one of the following recurrence relations has a long-term steady-state solution?

A $t_{n+1} = 2.5t_n - 5$ with $t_1 = 4$
B $t_{n+1} = -0.2t_n + 2$ with $t_1 = 5$
C $t_{n+1} = -3.4t_n + 28$ with $t_1 = 100$
D $t_{n+1} = \frac{3}{2}t_n - \frac{1}{2}$ with $t_1 = 40$
E $t_{n+1} = -\frac{5}{4}t_n - 7$ with $t_1 = 16$

7 Example 19 The sequence defined the recurrence relation $t_{n+1} = t_n + 4$, where $t_1 = 5$ is

A non-linear B arithmetic C geometric
D decreasing E alternating

8 Example 14 The graph below shows the first five terms of a sequence.



The recurrence relation that generates the terms of this sequence is

A $t_{n+1} = t_n + 4$, where $t_1 = 2$ B $t_{n+1} = 3t_n$, where $t_1 = 2$
C $t_{n+1} = t_n + 2$, where $t_1 = 4$ D $t_{n+1} = 6 - t_n$, where $t_1 = 4$
E $t_{n+1} = 4t_n$, where $t_1 = 2$

9 Example 20 Cameron invests $300 000 in an account that earns 5.5% per annum on the minimum yearly balance. At the end of each year, interest is paid into the account and Cameron withdraws $20 000 for his annual holiday. The money remaining in the account is reinvested each year.

If P_n is the amount of money invested at the start of the nth year, a recurrence relation that describes the way in which the amount of money in the account changes is

A $P_{n+1} = 0.055P_n - 20\,000$, where $P_1 = 300\,000$
B $P_{n+1} = 0.055P_n - 300\,000$, where $P_1 = 20\,000$
C $P_{n+1} = 5.5P_n - 20\,000$, where $P_1 = 300\,000$
D $P_{n+1} = 1.055P_n - 30\,0000$, where $P_1 = 20\,000$
E $P_{n+1} = 1.055P_n - 20\,000$, where $P_1 = 300\,000$

10 Example 20 An environmental authority has been monitoring the number of carp living in a river. Over time, it has been found that the number of carp living in the river decreases by 6% per year. Initially, the number of carp living in the river was estimated to be 10 000. A recurrence relation describing this situation is

A $C_n = -0.06C_{n-1}$, where $C_0 = 10\,000$ B $C_n = 0.94C_{n-1}$, where $C_0 = 10\,000$
C $C_n = 1.06C_{n-1}$, where $C_0 = 10\,000$ D $C_n = C_{n-1} + 0.06$, where $C_0 = 10\,000$
E $C_n = C_{n-1} - 0.06$, where $C_0 = 10\,000$

9780170250283

Short answer

11 Example 1 Find the first six terms of the sequence defined by the recurrence relation $t_{n+1} = 2t_n - 5$, where $t_1 = 2$.

12 Example 1 A sequence is defined by the recurrence relation $t_{n+1} = 0.8t_n + 5$, where $t_1 = 6$.
- a Find the values of t_2 and t_3.
- b What is the smallest value of n for which $t_n > 20$?

13 Example 6 The sequence 2, −2, 10, … is generated from a first order recurrence relation. Find this recurrence relation.

14 Example 9 The fifth term of an arithmetic sequence is 16 and the common difference is −6.
- a Find the first term of the sequence.
- b Write down the recurrence relation that describes this sequence.

15 Example 11 The nth term of a geometric sequence is given by the rule $t_n = 32\left(-\frac{1}{2}\right)^{n-1}$. Write a recurrence relation that defines this sequence.

16 Example 14
- a Plot the first five terms of the sequence $t_{n+1} = 0.4t_n - 6$, where $t_1 = 10$.
- b Describe the long-term solution of this sequence.

17 Example 16 A sequence is defined by the recurrence relation $t_{n+1} = 0.6t_n + 80$, where $t_1 = 10$. What value does this sequence approach in the long run?

18 Example 18 A car that was initially purchased at the start of 2015 for $45 000 depreciates at a rate of 15% per annum.
- a Write down a recurrence relation that gives the value of the car at the start of each year.
- b What is the value of the car at the start of 2017?
- c The car is replaced when its value first falls below $25 000. At the start of which year will the car be replaced?

Application

19 The Anderson family borrow $50 000 to install a new swimming pool. The interest they pay is calculated on the balance of the loan each month at a rate of 1% per month, compounded monthly. Each month, after the interest is calculated, they make a repayment of $1115.
- a The recurrence relation $A_{n+1} = kA_n + m$, where $A_1 = b$, models the amount left owing on the loan at the start of the nth month. Write down the values of k, m and b.
- b How much do the Andersons owe at the start of the
  - i second month
  - ii twentieth month?
- c At the start of which month does the amount owing on the loan first fall below $30 000?
- d At the start of the fortieth month, the Andersons make an extra repayment of $10 000. After this payment is made, how much is still outstanding on the loan?

20 At the start of 2014, a farmer introduces 2000 trout into his dam. Each year, it is estimated that the number of trout in his dam will increase by 20% and he allows his neighbours to fish in his dam, removing 300 trout each year.
- a Find a recurrence relation that gives the number of trout in the dam each year.
- b How many trout will be in the dam at the start of 2018?
- c At the start of which year will the number of trout in the dam first exceed 3000?
- d How many trout should the farmer allow to be removed each year if he wants the number of trout in his dam to remain constant?



Practice quiz





Shutterstock.com/ Tim Roberts Photography

5

TERMINOLOGY

adjacency matrix
bipartite graph
bridge
complete graph
connected graph
cycle
degree
directed graph (digraph)
edge
Eulerian graph
Euler's formula
face
Hamiltonian cycle
Hamiltonian path
isolated vertex
loop
multiple edges
path
planar graphs
semi-Eulerian graph
simple graph
spanning tree
subgraph
trail
tree
vertex (vertices)
walk
weighted graph

Prior learning

GRAPHS AND NETWORKS

UNDIRECTED GRAPHS

5.01 Basic concepts of a graph
5.02 Types of graphs
5.03 Applications of graphs
5.04 The adjacency matrix
5.05 Planar graphs and Euler's formula
5.06 Walks, paths and trails
5.07 Eulerian and semi-Eulerian graphs
5.08 Hamiltonian paths and cycles
5.09 Trees and spanning trees
5.10 Weighted graphs and minimum spanning trees
Chapter summary
Chapter review

THE DEFINITION OF A GRAPH AND ASSOCIATED TERMINOLOGY

- explain the meanings of the terms: graph, edge, vertex, loop, degree of a vertex, subgraph, simple graph, complete graph, bipartite graph, directed graph (digraph), arc, weighted graph, and network (ACMGM078)
- identify practical situations that can be represented by a network, and construct such networks; for example, trails connecting camp sites in a National Park, a social network, a transport network with one-way streets, a food web, the results of a round-robin sporting competition (ACMGM079)
- construct an adjacency matrix from a given graph or digraph. (ACMGM080)

PLANAR GRAPHS

- explain the meaning of the terms: planar graph, and face (ACMGM081)
- apply Euler's formula, $v + f - e = 2$, to solve problems relating to planar graphs. (ACMGM082)

PATHS AND CYCLES

- explain the meaning of the terms: walk, trail, path, closed walk, closed trail, cycle, connected graph, and bridge (ACMGM083)
- investigate and solve practical problems to determine the shortest path between two vertices in a weighted graph (by trial-and-error methods only) (ACMGM084)
- explain the meaning of the terms: Eulerian graph, Eulerian trail, semi-Eulerian graph, semi-Eulerian trail and the conditions for their existence, and use these concepts to investigate and solve practical problems; for example, the Königsberg Bridge problem, planning a garbage bin collection route (ACMGM085)
- explain the meaning of the terms: Hamiltonian graph and semi-Hamiltonian graph, and use these concepts to investigate and solve practical problems; for example, planning a sight-seeing tourist route around a city, the travelling-salesman problem (by trial-and-error methods only). (ACMGM086)

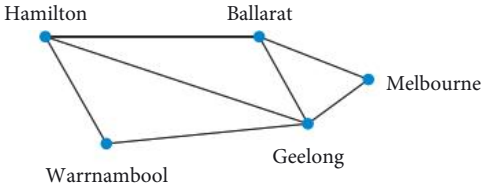
TREES AND MINIMUM CONNECTOR PROBLEMS

- explain the meaning of the terms tree and spanning tree identify practical examples (ACMGM101)
- identify a minimum spanning tree in a weighted connected graph either by inspection or by using Prim's algorithm (ACMGM102)
- use minimal spanning trees to solve minimal connector problems; for example, minimising the length of cable needed to provide power from a single power station to substations in several towns. (ACMGM103) AC

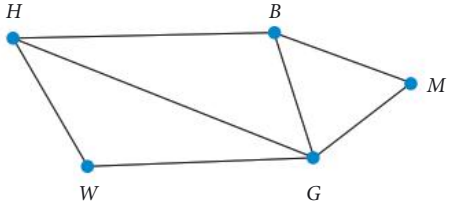
5.01 BASIC CONCEPTS OF A GRAPH

A graph is a diagram that consists of points (called **vertices**) that are connected by lines (called **edges**). Networks is the practical application of graphs and are a part of everyday life. The roads we travel on, the pipes that give us clean drinking water and the electricity cables buzzing above are all examples of networks.

The map below shows the roads connecting towns in the south-western region of Victoria.



This road network can be converted into a graph. The graph allows you to see how the towns (vertices) are connected by roads (edges). You will notice that Geelong is directly connected to every other town, however Melbourne is only directly connected to Ballarat and Geelong.



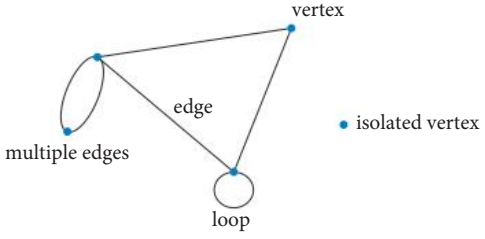
The vertices in a graph are labelled by single letters, in this case the first letter in the name of each town.

IMPORTANT

A graph is a diagram that consists of a set of vertices (points) that are connected by a set of edges (lines).

Two vertices that are connected by one or more edges are referred to as adjacent vertices. When two or more edges connect the same pair of vertices, the graph is said to have **multiple edges**. A **loop** is an edge that connects a **vertex** to itself. A vertex that is not connected to any other vertex in a graph is referred to as an **isolated vertex**.

The diagram below illustrates the different components of a graph.

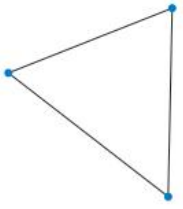


Example 1

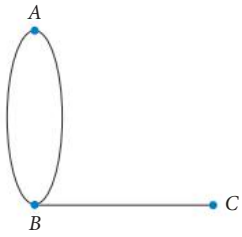
For each graph below, identify the number of

i vertices ii edges

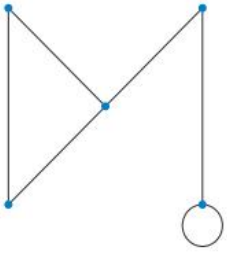
a



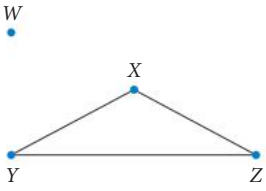
b



c



d



Solution

a There are three points and three lines connecting these points.
- i V = 3
- ii E = 3

b There are three points and three lines connecting these points.
- i V = 3
- ii E = 3

c There are five points, and there are five lines connecting pairs of points, and one loop connecting a point to itself.
- i V = 5
- ii E = 6

d There are four points and three lines connecting these points.
- i V = 4
- ii E = 3

Example 2

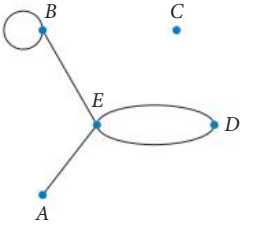
Consider the graph below.

a Name the isolated vertex.

b Which vertex is adjacent to A?

c The loop in the graph connects which vertex to itself?

d Which pair of vertices is connected by multiple edges?



9780170250283

Solution

a Vertex C is not connected to any other vertex. Vertex C.

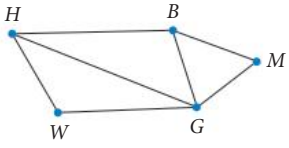
b Vertex A is connected to vertex E by an edge. Vertex E.

c There is a loop at vertex B. Vertex B.

d There are two edges connecting vertex E with vertex D. Vertex E and vertex D.

Example 3

Answer the following questions using the following map of south-west Victoria.



KEY:
H = Hamilton
W = Warrnambool
B = Ballarat
G = Geelong
M = Melbourne

a How many edges are directly connected to Hamilton?
b Which towns are adjacent to Melbourne?
c Which town is directly connected to every other?
d Which town is not directly connected to Ballarat?

Solution

a Ballarat, Warrnambool and Geelong are all connected to Hamilton by one edge each. 3 edges.

b Ballarat and Geelong are connected to Melbourne by an edge. Ballarat and Geelong.

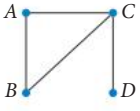
c Geelong is connected by one edge to every other town. Geelong.

d Warrnambool is not adjacent to Ballarat. Warrnambool.

Sometimes you will be required to draw a graph from given information.

Before drawing a graph, it is important to identify the vertices and edges. Vertices are usually labelled with letters, with the edges being labelled using the names of the two vertices that they connect.

The graph with vertices = {A, B, C, D} and edges = {AB, AC, BC, DC} is illustrated on the right.



This graph can be drawn differently so long as the edges are connecting the same vertices.

Example 4

Roads connect the following towns:
Tintun and Herry
Herry and Flander's Field
Flander's Field and Andervale
Tintun and Andervale

a List the vertices.
b List the edges.
c Draw a graph representing this information.

Solution

a There are four towns, therefore there are four vertices. These vertices are labelled using the first letter of each town.
The vertices are Tintun (T), Andervale (A), Herry (H) and Flander's Field (F).

$V = \{T, A, H, F\}$ ← This is referred to as the vertex set

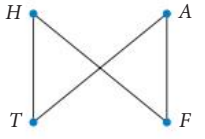
b The roads become the edges in the graph.
The edges are between:
Tintun (T) and Herry (H)
Herry (H) and Flander's Field (F)
Flander's Field (F) and Andervale (A)
Tintun (T) and Andervale (A).

$E = \{TH, HF, FA, TA\}$ ← This is referred to as the edge set

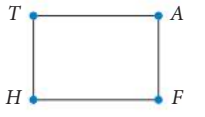
c Draw a vertex for each town using the letters H, A, T and F to represent each town.



Draw a line connecting:
T and H
H and F
F and A
T and A



Alternatively, this graph shows the same information.



9780170250283

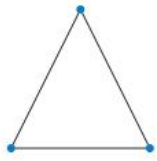
EXERCISE 5.01 Basic concepts of a graph

Concepts and techniques

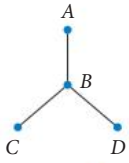
1 Example 1 For each graph below, identify the number of

i vertices ii edges

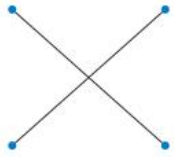
a



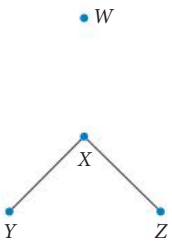
b



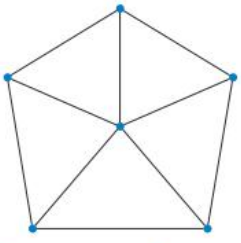
c



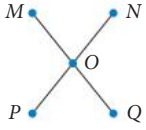
d



e



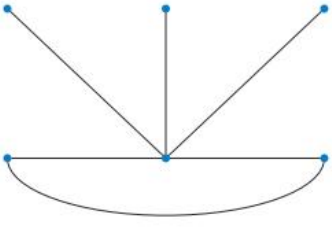
f



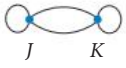
g



h

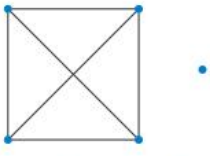


i

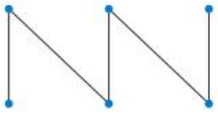


2 Which one of the following graphs does not contain five edges and six vertices?

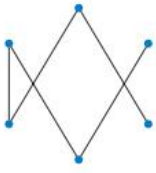
A



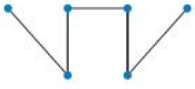
B



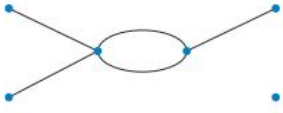
C



D



E



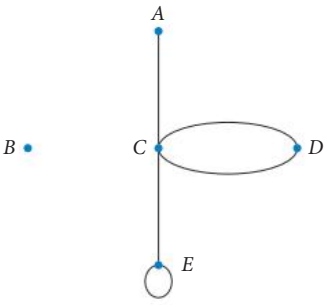
3 Example 2 Consider the graph below.

a Name the isolated vertex.

b Which vertex is adjacent to A?

c The loop in the graph connects which vertex to itself?

d Which pair of vertices is connected by multiple edges?



4 Which of the following vertices is not adjacent to B?

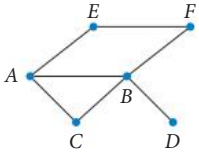
A E

B F

C A

D C

E D



5 Create a graph for each of the following.

a V = {A, B, C} and E = {AB, BC}

b V = {A, B, C, D, E} and E = {AD, BD, CD, DE}

c V = {A, B, C, D, E} and E = {AB, AD, AE, BC, BE, CD, CE, DE}

d V = {A, B, C, D, E} and E = {AB, AC, BC, CD, CE, DE}

e V = {A, B, C, D} and E = {AB, AC, AD, BC, CD}

f V = {A, B, C, D, E} and E = {AB, BC, CD, DE, EA}

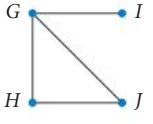
g V = {W, X, Y} and E = {WX, XY, YY}

h V = {F, G, H, I} and E = {FG, FH, HG}

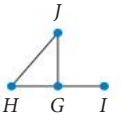
i V = {S, T, U, V} and E = {TU, SV, VT, VT}

6 Which one of the following graphs does not represent: V = {G, H, I, J} and E = {GH, GI, GJ, JH}?

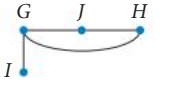
A



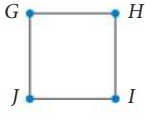
B



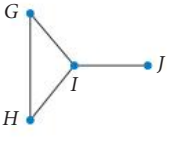
C



D



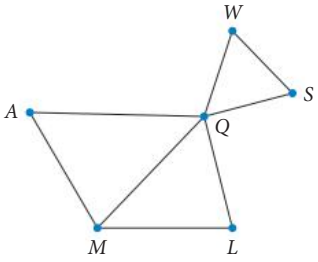
E



9780170250283

Communication and reasoning

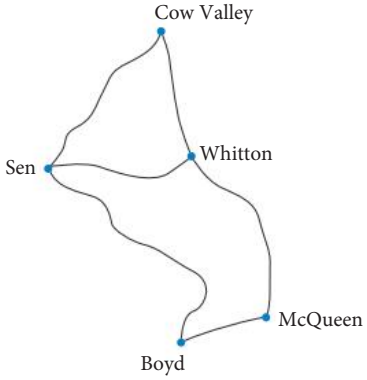
7 Example 3 Answer the following questions using the following map:



KEY:
A = Aphatown
M = Mullen
L = Long's Landing
Q = Queensvale
S = Springfield
W = Walter

a How many edges are directly connected to Mullen?
b Which towns are adjacent to Long's Landing?
c Which town is directly connected to every other?
d Which towns are not directly connected to Alphatown?

8 Which of the following towns is not directly connected to Whitton?
A Sen
B All towns are connected to Whitton
C Cow Valley
D McQueen
E Boyd



9 Example 4 Roads connect the following towns:

Aberjack to Rollington
Aberjack to Bonne
Grafter to Rollington
Grafter to Menzies
Menzies to Bonne

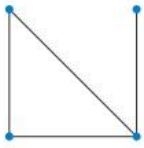
a List the vertices
b List the edges
c Draw a graph representing this information.

10 James wants to travel from his home in Sunvale to a winery in Praton.
a Construct a graph representing the following road network between five towns.
There are roads connecting:
Bundoon and Georgetown
Georgetown and Tinnit
Tinnit and Sunvale
Bundoon and Sunvale
Georgetown and Praton
b Describe the two different routes James can take to get from Sunvale to Praton.

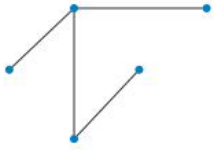
5.02 TYPES OF GRAPHS

All graphs are made up of vertices and edges. The type of graph depends on the composition of the vertices and edges. The different types of graphs are outlined below:

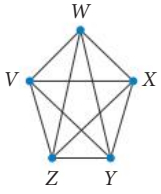
A **simple graph** contains no loops or multiple edges.



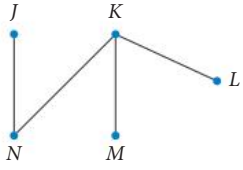
A **connected graph** contains no isolated vertices and therefore every vertex is **reachable** from every other vertex, i.e., there is a path between any two vertices of the graph.



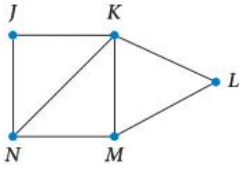
A **complete graph** has every vertex connected to every other vertex. A complete graph, with n vertices has $\frac{n(n-1)}{2}$ edges.



A **subgraph** is a graph that is part of another graph with no new edges or vertices.

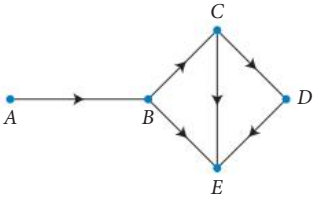


Subgraph

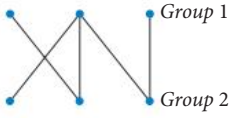


Original graph

A **directed graph** (**digraph**) contains edges that are directed. Arrows are used to indicate direction.



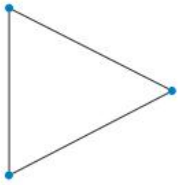
A **bipartite graph** is a graph whose vertices can be split into two distinct groups so that each edge connects each vertex in the first group with a vertex in the second group.

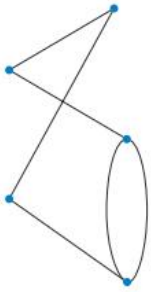


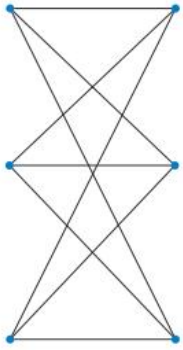
9780170250283

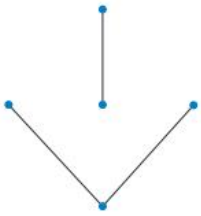
Example 5

Define each of the following graphs as either simple, connected, complete, bipartite or as a combination of these.

a 

b 

c 

d 

Solution

a	This graph contains no loops or multiple edges.	Simple.
	Every vertex is connected to every other vertex.	Complete.
	You can reach every vertex from another vertex.	Connected.
	Summarise the graph.	This graph is simple, complete and connected.
b	Every vertex in this graph is reachable from every other vertex.	This graph is connected.
c	The vertices in this graph are separated into two groups with edges connecting one group to the other.	Bipartite.
	You can reach every vertex from another vertex.	Connected.
	Summarise the graph.	This graph is a connected bipartite graph.
d	This graph contains no loops or multiple edges.	This graph is simple.

Example 6

How many edges would a complete graph have if it contained:

a 5 vertices? **b** 10 vertices? **c** 50 vertices?

Solution

a Substitute $n = 5$ into $\frac{n(n-1)}{2}$ to find the number of edges.

$$\text{Number of edges} = \frac{5(5-1)}{2}$$
$$= \frac{5 \times 4}{2}$$
$$= 10$$

A complete graph with 5 vertices will have 10 edges.

b Substitute $n = 10$ into $\frac{n(n-1)}{2}$ to find the number of edges.

$$\text{Number of edges} = \frac{10(10-1)}{2}$$
$$= \frac{10 \times 9}{2}$$
$$= 45$$

A complete graph with 10 vertices will have 45 edges.

c Substitute $n = 50$ into $\frac{n(n-1)}{2}$ to find the number of edges.

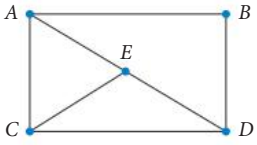
$$\text{Number of edges} = \frac{50(50-1)}{2}$$
$$= \frac{50 \times 49}{2}$$
$$= 1225$$

A complete graph with 50 vertices will have 1225 edges.

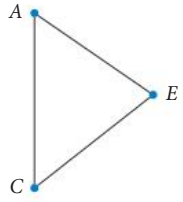
9780170250283

Example 7

Which of the following graphs is not a subgraph of the graph below?



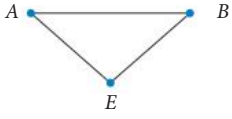
A



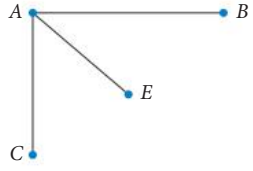
B



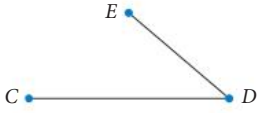
C



D



E



Solution

The original graph contains the following edges:
E = {*AB, AC, AE, BD, CD, CE, DE*}

Option C has an edge that is not included in the original graph and therefore this is not a subgraph. — The edge, *BE*, has been added.

Choose the correct response. — C.

EXERCISE 5.02 Types of graphs

WS Types of graphs

Concepts and techniques

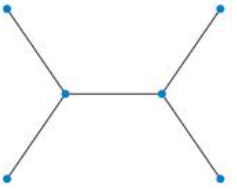
1 Match the following terms with their definition.

a Connected graph
b Simple graph
c Complete graph
d Bipartite graph
e Directed graph
f Subgraph

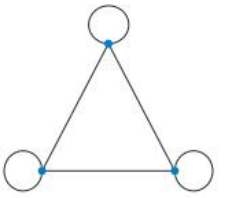
i This graph is a part of a larger graph.
ii Every vertex in this graph is connected to every other vertex.
iii This graph is separated into two groups with edges connecting each group.
iv You can go reach every vertex in this graph from another vertex.
v This graph contains no loop or multiple edges.
vi The edges of this graph show the direction you travel along the edge.

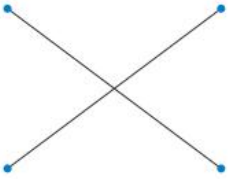
2 Example 5 Which of the following graphs are:

i simple? ii connected? iii complete? iv bipartite?

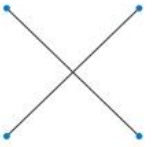
a 

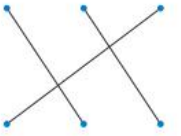
b 

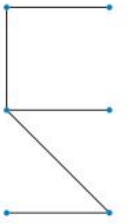
c 

d 

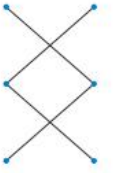
3 Which of the following graphs is not an example of a bipartite graph?

A 

B 

C 

D 

E 

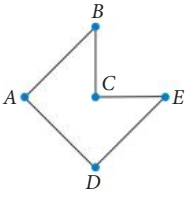
4 Example 6 How many edges would a complete graph have if it contained:

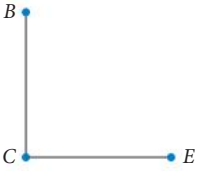
a 3 vertices? b 12 vertices? c 24 vertices? d 99 vertices?

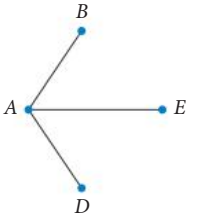
5 How many vertices would a complete graph have if it contained:

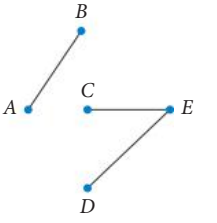
a 6 edges? b 190 edges? c 946 edges? d 500 500 edges?

6 Example 7 Which of the following graphs is not a subgraph of the graph below?

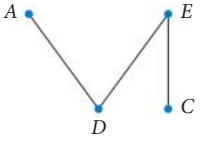


A 

B 

C 

D 

E 

9780170250283

Reasoning and communication

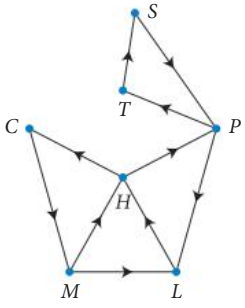
7 In 2013, four teams played in a round robin competition where each team played every other team once. In the previous year, there were six teams in the round robin competition. Knowing that a complete graph with n vertices has $\frac{n(n-1)}{2}$ edges, calculate:

a the number of matches played in 2012 **b** the number of matches played in 2013.

8 Two tennis clubs played in a grand final. Each team had three players and every player played against each player from the other club. This situation can be graphed using a:

A complete graph **B** bipartite graph **C** connected graph
D digraph **E** subgraph

9 A series of one-way streets connect the town of Marough as shown in the digraph below.



KEY:
S = Shopping centre
T = Train station
P = Police station
L = Library
H = Hospital
M = Market
C = Creek

a Is it possible to travel from the train station to every other location without passing a location more than once?
b Describe the two routes from the creek to the police station.
c Is it possible to get from the hospital to the library without passing the market?

10 A bipartite graph with six vertices would best represent which of the following?
A The allocation of three different jobs to three workers.
B A six-team soccer competition where all teams play each other once.
C The journey of a salesman visiting six businesses.
D A network of electricity cables within a house.
E A network of connecting roads in a small town.

5.03 APPLICATIONS OF GRAPHS

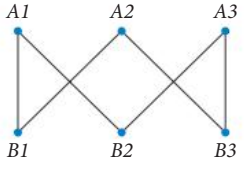
When using a graph to represent a practical situation, it is important to label the vertices either with a name, number or a letter (make sure you use a KEY if the letter abbreviates a name).

The following bipartite graph illustrates two tennis teams, A and B, and the matches that will be played.

Player $A1$ will play $B1$ and $B2$.

Player $A2$ will play $B1$ and $B3$.

Player $A3$ will play $B2$ and $B3$.



Example 8

A local football club is holding a weekend tournament with five teams. Each team is to play two other teams according to the following roster.

	Ants	Bears	Crocs	Devils	Eels
Ants		✔	✔		
Bears	✔				✔
Crocs	✔			✔	
Devils			✔		✔
Eels		✔		✔	

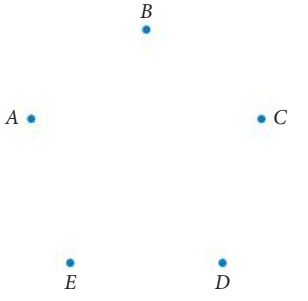
Represent this information using a graph diagram.

Solution

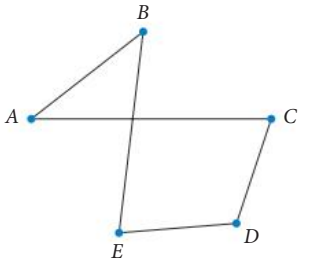
Write your KEY.

KEY
A = Ants
B = Bears
C = Crocs
D = Devils
E = Eels

There are five teams, so draw five vertices, one representing each team.



Draw an edge to represent each game between two teams. The edge set for this graph is:
E = {*AB, AC, BE, CD, DE*}.



9780170250283

Example 9

Two chess clubs are playing each other in a grand final. Each team consists of four players and each player must play each player from the opposing team.

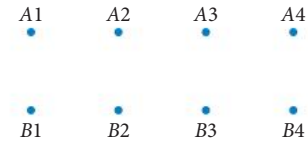
Represent this information using a graph.

Solution

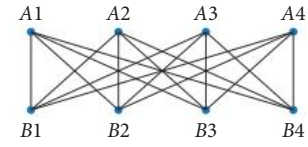
Draw two lines of four vertices to represent each team.



Label the vertices to differentiate the teams (A and B) and the players (1, 2, 3, 4).

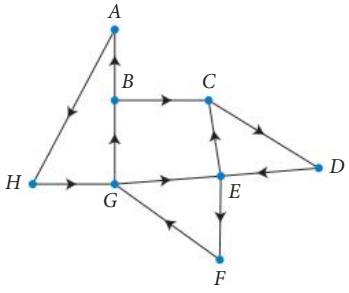


This is an example of a complete bipartite graph so use edges to connect each player in one team to every player on the opposing team.



Example 10

The following directed graph represents a monorail network in a city.



Alexis wants to travel from stop A to stop D without passing the same stop more than once.

a What are the two routes she could take?

b If the cost of the ticket depends on the number of stops travelled to, which route is cheaper?
Carly accidently missed her stop at B and has ended up at stop C.

c Can Carly catch the monorail directly back to stop B without passing any other stops?

d What is the route Carly will have to take to return to stop B?

Solution

a Each edge can only go in one direction, indicated by the arrow on the edge. Find routes starting at A and ending at D.

Route 1: $A - H - G - E - C - D$
Route 2: $A - H - G - B - C - D$

b Each route passes five stops, therefore the cost is the same.

Route 1: 5 stops
Route 2: 5 stops
Both routes have 5 stops, therefore their costs are the same.

c The edge connecting stop *B* and stop *C* is going in one direction from *B* to *C*, therefore she can not go directly back.

No, she cannot go directly back to *B* without passing any other stops.

d She will have to travel from stop *C*, to stop *D*, to stop *E*, to stop *F*, to stop *G*, and finish at stop *B*.

Route: *C* – *D* – *E* – *F* – *G* – *H* – *B*

EXERCISE 5.03 Applications of graphs

Reasoning and communication

1 Example 8 Six teams are competing in a local netball competition. Each team must play three games according to the following roster.

	Giants	Hornets	Icebergs	Joeys	Koalas	Lions
Giants		✔	✔	✔		
Hornets	✔				✔	✔
Icebergs	✔			✔	✔	
Joeys	✔		✔			✔
Koalas		✔	✔			✔
Lions		✔		✔	✔	

Draw a graph representing this information.

2 A local soccer club is holding a round robin tournament where every team plays every other team once. There are five teams in the competition.
- a What type of graph is this?
- b Without constructing a graph, how many games will be played?
- c Check your answer to part **b** by constructing the graph representing this tournament.

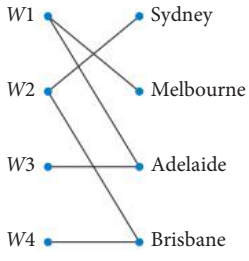
3 At a maths conference, a group of ten professors shook hands with one another. The number of handshakes was:

A 10 **B** 20 **C** 45 **D** 90 **E** 100

4 Example 9 Two badminton teams are playing for a spot in the grand final. Each team consists of three players and each player plays every player on the opposing team.

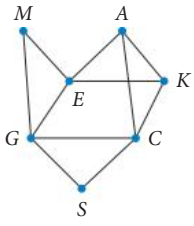
Construct a graph representing the matches to be played.

5 A computer hardware company has four warehouses to provide parts for four cities throughout Australia. The following bipartite graph represents which warehouses can provide for which cities.
- a Which cities does warehouse 1 (*W*1) provide parts for?
- b Which warehouses can provide parts for Brisbane?
- c Draw a graph if each city was to be matched with one warehouse only.



9780170250283

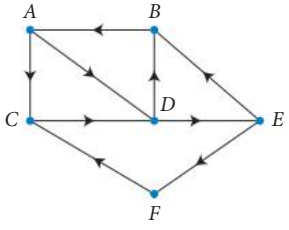
6 Emily has received a friend request on facebook from Sally. To find out how she knows Sally, she has drawn the following graph with each edge representing a friendship between the two people.



KEY:
E = Emily
S = Sally
M = Mark
G = George
A = Alison
K = Kath
C = Celeste

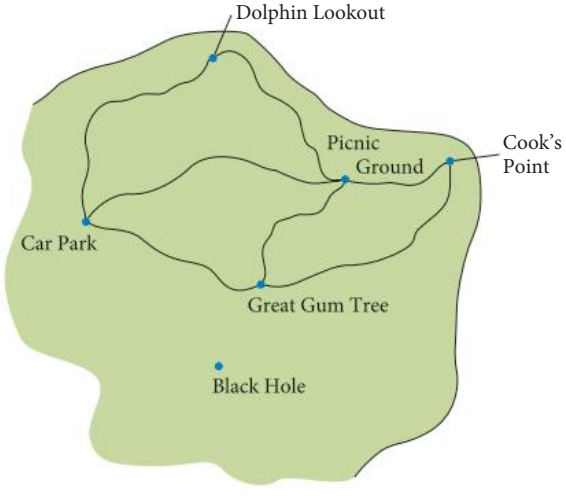
a Who is friends with both Emily and Sally?
b Who is Emily not friends with?
c If Emily wanted to be introduced to Sally, which friend should she ask to make this introduction?

7 Example 10 The following graph represents a series of one-way streets.



Rory wants to travel by car from his apartment building at A to the cinemas at F.
a Describe a route Rory would have to take to travel to the cinemas by car.
b What is the shortest route, determined by the least number of edges travelled on, for Rory to return home after his trip to the cinemas?
c i Is it possible for Rory to travel to each vertex only once if he was to start and finish at his house?
ii If no, which vertex/vertices does Rory travel to more than once?
d Is it possible for Rory to start at any vertex and travel to every other vertex only once in one trip? (He doesn't need to return to the starting vertex).

8 The following map illustrates walking trails in a national park.
a Which location is not connected to any other location?
b Which locations are not adjacent to the Car Park?
c Describe the different routes that visitors can take if they start at the Car Park and want to go to the Picnic Ground.
d Is it possible for visitors to take a route where they do not go back along a trail they have already walked, that doesn't include the Picnic Ground? Explain your answer.
e If visitors wanted to walk along every trail only once, where should they start and finish?



9 The Metro Hockey Centre is playing a round robin event where each team plays every other team once. Initially there were seven clubs competing and now two extra teams have joined. The number of extra matches to be played is:

A 2 B 9 C 15 D 18 E 36

5.04 THE ADJACENCY MATRIX

In a graph, the **degree** of a vertex is the number of edges linked to that vertex, with loops counted twice.

IMPORTANT

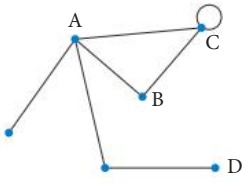
For any graph,

Degree sum = 2 × number of edges

Example 11

a Determine the degree of the labelled vertices in the following graph.

b Calculate the degree sum of this graph.



Solution

a Vertex *A*: four edges are connecting vertex *A* with other vertices.

Degree of *A* = 4

Vertex *B*: two edges are connecting vertex *B* with other vertices.

Degree of *B* = 2

Vertex *C*: two edges are connecting vertex *C* with other vertices and vertex C has one loop.

Degree of *C* = 2 + 2
= 4

Vertex *D*: one edge connects vertex *D* with another vertex.

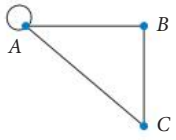
Degree of *D* = 1

b There are seven edges in this graph.

Degree sum = 2 × number of edges
= 2 × 7
= 14

Graphs can be represented using an **adjacency matrix**. A graph containing *n* vertices can be represented using an $n \times n$ matrix. To represent the graph, the names of the vertices are above the columns and to the left of the rows. The number of edges connecting the vertices is placed at the intersection of the corresponding row and column.

Remember that a loop is an edge that has both ends connected to the same vertex. Even though a loop counts for 2 when calculating the degree of a vertex, when you complete an adjacency matrix for any graph, the loop only counts as 1 edge. In the graph below, there is a loop at vertex A.



$$\begin{array}{c} \\ A \\ B \\ C \end{array}\begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \end{array}$$

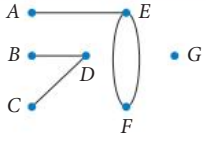
IMPORTANT

When representing a graph using an adjacency matrix you must remember the following:

- a '0' represents no edge connecting the vertices
- a '1' represents 1 edge connecting the vertices
- a number greater than 1 is used when there are multiple edges connecting the same vertices
- a '1' at the intersection for the same vertex indicates a loop

Example 12

Represent the following graph using an adjacency matrix.



Solution

Set up a blank matrix, placing the vertex names across the top and to the left.

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \\ G \end{array}\begin{array}{c} \begin{array}{ccccccc} A & B & C & D & E & F & G \end{array} \\ \begin{bmatrix} & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \end{bmatrix} \end{array}$$

Look at each vertex individually, starting with vertex A. Vertex A is connected to vertex E only. Therefore we place a 1 at the intersection of vertex A and vertex E.

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \\ F \\ G \end{array}\begin{array}{c} \begin{array}{ccccccc} A & B & C & D & E & F & G \end{array} \\ \begin{bmatrix} & & & & 1 & & \\ & & & & & & \\ & & & & & & \\ & & & & & & \\ 1 & & & & & & \\ & & & & & & \\ & & & & & & \end{bmatrix} \end{array}$$

Vertex B is connected to vertex D only. Therefore we place a 1 at the intersection of vertex B and vertex D.

	A	B	C	D	E	F	G
A					1		
B				1			
C							
D		1					
E	1						
F							
G							

Vertex C is connected to vertex D only. Therefore we place a 1 at the intersection of vertex C and vertex D.

	A	B	C	D	E	F	G
A					1		
B				1			
C				1			
D		1	1				
E	1						
F							
G							

Vertex E is connected to vertex F with two edges, therefore we place a 2 at the intersection of vertex E and vertex F.

	A	B	C	D	E	F	G
A					1		
B				1			
C				1			
D		1	1				
E	1					2	
F					2		
G							

Each edge has now been represented. Complete the matrix by placing a 0 in all unoccupied places.

	A	B	C	D	E	F	G
A	0	0	0	0	1	0	0
B	0	0	0	1	0	0	0
C	0	0	0	1	0	0	0
D	0	1	1	0	0	0	0
E	1	0	0	0	0	2	0
F	0	0	0	0	2	0	0
G	0	0	0	0	0	0	0

9780170250283

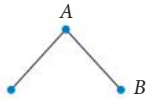
EXERCISE 5.04 The adjacency matrix

Adjacency matrices

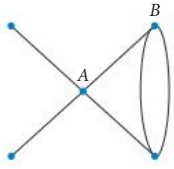
Concepts and techniques

1 Example 11 Determine the degree of the labelled vertices in each graph.

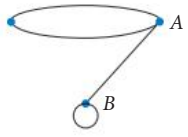
a



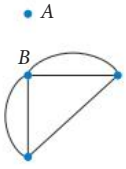
b



c



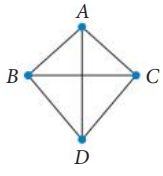
d



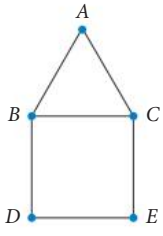
2 For each graph

i determine the degree of each vertex ii calculate the degree sum.

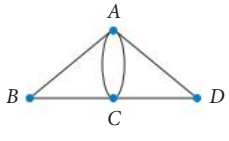
a



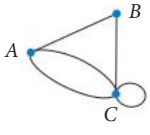
b



c



d



3 What do you notice about your answers to question **2**, parts **i** and **ii**?

4 A connected graph has six vertices.

Four of these vertices are of even degree and two are of odd degree. One extra edge is added, joining two of the existing vertices.

In the resulting graph, it is not possible to have six vertices where

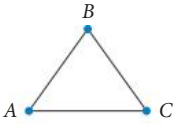
A all are of even degree
B five vertices are of even degree and one is of odd degree
C four vertices are of odd degree and two are of even degree
D four vertices are of even degree and two are of odd degree
E all are connected.

5 For each of the following graphs,

i calculate the degree sum

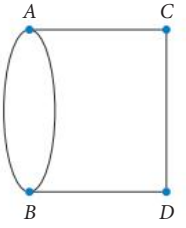
ii calculate the sum of all the entries in the corresponding adjacency matrix.

a



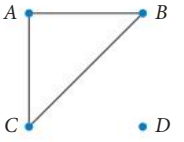
$$\begin{array}{c} \\ A \\ B \\ C \end{array}\begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \end{array}$$

b



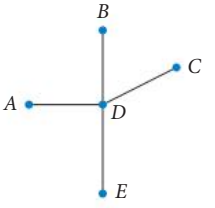
$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array}\begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

c



$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array}\begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

d

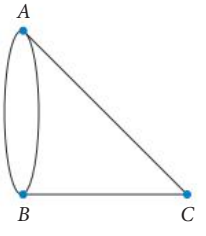


$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array}\begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

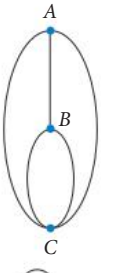
6 What do you notice about your answers to question **4**, parts **i** and **ii**?

7 Example 12 Represent each of the following graphs using an adjacency matrix.

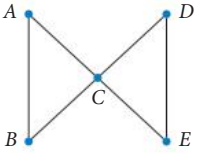
a



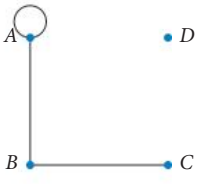
b



c



d



9780170250283

8 Use the following adjacency matrices to draw a corresponding graph

a

	A	B	C	D
A	0	1	1	1
B	1	0	1	0
C	1	1	0	1
D	1	0	1	0

b

	A	B	C	D
A	0	1	1	0
B	1	0	1	0
C	1	1	0	1
D	0	0	1	0

c

	A	B	C	D	E
A	0	2	0	1	0
B	2	0	0	0	1
C	0	0	0	1	0
D	1	0	1	0	1
E	0	1	0	1	0

Reasoning and communication

9 For parts **a** and **b** below, do the following.

i Use the adjacency matrix to draw the corresponding graph.

ii Calculate the degree sum.

iii Calculate the total number of entries in the adjacency matrix.

iv Does the degree sum equal the total number of entries in the adjacency matrix?

a

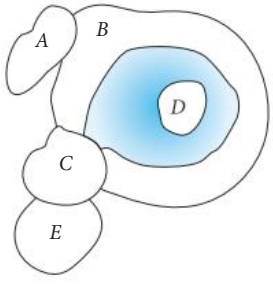
	A	B	C	D
A	1	1	0	0
B	1	0	1	0
C	0	1	0	0
D	0	0	0	0

b

	A	B	C	D	E
A	0	1	1	0	0
B	1	0	1	0	0
C	1	1	0	1	0
D	0	0	1	0	1
E	0	0	0	1	0

c When does the degree sum of a graph equal the total of the entries in the adjacency matrix?

10 The city of Nerak is divided into five suburbs labelled as A to E on the map below. A lake, which is in the middle of the city, is shaded on the map.



The adjacency matrix on the right represents the number of land borders between the towns.

	A	B	C	D	E
A	0	1	0	0	0
B	1	0	2	0	0
C	0	2	0	0	1
D	0	0	0	0	0
E	0	0	1	0	0

a Explain why all the values for vertex D are zero.

b The graph below represents the above map. The vertices represent the suburbs and the edges represent the borders. The diagram is incomplete.

i Which edge is missing?

ii Which vertex is missing?



5.05 PLANAR GRAPHS AND EULER'S FORMULA

A **planar graph** is a graph that has no edges that cross. A graph may appear non-planar (as in Diagram A below) but can be redrawn without any of the edges intersecting (as shown below in Diagram B).

Diagram A

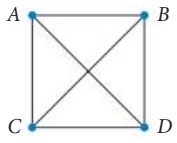
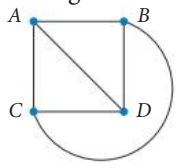


Diagram B

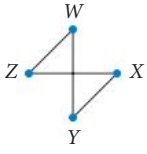


To determine whether a graph is planar or not, it is a good idea to imagine the edges as pieces of string that have been tied to nails on a board. If two (or more) pieces of string are crossed, then think about whether you can move a piece of string so that they are no longer crossed.

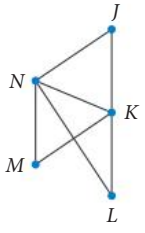
Example 13

Which of the following graphs are planar?

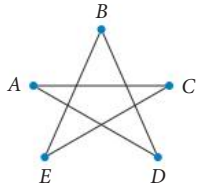
a



b

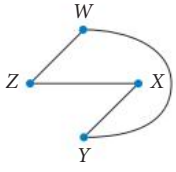


c



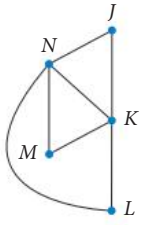
Solution

a Edges *WY* and *XZ* cross. The edge *WY* can be redrawn so as not to cross the edge *XZ*.



Therefore this graph is planar.

b Edges *KM* and *LN* cross. The edge *LN* can be redrawn so as not to cross the edge *KM*.



Therefore this graph is planar.

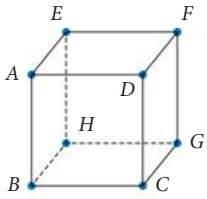
c There are five edges that cross. This graph cannot be redrawn to be made planar.

Therefore this graph is non-planar.

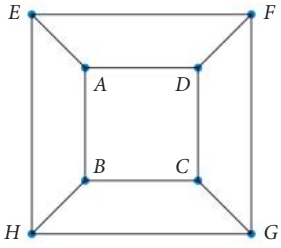
9780170250283

An application of planar graphs is the conversion of a three-dimensional solid into a two-dimensional projection of that shape.

You would be familiar with a cube (as shown to the right).



This cube can be redrawn as a planar graph as shown to the right.



In the graph, the edges are still connecting the same vertices; it has just been drawn in two-dimensional form. Therefore, a cube (or any prism for that matter) can be redrawn as a planar graph.

A cube has six sides–also known as **faces**. The faces in a planar graph are the regions enclosed by edges. You will notice that the two-dimensional representation of the cube only appears to have five faces. When determining the number of faces in a planar graph, you need to also consider the region outside of the graph as this represents the unexposed side of the cube.

IMPORTANT

For a planar graph **Euler's formula** applies:

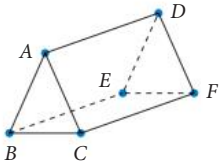
vertices = edges – faces + 2

$$V = E - F + 2$$

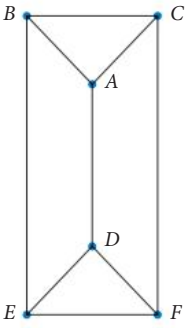
Euler's formula can be transposed to make E or F the subject if necessary.
$F = E - V + 2$
or
$E = V + F - 2$

Example 14

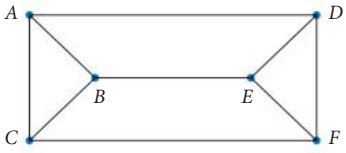
Below is a three-dimensional triangular prism. Which of the following is not a two-dimensional representation of this prism?



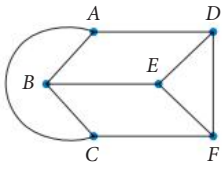
A



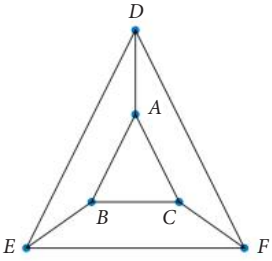
B



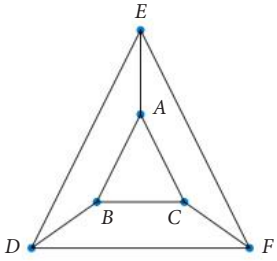
C



D



E



Solution

Each vertex is adjacent to three other vertices. Look at each vertex to see whether they are connected to the same vertices.

Vertex *A* is connected to vertices *B*, *C* and *D*; Vertex *B* is connected to vertices *C*, *A* and *E*; etc.

Options A, B, C and D are all representations of the original triangular prism.

Option E is a triangular prism but it has edges adjacent to the wrong vertices, i.e. vertex *A* is adjacent to vertex *E*, whereas in the original triangular prism vertex *A* is not adjacent to vertex *E*.

Option E.

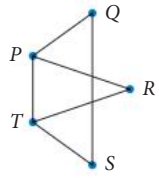
9780170250283

Example 15

Show that the following graph is planar by:

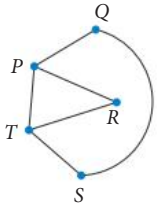
a redrawing the graph so that no edges intersect

b using Euler's formula.

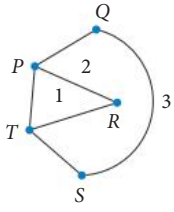


Solution

a Edge QS is crossing edge PQ and edge RT, therefore move edge QS to the outside of vertex R.



b There are three faces on this planar graph.



There are 5 vertices, 6 edges and 3 faces.

$V = 5$

$E - F + 2 = 6 - 3 + 2$

$= 5$

$V = E - F + 2$

Therefore the graph is planar.



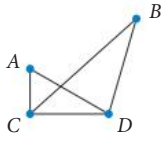
EXERCISE 5.05 Planar graphs and Euler's formula

Planar graphs

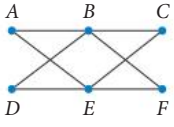
Concepts and techniques

1 Redraw the following graphs to show that they are planar.

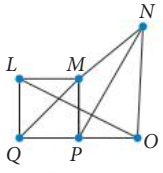
a



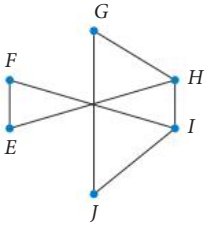
b



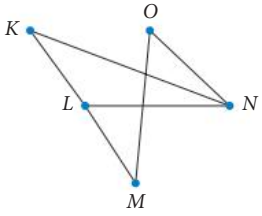
c



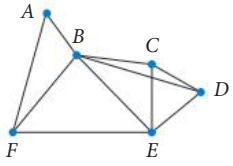
d



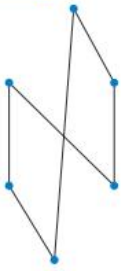
e

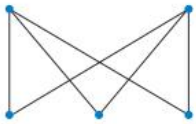


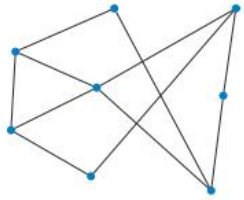
f

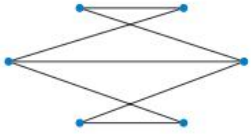


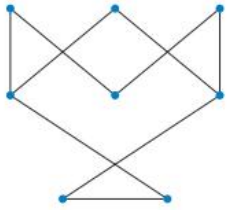
2 Example 13 Determine which of the following graphs are planar.

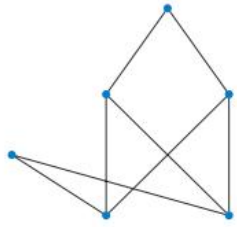
a 

b 

c 

d 

e 

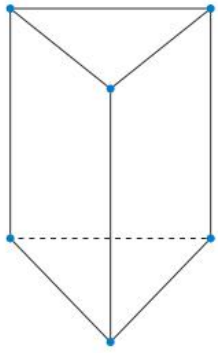
f 

3 The adjacency matrix below represents a planar graph with four vertices. The number of faces on the planar graph is

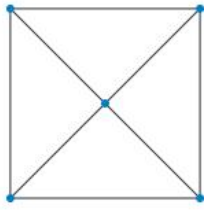
A 1 **B** 2 **C** 3 **D** 4 **E** 5

	A	*B*	*C*	*D*
A	0	1	1	1
B	1	0	1	0
C	1	1	0	1
D	1	0	1	0

4 Example 14 Represent the following triangular prism as a planar graph.



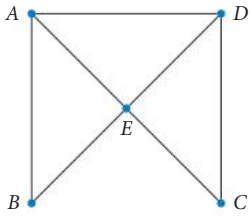
5 The following planar graph represents which three-dimensional shape?



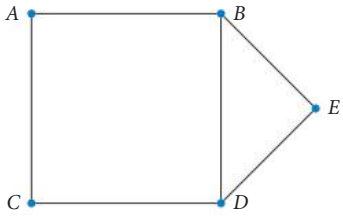
9780170250283

6 Which one of the following is a planar graph with five vertices and two faces?

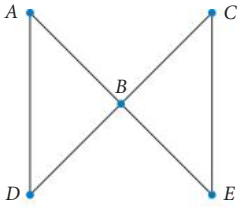
A



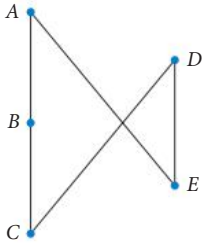
B



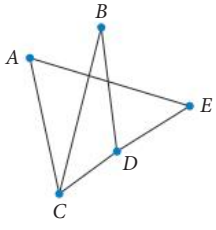
C



D



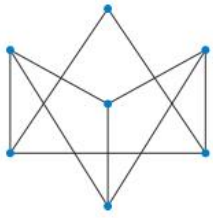
E



7 Example 15 Prove that the following graph is planar by:

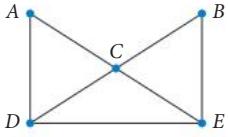
a redrawing the graph so that no edges intersect

b using Euler's formula.

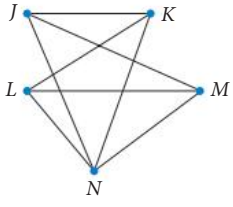


8 Euler's formula, relating vertices, faces and edges, does not apply to which one of the following graphs?

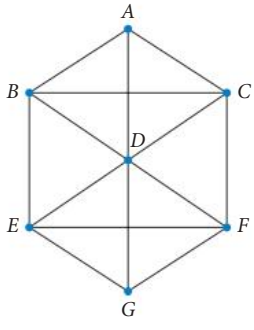
A



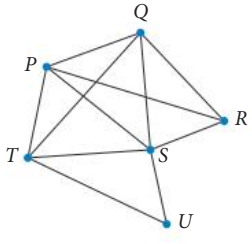
B



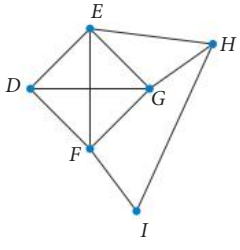
C



D



E



Reasoning and communication

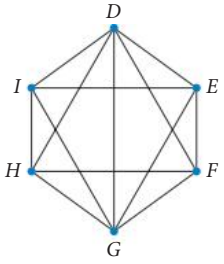
9 Given the following information, decide whether or not the graph is planar by using Euler's formula.

- **a** Six vertices, nine edges, and five faces.
- **b** Ten vertices, 14 edges, and six faces.
- **c** Nine vertices, 12 edges and four faces.

10 Draw each planar graph determined in question **9**.

11 The following graph is non-planar. Which edge would need to be removed to make it planar?

- **A** *EI*
- **B** *DI*
- **C** *DG*
- **D** *GI*
- **E** *EG*



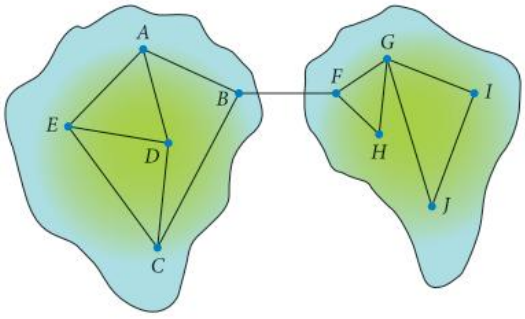
5.06 WALKS, PATHS AND TRAILS

A **walk** is the sequence of vertices such that from each of its vertices there is an edge to the next vertex in the sequence. A walk can include repeated vertices or repeated edges.

A **path** is a type of walk where all the edges and vertices are different. A path that starts and finishes at different vertices is defined as an open path, whereas a path that starts and finishes at the same vertex is said to be a closed path or **cycle**.

A cycle is also known as a circuit.

Consider the following graph as a basis for some definitions.



West Island East Island

An open walk is a sequence of vertices where you start and finish at a different vertex, for example, (*A*, *E*, *C*, *D*, *C*, *B*, *F*). On the other hand, a closed walk must start and finish at the same vertex, for example, (*A*, *E*, *C*, *D*, *A*).

A **trail** is a special type of walk where you can only travel along an edge once. In the open walk example above, edge *CD* was covered twice. If we wanted to visit the same vertices but only travel along an edge once, then we could follow this sequence of vertices: (*A*, *E*, *D*, *C*, *B*, *F*), which would be considered an open trail. A closed trail must start and finish at the same vertex but must only travel along an edge once.

9780170250283

A **bridge** is an edge that is vital to keeping the graph connected, i.e. if a bridge was removed, then the graph would no longer be connected. The two islands are connected by edge *BF*, which is defined as a bridge.

Example 16

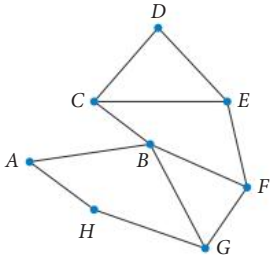
Using the following graph, define the following walks as either open or closed.

a (*A*, *B*, *C*, *E*, *F*, *B*, *G*)

b (*D*, *E*, *F*, *G*, *H*, *A*)

c (*G*, *F*, *E*, *C*, *B*, *G*)

d (*C*, *B*, *A*, *H*, *G*, *F*, *E*, *C*)



Solution

a This walk starts at vertex *A* and finishes at vertex *G*. — This is an open walk.

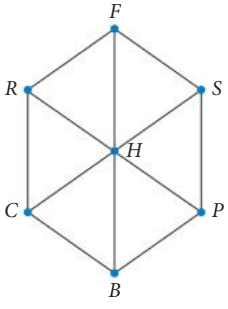
b This walk starts at vertex *D* and finishes at vertex *A*. — This is an open walk.

c This walk starts at vertex *G* and ends at vertex *G*. — This is a closed walk.

d This walk starts at vertex *C* and ends at vertex *C*. — This is a closed walk.

Example 17

The following graph outlines the locations of the main attractions at a fair.



KEY:
F = Front gate
R = Rollercoaster
C = Clowns
H = Hot dog stand
P = Pillow
S = Slippery slide
B = Back gate

a Claire wishes to visit every attraction once and then return to the front gates. Define a closed trail she could complete.

b Claire's friend, Andrew, arrived at the back gate. He too wants to visit every attraction once but finish at the front gate to meet Claire. Define an open trail that he could complete.

Solution

a Claire must start at *F* and visit every vertex once, except for the back gate, *B*, and then return to *F*.

(*F*, *R*, *C*, *H*, *P*, *S*, *F*)
or the reverse
(*F*, *S*, *P*, *H*, *C*, *R*, *F*)

b Andrew must start at *B* and visit every vertex once and finish at vertex *F*.

(*B*, *C*, *R*, *H*, *P*, *S*, *F*)
or
(*B*, *P*, *S*, *H*, *C*, *R*, *F*)

EXERCISE 5.06 Walks, paths and trails

Concepts and techniques

1 Complete the following sentences by filling in the missing words.

a Another name for a closed walk is a _____.

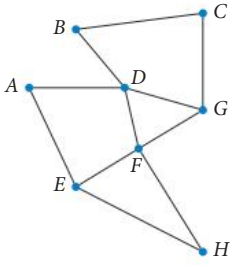
b An open walk must start and finish at _____ vertices.

c A path must not contain _____ vertices.

d A _____ is an edge that if removed would make the graph disconnected.

e A trail must go to _____ vertex only once.

2 Example 16 Using the following graph, define the following walks as either open or closed.



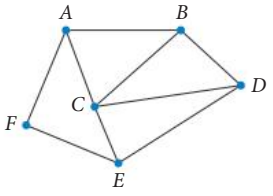
a (*A*, *D*, *B*, *C*, *G*, *F*)

b (*C*, *G*, *D*, *F*, *G*, *C*)

c (*H*, *E*, *F*, *G*)

d (*E*, *A*, *D*, *G*, *F*)

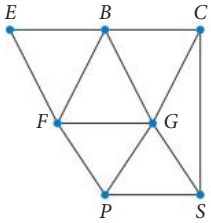
3 For the following graph, without returning to a vertex already used, the total number of paths connecting *A* and *C* is:



A 3 **B** 4 **C** 5 **D** 7 **E** 8

Reasoning and communication

4 Example 17 The following graph outlines the different stalls at a market.

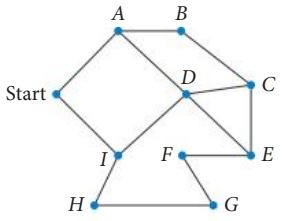


KEY:
E = Entrance
F = Florist
B = Butcher
C = Car park
G = Grocer
P = Plants
S = Snacks

a Deanna enters the market at the entrance and wishes to visit every stall once (excluding the car park) and then return to the entrance. Define a cycle she could complete.

b Deanna's friend, Chris, entered from the car park. He too wants to visit every attraction once but finish at the entrance to meet Deanna. Define an open trail that he could complete.

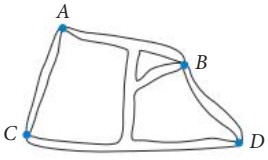
5 Cheryl is planning an orienteering course for a local school group. She wants the students to complete several courses around the locations outlined on the following graph.



a Outline a walk that takes students to:
   i every location and then back to the start
   ii three locations and then back to the start
   iii every location only once and ending at location *D*.

b Classify the walks defined in part **a** as either paths or cycles.

6 Vivian, Elina, Nancy and Helen each live in a different town. The following is a map of the roads that link the four towns *A*, *B*, *C* and *D*.



a How many roads connect *A* and *C*?

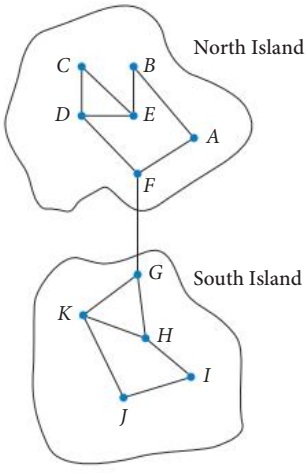
b Nancy lives in town *B*. One day she forgot her purse. Was she able to return to her town without having to pass through any other town? Explain why.

c Draw a graph representing this map.

d Vivian lives in town *A*. Can she reach her friends in the other towns directly without having to pass through another town? Explain why.

e Elina lives in town *C*. When Vivian went to visit Elina, Elina also went to visit Vivian but they never crossed paths. Explain why.

7 Patricia is touring two islands off the coast of Queensland. She is staying on the North Island.



a On her first day, she wishes to complete a closed path of the North Island, starting and finishing at her accommodation, *A*. Define the closed path that Patricia should take.

b On her second day, Patricia wishes to visit the South Island. She must travel from her accommodation, *A*, from the North Island to the South Island. If Patricia wishes to visit every location on the South Island and then return to her accommodation, define the closed walk Patricia should take.

c Explain the difference between a closed path and a closed walk using your answers to parts **a** and **b**.

5.07 EULERIAN AND SEMI-EULERIAN GRAPHS

A connected graph is said to be **Eulerian** if it contains a closed trail, i.e., it starts and finishes at the same vertex and includes every edge once only. A connected graph is considered to be **semi-Eulerian** if it contains an open trail, i.e., it starts and finishes at different vertices and includes every edge once only. Eulerian and semi-Eulerian graphs may contain trails that include repeated vertices but not repeated edges.

IMPORTANT

For an **Eulerian graph** and a closed trail to exist, the following conditions must apply.

1 All vertices must be of even degree.

2 You must start and finish at the same vertex.

For a semi-Eulerian graph and an open trail to exist, the following conditions must apply.

1 Only two vertices can be of odd degree; every other vertex must be of even degree.

2 The trail must start at a vertex of odd degree and finish at the other vertex of odd degree.

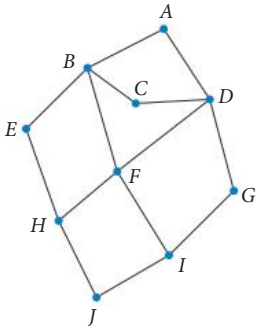
An open trail is also known as an Eulerian path and a closed trail is also known as an Eulerian cycle.

Example 18

For the graph on the right,

a decide whether it is Eulerian or semi-Eulerian.

b define the closed or open trail that exists.



Solution

a Determine the degree of every vertex.

$A = 2$
$B = 4$
$C = 2$
$D = 4$
$E = 2$
$F = 4$
$G = 2$
$H = 3$
$I = 3$
$J = 2$

A semi-Eulerian graph must have two vertices of odd degree. Every other vertex must be of even degree.

This graph is semi-Eulerian.

b The open trail must start at an odd vertex and finish at the other odd vertex.

H and *I* have an odd degree.

Starting at *H*, we must travel along every edge before returning to *I*. Vertices can be covered more than once.

(*H, E, B, A, D, C, B, F, D, G, I, F, H, J, I*)

9780170250283

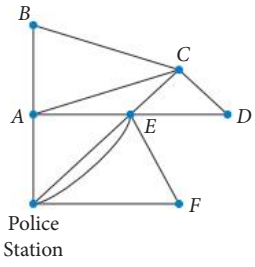
Example 19

Police are patrolling inner-city streets to prevent crime. Police officials must determine the best route for officers so that they do not walk along a street more than once. Officers must start and finish at the police station.

a Describe a route that the officers can take.

b Explain why this graph is Eulerian and not semi-Eulerian.

c If one of the roads connecting the police station with E was closed, would officers still be able to patrol every road? (They don't have to return to the police station).



Solution

a The officers must start at the station and cover each edge once.
(Station, A, B, C, A, E, C, D, E, F, Station, E, Station, E, Station)

b This graph contains no vertices of odd degree.
This graph contains a closed trail and therefore is Eulerian.

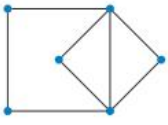
c Removing this edge will give two vertices an odd degree.
This graph will contain an open trail and therefore officers will be able to patrol every road.

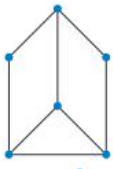
EXERCISE 5.07 Eulerian and semi-Eulerian graphs

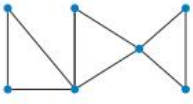
Eulerian graphs

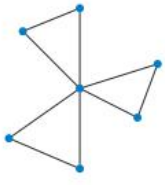
Concepts and techniques

1 Considering the conditions necessary for an Eulerian graph to exist, determine whether the following graphs are Eulerian or not by calculating the degree of each vertex.

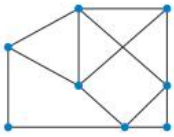
a 

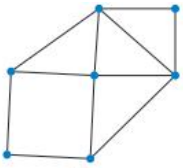
b 

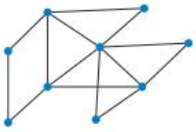
c 

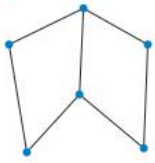
d 

2 Considering the conditions necessary for a semi-Eulerian graph to exist, determine whether the following graphs are semi-Eulerian or not.

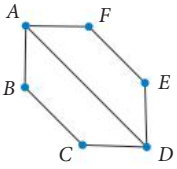
a 

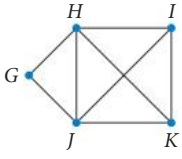
b 

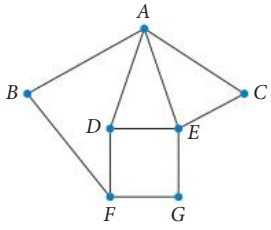
c 

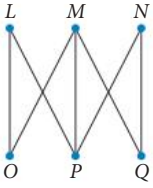
d 

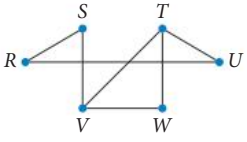
3 The following graphs are all semi-Eulerian. The open trail must start and finish at which two vertices?

a 

b 



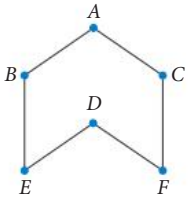
c 

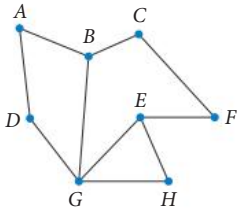
d 

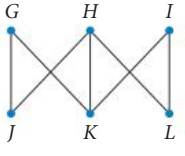
4 Example 18 For each graph below,

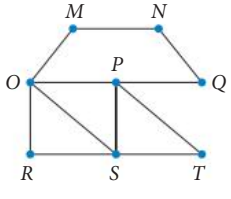
i decide whether it is Eulerian or semi-Eulerian

ii define the closed or open trail that exists.

a 

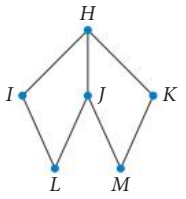
b 

c 

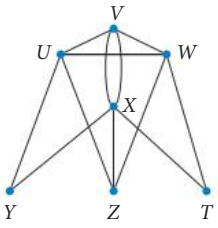
d 

9780170250283

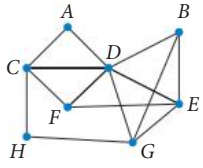
e



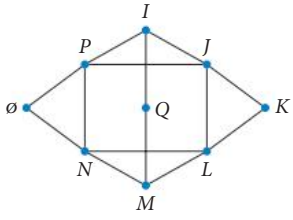
f



g



h

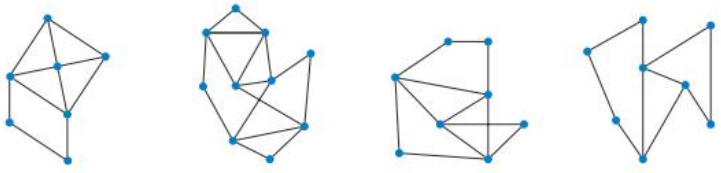


5 Where could an edge be placed to make the graph on the right Eulerian?

A *BD** **B** *GC** **C** *DF*

D *DG** **E** *EF*

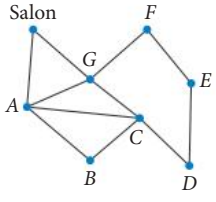
6 How many of the four following graphs are Eulerian?



A 0 **B** 1 **C** 2 **D** 3 **E** 4

Reasoning and communication

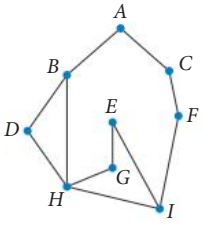
7 Example 19 Louise is distributing flyers for her hair salon. The graph below outlines the streets she must cover. So that she completes her task in the minimum time possible, she does not want to travel along a road more than once. She wants to return to her salon when she is finished.



a Describe a route that Louise could take.

b Is this graph Eulerian or semi-Eulerian?

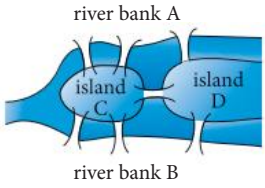
8 A salesman is travelling around a local suburb selling an Internet package to residents. In order to minimise his time, he doesn't want to travel along a street more than once. He doesn't have to finish at the same location that he started from.



a Where should the salesman start and finish?

b Describe a route that the salesman could take.

9 The Konigsberg Bridge Problem was the starting point for Eulerian graphs. People would attempt to cross each of the seven bridges once without crossing the same bridge twice.



a Represent the above map of the Konigsberg bridges as a graph.
b Explain why it is not possible to cross every bridge once and once only.
c Which edge/s could be removed in order to make this graph semi-Eulerian, hence allowing people to cross every bridge once?

10 A connected planar graph has five vertices F, G, H, I and J.

The degree of each vertex is given in the following table.

Vertex	Degree
F	5
G	4
H	3
I	3
J	3

Which one of the following statements regarding this planar graph is true?
A The addition of one further edge could make the graph semi-Eulerian.
B The graph is Eulerian.
C Euler's formula cannot be used.
D The sum of the degrees of the vertices equals 17.
E More than one open trail exists.

5.08 HAMILTONIAN PATHS AND CYCLES

A **Hamiltonian path** is a path that includes every vertex in a graph once only and starts and finishes at different vertices. A **Hamiltonian cycle** is a cycle that includes each vertex in a graph once only and starts and finishes at the same vertex.

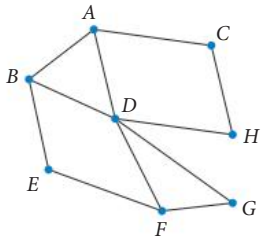
IMPORTANT

Unlike Eulerian graphs, to determine whether a Hamiltonian path or Hamiltonian cycle exists, you must use trial and error.

9780170250283

Example 20

List a Hamiltonian path for the graph below.



Solution

Start at any vertex.

Start at vertex A.

A Hamiltonian path must visit every vertex. Move from A to another vertex.

$A - B$

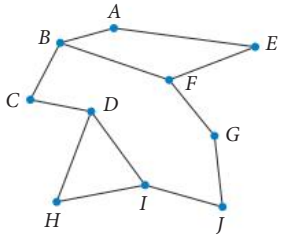
Keep moving around the graph so each vertex is only visited once. Stop when you have visited all vertices.

(A, B, E, F, G, D, H, C)

The choice of the starting vertex is arbitrary. If a Hamiltonian path cannot be found by starting at this vertex, try starting at another vertex.

Example 21

List a Hamiltonian cycle for the graph below.



Solution

Start at any vertex.

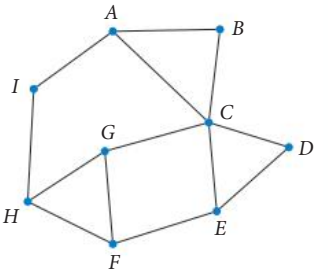
Start at vertex A.

Follow a path that travels to every vertex only once and then returns to A.

$(A, B, C, D, H, I, J, G, F, E, A)$

Example 22

A salesman is travelling to seven country towns to sell flour to the local bakeries. A map of the towns is shown below, with the vertices representing the towns and the edges representing the roads connecting the towns.



The salesman wants to travel from town *A* to visit every town once on his journey and then return back to town *A*.

a What is this walk an example of?

b Describe a route the salesman could take.

c List the edges that the salesman does not travel on.

Solution

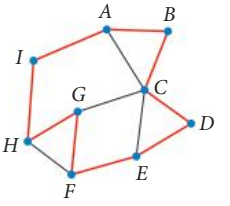
a The salesman must go to every vertex and return to the vertex he started at.

Hamiltonian cycle.

b The salesman must start at *A* and travel to every other town once only before returning to *A*.

(A, I, H, G, F, E, D, C, B, A)

c Highlight the path the salesman took.



The edges that are not highlighted are the edges the salesman did not travel on.

HF, *GC*, *CE* and *AC*.

WS

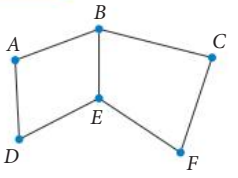
Hamiltonian paths and cycles

EXERCISE 5.08 Hamiltonian paths and cycles

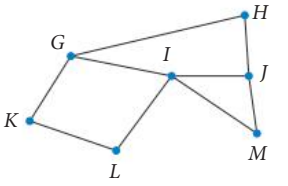
Concepts and techniques

1 Example 20 List a Hamiltonian path for each of the following graphs.

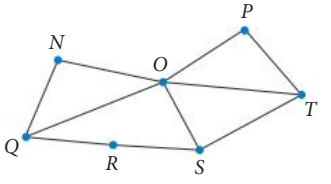
a



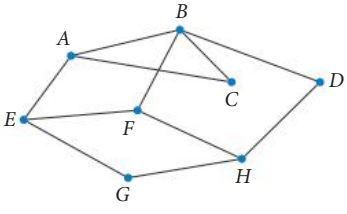
b



c



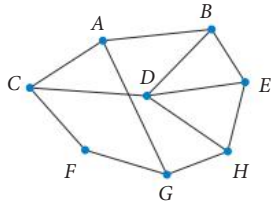
d



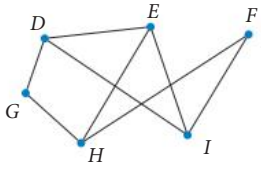
9780170250283

2 Example 21 List a Hamiltonian cycle for each of the following graphs.

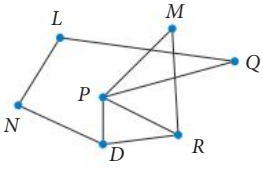
a



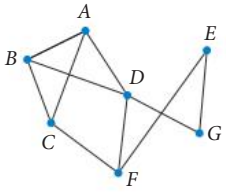
b



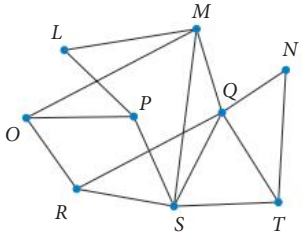
c



d

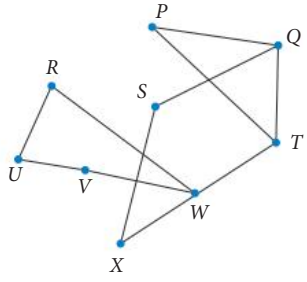


3 A Hamiltonian cycle for the graph on the right is



A $(L, P, O, M, Q, N, T, S, P, L)$
B $(R, Q, N, T, S, P, L, M, O)$
C $(Q, N, T, S, R, Q, M, L, P, O, M, Q)$
D $(M, O, R, Q, N, T, S, P, L, M)$
E $(L, P, O, R, Q, N, T, S, M)$

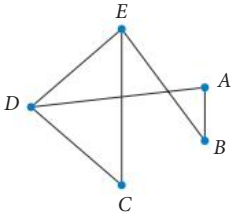
4 Which one of the following statements is true regarding the graph on the right?



A (P, Q, T, P) is a Hamiltonian cycle.
B Only one Hamiltonian path exists.
C (P, Q, T, P) is a closed trail.
D There are no cycles.
E The graph is a tree.

Reasoning and communication

5 Example 22 The Nixon family is travelling to five country towns to visit relatives. A map of the towns is shown on the right, with the vertices representing the towns and the edges representing the roads connecting the towns.



The family wants to travel from town A, visit every town once and then return back to town A.

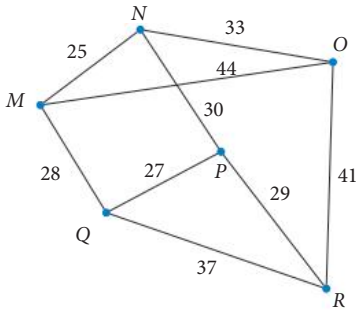
a Describe a route that the family could take.
b List the edge/s that the family does not travel on.

6 Merrick, Nottinham, Opple, Printone, Quetec and Rusher are six towns shown on the following graph. The numbered edges show how the towns are connected by roads and the distances between these towns in kilometres.

a In kilometres, what is the shortest distance between Nottinham and Rusher?

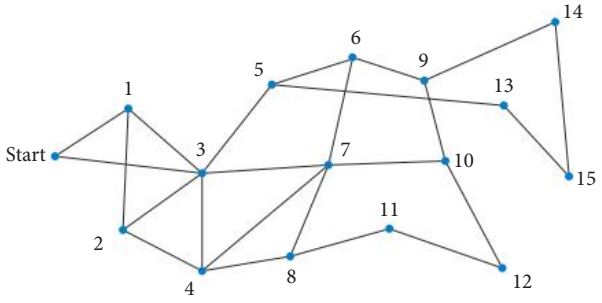
b How many different ways are there to travel between Nottinham and Rusher without passing through any town more than once? List the paths.

c A visiting politician wants to visit each town. What is the shortest distance that the politician has to travel if she starts at Merrick and finishes at Rusher?



7 The annual Mt. Crumpit orienteering race states that competitors must visit all 15 checkpoints before heading back to the starting point. Competitors will be penalised if they go to a checkpoint more than once.

List a possible cycle that competitors can take.

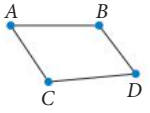


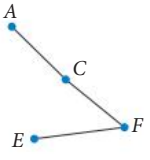
5.09 TREES AND SPANNING TREES

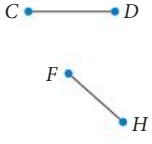
A **tree** is a connected graph that cannot contain any loops, multiple edges, or cycles. A **spanning tree** is a subgraph of another graph that includes all vertices and contains the same properties as a tree.

Example 23

For each of the following graphs, state whether or not it is a tree.

a 

b 

c 

d 

9780170250283

Solution

a Trees can not contain cycles. — This is not a tree.

b The graph is connected and contains no loops, multiple edges or cycles. — This is a tree.

c The graph is not connected. — This is not a tree.

d The graph is connected and contains no loops, multiple edges or cycles. — This is a tree.

Example 24

i Use the following adjacency matrices to draw a corresponding graph.

ii State whether the graph is a tree or not.

a

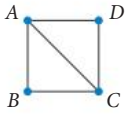
	A	B	C	D
A	0	1	1	1
B	1	0	1	0
C	1	1	0	1
D	1	0	1	0

b

	A	B	C	D	E
A	0	0	0	1	0
B	0	0	0	1	0
C	0	0	0	1	0
D	1	1	1	0	1
E	0	0	0	1	0

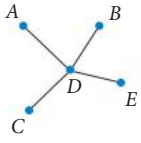
Solution

a Draw the four vertices and corresponding edges.



The tree contains a cycle. — This graph is not a tree.

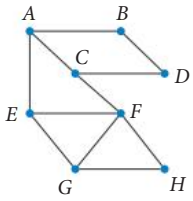
b Draw the five vertices and corresponding edges.



This graph does not contain any cycles, loops or multiple edges. — This graph is a tree.

Example 25

The following graph outlines the current water system between garden beds. Plumbers have noticed that water is being wasted as too many pipes are being used to connect the garden beds.



a Draw a spanning tree for this graph.

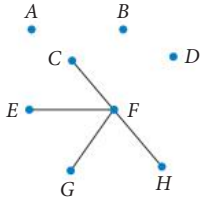
b According to the spanning tree drawn in part **a**, which edges should be removed?

Solution

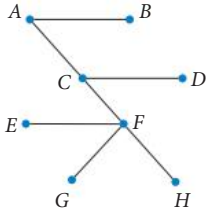
a Start at any vertex.

Vertex *F*.

Connect to every adjacent vertex.



Continue connecting the vertices, making sure that every vertex is connected and no cycles exist.



b Compare the graph drawn in part **a** with the original to determine which edges weren't covered.

AE, *BD*, *EG* and *GH*.

EXERCISE 5.09 Trees and spanning trees

Concepts and techniques

1 Example 23 For each of the following graphs, state whether or not it is a tree.

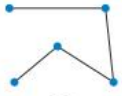
a



b



c



d



e



f



9780170250283

2 Example 24 i Use the following adjacency matrices to draw a corresponding graph.
ii State why the graph is not a tree.

a

	A	B	C
A	0	2	1
B	2	0	1
C	1	1	0

b

	A	B	C	D	E
A	0	2	0	1	0
B	2	0	0	0	1
C	0	0	0	1	0
D	1	0	1	0	1
E	0	1	0	1	0

c

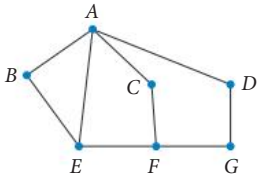
	A	B	C	D
A	0	1	1	0
B	1	0	1	0
C	1	1	0	0
D	0	0	0	0

d

	A	B	C	D
A	1	1	0	0
B	1	0	1	0
C	0	1	0	0
D	0	0	0	0

3 State whether each of the following statements is true or false.
a Trees can contain multiple edges or loops.
b A Hamiltonian path is an example of a tree.
c A tree must not contain any cycles.
d A spanning tree must connect all vertices.
e A spanning tree must be a connected graph.

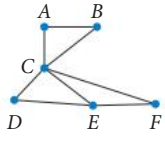
4 The minimum number of edges to be removed from this graph to create a spanning tree is



A 1 **B** 2 **C** 3
D 4 **E** 5

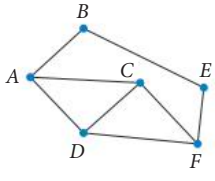
Reasoning and communication

5 Example 25 The graph on the right outlines the paths connecting buildings on a university campus. Undercover walkways are to be installed so that people will be able to walk from one building to another without being affected by the weather.



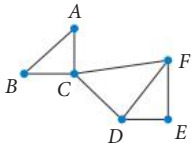
a Draw a spanning tree for this graph.
b According to the spanning tree drawn in part **a**, which edges should be removed?

6 Connor wishes to set up a surround sound system in his house. His amplifier is located at C in the graph below. Every room must be connected to the amplifier by wiring.



Draw a possible spanning tree, indicating the edges where the wiring will be placed.

7 Two trucks have been sent from a warehouse, C, to deliver fresh fruit to local supermarkets, A, B, D, E and F.



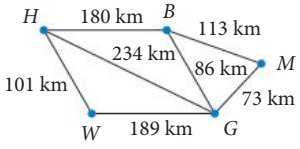
Draw a spanning tree showing their possible routes.

5.10 WEIGHTED GRAPHS AND MINIMUM SPANNING TREES

A **weighted graph** is a type of network where extra information such as distances, times or costs are labelled on the edges.

In the beginning of this chapter we looked at a map of southwest Victoria. The same graph is drawn below, but this time the distances are added to the edges.

This weighted graph now allows us to determine things such as the shortest distance from Melbourne to Hamilton – is it better to go via Geelong and Ballarat, or Geelong and Warrnambool? Let's look at the possible routes.



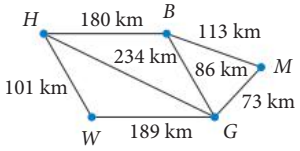
$M - G - H = 307$ km
$M - B - H = 293$ km
$M - G - W - H = 363$ km

Therefore, the shortest distance from Melbourne to Hamilton is 293 km, going via Ballarat.

To find the shortest route, we don't necessarily have to list each possible route. Sometimes it is much easier to look at the overall network to get a rough estimate of the shortest route.

Example 26

Using the map of south-west Victoria, decide which is the shortest route from Ballarat to Warrnambool.



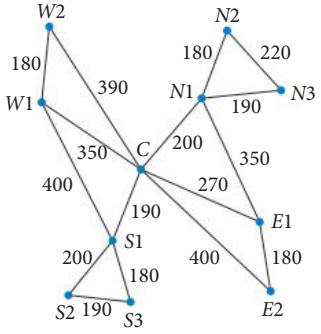
Solution

There are two routes that appear to be the shortest.	$B - G - W$ and $B - H - W$.
The route from Ballarat to Warrnambool via Geelong.	86 km + 189 km = 275 km
The route from Ballarat to Warrnambool via Hamilton.	180 km + 101 km = 281km
Write the answer.	The shortest route goes via Geelong.

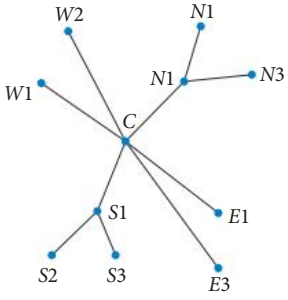
Spanning trees can be created from weighted graphs to find the minimum distance between all vertices – this is known as the minimum spanning tree. The minimum spanning tree has practical applications such as finding the minimum length of pipes required in a water system, or the minimum length of pathways required to connect buildings.

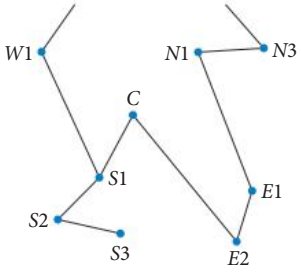
9780170250283

The map below outlines where certain buildings are within a university. The vertices represent the buildings and the edges represent existing paths (and their distances in metres).

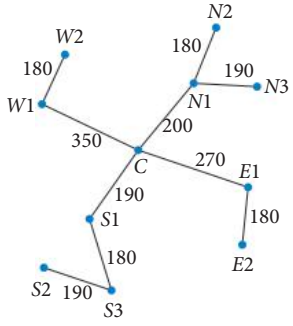


The maintenance team want to cover some of the paths so that the students stay dry when it is raining. Possible spanning trees could be:





The maintenance team want to keep costs down, therefore they need to find the spanning tree with the minimum distance – this is known as the minimum spanning tree.



We can see that the above spanning tree is also the minimum spanning tree (the minimum distance was found to be 2110 m), but there is also an algorithm we could use to determine the minimum spanning tree, known as Prim's algorithm.

IMPORTANT

Prim's Algorithm

Step 1: Choose any vertex within the network.

Step 2: Find the edge with the lowest value connecting this vertex to another vertex.

Step 3: Look at all edges connecting these two vertices and choose the edge that has the lowest value.

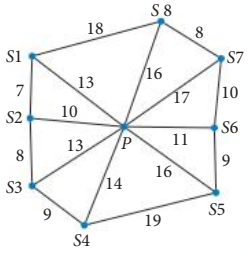
Step 4: Look at all the vertices covered so far and select the edge with the lowest value.

Step 5: Repeat step 4 until all the vertices in the graph are included in the tree.

Note that if there is more than one edge with the lowest value, then just choose either.

Example 27

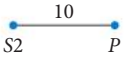
Edgar wishes to connect electricity to eight sites in a caravan park. The weighted graph on the right indicates the direct distance between each site (in metres) and the location of the electricity source (P).



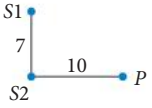
Use Prim's algorithm to find the least amount of electric cable that Edgar will need to use to connect electricity to all eight sites.

Solution

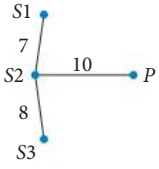
Start at P and select the edge with the shortest length.



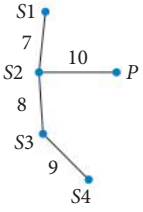
Then choose the shortest edge connecting either P or $S2$.



Choose the shortest edge connecting P, $S2$ or $S1$.

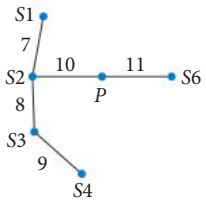


Choose the shortest edge connecting P, $S2$, $S1$ or $S3$.

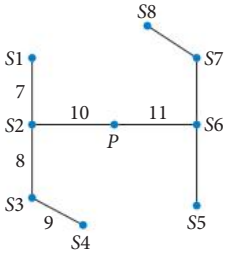


9780170250283

Choose the shortest edge connecting P, $S2$, $S1$, $S3$ or $S4$.



Continue this process until all vertices are connected.



Add all the edges together in the minimum spanning tree to get the minimum length of cable required.

$7 + 8 + 9 + 10 + 11 + 8 + 10 + 9$
$= 72$ m

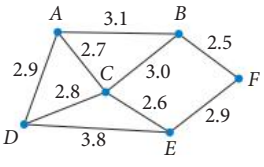
EXERCISE 5.10 Weighted graphs and minimum spanning trees

WS Minimum spanning trees

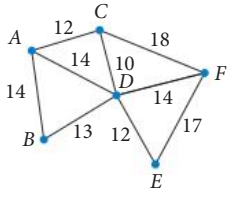
Concepts and techniques

1 For each of the following weighted graphs, determine the shortest distance from A to F.

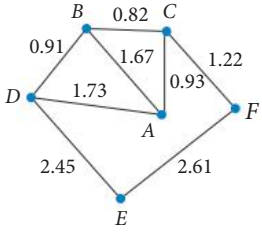
a



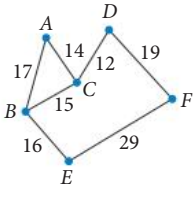
b



c

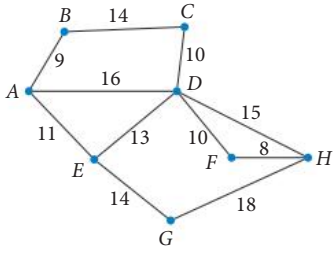


d

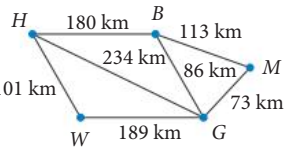


2 Referring to the network on the right, which of the following paths is the shortest?

A $C - B - A - D - H$
B $G - E - D - F - H$
C $A - D - F - H - G$
D $E - D - F - H - G$
E $D - H - G - E - A$

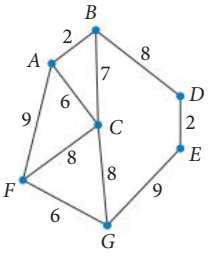


3 Example 26 Using the map of south-west Victoria, decide which is the shortest route from Melbourne to Hamilton.

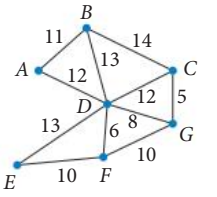


4 For each of the following graphs
i draw the minimum spanning tree.
ii calculate the total length of this minimum spanning tree.

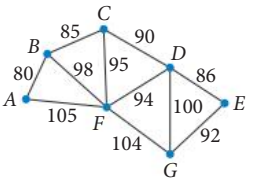
a



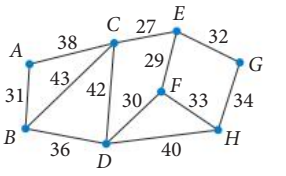
b



c



d



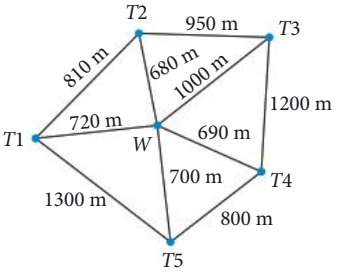
Reasoning and communication

5 The following table represents distances between six towns.

	Moonville	Sunnyvale	Daytown	Nightown	Cloudon	Windsor
Moonville	-	3.5 km	-	2.8 km	-	8.2 km
Sunnyvale	3.5 km	-	3.4 km	-	4.1 km	-
Daytown	-	3.4 km	-	5.3 km	-	5.9 km
Nightown	2.8 km	-	5.3 km	-	1.2 km	-
Cloudon	-	4.1 km	-	1.2 km	-	-
Windsor	8.2 km	-	5.9 km	-	-	-

a Draw a weighted graph representing this information.
b Find the shortest route from Moonville and Daytown.

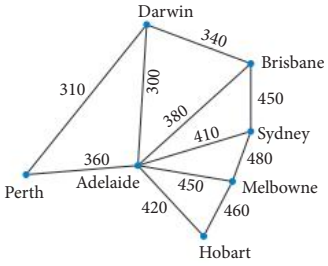
6 Example 27 Elijah wishes to connect plumbing to five toilet blocks in a national park. Use Prim's algorithm to determine the least amount of piping required to connect water from the water source (W) to the five toilet blocks.



9780170250283

7 Airsafe Airlines services five cities. To minimise the number of routes, they need to maximise the total number of passengers they can carry.

The following weighted graph illustrates the current services offered, with the edges representing routes and the quantities on the edges representing the maximum number of passengers that can travel on that route.



Find the maximum spanning tree that meets the airline's requirements and calculate the total number of passengers that can travel on these routes.

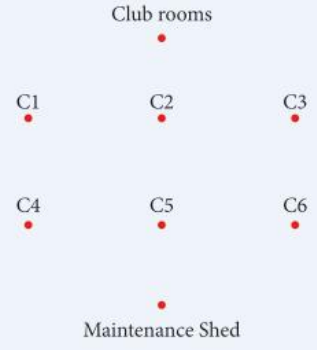
INVESTIGATION Tennis club lighting cable network

Springvale Tennis Club is holding a twilight tournament. At the moment, the courts have no lighting, therefore the club president, Therese, needs to connect electricity to each court.

Below is a table indicating the distances between each court.

A dash indicates that no overhead cables can be connected between the two locations.

	Court 6	Court 5	Court 4	Court 3	Court 2	Court 1	Maint. Shed
Club rooms	25 m	–	24 m	15 m	10 m	14 m	–
Maint. Shed	–	13 m	–	–	–	–	
Court 1	–	18 m	–	–	17 m		
Court 2	–	–	–	16 m			
Court 3	–	19 m	–				
Court 4	–	17 m					
Court 5	16 m						



1 Using the information in the above diagram, complete the graph showing the distances between each court.

2 Therese doesn't want any cables crossing over. What is the minimum number of cables that need to be removed so that no cables cross over? Which cables are these?

3 Therese has now realised that she doesn't need all the cables in order to have electricity connected to all locations. She needs help reviewing the original network to calculate the minimum length of cable required to provide electricity to all locations.

   a Use Prim's algorithm to find the shortest length of cable required to get electricity to every location. Draw this as a minimum spanning tree.

   b What is the length of this minimum spanning tree?

4 The electrician has found that there isn't enough electricity being generated from the clubrooms to light all of the courts at once. He has installed an electricity generator at the maintenance shed. The generator has enough electricity to provide lighting for two courts.

   a Which two courts should be connected to this generator in order to minimise the total cable length used overall?

   b Draw a network illustrating how the cables can be connected to the courts from the clubrooms and maintenance shed to minimise the total length of cable used.

   c How much cable has been saved?



Shutterstock.com/Chad McDermott

CHAPTER SUMMARY
UNDIRECTED GRAPHS

5

- A graph is a diagram that consists of a set of vertices (points) that are connected by a set of edges (lines).
- When two or more edges connect the same pair of vertices, the graph is said to have multiple edges. A loop is an edge that connects a vertex to itself. A vertex that is not connected to any other vertex in a graph is referred to as an isolated vertex.
- There are many different types of graphs, depending on the composition of vertices and edges. These graphs include: a simple graph, connected graph, complete graph, subgraph, directed graph, and a bipartite graph.
- An adjacency matrix is a way of representing graphs. The number of edges connecting the vertices is placed at the intersection of the corresponding row and column.
- A planar graph is a graph that has no edges that cross. A planar graph is defined by Euler's formula:

  vertices = edges – faces + 2

  $V = E - F + 2$
- A walk is the sequence of vertices such that from each vertex there is an edge to the next vertex in the sequence. A walk can include repeated vertices or repeated edges.
- A path is a type of walk where all the edges and vertices are different. A path that starts and finishes at different vertices is defined as an open path, whereas a path that starts and finishes at the same vertex is said to be a closed path or cycle.
- A walk is a sequence of vertices where each vertex shares an edge with the next vertex in the sequence. A walk can include repeated vertices or repeated edges. An open walk is a sequence of vertices where you start at one vertex and finish at a different one, whereas, a closed walk must start and finish at the same vertex.
- A trail is a special type of walk where you can only travel along an edge once. An open trail must travel along every edge only once and a closed trail must travel along every edge only once and return to the starting vertex.
- A connected graph is said to be Eulerian if it contains a closed trail, i.e., it starts and finishes at the same vertex, and it includes every edge once only. A connected graph is called semi-Eulerian if it contains an open trail, i.e., it starts and finishes at different vertices and it includes every edge once only. Eulerian and semi-Eulerian graphs may contain trails that include repeated vertices but not repeated edges.
- A Hamiltonian path is a path that includes every vertex in a graph once only and starts and finishes at different vertices. A Hamiltonian cycle is a cycle that includes each vertex in a graph once only and starts and finishes at the same vertex.
- A tree is a connected graph that cannot contain any loops, multiple edges, or cycles. A spanning tree is a subgraph of another graph that includes all vertices and contains the same properties as a tree.
- A weighted graph is a type of network where extra information such as distances, times or costs are labelled on the edges.

- Spanning trees can be created from weighted graphs to find the minimum distance between all vertices, this is known as the minimum spanning tree.
- Prim's algorithm is used to determine the minimum spanning tree.

Step 1: Choose any vertex within the network.

Step 2: Find the edge with the lowest value connecting this vertex to another vertex.

Step 3: Look at all edges connecting these two vertices and choose the edge that has the lowest value.

Step 4: Look at all the vertices covered so far and select the edge with the lowest value.

Step 5: Repeat step 4 until all the vertices in the graph are included in the tree.

Note that if there is more than one edge with the lowest value, then just choose either.

9780170250283

CHAPTER REVIEW
UNDIRECTED GRAPHS
5

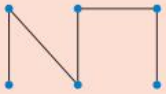
Multiple choice

1 Example 1 The number of vertices and edges in the following graph is:



A Edges = 5 vertices = 5
B Edges = 5 vertices = 8
C Edges = 7 vertices = 5
D Edges = 8 vertices = 5
E Edges = 16 vertices = 5

2 Example 5 The following graph is considered to be

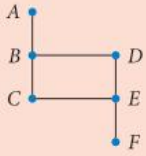


A bipartite
B simple
C complete
D Eulerian
E directed

3 Example 6 A complete graph with 20 vertices contains:

A 10 edges
B 20 edges
C 190 edges
D 200 edges
E 210 edges

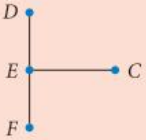
4 Example 7 Which of the following graphs is not a subgraph of the graph directly below?



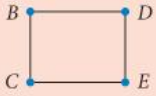
A



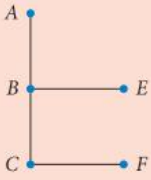
B



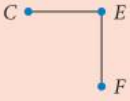
C



D



E

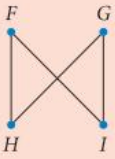


5 Example 8 A local soccer club is holding a tournament. There are five teams and they have to play two other teams according to the following roster.

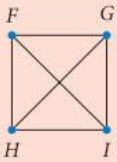
	Frogs	Goannas	Hippos	Iguanas
Frogs		✔		✔
Goannas	✔		✔	
Hippos		✔		✔
Iguanas	✔		✔	

Which of the following graphs represent the roster above?

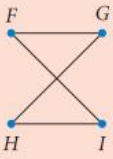
A



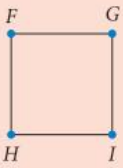
B



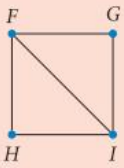
C



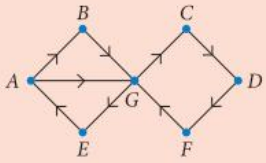
D



E

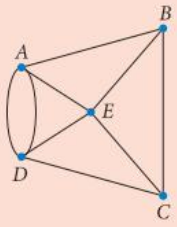


6 Example 10 In the following directed graph, which path should you follow to get from D to A?



A (A, B, G, C, D)
B (A, G, C, D)
C (D, F, A)
D (D, C, B, A)
E (D, F, G, E, A)

7 Example 12 The following graph can be represented by which one of the following adjacency matrices?



A

	A	B	C	D	E
A	0	1	0	2	1
B	1	0	1	0	1
C	0	1	0	1	1
D	2	0	1	0	1
E	1	1	1	1	0

B

	A	B	C	D	E
A	2	1	0	2	1
B	1	0	1	0	1
C	0	1	0	1	1
D	2	0	1	0	1
E	1	1	1	1	0

C

	A	B	C	D	E
A	0	1	0	1	1
B	1	0	1	0	1
C	0	1	0	1	1
D	1	0	1	0	1
E	1	1	1	1	0

D

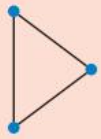
	A	B	C	D	E
A	1	1	0	2	1
B	1	0	1	0	1
C	0	1	0	1	1
D	2	0	1	1	1
E	1	1	1	1	0

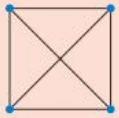
E

	A	B	C	D	E
A	0	1	1	1	1
B	1	0	1	1	1
C	1	1	0	1	1
D	1	1	1	0	1
E	1	1	1	1	0

9780170250283

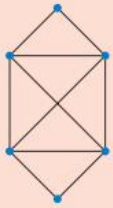
8 Example 13 Which one of the following graphs is not planar?

A 

B 

C 

D 

E 

9 Example 16 Which one of the following is an example of a closed path?

A A salesman visits five businesses, starting and finishing at different businesses.

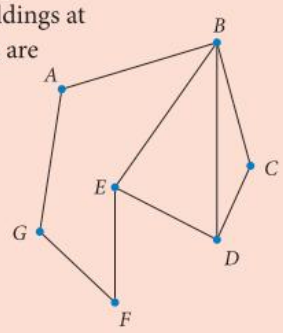
B Mary travels from Melbourne to Mornington to visit a friend and returns home by taking a different route.

C A gym instructor holds a circuit class where each participant trains on each piece of equipment once.

D Five soccer teams play each other in a round robin.

E A walking track that has a dead-end.

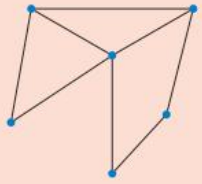
10 Example 17 Margaret needs to place flyers on certain buildings at her university. The buildings and their connecting paths are shown in the diagram below.



Her walk, (A, G, F, E, D, C, B), is an example of a

A closed trail
B cycle
C Eulerian path
D open trail
E planar graph.

11 Example 18 The following graph is considered semi-Eulerian because



A all vertices are of even degree

B two vertices are of odd degree

C the graph contains 6 vertices

D you can travel to each vertex

E a path can contain all vertices once only.

12 Example 20 A Hamiltonian path for the graph directly below could be

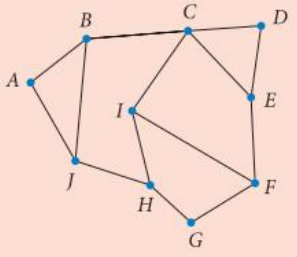
A $(I, H, G, F, E, D, C, B, A, J)$

B $(A, B, C, D, E, F, G, H, J)$

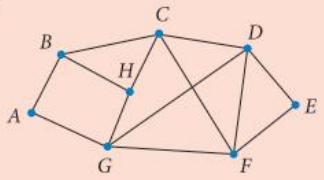
C $(I, H, J, A, B, C, D, E, F, G, H)$

D (F, G, H, I, C, D, E, F)

E (J, B, C, D, E, F, G, H)



13 Example 21 Which of the following is not a Hamiltonian cycle for the graph on the right?

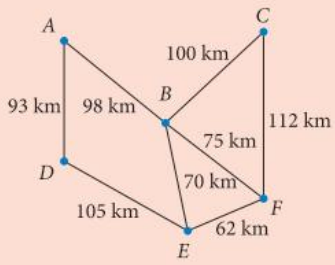


A $(A, B, H, C, D, E, F, G, A)$
B $(H, B, A, G, F, E, D, C, H)$
C $(H, B, C, D, E, F, G, A, H)$
D $(F, C, H, B, A, G, D, E, F)$
E $(E, F, G, A, B, H, C, D, E)$

14 Example 23 Which one of the following statements is true?

A Trees are complete graphs.
B Trees can contain multiple edges.
C Trees must be a subgraph.
D Trees must not contain any cycles.
E An Eulerian graph is also a tree.

15 Example 26 The following graph represents six towns and the roads connecting them. The distance between the towns (via the roads) is written on the edges.

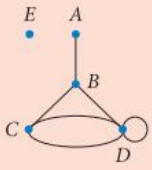


The difference between the longest route and the shortest route between town A and town F (without passing a town more than once) is

A 137 km
B 173 km
C 307 km
D 310 km
E 480 km

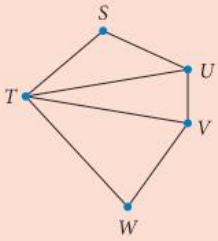
Short answer

16 Example 2 Consider the graph on the right.



a Name the isolated vertex.
b Which vertex is adjacent to A?
c The loop in the graph connects which vertex to itself?
d Which pair of vertices is connected by multiple edges?

17 Example 3 The graph illustrates five towns and the roads connecting them.



a How many edges are connected to town V?
b Which towns are adjacent to town W?
c Which town is directly connected to every other?
d Which town is not directly connected to town U?

18 Example 4 Paths connect the following buildings.
North 1 connects with North 2, North 3, East 1, South 1 and West 1.
East 1 connects with East 2.
South 2 connects with South 1 and West 1.
West 1 connects with West 2.
a List the vertices.
b List the edges.
c Draw a graph representing this information.

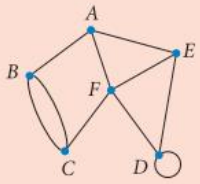
9780170250283

19 Example 9 Two golf teams are playing each other in a competition. Each team consists of four players. Each player must play each player from the opposing team. Represent this information using a graph.

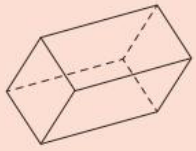
20 Example 11

a Determine the degree of each vertex in the above graph.

b Calculate the degree sum of this graph.



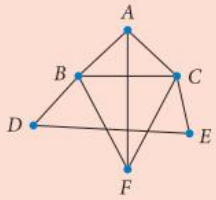
21 Example 14 Represent the following three-dimensional prism as a planar graph.



22 Example 15 Show that the graph on the right is planar by:

a redrawing the graph so that no edges intersect

b using Euler's formula.

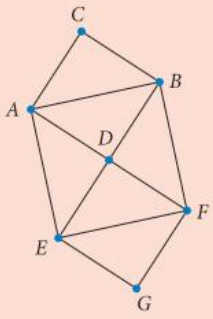


23 Example 19 The graph on the right outlines the roads that Trashtalk Garbage Collection services. The garbage truck needs to collect rubbish in every street, ideally without having to travel along a road more than once. The truck must finish at C, the collection yard.

a Describe a route that the truck can take.

b Explain why this graph is Eulerian and not semi-Eulerian.

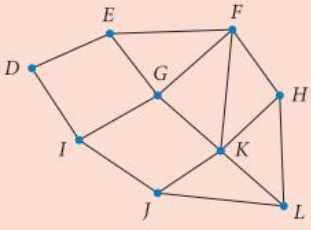
c The company has decided that the truck will continue onto its next route without returning to the collection yard. Will the truck be able to continue from its last street without returning to the collection yard? Explain your answer.



24 Example 22 An electrician needs to check the lighting in several rooms within an office building. The following graph outlines the location of the offices (vertices) and the paths connecting them (edges).

a Describe the route that the electrician must take to check every room and then return to the room that he started at.

b What is this route an example of?

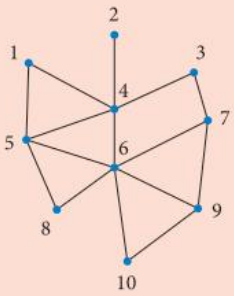


25 Example 26

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

a Draw the graph that corresponds to the above matrix.
b Explain why the graph drawn in part **a** is not a tree.
c Draw a spanning tree for the graph in part **a**.
d Create the adjacency matrix for the spanning tree drawn in part **c**.

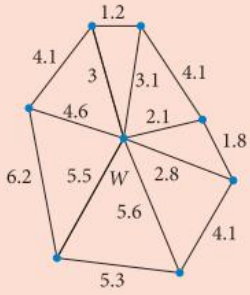
26 Example 25 Maintenance men at a local school have realised that students are getting wet because there are no covered walkways. The following graph outlines the location of the paths and how they connect the classrooms.



a Draw a spanning tree for this graph, showing a possible option for covered walkways.
b According to the spanning tree drawn in part **a**, which paths will not be covered?

27 Example 27 Sophie is installing a water system in her garden. The following graph indicates the location of the garden beds and the distance (in metres) between them.

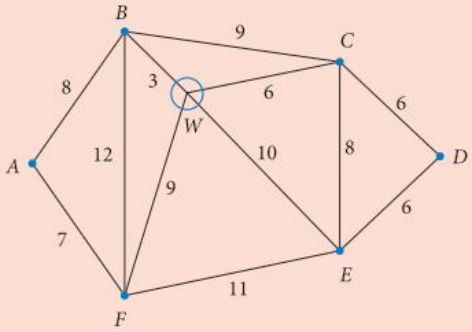
Use Prim's algorithm to help Sophie find the minimum length of pipe required to connect all garden beds to the water source (W).



9780170250283

Application

28 The graph on the right shows the location of a warehouse, W. This warehouse supplies timber to six factories A, B, C, D, E and F. The number along the edges indicates the length (in kilometres) of the connecting roads.



a What is the degree of vertex W?

b A delivery van is at vertex A. It must make a delivery to vertex E before returning to the warehouse. Determine the minimum distance travelled on this journey.

c A salesman wants to leave the warehouse and visit every factory once before returning to the warehouse.

i What is the mathematical term used to describe the route he is going to take?

ii Define a route that he could take.

d The company wishes to lease an office which is adjacent to two of the factories.

i What impact will these two extra roads have on the degree of the vertices?

The owner wishes to check for any competitors along the route.

ii Will the owner have to complete an open or closed trail if she wants to travel along every edge only once?

Practice quiz





6

TERMINOLOGY

average percentage method
bivariate data
centring
cyclic pattern
independently
irregular fluctuations
forecasting
least-squares regression
median smoothing
moving averages
negative secular trend
outliers
positive secular trend
random variation
seasonal index
seasonal pattern
seasonality
short term fluctuations
smoothing
systematic
time series
time series data
time series plot
trend
unsystematic

Prior learning

TIME SERIES ANALYSIS

TIME SERIES ANALYSIS

6.01 Constructing and describing time series plots
6.02 Smoothing time series data using moving averages
6.03 Smoothing time series data using medians
6.04 Adjusting for seasonality
6.05 Fitting a least-squares regression line to time series data
Chapter summary
Chapter review

DESCRIBING AND INTERPRETING PATTERNS IN TIME SERIES DATA

- construct time series plots (ACMGM087)
- describe time series plots by identifying features such as trend (long term direction), seasonality (systematic, calendar-related movements), and irregular fluctuations (unsystematic, short term fluctuations), and recognise when there are outliers; for example, one-off unanticipated events. (ACMGM088)

ANALYSING TIME SERIES DATA

- smooth time series data by using a simple moving average, including the use of spreadsheets to implement this process (ACMGM089)
- calculate seasonal indices by using the average percentage method (ACMGM090)
- deseasonalise a time series by using a seasonal index, including the use of spreadsheets to implement this process (ACMGM091)
- fit a least-squares line to model long-term trends in time series data. (ACMGM092)

THE DATA INVESTIGATION PROCESS

- implement the statistical investigation process to answer questions that involve the analysis of time series data. (ACMGM093) AC

6.01 CONSTRUCTING AND DESCRIBING TIME SERIES PLOTS

Data observed and recorded at successive intervals of time is referred to as **time series data**. In time series, the independent or explanatory variable is time, which increases in regular increments, such as hours, weeks, months or years. Graphs drawn with time placed on the horizontal axis are called time series plots. The main objective of the time series is to investigate how some variable quantity changes with time.

Time series can be plotted and identified by features such as **trend**, **seasonality**, and **irregular fluctuations**. **Outliers** that represent one-off unanticipated events can occur.

Example 1

Create a time series plot using this table showing births by year in Texas from 1910 to 1930.

Year	Total live births
1910	41 104
1911	40 623
1912	39 200
1913	29 239
1914	41 300
1915	45 601
1916	45 122
1917	40 123
1918	47 563

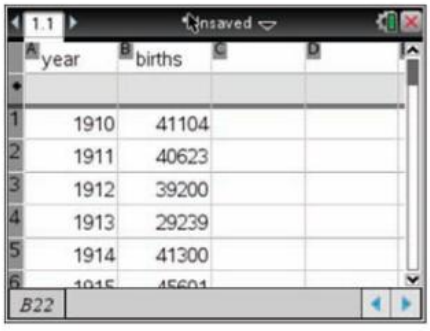
9780170250283

Year	Total live births
1919	48 323
1920	48 124
1921	39 576
1922	45 623
1923	48 789
1924	49 792
1925	50 124
1926	50 678
1927	51 237
1928	53 456
1929	54 789
1930	53 627

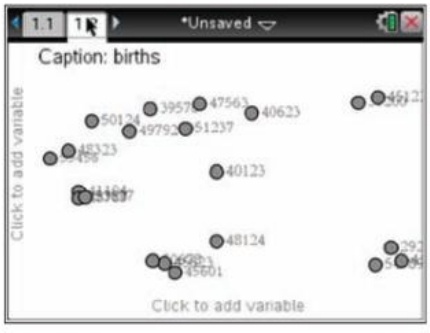
Solution

TI-Nspire CAS

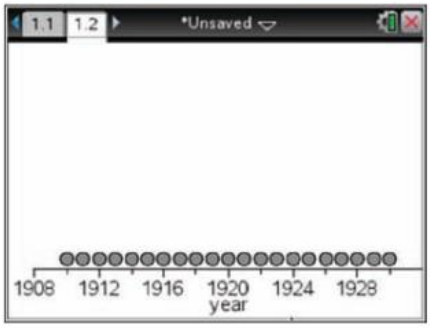
Open a New Document with a Lists & Spreadsheet page.
Enter the data as shown.



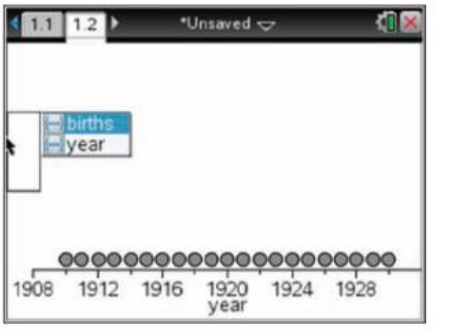
Add a Data & statistics page.

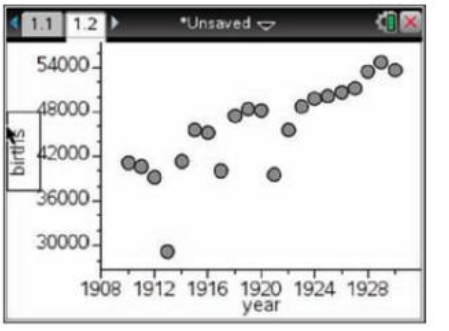


Move the cursor to the 'Click to add variable' box at the bottom of the screen and select year.

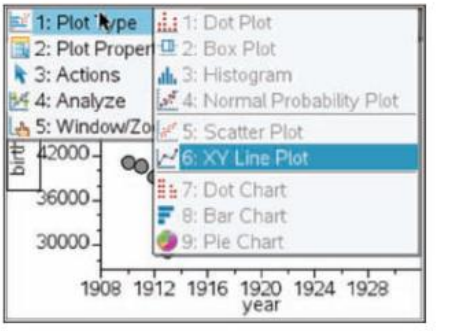


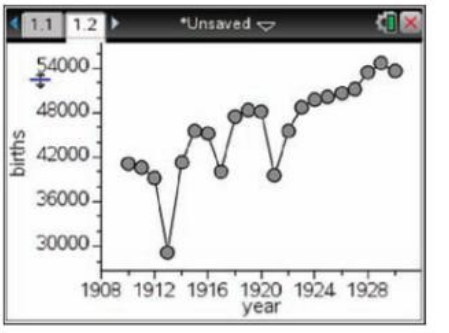
Move the cursor to the 'Click to add variable' box at the left of the screen and select births.





Press menu then 1; Plot Type then 6: XY Line Plot.



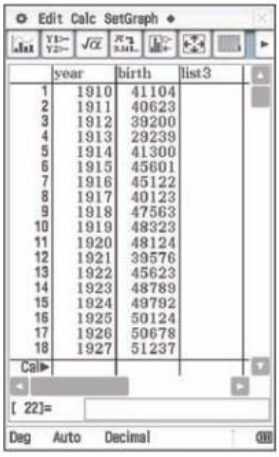


9780170250283

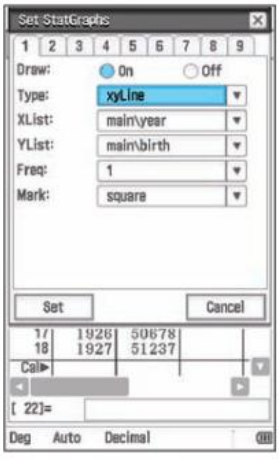
ClassPad

Use the **Statistics** application.

Enter the data as shown.

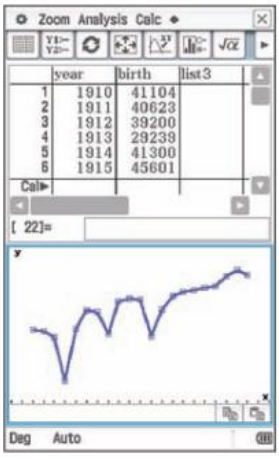


Tap **SetGraph**. Make sure that **StatGraph1** is checked, then tap **Setting…** and complete the screen as shown.
Tap Set.



Tap to draw the graph.

Tap Resize to view just the graph screen.



A time series plot can be used in many areas. Daily minimum (or maximum) temperatures recorded by the weather bureau are examples of time series. Other examples are: the annual sales of a particular brand of shoe in a shoe shop; or the annual production of fluorescent cycling gear in Queensland over a period of years. Once a time series plot is displayed, a description of the plot's characteristic features can be created.

Time series plots are described by identifying features such as trend (long-term direction), seasonality (**systematic**, calendar-related movements), and irregular fluctuations (**unsystematic**, short-term fluctuations) or outliers. The time series plot shown in Example **1** shows an upward trend, but also some dips that could be described by special events.

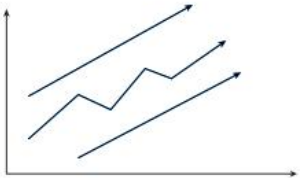
IMPORTANT

Describing a time series plot

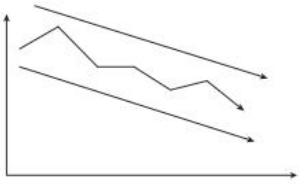
A **time series** can have one or more of the following four components.

- trend
- seasonal pattern
- cyclic pattern
- random variation, also known as random pattern

A **trend** exists if there is a long-term increase or decrease in the dependent variable as time passes. In Example **1**, **Births by year in Texas from 1910 to 1930**, the births steadily increase over the years. The time series plot shows that the trend of the time series data over the period 1910 to 1930 is moving in an overall positive direction. We say that the time series plot has a **positive secular trend**. A simplified version of a **positive secular trend** is illustrated on the right, showing overall positive movement.



A time series plot can also show an overall negative movement. This means that the **time series data** has a **negative secular trend**. A simplified version of a **negative secular trend** is illustrated on the right, showing overall negative movement.



To decide which type of secular trend is the most applicable, make sure that there are enough observations in the plot for any overall movement to emerge.

Example 2

The number of deaths due to polio in an overseas country has decreased steadily from 1950 to 2000, due to an organised vaccine program run by the International Red Cross.

a Use the table below to sketch a time series plot.

Year	Total polio deaths
1950	10 104
1960	10 112
1970	8900
1980	7890
1990	7500
2000	6990

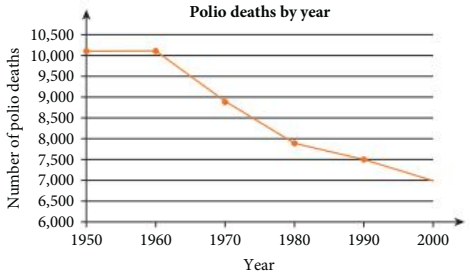
b Determine if the plot shows a positive secular trend, or a negative secular trend.

9780170250283

Solution

a Construct a time series plot by plotting the points (1950, 10 104), (1960, 10 112), etc.

A CAS calculator can be used to construct the time series plot.



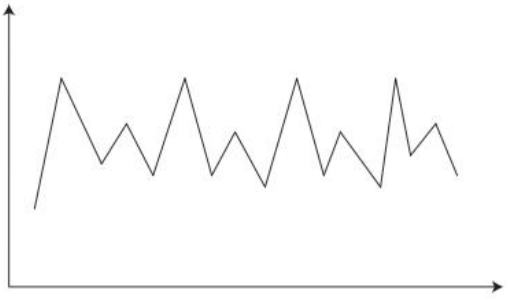
b Determine if the plot shows a positive secular trend, or a negative secular trend.

Write the answer.

The plot is showing overall negative movement.

The time series plot shows a negative secular trend.

A **seasonal pattern** arises when the actual seasons of the year affect sales or production of goods. In such a context, there will be a similar pattern of regular peaks and troughs at corresponding months or quarters of successive years. For example, the sales figures of ice-cream in a beach resort is high each summer, and the time series repeats this pattern each year after a regular interval of time.



A simplified version of a **seasonal variation** is illustrated top right, showing peaks at regular intervals – these are the summer quarters in each year.

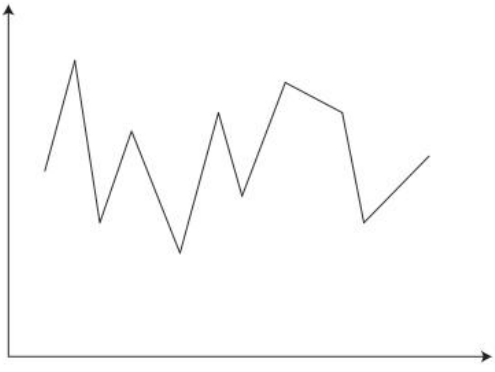
A **cyclic pattern** exists when significant peaks and troughs occur at irregular intervals, usually with a variation in the amplitudes of the peaks and troughs.



For example, the earthquakes recorded each year show significant peaks and troughs, but at unpredictable times. The main difference between cyclic and seasonal patterns is the unpredictability of when peaks and troughs occur.

A simplified version of a **cyclic pattern** is illustrated middle right, showing peaks that are irregular.

Random variations or **random patterns** reflect the unpredictability common in all statistical data. In a **time series plot**, a random pattern would not show predictable peaks and troughs, nor would there be significant peaks and troughs at unpredictable times. Instead, there will be random fluctuations about a relatively stable mean.



A simplified version of a **random pattern** is illustrated above, showing random peaks and troughs.

Time series data

EXERCISE 6.01 Constructing and describing time series plots

Concepts and techniques

1 Example 1 An example of time series data in a table is displayed below using births in the second column and time (in years) in the first column. Create a time series plot using the data listed.

Births by year in Minnesota from 1910 to 1920

Year	Total live births
1910	11 124
1912	12 000
1911	12 200
1913	13 329
1914	13 500
1915	14 005
1916	14 100
1917	14 127
1918	15 163
1919	15 323
1920	16 224

b Describe if the plot shows a positive secular trend, or a negative secular trend.

2 Example 2 The number of deaths in a particular city centre has decreased steadily from 1950 to 2000, due to increased police presence on the streets at night.

a Use the table below to sketch a time series plot.

Year	Total deaths
1950	104
1960	102
1970	89
1980	78
1990	75
2000	65

b Determine if the plot shows a positive secular trend, or a negative secular trend.

3 For each of the following time series situations, which of the components listed is most likely to apply?

a The sales of a new popular ice-cream flavour in a beach resort.

A positive secular trend B negative secular trend C seasonal pattern
D cyclic pattern E random pattern

b The average monthly maximum temperatures in Melbourne over a year.

A positive secular trend B negative secular trend C seasonal pattern
D cyclic pattern E random pattern

9780170250283

c The average monthly rainfall in the Colac Otway rainforest over a year.
A positive secular trend B negative secular trend C seasonal pattern
D cyclic pattern E random pattern

d The daily distances covered by a taxidriver in Sydney over a certain period of time.
A positive secular trend B negative secular trend C seasonal pattern
D cyclic pattern E random pattern

e The share price of a bank, measured weekly over a year.
A positive secular trend B negative secular trend C seasonal pattern
D cyclic pattern E random pattern

f The number of daily births in a particular country hospital over a period of 10 days.
A positive secular trend B negative secular trend C seasonal pattern
D cyclic pattern E random pattern

g The monthly sales data for umbrellas over the years.
A positive secular trend B negative secular trend C seasonal pattern
D cyclic pattern E random pattern

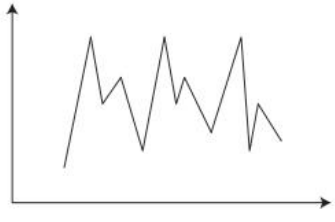
h The cost of living measured over the years.
A positive secular trend B negative secular trend C seasonal pattern
D cyclic pattern E random pattern

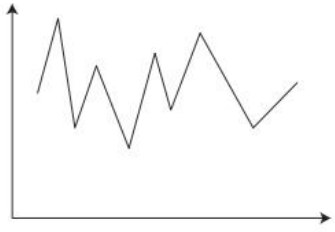
4 The attendance at a holiday resort, over 3 years is shown below.

Time period	2010 Jan–Mar	2010 Apr–June	2010 July–Sept	2010 Oct–Dec	2011 Jan–Mar	2011 Ap–June	2011 July–Sept	2011 Oct–Dec	2012 Jan–Mar	2012 Apr–June	2012 July–Sept	2012 Oct–Dec
Attend in 1000s	12	8	6	11	13	9	5	14	15	9	8	13

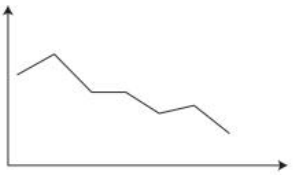
a Plot the data on suitable axes.
b Comment on any trend shown.

5 For each of the following time series plots, which of the components listed is most likely to apply?

a

A positive secular trend B negative secular trend C seasonal pattern
D cyclic pattern E random pattern

b

A positive secular trend B negative secular trend C seasonal pattern
D cyclic pattern E random pattern

c



A positive secular trend B negative secular trend C seasonal pattern
D cyclic pattern E random pattern

d



A positive secular trend B negative secular trend C seasonal pattern
D cyclic pattern E random pattern

Reasoning and communication

6 Draw a time series plot of the following data and describe the time series plot in terms of its trend, seasonal, cyclic or random patterns.

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

7 The data in the graph represents the number of serious injuries from bicycle accidents which took place in Tasmania over a period of 24 months, starting at January, (2010–2011).



a Is there a secular trend in the time series plot? If so, what kind?

b Is there a seasonal, cyclic or random pattern present in time series plot?

8 The sales of life jackets (in thousands) by a company making and selling safety gear are shown below for the Years 2000–2012. Plot the data and identify any trends and outliers.

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Sales	17.2	18.1	18.6	19.3	20.2	20.8	21.6	22.3	18.5	22.6	24.4	25.1

9 The data below shows the number of births (in thousands) in Australia from 2009–2011. Plot the data and identify any trends.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2009	23.9	23.2	25.1	24.6	24.3	24.1	25.2	24.1	25.2	25.6	23.6	24.4
2010	24.5	23.1	26.3	24.3	25	24.7	24.9	24.6	24.7	24.5	22.9	23.1
2011	24.6	23.2	26.1	24.4	25.1	24.6	24.4	24.7	24.9	24.6	23.7	23.8

9780170250283

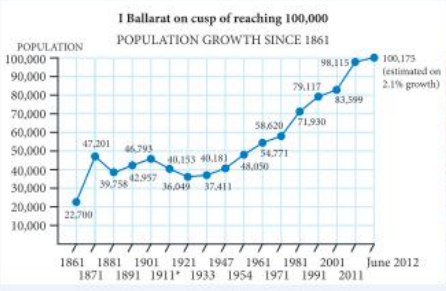
6.02 SMOOTHING TIME SERIES DATA USING MOVING AVERAGES

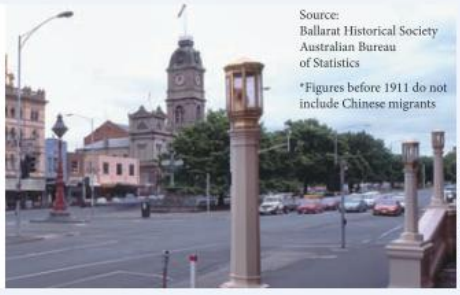
Smoothing time series data is a technique used to level out fluctuations and hence show a clearer picture of an overall trend. When a trend line is found for a set of **time series data**, it is important to be able to rely on the accuracy of the prediction gained from the trend line. Time series data can fluctuate from point to point; meaning that at times a trend line cannot accurately predict the future if there is a large variation in how the data moves.

INVESTIGATION Population growth in Ballarat, Victoria

Ballarat's population is reported to be on the cusp of 100 000, according to data released in 2011 by the Australian Bureau of Statistics. Almost 2000 people made Ballarat their home between 2009 and 2010, the 14th fastest growth rate of any local government area in Victoria. At 30 June 2010, an estimated 96 000 lived in Ballarat, up from 94 000 in 2009 and an estimated 92 000 in 2008. If the trend continued, Ballarat should have reached a six-figure population by June 2012.

Using the graph below, investigate the type of time series represented and attempt to find a prediction for the population in 2013, 2014, 2015, …





Source: Jordon Oliver, "Ballarat Population nears 100,000" *The Courier*, 22 April 2011

To smooth time series data, two techniques are used:

- smoothing using **moving averages** (means)
- **median smoothing**

An example to demonstrate this is given below.

The following table shows the number of students in a Year 12 Mathematics class over the period 2003 to 2012.

IMPORTANT

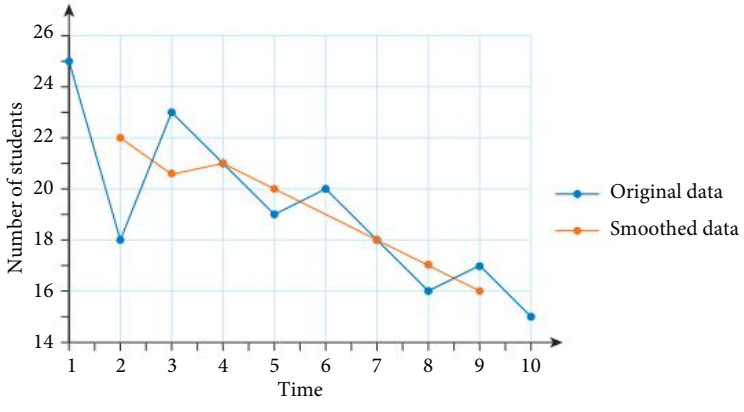
Smoothing using **moving averages** takes a fixed number of data points and averages them to give a smoothed data point.

Year, t	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Number of students, y	25	18	23	21	19	20	18	16	17	15

We will smooth this data to determine the long-term trend in the number of students in a Year 12 Mathematics class. To use a three-point moving average, we need three successive data points to create a smoothed data value. Moving through the data by always using three points at a time means we end up with a new set of smoothed data.

Year, t	Number of students, y	Three-point moving average
1	25	
2	18	$\frac{25+18+23}{3}=22$
3	23	$\frac{18+23+21}{3}=20.67$
4	21	$\frac{23+21+19}{3}=21$
5	19	$\frac{21+19+20}{3}=20$
6	20	$\frac{19+20+18}{3}=19$
7	18	$\frac{20+18+16}{3}=18$
8	16	$\frac{18+16+17}{3}=17$
9	17	$\frac{16+17+15}{3}=16$
10	15	

We can then draw a time series plot with the original and the smoothed data, using the three-point moving average, to see the visual effect that the smoothing has on the data. The trend becomes clear.



This is called a 3-point moving average. Notice that in this method we lose two data points as we do not have the before and after points to use for the 3-point data smoothing.

Example 3

The introduction of speed cameras in Victoria reduced the number of people killed in road accidents over the period 1990 to 2000. This data is tabulated below.

Year, t	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Number of road deaths, y	225	240	201	192	185	160	172	127	132	101	100

a Complete a table using the method of 3-point moving averages, using $t = 1$ to represent 1990 to $t = 11$ to represent 2000.

b On separate axes, graph the original data and the smoothed data.

c What does the graph of the smoothed data suggest about the trend in the data?

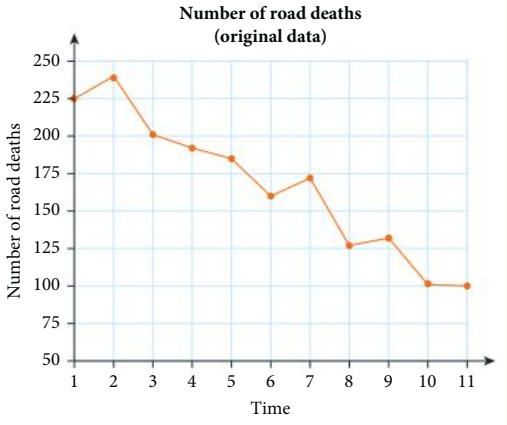
Solution

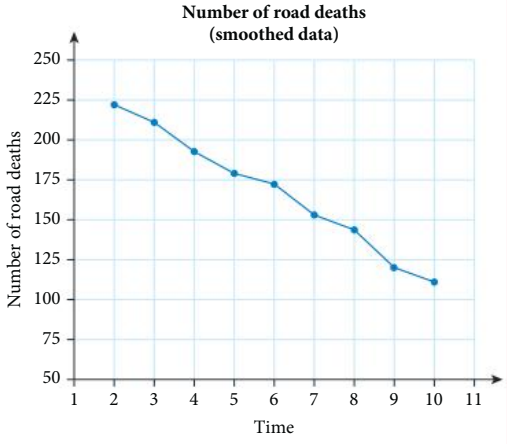
a Set up a table with 3 columns to list Year, t, in list 1, Number of road deaths, y, in list 2 and calculations for three-point moving averages in list 3.

Year, t	Number of road deaths, y	Three-point moving average
1	225	
2	240	$\frac{225+240+201}{3} = 222$
3	201	$\frac{240+201+192}{3} = 211$
4	192	$\frac{201+192+185}{3} = 192.67$
5	185	$\frac{192+185+160}{3} = 179$
6	160	$\frac{185+160+172}{3} = 172.33$
7	172	$\frac{160+172+127}{3} = 153$
8	127	$\frac{172+127+132}{3} = 143.67$
9	132	$\frac{127+132+101}{3} = 120$
10	101	$\frac{132+101+100}{3} = 111$
11	100	



b Graph the original data and the smoothed data on separate axes.





c Write the answer.

The graph of the smoothed data suggests a negative secular trend.

Another way to smooth data is to use a two or four-point moving average with **centring**. Smoothing with an odd number of points gives a smoothed data point which is located at the middle time, that is, the time value of the data point that corresponds to one of the original times.

Year, t	Number of road deaths, y	Three-point moving average
1	225	
2	240	$\frac{225+240+201}{3} = 222$
3	201	

Centred time value

Smoothed data value for $t = 2$

In Example **3**, using the three-point moving average, the first three points are centred on $t = 2$.

In contrast, when an even number of points is used to smooth the data, there is a problem finding the centred time value. The centre of the set of points is not one of the original points. In this situation, we do another step by taking a 2-moving average of the already smoothed values, so they line up with original time values. This is called **centring**.

We again consider the data from Example **3**, but this time we use a four-point moving average.

9780170250283

Year, t	Number of road deaths, y	Four-point moving average	4-moving average with centring
1	225		
2	240		
		$\frac{225+240+201+192}{4}=214.5$	
3	201		$\frac{214.5+204.5}{2}=209.5$ ← smoothed data value for $t=3$
		$\frac{240+201+192+185}{4}=204.5$	
4	192		$\frac{204.5+184.5}{2}=194.5$
		$\frac{201+192+185+160}{4}=184.5$	
5	185		$\frac{184.5+177.25}{2}=180.875$
		$\frac{192+185+160+172}{4}=177.25$	
6	160		$\frac{177.25+161}{2}=169.125$
		$\frac{185+160+172+127}{4}=161$	
7	172		$\frac{161+147.75}{2}=154.375$
		$\frac{160+172+127+132}{4}=147.75$	
8	127		$\frac{147.75+133}{2}=140.375$
		$\frac{172+127+132+101}{4}=133$	
9	132		$\frac{133+115}{2}=124$
		$\frac{127+132+101+100}{4}=115$	
10	101		
11	100		



Example 4

a Perform a centred four-point moving average to smooth the data given on the number of births per month over a calendar year in a small country hospital.

b Graph the smoothed data in a time series plot.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
Number of births	10	12	6	5	22	18	13	7	9	10	8	15

Solution

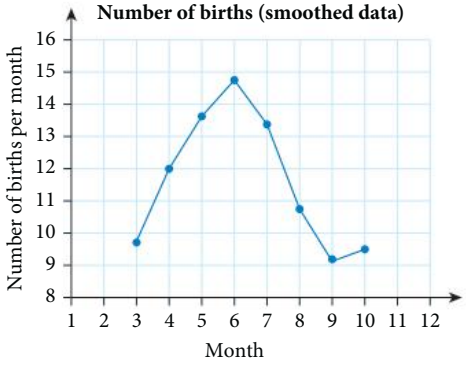
a Set up a table with 4 columns for centring of 4-point moving average.

Month, t	Number of births, y	Four-point moving average	4-moving average with centring
Jan	10		
Feb	12		
		$\frac{10+12+6+5}{4}=8.25$	
Mar	6		$\frac{8.25+11.25}{2}=9.75$
		$\frac{12+6+5+22}{4}=11.25$	
Apr	5		$\frac{11.25+12.75}{2}=12$
		$\frac{6+5+22+18}{4}=12.75$	
May	22		$\frac{12.75+14.5}{2}=13.625$
		$\frac{5+22+18+13}{4}=14.5$	
June	18		$\frac{14.5+15}{2}=14.75$
		$\frac{22+18+13+7}{4}=15$	
July	13		$\frac{15+11.75}{2}=13.375$
		$\frac{18+13+7+9}{4}=11.75$	
Aug	7		$\frac{11.75+9.75}{2}=10.75$
		$\frac{13+7+9+10}{4}=9.75$	
Sept	9		$\frac{9.75+8.5}{2}=9.125$
		$\frac{7+9+10+8}{4}=8.5$	

9780170250283

Month, t	Number of births, y	Four-point moving average	4-moving average with centring
Oct	10		$\frac{8.5+10.5}{2}=9.5$
		$\frac{9+10+8+15}{4}=10.5$	
Nov	8		
Dec	15		

c Graph the smoothed data.



EXERCISE 6.02 Smoothing time series data using moving averages

Smoothing time series data using moving averages

Concepts and techniques

1 Example 3 The three-point moving average can be found by

A using three successive data points to create a smoothed data value
B using five successive data points to create a smoothed data value
C using four successive data points to create a smoothed data value
D using three successive data points with centring to create a smoothed data value
E using three successive data points with centring to create a smoothed data value

2 Example 3 The data below shows the number of female students in a Year 12 Calculus class over the period 1960 to 2010, shown in five-year periods.

Year, t	1960	1965	1970	1975	1980	1985	1990	1995	2000	2005	2010
Number of female students, y	2	1	3	5	9	10	4	6	9	12	15

a Use a three-point moving average to smooth the data.
b Comment on the trend of the smoothed data.

3 The data set below shows English examination percentage results for a student over six years of schooling.

Year, t	1	2	3	4	5	6
Student's result, y%	94	54	35	61	88	75

Find

a the 3-point moving average y value for $t = 2$
b the 3-point moving average y value for $t = 4$
c the 3-point moving average y value for $t = 5$
d the 5-point moving average y value for $t = 3$
e the 2-point moving average y value for $t = 3$
f the 2-point moving average y value for $t = 5$
g the 4-point moving average y value for $t = 3$
h the 4-point moving average y value for $t = 4$

4 Example 4 a Use a centred four-point moving average to smooth the number of sales of textbooks from 2000 to 2012, matching the years, t.

Year, t	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Sales of textbooks, y	2250	2230	2000	2010	1990	3000	2045	2989	3000	1950	2120	2255	2297

b Graph the smoothed data in a time series plot.

5 The entry numbers at the local swimming pool over a 3-month summer period have declined over recent years and the council is deciding whether to close the pool or not. The data they have collected is below.

Year, t	2004	2005	2006	2007	2008	2009	2010	2011	2012
Entries at pool, y	1000	889	768	876	556	432	300	498	228

a Use a two-point moving average, with centring, to smooth the number of entries recorded at the local swimming pool over a 3-month summer period from 2004 to 2012.
b Graph the smoothed data in a time series plot.
c Comment on the trend to decide if the councillors need to be worried about attendance numbers.

Reasoning and communication

6 The rainfall for 12 consecutive days is given below.

Day, t	Rainfall, y (mm)
1	12
2	14
3	15
4	13
5	11
6	10
7	16
8	15
9	14
10	12
11	15
12	19

a Draw a time series plot of the data.
b Complete the table using three-point and five-point moving averages to smooth the data.
c Plot the three-point and five-point smoothed values on the original time series plot.

Find

d the 3-point moving average y value for $t = 2$
e the 3-point moving average y value for $t = 10$
f the 5-point moving average y value for $t = 5$
g the 5-point moving average y value for $t = 9$
h What conclusion can be made about three and five-point moving averages in terms of smoothing?

9780170250283

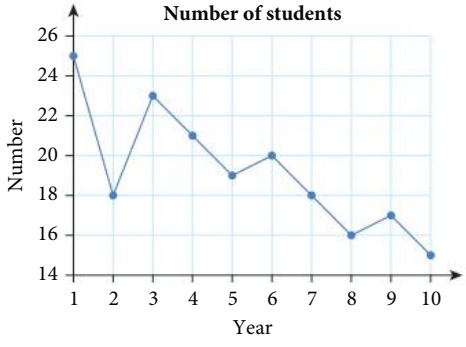
6.03 SMOOTHING TIME SERIES DATA USING MEDIANS

Median smoothing uses a similar technique to moving averages; using the median values instead of the average (mean) values. The process for using an odd number of points is very similar. An even number of points, as before, needs the method of **centring** for median smoothing. A graphical as well as a numerical approach can be used.

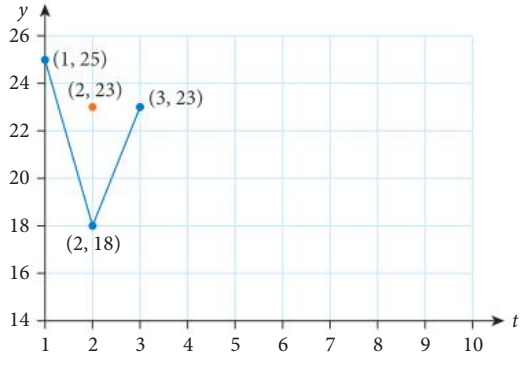
To demonstrate median smoothing, we use again the example of the number of students in a Year 12 Mathematics class over the period 2003 to 2012.

Year, t	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Year, t	1	2	3	4	5	6	7	8	9	10
Number of students, y	25	18	23	21	19	20	18	16	17	15

To use three-point median smoothing, we use three successive data points to create a smoothed data value. The times series plot is shown.

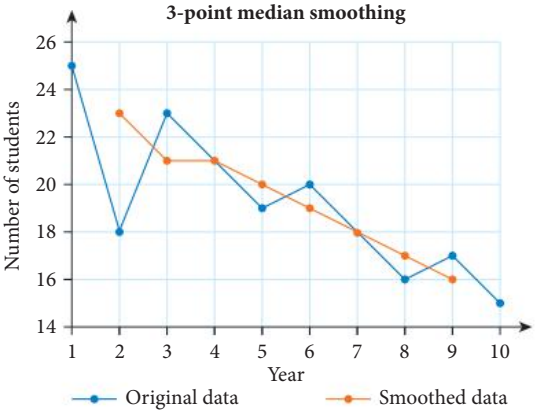


We can smooth the data graphically without formal calculation. Consider the first three points in the times series plot: (1, 25), (2, 18) and (3, 23).



Horizontally, the median is at the middle t value ($t = 2$). Vertically, the median is at the middle y value ($y = 23$). Place a point at these values.

Continue in this way, taking the next three points (corresponding to $t = 2, 3, 4$). When all the smoothed points have been found, join them to produce the smoothed time series plot. The result is shown.



We could also have smoothed the data using a numerical approach.

Year, t	Number of students, y	Three-point moving median
1	25	
2	18	median of 25, 18, 23 = 23
3	23	median of 18, 23, 21 = 21
4	21	median of 23, 21, 19 = 21
5	19	median of 21, 19, 20 = 20
6	20	median of 19, 20, 18 = 19
7	18	median of 20, 18, 16 = 18
8	16	median of 18, 16, 17 = 17
9	17	median of 16, 17, 15 = 16
10	15	

This gives the following list of number pairs.

(2, 23), (3, 21), (4, 21), (5,20), (6, 19), (7,18), (8, 17), (9, 16)

To perform median smoothing for an even number of points, we need to centre the data by taking the average of each pair of smoothed points. This allows us to more easily plot a trend line and use it to make predictions.

9780170250283

Example 5

An ice-cream van by the beach is trying to predict its sales figures for 2003. The manager of the van has his sales figures for the years 1993 to 2002.

Year, t	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Sales, y ($1000s)	7.3	13.8	11.4	11.2	16	12.2	10.9	15.3	12.7	16

a Perform a centred four-point moving median smoothing.
b Graph the original data and smoothed data in a time series plot.
c Find the equation of the least-squares regression line that fits the smoothed data. Write the constants correct to 3 decimal places.
d Use the equation of the trend line to predict sales for 2003.

Solution

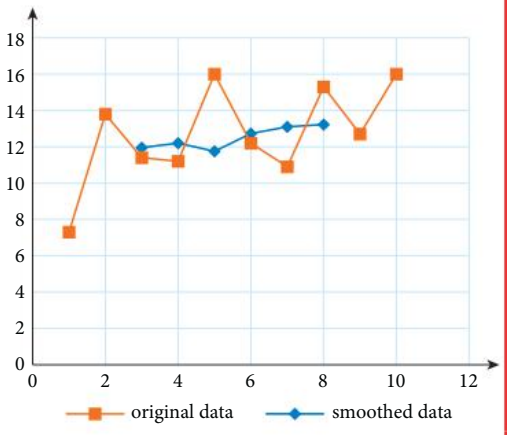
a Set up a four-point moving median table.
Let 1993 be Year 1, 1994 be Year 2 and so on.
Find the four-point moving medians.
Find the centred four-point moving medians.

Year, t	Sales, y ($1000)	Four-point moving median	Centred
1	7.3		
2	13.8		
		Median of (7.3, 13.8, 11.4, 11.2) = 11.3	
3	11.4		(11.3 + 12.6) ÷ 2 = 11.95
		Median of (13.8, 11.4, 11.2, 16) = 12.6	
4	11.2		(12.6 + 11.8) ÷ 2 = 12.2
		Median of (11.4, 11.2, 16, 12.2) = 11.8	
5	16		(11.8 + 11.7) ÷ 2 = 11.75
		Median of (11.2, 16, 12.2, 10.9) = 11.7	
6	12.2		(11.7 + 13.75) ÷ 2 = 12.725
		Median of (16, 12.2, 10.9, 15.3) = 13.75	



Year, t	Sales, y ($1000)	Four-point moving median	Centred
7	10.9		(13.75 + 12.45) ÷ 2 = 13.1
		Median of (12.2, 10.9, 15.3, 12.7) = 12.45	
8	15.3		(12.45 + 14) ÷ 2 = 13.225
		Median of (10.9, 15.3, 12.7, 16) = 14	
9	12.7		
10	16		

b Construct a time series plot using the original data.
Plot the smoothed data on the same plot.



c Use the smoothed data.

Year, t	Sales, y
3	11.95
4	12.2
5	11.75
6	12.725
7	13.1
8	13.225

Use a CAS calculator to find the equation of the least-squares regression line with Year (t) as the independent variable and Sales (y) as the dependent variable.

Write the equation of the least-squares regression line with constants correct to 3 decimal places.

$y = 0.287t + 10.912$

9780170250283

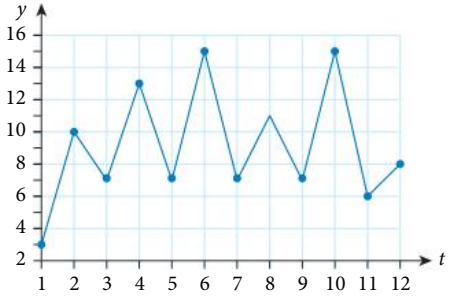
d	Determine the value of t for the year 2003.	Since $t = 1$ for 1993, $t = 11$ for 2003.
	Substitute $t = 11$ into the least-squares regression line equation.	$y = 0.287t + 10.912$ $= 0.287 \times 11 + 10.912$ $= 14.069$
	Write the answer, remembering that sales are given in \$1000s.	The sales for 2003 are predicted to be $14.069 \times \$1000 = \$14\,069$.

EXERCISE 6.03 Smoothing time series data using medians

Smoothing time series data using medians

Concepts and techniques

1 Consider the time series plot below with the starting point at (1, 3), ending at (12, 8).



a The number of points represented in the graph is

A 4 **B** 6 **C** 8 **D** 10 **E** 12

b The number of line segments represented in the graph is

A 3 **B** 5 **C** 7 **D** 9 **E** 11

c The time series plot has a

A positive secular trend **B** negative secular trend **C** seasonal pattern
D cyclic pattern **E** random pattern

d The median of the first 3 points is

A (2, 7) **B** (2, 3) **C** (6.5, 9) **D** (11, 8) **E** (10, 15)

e The median of the last 3 points is

A (2, 11) **B** (2, 3) **C** (6.5, 9) **D** (11, 8) **E** (10, 15)

f The median of the middle 4 points is

A (2, 11) **B** (2, 3) **C** (6.5, 9) **D** (11, 8) **E** (10, 15)

2 List the points displayed in the graph in question **1**.

3 Copy the graph in question **1** and draw a smoothed time series plot using the graphical approach to three-point median smoothing.

4 Example 5 For the data below, what would the smoothed value be at $t = 3$, if a four-point moving median with centring was used?

Day, t	1	2	3	4	5
Interest rate (%)	3.2	4.1	4.6	3.25	4.5

5 For the data below, what would the smoothed value be at

a $t = 3$ b $t = 4$ c $t = 5$

if a three-point moving median was used?

Day, t	1	2	3	4	5	6
Entry to museum	320	420	460	325	450	421

Reasoning and communication

6 A political party is trying to predict its voting numbers in Victoria for 2013. The party has polling figures for the years 1993 to 2012.

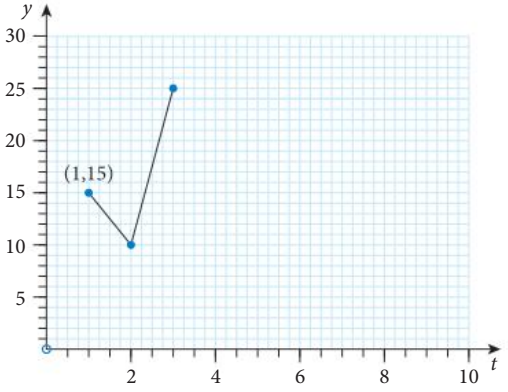
Year, t	1993	1994	1995	1996	1997	1998	1999	2000	2001	2012
votes, y (% per pop'n)	71	42	12	20	16	62	51	32	52	22

Using a four-point moving median with **centring**, smooth the data and use it to predict polling figures for 2013.

7 Consider the data in the following table, which represents the number of births each month in a city hospital.

Month, t	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
Births, y	15	10	25	20	25	27	21	16	21	24

Using $t = 1$ for March, and so on, the first three terms of the time series plot are shown below.



a Copy and complete the time series plot.

b Without formal calculation, smooth the time series plot using a three-point moving median.

c Comment on the trend of the smoothed data.

8 A dress shop is trying to predict its sales figures for 2013. The accountant for the shop views her sales figures for the years 2005 to 2012.

Year, t	2005	2006	2007	2008	2009	2010	2011	2012
Sales, y ($100 000s)	10.6	12.2	15.8	12.8	10.6	16.3	12.7	16

a Perform a centred four-point moving median smoothing.

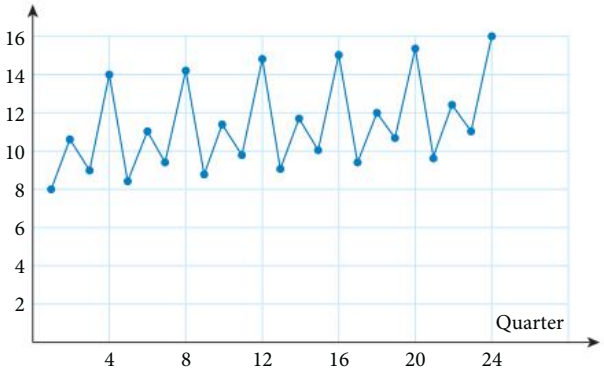
b Graph the original data and smoothed data in a time series plot.

c Find the equation of the least-squares regression line that fits the smoothed data. Write the constants correct to 3 decimal places.

d Use the equation of the trend line to predict sales for 2013.

6.04 ADJUSTING FOR SEASONALITY

Seasonal data shows a regular cyclic pattern that repeats itself at regular intervals and is of similar intensity. In the example shown, there is also a slight upward trend. In the given time series plot, the quarterly time series has a regular pattern that repeats itself every four quarters.



Seasonal indices can be calculated by using a method called the **average percentage method**. Seasonal indices are used to deseasonalise a time series.

Seasonal factors are things such as the weather, the day of the week, the month of the year, or the quarter of the year. For example, ice-cream sales are greater in summer than in winter, retail sales increase during December and the post-Christmas period. To smooth seasonal data, using the methods of a moving average or median smoothing may not be effective because smoothing two consecutive points might connect two data points with significant seasonal differences.

Deseasonalisation is a form of smoothing, which takes out the seasonal effects of the data so that a trend line can be fitted and long-term trends can be predicted.

IMPORTANT

Steps to deseasonalising data

1. Find the average for each year (= **yearly average**).
2. Divide each month by the yearly average.
3. Find the average of each month (= **seasonal index** for each month).
4. Divide original data by its seasonal index.

The formula for calculating deseasonalised values is:

$$\text{Deseasonalised value} = \frac{\text{Actual value}}{\text{Seasonal index}}$$

To explain this process, we can use data provided in Example **6** as quarterly figures.

Step 1: Calculate the yearly average. Calculate the totals for each year, and find the average for each year by dividing each total by 4.

Step 2: Calculate the quarterly proportions. Divide each quarterly sales figure by the corresponding yearly average to obtain quarterly proportions.

Step 3: Calculate the seasonal indices (SI). Seasonal indices (SI) are obtained by finding the average of the quarterly proportions.

Step 4: Deseasonalise the original time series data using the formula

$$\text{Deseasonalised value} = \frac{\text{Actual value}}{\text{Seasonal index}}$$

Example 6

The quarterly sales figures (number of cars sold) were recorded by a car sales yard for 2010 to 2012.

Year	Sales figures quarter 1	Sales figures quarter 2	Sales figures quarter 3	Sales figures quarter 4
2010	5	7	9	3
2011	4	8	9	4
2012	5	9	10	5

a Calculate the seasonal index for each quarter.
b Deseasonalise the data.
c Plot the original and the deseasonalised data on the same set of axes.

Solution

a **Step 1:** Calculate the yearly average. Calculate the totals for each year, and find the average for each year by dividing each total by 4.

Year	Q1	Q2	Q3	Q4	Yearly average
2010	5	7	9	3	$\frac{5+7+9+3}{4}=6$
2011	4	8	9	4	$\frac{4+8+9+4}{4}=6.25$
2012	5	9	10	5	$\frac{5+9+10+5}{4}=7.25$

Step 2: Calculate the quarterly proportions. Divide each quarterly sales figure by the corresponding yearly average to obtain quarterly proportions. Give answers correct to 4 decimal places.

Year	Q1	Q2	Q3	Q4
2010	$\frac{5}{6}=0.8333$	$\frac{7}{6}=1.1667$	$\frac{9}{6}=1.5000$	$\frac{3}{6}=0.5000$
2011	$\frac{4}{6.25}=0.6400$	$\frac{8}{6.25}=1.2800$	$\frac{9}{6.25}=1.4400$	$\frac{4}{6.25}=0.6400$
2012	$\frac{5}{7.25}=0.6897$	$\frac{9}{7.25}=1.2414$	$\frac{10}{7.25}=1.3793$	$\frac{5}{7.25}=0.6897$

Step 3: Calculate the seasonal indices (SI). Seasonal indices (SI) are obtained by finding the average of the quarterly proportions.

Year	Q1	Q2	Q3	Q4
2010	0.8333	1.1667	1.5000	0.5000
2011	0.6400	1.2800	1.4400	0.6400
2012	0.6897	1.2414	1.3793	0.6897
Total	2.1630	3.6881	4.3193	1.8297
SI	$\frac{2.1630}{3}=0.7210$	$\frac{3.6881}{3}=1.2294$	$\frac{4.3193}{3}=1.4398$	$\frac{1.8297}{3}=0.6099$

Give a summary table for Step 3.

Year	Q1	Q2	Q3	Q4
Seasonal Index	0.7210	1.2294	1.4398	0.6099

9780170250283

b Deseasonalise the data. This is Step 4.
Deseasonalise the original time series data using the formula

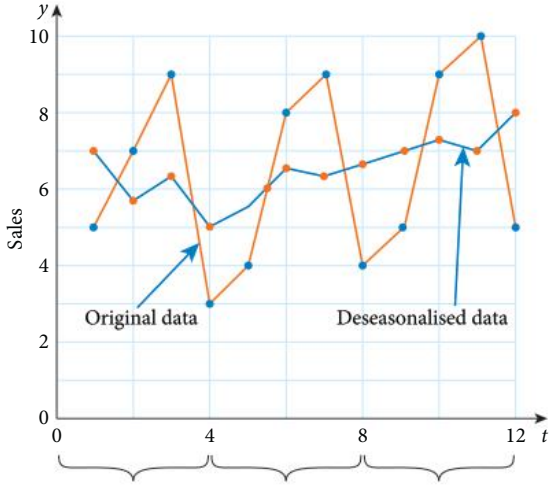
$$\text{Deseasonalised value} = \frac{\text{Actual value}}{\text{Seasonal index}}$$

Year	Q1	Q2	Q3	Q4
2010	$\frac{5}{0.7210} = 6.94$	$\frac{7}{1.2294} = 5.69$	$\frac{9}{1.4398} = 6.25$	$\frac{3}{0.6099} = 4.92$
2011	$\frac{4}{0.7210} = 5.55$	$\frac{8}{1.2294} = 6.51$	$\frac{9}{1.4398} = 6.25$	$\frac{4}{0.6099} = 6.56$
2012	$\frac{5}{0.7210} = 6.94$	$\frac{9}{1.2294} = 7.32$	$\frac{10}{1.4398} = 6.95$	$\frac{5}{0.6099} = 8.20$

Give a summary table for Step 4.

Year	Q1	Q2	Q3	Q4
2010	6.94	5.69	6.25	4.92
2011	5.55	6.51	6.25	6.56
2012	6.94	7.32	6.95	8.20

c Plot the original and the deseasonalised time series data.
Using $t = 1$ to represent the first quarter of 2003, and so on, the original and deseasonalised time series data are shown in the plot on the right.



The seasonal index can give us information about the performance in a particular season. If the performance in each season of a year was the same, then the SI for each quarter would equal 1. Once we have the seasonal index, we can add this up to make sure that we have a total of 4 for four quarters. Using data from Example **6**, the total is 4 as expected.

(We can expect some rounding errors.)

Year	Q1	Q2	Q3	Q4	total
SI	0.7210	1.2294	1.4398	0.6099	4.0001

This can be interpreted as

- 0.7210 tells us that sales in the first quarter are typically 72% of what is expected.
- 1.2294 tells us that sales in the second quarter are typically 123% of what is expected.
- 1.4398 tells us that sales in the third quarter are typically 144% of what is expected.
- 0.6099 tells us that sales in the fourth quarter are typically 61% of what is expected.

The higher the seasonal index, the higher the performance of the corresponding quarter compared to the average quarterly value. The lower the seasonal index, the lower the performance compared to the average. In the car sales example, the seasonal index of 143.98% for the third quarter indicates a performance of 43.98% above the average.

Example 7

A distributor of computers predicts that its deseasonalised quarterly sales will be $100 000 for each quarter. Predict its actual quarterly sales given the following seasonal indices: summer 74%; autumn 110%; winter 120%; spring 96%.

Solution

Check that these percentages equal 400% as expected.

$74 + 110 + 120 + 96 = 400$

Quarter	Seasonal index %	Seasonal index	Actual value SI × 100 000
summer	74	0.74	74 000
autumn	110	1.10	110 000
winter	120	1.20	120 000
spring	96	0.96	96 000

Write the answer. The final column gives the actual quarterly sales.

Quarter	Actual quarterly sales
summer	$ 74 000
autumn	$110 000
winter	$120 000
spring	$ 96 000

WS Seasonality

EXERCISE 6.04 Adjusting for seasonality

Concepts and techniques

1 Example 6 A manufacturer of televisions has determined the quarterly proportions shown, describing the shop's projected quarterly sales figures for three years.

Year	Q1	Q2	Q3	Q4
2002	0.71	0.86	1.08	1.35
2003	0.66	0.85	1.23	1.26
2004	0.55	0.91	1.45	1.09

a The seasonal index for the first quarter is

A 1.23 B 0.71 C 0.87 D 1.25 E 0.64

b The seasonal index for the second quarter is

A 1.23 B 0.71 C 0.87 D 1.25 E 0.64

9780170250283

c The seasonal index for the third quarter is

A 1.23 B 0.71 C 0.87 D 1.25 E 0.64

d The seasonal index for the fourth quarter is

A 1.23 B 0.71 C 0.87 D 1.25 E 0.64

e The sum of the seasonal indices for the four quarters is

A 4 B nearly 4 C 16 D 8 E 400

2 The quarterly sales figures for the number of king-sized beds sold were recorded by a furniture shop for 2011 to 2013.

Year	Sales figures quarter 1	Sales figures quarter 2	Sales figures quarter 3	Sales figures quarter 4
2011	11	8	1	3
2012	9	9	3	1
2013	4	9	1	7

a Calculate the seasonal index for each quarter.

b Deseasonalise the data.

c Plot the original and the deseasonalised time series data.

3 Every three months, the Reserve Bank reviews the official interest rate and changes it if necessary. The bank has released the following quarterly proportions describing the quarterly interest rates over a three-year period.

Year	Q1	Q2	Q3	Q4
2009	1.18	0.86	1.10	0.86
2010	0.21	1.15	2.20	0.44
2011	0.54	1.35	1.45	0.66

Calculate the seasonal index for each quarter.

4 Example 7 A distributor of calculators predicts its deseasonalised quarterly sales to be $8 000 for each quarter. Predict its actual quarterly sales given the following seasonal indices: summer 32%; autumn 125%; winter 130%; spring 113%.

5 Actual quarterly sales of calculators are 830. The seasonal index for that quarter is 0.95. What is the deseasonalised figure?

6 At a certain hotel, the deseasonalised occupancy figure for the quarter is 2010. The seasonal index for that quarter is 0.75. What was the actual number of guests in the hotel?

Reasoning and communication

7 The Maths Head of Faculty wants to analyse the number of absentees in her classes. She listed the daily average number of absentees among Year 12 students for each term over a three-year period.

Year	Q1	Q2	Q3	Q4
2009	3	8	11	6
2010	3	7	22	3
2011	4	6	15	10

a Calculate the seasonal index for each term.
b Deseasonalise the time series.
c On the same axes, plot both the original data and the deseasonalised data.
d What can you infer from the graph about the absenteeism?

8 A household has examined the number of hours its children have spent in 'screen time' over a two-year period from January 2010 to December 2011.

	Month	Screen time	Month	Screen time
Jan	1	13.2	13	21
Feb	2	15.6	14	34.5
Mar	3	17.7	15	21
Apr	4	10.2	16	32.6
May	5	11.5	17	4.2
June	6	31.6	18	21.1
July	7	32.1	19	24.5
Aug	8	34.5	20	20
Sept	9	3.6	21	3.7
Oct	10	5.7	22	4.5
Nov	11	6.8	23	2.2
Dec	12	20.2	24	32

a Plot the time series of the original data.
b Calculate the seasonal index for each month.
c Deseasonalise the time series data.
d On the same axes, plot both the original data and the deseasonalised data.
e What can you infer from the graph about the number of hours the children spent using 'screen time'?

6.05 FITTING A LEAST-SQUARES REGRESSION LINE TO TIME SERIES DATA

Time series data can be analysed using a least-squares line so that we can model long-term trends. When using the method of least-squares, a calculator is the most effective way of finding the equation of the line of best fit. A CAS calculator can be used to find the equation of the linear regression line.

IMPORTANT

A straight line can be fitted to a time series plot using the **least-squares method**.

The trend line can be used to predict future behaviour over time. A trend exists if there is a long-term increase or decrease in the dependent variable as time passes.

If there appears to be a linear trend, then a regression line can be fitted.

9780170250283

The linear model (regression equation) can be used to **forecast** future values (make predictions).

When data exhibits seasonality, we may need to deseasonalise the data before doing linear regression. Remember that when using deseasonalised data to make predictions, the seasonality has been factored out. To be 'meaningful' this result must be 'seasonalised' by multiplying by the appropriate SI (seasonal index).

IMPORTANT

Using a trend line fitted to a times series plot to make predictions about future values is known as **forecasting**.

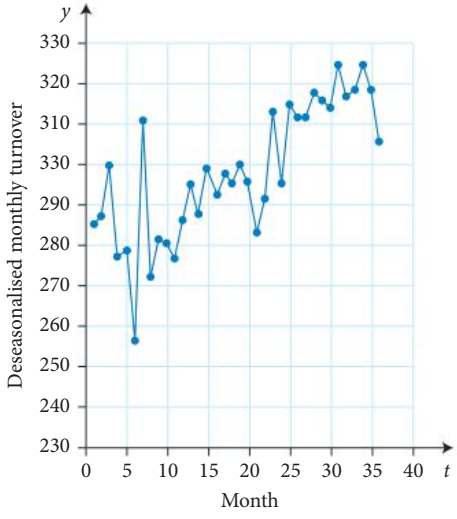
Example 8

The following deseasonalised data represents the monthly turnover, in $millions, in a department store over a period of 4 years.

Year	Jan	Feb	Mar	Apr	May	Jun
2011	285.4	286.9	299.3	277.2	278.6	256.7
2012	294.8	288.3	298.9	292.4	297.3	295.5
2013	314.8	311.8	311.8	317.7	315.6	313.9
SI	0.8409	0.7040	0.8740	0.9348	0.9657	0.9989

Year	Jul	Aug	Sep	Oct	Nov	Dec
2011	311.3	272.4	281.7	280.4	277.1	287.1
2012	300.2	295.5	283.4	291.6	312.9	295.7
2013	324.5	316.7	318.6	324.4	318.4	305.6
SI	0.8693	0.8338	0.8725	0.9783	1.1638	1.9641

The deseasonalised data is graphed on the axes below. On the horizontal axis, $t = 1$ represents Jan 2011, $t = 2$ represents Feb 2011, and so on.



a Determine the equation of the least-squares regression line for the deseasonalised time series data. Write constants correct to 3 decimal places.

b Plot the least-squares regression line on the time series plot.

c Use this line to estimate the turnover for August of 2012, in millions of dollars. Answer correct to one decimal place.

d Use the line to forecast the turnover for May of 2014, in millions of dollars. Answer correct to one decimal place.

Solution

a Use a CAS calculator to find the equation of the least-squares regression line with Month (t) as the independent variable and Deseasonalised monthly turnover (y) as the dependent variable.

Write the equation of the least-squares regression line with constants correct to 3 decimal places.

$y = 1.242\,t + 275.204$

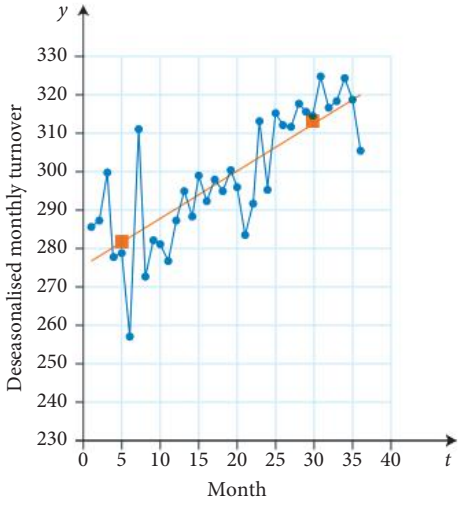
b Find the coordinates of one point on the line. Select a value of t and calculate the corresponding value of y.

$t = 5$
$y = 1.242 \times 5 + 275.204$
$= 281.414$
One point on the line is (5, 281.414).

Find the coordinates of a second point on the line. Select a value of t, not too close to the first, and calculate the corresponding value of y.

$t = 30$
$y = 1.242 \times 30 + 275.204$
$= 312.464$
Another point on the line is (30, 312.464).

Plot these two points on the time series plot and rule a straight line through them.



c Determine the month number corresponding to August of 2012. Substitute $t = 32$ into the least-squares regression line equation.

August of 2012 is $t = 32$
$y = 1.242\,t + 281.414$
$= 1.242 \times 32 + 281.414$
$= 321.158$
The estimated deseasonalised turnover for August of 2012 is \$321.2 million.

This figure needs to be converted to the seasonalised turnover.
Multiply by the seasonal index for August.

Seasonal index for August is 0.8338.
Multiply \$321.2 million by 0.8338.
$321.2 \times 0.8338 =$ \$267.78 million

d Determine the month number corresponding to May of 2014. Substitute $t = 53$ into the least-squares regression line equation.

May of 2014 is $t = 53$
$y = 1.242\,t + 275.204$
$= 1.242 \times 53 + 275.204$
$= 341.03$
The deseasonalised turnover for 2014 is forecasted to be \$341.0 million.

This figure needs to be converted to the seasonalised turnover.
Multiply by the seasonal index for May.

Seasonal index for May is 0.9657.
Multiply \$341.0 million by 0.9657.
$341.0 \times 0.9657 =$ \$329.30 million.

9780170250283

EXERCISE 6.05 Fitting a least-squares regression line to time series data

WS Least-squares regression line

Reasoning and communication

1 Staff in a small country hospital are hoping that extra care of the elderly in their community has decreased the number of deaths recorded over a period of 3 years. In this question, use the time series data in the table below with the number of deaths in the columns Q1, Q2, Q3 and Q4 and time (in quarters) in the first column.

Year	Q1	Q2	Q3	Q4
2000	2	4	10	5
2001	1	4	11	7
2002	1	1	7	1

a Create a time series plot using the data given.
b Find the least-squares regression line using technology.
c State the equation of the trend line.
d Predict the number of deaths in the small country hospital in the first quarter of 2003.

2 Example 8 a Comment on why the data in question **1** may need to be deseasonalised.
b Deseasonalise the data in the table in question **1**.
c Create a time series plot using the data from the table and the deseasonalised data from part **b**.
d Using technology, find the least-squares regression line on the deseasonalised data.
e State the equation of the trend line for the deseasonalised data.
f Predict the number of deaths in the small country hospital in the first quarter of 2003 by using the deseasonalised data.
g Which prediction is more meaningful? Your answer to question **1d** or **2f**?

3 The data shows the number of students enrolled in a certain type of school (in thousands) in Australia over the period 1992–2001.

Year	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
No. of students	525	539	545	556	581	596	600	603	600	614

a Plot the time series.
b Comment on the data.
c Fit a least-squares regression trend line to the data using 1992 as year 1. Add it to your plot in part **a** and interpret its slope.
d Use the trend line equation to predict the number of students enrolled in a certain type of school in Australia in 2008.

4 a The table below shows the deseasonalised computer sales of a company over three years.

No. of sales deseasonalised	Q1	Q2	Q3	Q4
Year 1	53	51	54	55
Year 2	64	64	61	63
Year 3	67	69	68	66

Use a least-squares regression to fit a trend line to the data.

b Use this trend equation for computer sales, together with the seasonal indices given to forecast the sales of computers in the fourth quarter of year 4.

SI	0.90	0.81	1.11	1.18

5 The table below shows the percentage of total retail sales that were made in a department store over an 11-year period.

Sales (%)	12.3	12.0	11.7	11.5	11.0	10.5	10.6	10.7	10.4	10.0	9.4
Year	1	2	3	4	5	6	7	8	9	10	11

a Construct a time series plot.

b Comment on the data.

c Fit a trend line to the time series plot, find its equation and interpret its slope.

d Draw in the trend line on your time series plot

e Use the trend line to forecast the % of retail sales which will be made by the department store in year 15.

6 The following deseasonalised data represents the monthly turnover, in dollars, by a small roadside van over a period of 2 years.

Year	Jan	Feb	Mar	Apr	May	Jun
2010	185.4	286.9	199.3	177.2	178.6	256.7
2011	194.8	288.3	198.9	192.4	197.3	295.5
SI	1.8102	0.7040	0.7740	0.7348	0.8657	0.8989

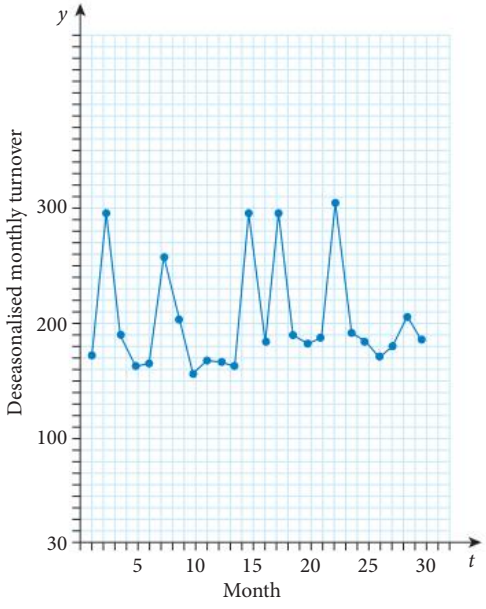
Year	Jul	Aug	Sep	Oct	Nov	Dec
2010	211.3	172.4	181.7	180.4	177.1	287.1
2011	200.2	195.5	183.4	191.6	212.9	195.7
SI	0.7693	1.1338	1.0725	0.9789	1.2278	1.0641

The deseasonalised data is graphed on the axes on the right. $t = 1$ represents Jan 2010, $t = 2$ represents Feb 2010, and so on.

a Determine the equation of the least-squares regression line for the deseasonalised time series data. Write constants correct to 3 decimal places.

b Plot the least-squares regression line on the time series plot.

c Use this line to estimate the turnover for May of 2012.



9780170250283

The following information relates to questions **7**, **8** and **9**.

The given deseasonalised data represents the weekly turnover, in dollars, by an ice-cream van over a period of 6 weeks in the summer season. The 6 weeks of the summer season officially start on Boxing Day.

Week	1	2	3	4	5	6
Summer 2012	85.2	186.9	200	77.2	78	100
Summer 2013	194	188.3	98.9	192	197	150
SI	1.8102	0.7040	1.7740	0.6000	0.8000	0.3118

7 Sketch a graph of the deseasonalised data where $t = 1$ represents week 1 in summer 2012, $t = 2$ represents week 2 in summer 2012, and so on.

8 **a** Plot the points (3, 200) and (12, 150) on your graph in question **7**.

b Find the equation of the linear regression line between these two points and sketch the line on your graph in question **7**.

c Use this equation to predict the sales in the week beginning on Boxing Day 2014.

9 **a** Find the linear regression line equation using technology.

b Use the equation found in part **a** to predict sales in the week beginning on Boxing Day in 2014.

10 The manager of a small country bank is trying to increase the number of customers who learn to do their banking online. In this question use the time series data in the table below, with the number of people learning to do their banking online in the columns Q1, Q2, Q3 and Q4 and time (in quarters) in the first column.

Year	Q1	Q2	Q3	Q4
2010	20	10	10	5
2011	12	14	10	20
2012	20	25	30	30

a The equation of the least-squares regression line using technology is

A $y = 20t + 30$ **B** $y = -20t + 30$ **C** $y = 0.398t^2 - 3.456t + 18.091$

D $y = 0.325t + 0.920$ **E** $y = 1.713t + 6.030$

b Using the trend line from part **a**, the number of people learning to do their banking online in the third quarter of 2013 is closest to

A 30 **B** 32 **C** 10 **D** 20 **E** 22



6 CHAPTER SUMMARY
TIME SERIES ANALYSIS

- Data observed and recorded at successive intervals of time is referred to as **time series data**.
- Time series can be plotted and identified by features such as trend, seasonality, and irregular fluctuations.
- A time series can have one or more of the following four components.
  - trend
  - seasonal pattern
  - cyclic pattern
  - random variation, also known as random pattern
- A **trend** exists if there is a long-term increase or decrease in the dependent variable as time passes.
- A **seasonal pattern** arises in the context of economic time series models where the actual seasons of the year affect sales or production of goods.
- A **cyclic pattern** exists when significant peaks and troughs occur at irregular intervals, usually with variation in the amplitudes of the peaks and troughs.
- **Random variation** or **random pattern** reflects the unpredictability common in all statistical data.
- Smoothing time series data is a technique used to level out fluctuations and hence show a clearer picture of an overall trend.
- To smooth time series data, two techniques are used.
  - smoothing using moving averages (means)
  - median smoothing
- Smoothing using **moving averages** takes a fixed number of data points and averages them to give a smoothed data point.
- A way to smooth data is to use a two or four-point moving average with **centring.** Smoothing with an odd number of points gives a smoothed data point, which is located at the middle time.
- **Median smoothing** uses the median values instead of the average (mean) values. An even number of points needs the method of **centring** for median smoothing.
- Seasonal indices can be calculated by using a method called the average percentage method. Seasonal indexes are used to deseasonalise a time series.
- Deseasonalisation is a form of smoothing, which takes out the seasonal effects of the data so that a trend line can be fitted and long-term trends can be predicted.

9780170250283

- **Steps to deseasonalising data**
  - Find the average for each year (= **yearly average**)
  - Divide each month by the yearly average
  - Find the average of each month (= **seasonal index** for each month)
  - Divide the original data by its seasonal index

The formula for calculating deseasonalised values is:

$$\text{Deseasonalised value} = \frac{\text{Actual value}}{\text{Seasonal index}}$$

- Time series data can be analysed using a least-squares line so that we can model long-term trends. When using the method of least-squares, a calculator is the most effective way of finding the equation of the line of best fit.
- A straight line can be fitted to a time series plot using one of the following methods.
  - by eye
  - three median method
  - least-squares method.
- The linear model (regression equation) can be used to **forecast** future values (make predictions).
- Using a trend line fitted to a times series plot to make predictions about future values is known as **forecasting**.

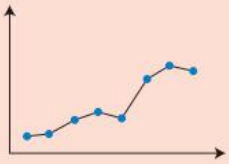
6 CHAPTER REVIEW
TIME SERIES ANALYSIS

Multiple choice

The table below for the 'Deaths by year in Ballarat from 1910 to 1917' contains information to be used in questions **1**, **2** and **3**.

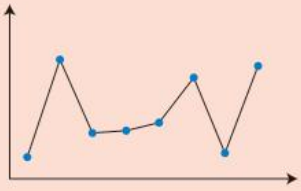
Year	Total deaths
1910	104
1911	123
1912	200
1913	239
1914	204
1915	423
1916	500
1917	469

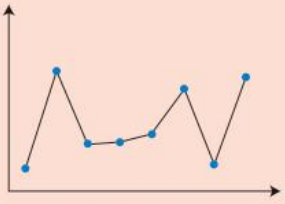
1 Example 1 The time series plot would look like

A 

B 

C 

D 

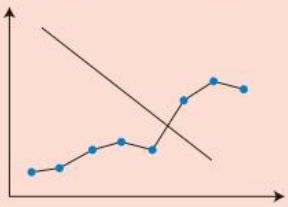
E 

2 Example 2 The time series plot shows a

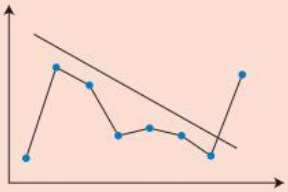
A positive secular trend B negative secular trend C seasonal pattern
D cyclic pattern E random variation

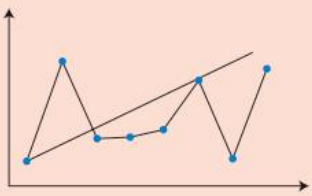
9780170250283

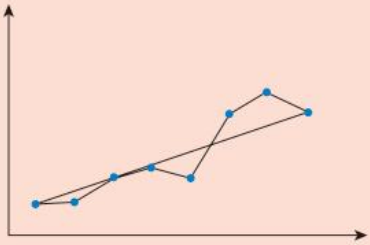
3 A trend line drawn on the time series plot would look like

A 

B 

C 

D 

E 

4 The following data represents the number of employees at a pie factory.

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014
Number of employees	840	820	830	790	770	780	750	740	730

If the data is smoothed using a three point moving average, the smoothed value for 2010 is

A 757 B 760 C 767 D 770 E 780

5 The average yearly temperatures in Cairns over a year represents a

A positive secular trend B negative secular trend C seasonal pattern
D cyclic pattern E random pattern

6 The daily distances covered by a taxidriver in Melbourne over a certain period of time represents a

A positive secular trend B negative secular trend C seasonal pattern
D cyclic pattern E random pattern

7 Examples 6, 7 The quarterly sales figures ($1000s) in 2000 and seasonal indices for a certain company are listed in the table below.

Quarter	Q1	Q2	Q3	Q4
Sales ($1000s)	62	72	69	57
Seasonal index (%)	92.2	106.5	114.2	?

The seasonal index for Q4 is

A 78.6 B 80.7 C 82.5 D 87.1 E 92.3

8 The seasonal index each quarter for the tourist numbers in Queensland is given in the table below.

Quarter	Q1	Q2	Q3	Q4
Seasonal index	1.30	0.75	0.85	1.10

The seasonal index row should add up to

A 4 B 40 C 400 D 0.4 E 4%

9 The seasonal index each quarter for the tourist numbers in Queensland is given in the table below.

Quarter	Q1	Q2	Q3	Q4
Seasonal index	1.30	0.75	0.85	1.10

If the deseasonalised figure for Q1 is 25 600, then the actual number of tourists in Q1 is

A 19 692 B 23 273 C 28 160 D 32 028 E 33 280

10 The first quarter seasonal index for a set of sales figures is 0.79. In the first quarter of 2010, the sales figures were 2350 items. The deseasonalised figure is closest to

A 1860 B 2150 C 2280 D 2350 E 2970

Short answer

11 Example 3 Refer again to the table at the start of this exercise on page 350.

a Use the method of 3-point moving averages to smooth the data, using $t = 1$ to represent 1910 through to $t = 8$ to represent 1917.

b Using the smoothed data, what would be the predicted number of deaths in Ballarat in 1911?

12 Example 4 a For the data in the table below, use a two-point moving average, with centring, to smooth the number of tickets sold at the local cinema over a 12-month period in 2010.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
Tickets sold	100	120	60	50	220	180	130	70	90	100	80	150

b What is the smoothed number of tickets sold at the local cinema in August?

13 a For the data in question **12**, use a four-point moving average, with centring, to smooth the number of tickets sold at the local cinema over a 12-month period in 2010.

b What is the smoothed number of tickets sold at the local cinema in August?

14 Example 5 A coffee vendor is trying to predict its sales figures for 1999. The manager of the van has his sales figures for the years 1993 to 1998.

Year, t	1993	1994	1995	1996	1997	1998
Sales, y ($'000)	73	138	114	112	160	122

Using a four-point moving median, with **centring**, smooth the data and use it to predict sales for 1999.

9780170250283

Application

15 The data below lists attendances at a local ballet school for January to December from 1991 and 1992. Set up a table using 3-point moving average, 4-point moving average smoothing and 4-point moving average centred.

		Data
1991	J	45
	F	39
	M	70
	A	30
	M	34
	J	29
	J	44
	A	47
	S	41
	O	35
	N	35
	D	53

		Data
1992	J	34
	F	33
	M	34
	A	47
	M	30
	J	29
	J	23
	A	26
	S	28
	O	38
	N	30
	D	40

16 Examples 8, 9 The following deseasonalised data represents the monthly turnover, in $million, in a department store during 2010 and 2011.

Year	Jan	Feb	Mar	Apr	May	Jun
2010	26.1	27.3	26.5	27.8	24.7	28.9
2011	28.4	28.9	29.3	27.2	28.6	26.7

Year	Jul	Aug	Sep	Oct	Nov	Dec
2010	22.8	21.4	28.8	27.6	25.4	26.7
2011	31.3	22.4	28.7	28.4	27.1	27.1

Practice quiz

a Fit a least square regression line for the deseasonalised data.
b Use the line to estimate the turnover for January of 2012.
c Use the line to forecast the turnover for January of 2013.

9780170250283

MIXED REVISION
CHAPTERS 4 • 5 • 6

2

Multiple choice

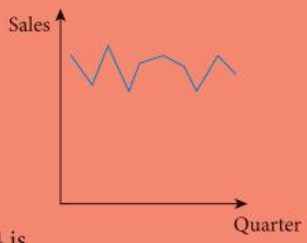
1 The sum of the degrees of all the vertices in the graph below is:



A 6
B 9
C 12
D 15
E 18

2 The sequence generated by the recurrence relation $t_{n+1} = t_n - 3$, where $t_1 = 10$ is
A a decreasing geometric sequence
B an increasing geometric sequence
C a decreasing arithmetic sequence
D an increasing arithmetic sequence
E a sequence which is neither arithmetic nor geometric.

3 The time series graph shown displays



A random fluctuations only
B seasonal variation
C cyclic variation
D a trend and random fluctuations
E a trend and cyclic fluctuations.

4 The solution to the difference equation $t_{n+1} - 5t_n = 0$, where $t_1 = 4$ is
A $t_n = 4(-5)^{n-1}$
B $t_n = 4(5)^{n-1}$
C $t_n = -5(4)^{n-1}$
D $t_n = 5(4)^{n-1}$
E $t_n = 4\left(\frac{1}{5}\right)^{n-1}$

5 A connected planar graph has 12 edges. This graph could have:
A 5 vertices and 6 faces
B 5 vertices and 8 faces
C 6 vertices and 8 faces
D 6 vertices and 9 faces
E 7 vertices and 9 faces.

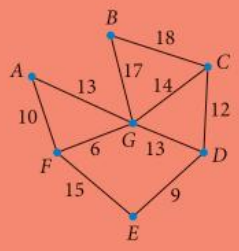
6 The seasonal indices for the sales of dresses by a factory are given in the table below.

Quarter	Q1	Q2	Q3	Q4
SI	0.94	1.13	1.07	0.86

The actual number of dresses sold in the third quarter was 547. The seasonally adjusted sales figure is closest to
A 470 B 484 C 511 D 547 E 585

7 The shortest path from A to E is:

A 6
B 10
C 16
D 25
E 34



8 A geometric sequence has a common ratio of –2. The third term in the sequence is –20. A recurrence relation that describes this sequence is

A $t_{n+1} = t_n - 2$, where $t_1 = -20$
B $t_{n+1} = -2t_n$, where $t_1 = -20$
C $t_{n+1} = t_n - 2$, where $t_1 = -5$
D $t_{n+1} = -2t_n$, where $t_1 = -5$
E $t_{n+1} = -2t_n$, where $t_1 = -80$

9 The seasonal indices of the sales from a retail outlet are given in the table.

Quarter	Index
1	0.8
2	1.0
3	1.3
4	0.9

For the first quarter of this year, the deseasonalised sales were $48 000. The actual sales figures were

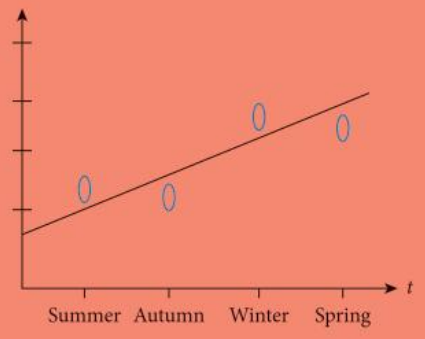
A $38 400
B $43 200
C $48 000
D $60 000
E $62 400

Short answer

Questions **1** and **2** relate to the following information
The time series graph below represents the deseasonalised sales data of surfboards at a surf outlet in a coastal town. A line of best fit has been drawn and its equation is given.

2010 surfboard sales deseasonalised $S = 3.9 + 3.1t$, where t is the time in quarters.

Season	Summer	Autumn	Winter	Spring
SI	?	0.64	0.87	1.23

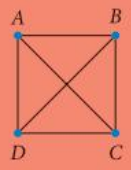


1 What would the seasonal index for summer approximately be?

2 What would be the best estimate of the number of surfboards the outlet will sell in winter?

9780170250283

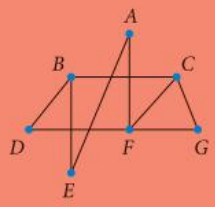
3 Redraw the graph on the right so that it is planar.



4 Write a recurrence relation to describe each of the following sequences.
   a $t_n = -5n + 4$
   b $t_n = 4(1.6)^{n-1}$

5 a Plot the first five terms of the sequence $t_{n+1} = 0.5t_n - 4$, where $t_1 = 16$.
   b Describe the long-term solution of this sequence.

6 Determine a Hamiltonian path for the graph drawn on the right that starts at G and ends at E.



Application

1 In a local village, there are four friends: Chelsea, Evelyn, Doreen and Fay. The friends' houses are connected by paths as follows:
- Chelsea and Evelyn are connected by one path
- Evelyn is connected to Doreen by one path
- Doreen and Chelsea are connected by two paths
- Fay is connected to both Evelyn and Doreen by one path

   a Draw a network diagram with the vertices representing each friend and the edges representing the paths connecting them.
   b Create an adjacency matrix representing the network diagram from part **a**.
   c If the girls wanted to follow an Eulerian path, where should they start?
   d Describe the Eulerian path the girls should take.

2 In 2010, the wild horse population in Region A was 400 horses and decreasing by 10% each year. The wild horse population in Region B was 420 horses and increasing by 15% each year.
   a A recurrence relation that gives the number of wild horses in Region A in the nth year after 2010 is $A_{n+1} = 0.9A_n$, where $A_0 = 400$. Find a recurrence relation in the form $B_{n+1} = rB_n$, where $B_0 = k$ that gives the number of wild horses in Region B in the nth year after 2010.
   b Determine the wild horse population in each region in 2011.
   c i In what year will the wild horse population in Region B first exceed 800?
      ii How many horses are there?
      iii What is the wild horse population in Region A at this time? Answer correct to the nearest whole number.

   When the wild horse population in Region B first exceeds 800, 120 horses from this region are relocated to Region A each year.

d A recurrence relation of the form $P_{n+1} = rP_n + d$, where $P_0 = a$, describes the population in Region B, n years after the first 120 horses have been located. Write down the values of r, d and a.

e Describe what happens to the population of horses in Region B in the long term.

f The wild horse population in Region A, n years after the first 120 horses have been located, is given by $H_{n+1} = 0.9H_n + 120$, where $H_0 = 236$. In the long term, the horse population in Region a reaches a steady state. What is the size of the horse population in Region A in the long run?

3 At the start of 2014, Maria's salary was $80 000. She negotiated a salary contract for the next five years so that her salary increases by 5% each year.

a Use a recurrence relation to model this situation.

b Graph the information.

c What is Maria's annual salary for the year 2018?

d In which year will Maria's salary first exceed $100 000?

e How much does Maria earn in total between 1 January 2014 and 31 December 2019?

4 The following data is a record of blanket sales for "The Rug-Up Company". Table 1 gives the number of blanket sales in each of the four seasons for the years 2010 to 2012.

Table 1: Blankets sales for 2010–2012 per quarter

Year/season	Summer	Autumn	Winter	Spring	Average
2010	446	1085	1241	920	923
2011	541	1180	1356	1033	1028
2012	659	1234	1450	1299	
SI	0.52	1.15		1.03	

a What is the average number of sales per season for the year 2012? Answer correct to the nearest whole number.

b What is the seasonal index for Winter?

c What does the seasonal index of 0.52 indicate about the sale of blankets during Summer?

The number of blanket sales is shown in the time-series plot below.



d Describe the time-series.

9780170250283

5 The 'Rug-Up Company' would like to predict sales for the four seasons in the year 2013. To do this, they deseasonalise the data as shown in Table 2 below.
Table 2: Deseasonalised data

Year/season	Summer	Autumn	Winter	Spring
2010		943	955	893
2011	1040		1043	1005
2012	1267	1151	1115	1261

a Complete the table for the Sumnmer of 2010 and the Autumn of 2011.

b The time-series plot below shows the deseasonalised data the for blanket sales. On this time-series plot, find the equation of the least squares regression line.



6 a For 'The Rug-Up Company' assume the trendline is given by $N = 839 + 32x$ (x = time in quarters). Use this trendline to calculate the value of N for winter of 2013.

b Using this value, estimate the number of blanket sales in winter of 2013.

7 Let P denote a planar graph with n vertices.

a If T denotes a spanning tree of P, specify:

i the number of vertices in T

ii the number of edges in T

iii the number of faces in T.

b Hence, verify that the graph T is connected by using Euler's formula.


idential
lend



7

TERMINOLOGY

buying on terms
compound interest
compounding periods
effective interest rate
exponential
flat-rate interest
future value
hire-purchase
interest
nominal interest rate
p.a.
present value
principal
repayments
simple interest

Prior learning

LOANS, INVESTMENTS AND ANNUITIES

LOANS AND INVESTMENTS

7.01 Simple interest review
7.02 Hire-purchase
7.03 Compound interest
7.04 The compound interest formula
7.05 Interest graphs
7.06 Effective interest rate
Chapter summary
Chapter review

COMPOUND INTEREST LOANS AND INVESTMENTS

- use a recurrence relation to model a compound interest loan or investment, and investigate (numerically or graphically) the effect of the interest rate and the number of compounding periods on the future value of the loan or investment (ACMMGM094)
- calculate the effective annual rate of interest and use the results to compare investment returns and cost of loans when interest is paid or charged daily, monthly, quarterly or six-monthly (ACMMGM095)
- with the aid of a calculator or computer-based financial software, solve problems involving compound interest loans or investments; for example, determining the future value of a loan, the number of compounding periods for an investment to exceed a given value, the interest rate needed for an investment to exceed a given value. (ACMMGM096) AC

7.01 SIMPLE INTEREST REVIEW

Simple interest is **interest** calculated as a percentage of the **principal** (the amount of money invested or borrowed). When borrowing money it can also be referred to as **flat-rate interest**.

IMPORTANT

Simple interest can be calculated using the formula:

$$I = \frac{Prn}{100}$$

where

I = simple interest

P = principal

r = interest rate per period, as a percentage

n = number of periods

Example 1

Find the quarterly interest paid when Hassan invests $16 000 in bonds that pay 7.5% p.a. simple interest.

p.a. means per annum which is 'per year'

Solution

Write the values of the known variables. Quarterly interest means n is $\frac{1}{4}$ of a year.

$P = 16\,000, r = 7.5, n = \frac{1}{4}$

Write the simple interest formula.

$I = \frac{Prn}{100}$

Substitute into the formula.

$= \frac{16\,000 \times 7.5 \times \frac{1}{4}}{100}$

$= 300$

Write the answer.

Hassan is paid $300 in interest each quarter.

9780170250283

Once the simple interest is calculated, the amount to which the investment has grown can be found by adding the interest to the principal.

IMPORTANT

$A = I + P$

or

$A = \frac{Prn}{100} + P$

Example 2

Sasha invests $4000 in an account earning 5.6% p.a. simple interest.

a How much interest will she earn in $3\frac{1}{2}$ years?

b What amount will her investment grow to in $3\frac{1}{2}$ years?

Solution

a Write the values of the known variables. $P = 4000, r = 5.6, n = 3\frac{1}{2}$

Write the simple interest formula. $I = \frac{Prn}{100}$

Substitute into the formula.
$$= \frac{4000 \times 5.6 \times 3\frac{1}{2}}{100}$$
$$= 784$$

Write the answer. Sasha will earn $784 interest.

b The principal and the interest are needed to find the value of the investment. $I = 784, P = 4000$

Write the formula. $A = I + P$

Substitute into the formula.
$$= 784 + 4000$$
$$= 4784$$

Write the answer. Her investment will grow to $4784.

If the amount of interest is known, the simple interest formula can also be used to find one of the other values: principal, interest rate or time. Substitute the values of the three known variables into the simple interest formula and solve the resulting equation for the unknown.

Example 3

An account is advertised as having 7.2% p.a. interest and paying interest every fortnight. What amount would you have to invest in order to earn $150 per fortnight in interest?

Solution

Write the values of the known variables. $I = 150, r = 7.2, n = \frac{1}{26}$

Fortnightly interest means n is $\frac{1}{26}$ of a year.

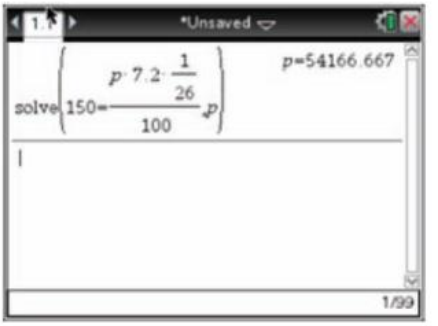
Write the simple interest formula. $I = \frac{Prn}{100}$

Substitute into the formula.

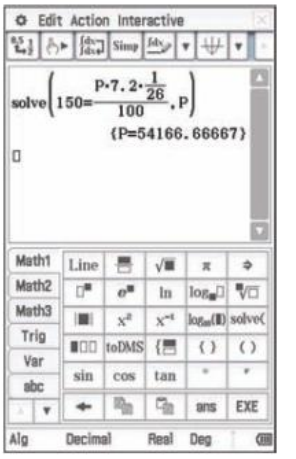
$$150 = \frac{P \times 7.2 \times \frac{1}{26}}{100}$$

Solve the equation to find P.

TI-Nspire CAS



ClassPad



$P = 54\,166.666...$
$\approx 54\,166.67$

Write the answer.

In order to earn \$150 per fortnight in interest, you would have to invest \$54 166.67.

Applications of simple interest

EXERCISE 7.01 Simple interest review

Concepts and techniques

1 Example 1 Find the amount of simple interest paid by Flynn if he borrows:
- a \$14 000 at 8.3% p.a. for 4 months
- b \$6000 at $1\frac{1}{3}$% per month for 3 years
- c \$5600 at 4.5% p.a. for $3\frac{1}{2}$ years
- d \$29 620 at 0.47% per month for two months.

2 Paula invests \$105 000 at 7.5% p.a. simple interest. Calculate the interest she will earn if she invests it for $2\frac{1}{2}$ years.

3 Example 2 Stephen borrows \$45 000 at 11.5% p.a. simple interest.
- a Calculate the interest he will pay if he repays the loan in 5 years.
- b Calculate the total amount he will repay.

9780170250283

4 Calculate the amount Hayley's investment will grow to if she invests:
   a $124 000 at $7\frac{1}{4}$% p.a. for 9 months
   b $16 500 at 0.96% per month for 2 years and 4 months
   c $45 000 at 0.0245% per day for 37 days
   d $1250 at 8.7% p.a. for 136 days.

5 $6000 is invested for two years at a simple interest rate of 3% p.a. The total amount of interest earned by the investment is:
   A $90 B $180 C $360 D $6180 E $6360

6 Paul borrows $72 000 for six years at a simple interest rate of 7.25% per annum. The total amount he has to repay is:
   A $5220 B $31 320 C $72 870 D $77 220 E $103 320

7 Example 3 What principal would earn $6000 in interest if it is invested for eight years at 7.5% p.a.?

8 Shin deposited $9000 into an investment account. Three years later, at the end of the investment period, he received $11 500.
   a How much interest did he earn?
   b What was the interest rate per annum? Answer as a percentage correct to one decimal place.

9 For how long will $30 000 need to be invested to earn $625 in simple interest if the interest rate is 2.5% p.a.?
   A 5 months B 8 months C 10 months
   D 19 months E 25 months

10 Rebecca used a short term loan of $500 to buy some new clothes. She repaid the loan after 5 weeks with interest charged at 13% p.a. simple interest. How much did she have to pay back at the end of the 5 weeks?

11 Sophia invests $24 000 in bonds paying 5.7% p.a. interest with monthly payments. How much does she receive each month?

12 Max deposited $16 000 in an account and ended up with $17 904 after 21 months.

   What was the interest rate per quarter?

13 Rylee earned $1825 in simple interest on an investment of $100 000 with a term of 117 days. Calculate the daily interest rate as a percentage, correct to 4 decimal places.

Reasoning and communication

14 Toby invested $12 000 in an account that earned 7.5% p.a. simple interest. Show that the amount of simple interest could have been calculated from the formula $I = 900n$.

15 Isobel is going to invest $6000 for 2 years. She has two investment options available. The first is investing the whole amount at 12% p.a. simple interest and the second involves investing $2000 at 13% p.a. simple interest and the rest at 11% p.a. simple interest. Which investment would give the best return? Explain.



16 Erin has a credit card that charges simple interest of 19.6% p.a. with no interest-free period. She bought a surfboard for $649 on 12 November and paid her account in full on 3 January. Interest is charged daily from the date of purchase to the date of payment inclusive.
   a How many days did she pay interest on the purchase?
   b Calculate the interest charged.
   c Find the amount she paid.

17 Pascal borrowed $24 000 at 6.5% p.a. flat-rate interest to buy a car.
   a If the loan was for 5 years, how much interest did he have to pay?
   b What was the total amount he had to repay?
   c If he paid the loan back in fortnightly instalments, what was the amount of each instalment? Answer correct to the nearest cent.

18 Renee took out a personal loan for $8500 at 7.25% p.a. simple interest to furnish her house. She paid the loan off in monthly instalments over 3 years. Calculate the amount of each instalment. Answer correct to the nearest cent.

7.02 HIRE-PURCHASE

Hire-purchase means 'paying off' an expensive item over time, sometimes after a deposit is made. If the customer does not keep up with the regular payments or follow the terms and conditions of the plan, higher interest may be charged or the item repossessed (taken back). This is sometimes referred to as **buying on terms.**

By calculating the total amount paid, the interest rate that has been charged can be calculated.

Example 4

A refrigerator is advertised for $1200 cash or on terms at 10% deposit and $57 per month for 2 years. If buying on terms:
a how much more will the refrigerator cost than if bought for cash?
b what is the flat rate of interest charged per annum?

Solution

a Calculate the total cost.

$$\begin{aligned}\text{Total cost} &= \text{deposit} + 24 \text{ monthly repayments}\\ &= 10\% \times \$1200 + \$57 \times 24\\ &= \$120 + \$1368\\ &= \$1488\end{aligned}$$

Find the difference in costs between buying on terms and paying in cash.

$$\begin{aligned}\text{Difference} &= \$1488 - \$1200\\ &= \$288\end{aligned}$$

This is the interest on the loan.

Write the answer.

The refrigerator will cost $288 more if bought on terms.

9780170250283

b The principal is the amount borrowed. It is the difference between the deposit and the cash price.

$$\text{Principal} = \$1200 - \$120 = \$1080$$

Write the values of the known variables.

$$I = 288,\ P = 1080,\ n = 2$$

Write simple interest formula.

$$I = \frac{Prn}{100}$$

Substitute into the formula and then solve the equation to find r.

$$288 = \frac{1080 \times r \times 2}{100}$$

Multiply by 100.

$$28\,800 = 1080 \times r \times 2$$

Divide by (1080×2).

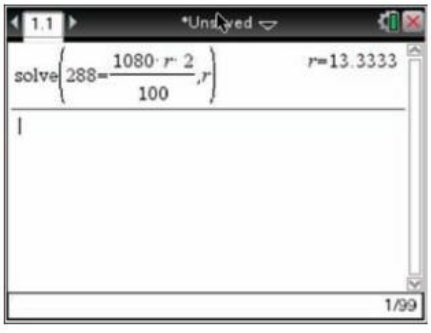
$$r = \frac{28\,800}{(1080 \times 2)}$$

$$= 13\frac{1}{3}$$

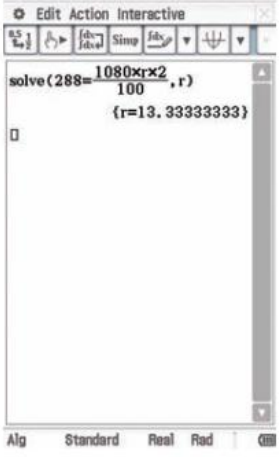
$$= 13$$

Alternatively solve using a CAS.

TI-Nspire CAS



ClassPad



Write the answer.

The flat rate of interest charged is $13\frac{1}{3}$% p.a.

A deferred payment plan is when there are no **repayments** until a later date. When using a deferred payment plan it is important to try to save in order to pay off the item by the end of the agreed period, otherwise interest is charged on the amount still owed from the date of purchase. This can be a very expensive option.

Example 5

Two competing retail stores offered different deferred payment plans to customers wanting to buy a dishwasher for $1350.





'Unpaid balance' and 'outstanding amounts' both mean the remaining part of the loan still owing.

a Which plan should Peter choose if he wants more time to save up enough money to pay off the dishwasher before any interest is charged?

b Souki chose the Krazy Kitchens plan. She paid $800 on the date of purchase but had paid no more when the 300 days were over. How much does she owe now?

Solution

a Determine the number of days in two years to compare the lengths of the plans.

$2 \text{ years} = 2 \times 365$
$= 730 \text{ days}$
Peter should choose Awesome Appliances, because its deferred payment period of 2 years is longer than Krazy Kitchens' plan of 300 days.

b Calculate the outstanding amount.
The $800 deposit needs to be subtracted from the cost of the dishwasher.

$\text{Outstanding amount} = \$1350 - \$800$
$= \$550$

Write the values of the known variables.

$P = 550, r = 0.11, n = 300$

Calculate the amount of interest using the simple interest formula.

$$I = \frac{Prn}{100}$$
$$= \frac{550 \times 0.11 \times 300}{100}$$
$$= \$181.50$$

The outstanding amount is charged interest from the date of purchase, so for 300 days.

The total owing is the sum of the amount outstanding and the interest.

$\text{Total owing} = \$550 + \$181.50$
$= \$731.50$

Write the answer.

She now owes $731.50.

9780170250283

EXERCISE 7.02 Hire-purchase

Hire-purchase

Reasoning and communication

1 Example 4 Del and Barry bought an old cottage for $11 000 for their farm. They paid 10% deposit and monthly repayments of $360 for 4 years.
   a How much interest did they pay by buying on terms?
   b What was the annual interest rate they were charged when they paid on terms? Answer correct to one decimal place.

2 Scott purchased an engagement ring for $7560, paying $100 deposit and 7.2% p.a. interest for 18 months. The total amount paid for the ring was
   A $7460 B $7660 C $8056.80
   D $8265.68 E $8365.68

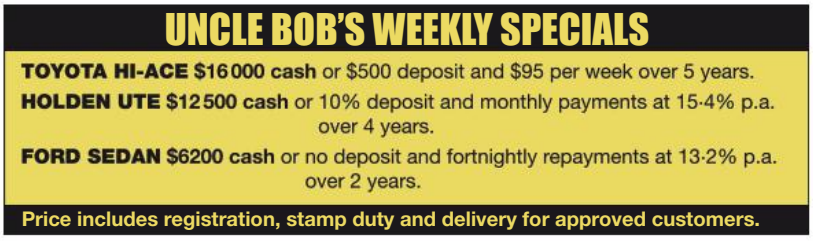
3 Chloe purchases new study furniture valued at $2500 on terms of 10% deposit and $151.25 per month for 18 months. The interest rate charged per annum was
   A 4.0% B 5.9% C 10.89% D 12.6% E 14%

4 Darby and Joan bought a second-hand motorhome for $19 500 to travel around Australia. There were two options for paying for the motorhome.

   Plan A: 20% deposit and 36 monthly instalments of $595
   Plan B: No deposit and monthly instalments over 4 years at 8.9% p.a. flat-rate interest.

   a By calculating and comparing, find which plan:
      i charges more interest
      ii has lower monthly instalments.
   b What other factors should Darby and Joan take into account when deciding between the two options?

5 Uncle Bob sells used cars on terms. Here are his weekly specials.



   a How much do you pay for the Ford sedan on Bob's terms?
   b How much would you save on the Holden ute by paying cash rather than term payments?
   c What is the flat rate of interest per annum, correct to one decimal place, charged on the Toyota Hi-Ace if making term payments?
   d Give a possible reason Bob charges different rates of interest for different cars.

6 Georgia bought a new outdoor setting valued at $4500. She was given two options.

   Option A: 5% deposit and repayments of $245.81 per month for 2 years.
   Option B: 15% deposit and repayments of $87.54 per week for 1 year.

   a Calculate the total paid for each option.
   b Which option should Georgia choose? Why?
   c Give one advantage of the other option that may influence Georgia's decision.

7 Indiana wants to put in a new kitchen for $17 000. The firm she is going with offers the following terms: 10% deposit and 14.6% p.a. flat-rate interest on the balance. Equal fortnightly instalments over 4 years.

If she chose to buy on terms, how much per fortnight would she pay?

8 Example 5 Two finance companies offered deferred payment plans and charged interest of 0.2% per day on outstanding amounts from the purchase date.

Lemon Loans: 10% deposit then nothing to pay for 12 months
Feelgood Finance: Buy now and pay nothing for 100 days

a Jayden bought a $1600 desk through Lemon Loans. Find:
  i the amount owing 12 months after purchase
  ii the amount owing after 396 days if only the deposit had been paid.

b Adison bought a computer for $1350 using Feelgood Finance. Calculate the amount if he repaid the debt in full after:
  i 90 days ii 101 days

c Give one positive and one negative feature of each plan.

9 Billy Joe bought a drum set for $1600 and was deciding between buying on terms or a deferred payment plan.

How much will he pay for the drum set if he:
a buys on terms?
b uses the deferred payment plan and pays the amount owing 2 years after purchase?
c uses the deferred payment plan, pays half the cost plus fees in 2 years, and the rest 6 months later?



Shutterstock.com

10 Ishmail bought a car navigation system (GPS) on terms. How much would he save by paying cash?



Shutterstock.com

11 Write down one advantage and one disadvantage of:
a paying cash for goods
b buying on terms
c using a deferred payment plan.

7.03 COMPOUND INTEREST

Most investments are **compound interest** investments. The interest is calculated at the end of a certain time period. The interest is then added onto the principal and then the interest for the next time period is calculated using this new principal. This continues for the full term of the investment.

Example 6

$3000 is invested at 5% p.a. compounded yearly. Calculate:

a the amount of the investment at the end of 4 years

b the compound interest earned after 4 years.

compounded yearly means the interest is added on each year and then the interest is calculated on the new total for the next year

Solution

a Write the values of the known variables for the first year.

$P = 3000, r = 5, n = 1$

Calculate the interest for the first year using $I = \frac{Prn}{100}$.

$$I = \frac{3000 \times 5 \times 1}{100}$$
$$= 150$$

The value of the investment at the beginning of the second year is the principal plus the interest.

$$P = 3000 + 150$$
$$= 3150$$

Write the values of the variables for the second year.

$P = 3150, r = 5, n = 1$

Calculate the interest for the second year.

$$I = \frac{3150 \times 5 \times 1}{100}$$
$$= 157.50$$

Calculate the value at the beginning of the third year.

$$P = 3150 + 157.50$$
$$= 3307.50$$

Write the values of the variables for the third year.

$P = 3307.50, r = 5, n = 1$

Calculate the interest in the third year.

$$I = \frac{3307.50 \times 5 \times 1}{100}$$
$$= 165.375$$

Calculate the value at the beginning of the fourth year.

$$P = 3307.50 + 165.375$$
$$= 3472.875$$

Write the values of the variables for the fourth year.

$P = 3472.875, r = 5, n = 1$

Calculate the interest in the fourth year.

$$I = \frac{3472.875 \times 5 \times 1}{100}$$
$$= 173.64375$$



	Calculate the value at the end of the fourth year.	$P = 3472.875 + 173.64375$ $= 3646.51875$ ≈ 3646.52
	Write the answer.	The amount of the investment at the end of 4 years is \$3646.52.
b	Interest = final amount of investment – initial principal.	$I = 3646.52 - 3000$ $= 646.52$
	Write the answer.	The compound interest earned after 4 years is \$646.52.

Compound interest is calculated on the amount originally invested together with any interest accumulated. The value of the investment at the end of each time period forms a geometric sequence which can be represented by a recurrence relation of the form $t_{n+1} = rt_n$, where $t_0 = a$.

In Example **6**, the value of the investment at the end of each year is summarised in the table below.

n	t_n
0	\$3000
1	$3000 + 3000 \times 0.05 = 3000 \times 1.05 =$ \$3150
2	$3150 + 3150 \times 0.05 = 3150 \times 1.05 =$ \$3307.50
3	$3307.50 + 3307.50 \times 0.05 = 3307.50 \times 1.05 =$ \$3472.875
4	$3472.875 + 3472.875 \times 0.05 = 3472.875 \times 1.05 =$ \$3646.51875

The value of the investment at the end of any year can be specified by a recurrence relation, since this is calculated using the value of the investment at the end of the previous year.

The sequence in the table above can be generated by the recurrence relation $t_{n+1} = 1.05 \times t_n$, where $t_0 = 3000$. The term t_n is the value of the investment at the end of the nth year and t_0 is the value of the initial deposit. The value of the investment at the end of the 8th year, for example, is t_8.

IMPORTANT

The value of a compound interest investment at the end of the nth compounding period is given by the recurrence relation

$t_{n+1} = rt_n$ where $t_0 = a$

t_n is the value of the investment at the end of the nth compounding period

a is the principal

$$r = 1 + \frac{\text{interest rate per compounding period}}{100}$$

Example 7

$65 000 is invested at 7.5% p.a. compounded yearly. Write down a recurrence relation that can be used to describe the value of the investment at the end of each year.

Solution

Define the variable.

Let t_n be the value of the investment at the end of the nth year.

The investment increases by 7.5% every year.

$$\begin{aligned} t_{n+1} &= t_n + 7.5\% \text{ of } t_n \\ &= t_n(1 + 0.075) \\ &= t_n \times 1.075 \end{aligned}$$

The initial value of the investment was the principal of $65 000.

$t_0 = 65\,000$

Write the difference equation.

The difference equation is:
$t_{n+1} = 1.075 \times t_n$, where $t_0 = 65\,000$

Example 8

Alan inherits $35 000 and decides to invest it in an account where he earns interest of 6.2% p.a. compounded annually.

a Write down a difference equation that can be used to describe the value of his investment at the end of each year.

b What will be the value of the investment after 7 years?

Solution

a Define the variable.

Let t_n be the value of the investment at the end of the nth year.

The investment increases by 6.2% every year.

$$\begin{aligned} t_{n+1} &= t_n + 6.2\% \text{ of } t_n \\ &= t_n(1 + 0.062) \\ &= t_n \times 1.062 \end{aligned}$$

The initial value of the investment was $35 000.

$t_0 = 35\,000$

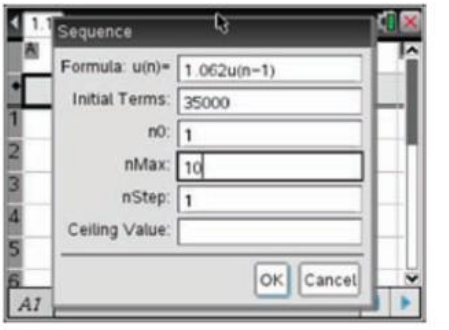
Write the difference equation.

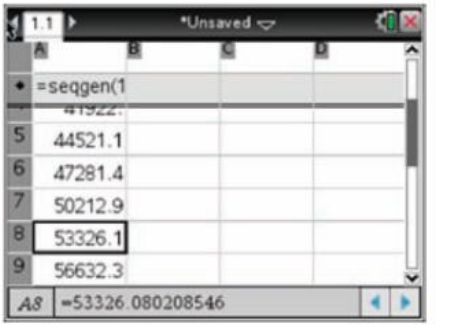
The difference equation is:
$t_{n+1} = t_n \times 1.062$, where $t_0 = 35\,000$



b The value of the investment at the end of the 7th year is required so we need t_7.
Use a CAS calculator to find t_7.

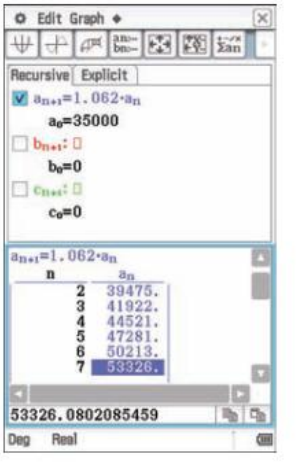
TI-Nspire CAS





t_7 is the term in the 8th row since t_0 is the term in row 1

ClassPad



Write the answer.

The value of the investment after 7 years will be approximately \$53 326.08.

9780170250283

EXERCISE 7.03 Compound interest

Concepts and techniques

1 Example 6 Sam invests $17 000 in an account earning 3% p.a. compounded annually for 2 years. Calculate:
- a the interest earned in the first year
- b the value of the investment at the end of the first year
- c the interest earned in the second year
- d the value of Sam's investment at the end of the two years
- e the total interest that Sam earned.

2 Petra invests $50 000 in an account earning 5.6% p.a. compounded annually for 3 years. Calculate:
- a the value of Petra's investment at the end of the three years
- b the interest Petra earned over the three year period.

3 Example 7 $12 000 is invested at 3.8% p.a. compounded yearly. Write down a recurrence relation that can be used to describe the value of the investment at the end of each year.

4 Write down a recurrence relation that can be used to describe the value of the investment at the end of each year for each of the following.
- a $38 000 is invested at 8% p.a. compounded yearly.
- b $125 000 is invested at 4.6% p.a. compounded yearly.
- c $15 600 is invested at 6.5% p.a. compounded yearly.
- d $25 000 is invested at $7\frac{1}{2}$ % p.a. compounded yearly.

5 The recurrence relation: $t_{n+1} = t_n \times 1.06$, where $t_0 = 15\,000$, represents the yearly value of an investment of:

A $6000 at 15% p.a. compounded yearly
B $15 000 at 10.6% p.a. compounded yearly
C $10 600 at 15% p.a. compounded yearly
D $15 000 at 6% p.a. compounded yearly
E $15 000 at 1.06% p.a. compounded yearly

6 Example 8 Paula wins $40 000 in lotto and decides to invest it in an account where she earns interest of 7.1% p.a. compounded annually.
- a Write down a recurrence relation that can be used to describe the value of her investment at the end of each year.
- b What will be the value of the investment after 10 years?

7 Joshua invests $45 000 in an account where he earns interest of $4\frac{1}{2}$% p.a. compounded annually.
- a Write down a recurrence relation that can be used to describe the value of his investment at the end of each year.
- b What will be the value of the investment after 5 years?



Reasoning and communication

8 $7000 is invested at 5.2% p.a. compounded yearly.
- **a** What will be the value of the investment after 6 years?
- **b** How much interest would be earned after 6 years?

9 Joseph invests $8000 at 6% p.a. compounded six-monthly.
- **a** What is the interest rate for 6 months?
- **b** What will be the value of the investment after 18 months?
- **c** How much interest would be earned after 18 months?

10 Bryson invests $63 000 at 8% p.a. compounded quarterly.
- **a** What will be the value of the investment after 5 years?
- **b** How much interest would be earned after 5 years?

7.04 THE COMPOUND INTEREST FORMULA

In the previous section, you saw that a compound interest investment can be represented by the recurrence relation $t_{n+1} = t_n \times \left(1 + \frac{\text{interest rate per compounding period}}{100}\right)$, where $t_0 = a$.

For example, if $\$P$ is invested at r% per time period, then $t_{n+1} = t_n \times \left(1 + \frac{r}{100}\right)$, where $t_0 = P$. We can then calculate the value of the investment at the end of each year.

	t_n
Start	P
1st year	$P\left(1+\frac{r}{100}\right)$
2nd year	$P\left(1+\frac{r}{100}\right)\times\left(1+\frac{r}{100}\right)=P\left(1+\frac{r}{100}\right)^2$
3rd year	$P\left(1+\frac{r}{100}\right)^2\times\left(1+\frac{r}{100}\right)=P\left(1+\frac{r}{100}\right)^3$
4th year	$P\left(1+\frac{r}{100}\right)^3\times\left(1+\frac{r}{100}\right)=P\left(1+\frac{r}{100}\right)^4$
. . .	
nth year	$P\left(1+\frac{r}{100}\right)^{n-1}\times\left(1+\frac{r}{100}\right)=P\left(1+\frac{r}{100}\right)^n$

This leads us to the general formula for a compound interest investment.

IMPORTANT

The compound interest formula is $A = P\left(1 + \frac{r}{100}\right)^n$

where

A (amount) = **future value** or final balance

P = principal or initial investment

r = percentage interest rate **per compounding period**

n = number of compounding periods

The interest earned can be found by subtracting the principal from the total amount of the investment, i.e., $I = A - P$.

When using the formula, it is important to realise that when interest is not compounded yearly, the values of r and n need to be adjusted accordingly. For example, if interest is compounding monthly, then r will be the interest rate per month and n will be the number of months.

Example 9

Luisa has the choice of investing her $11 000 in two different accounts.
Option A offers interest at a rate of 7.9% p.a. compounded yearly.
Option B offers interest at a rate of 7.7% p.a. compounded monthly.

a Calculate the total amount of the investment after 5 years using:
   i option A ii option B.

b What is the most interest that she can earn using one of these investment options?

Solution

a i Write the values of the known variables.

$P = 11\,000$, $r = 7.9$, $n = 5$

Write the compound interest formula.

$A = P\left(1 + \frac{r}{100}\right)^n$

Substitute the values of P, r and n into the formula.

$= 11000\left(1 + \frac{7.9}{100}\right)^5$

$= 16\,087.920...$

$\approx 16\,087.92$

Write the answer.

With option A the investment will be worth $16 087.92 after 5 years.



ii Interest is compounded monthly so find the monthly interest rate and the number of months in 5 years.

$r = 7.7 \div 12$
$= \frac{7.7}{12}$

$n = 5 \times 12$
$= 60$

Write the values of the known variables.

$P = 11\,000, r = \frac{7.7}{12}, n = 60$

Write the compound interest formula.

$A = P\left(1 + \frac{r}{100}\right)^n$

Substitute the values of P, r and n into the formula.

$A = 11000\left(1 + \frac{\frac{7.7}{12}}{100}\right)^{60}$

$= 16\,145.886...$
$\approx 16\,145.89$

With option B, the investment will be worth \$16 145.89 after 5 years.

b Compare the answers for part **a** and **b** to see which is higher.

Option B gives the best result.

Calculate the interest for option B by subtracting the principal from the value of the investment at the end of 5 years.

$I = A - P$
$= \$16\,145.89 - \$11\,000$
$= \$5145.89$

Write the answer.

The most interest she can earn is \$5145.89 with option B.

The compound interest formula can also be used to calculate the principal, interest rate or time by substituting the values of three known variables and solving the resulting equation.

Example 10

Patrick invests \$14 400 in an account earning interest that is compounded quarterly. In 6 years, the money grows to \$18 700. What was the annual interest rate for his investment? Answer as a percentage, correct to one decimal place.

Solution

Interest is compounded quarterly so find the number of quarters in 6 years.

$n = 6 \times 4$
$= 24$

Write the values of the known variables.

$A = 18\,700, P = 14\,400, n = 24$

Write the compound interest formula.

$A = P\left(1 + \frac{r}{100}\right)^n$

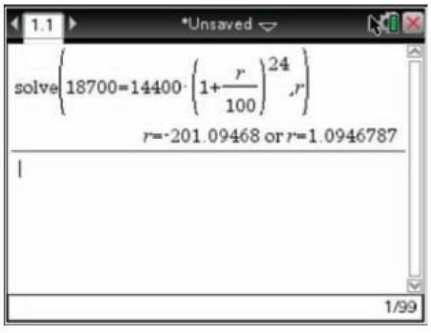
Substitute into the formula.

$18700 = 14400\left(1 + \frac{r}{100}\right)^{24}$

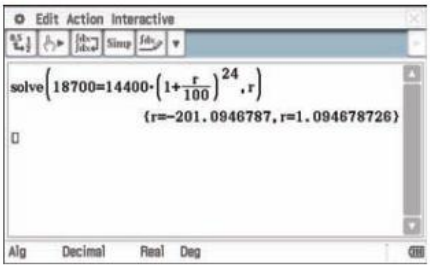
9780170250283

Solve for r.

TI-Nspire CAS



ClassPad

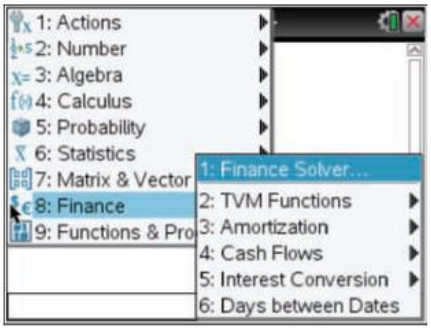


Select the positive solution.	$r = 1.0946 \ldots$ per quarter
r is the interest rate per quarter so multiply by 4 to find the annual interest rate.	$r = 1.0946 \ldots \times 4$ $= 4.3784 \ldots$
Round the answer.	≈ 4.4
Write the answer	The annual interest rate was 4.4%.

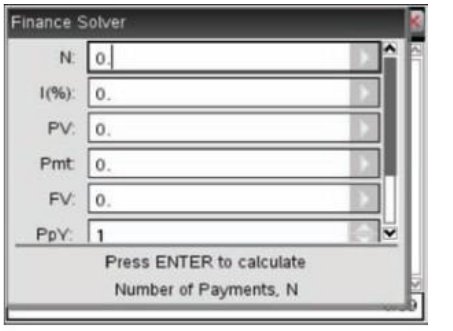
CAS calculators have inbuilt applications called the 'Finance Solver' on the TI-Nspire CAS or 'Financial Application' on the ClassPad. These financial applications can be used to solve compound interest problems.

TI-Nspire CAS

Open a New Document with a Calculator Page.
Press menu, 8 (Finance) and 1 (Finance Solver…).



The fields for the Finance Solver are defined as follows.
N is the number of time periods.
I% is the interest rate as a percentage per annum.
PV is the **present value** (when investing, this is negative as you are giving the bank this amount).
Pmt is the value of any payments being made (for compound interest investments where there is only the initial deposit, this is zero).
FV is the future value (this will be positive as this is the money you get back).
PpY is the number of payments per year.



CpY is the number of times in a year interest is compounded.
PpY and **CpY** take the same value for all compound interest calculations.
PmtAt is when the payment is made, at the 'beginning' or 'end' of the time period. Compound interest is paid at the end of the time period to leave PmtAt set to END.

When using the Finance solver, press tab to move between fields when entering data.

Once the data is entered, press tab to move the cursor to the field representing the unknown quantity and press enter.

ClassPad

Tap Menu followed by the Financial application.

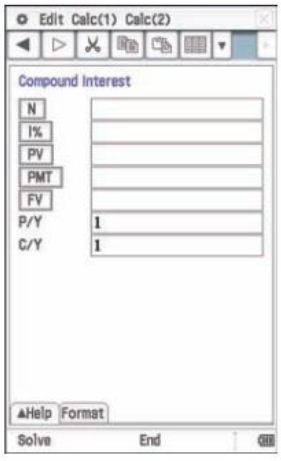


9780170250283

Tap **Calc(1)** on the top toolbar followed by **Compound Interest**.



The fields for the Compound Interest application are defined as follows.
N is the number of time periods.
I% is the interest rate as a percentage per annum.
PV is the present value (when investing, this is negative as you are giving the bank this amount).
PMT is the value of any payments being made (for compound interest investments where there is only the initial deposit, this is zero).
FV is the future value (this will be positive as this is the money you get back).
P/Y is the number of payments per year.
C/Y is the number of times in a year interest is compounded.
P/Y and **C/Y** take the same value for all compound interest calculations.
Tapping Help at the bottom left of the screen gives the definition for a selected field.



Once the data is entered, tap the unknown quantity to find its value.



Example 11

Carina invested $36 000 in a term deposit earning 8.1% p.a. compounding yearly.

a What is the value of her investment at the end of 5 years?

b How much interest did she earn?

Solution

a Write down values of the known quantities.
Interest is compounded yearly, so N is 5 and both PpY (P/Y) and CpY(C/Y) are 1.
The interest rate per annum is 8.1% so I% is 8.1.
$36 000 is being invested so PV is −36 000.

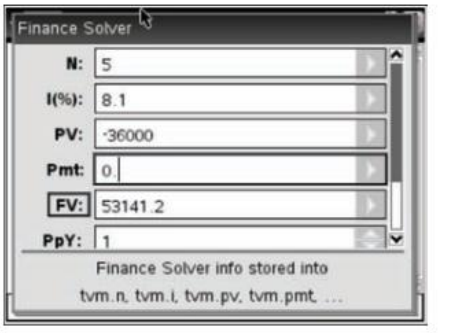
$N = 5$
$I\% = 8.1$
$PV = -36000$
$PMT = 0$
$FV = ?$
$P/Y = 1$
$C/Y = 1$
$PmtAt = END$

When solving problems using a finance solver, it is usual to write down the values entered in each field.

Enter the data in the finance solver or financial application.

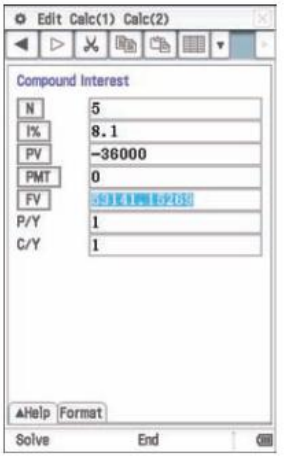
TI-Nspire CAS

Move the cursor to the data entry line for FV and press enter.



ClassPad

Tap Menu followed by the **Financial** application.
Tap **Calc(1)** on the top toolbar followed by **Compound Interest**.
Fill in the fields on the screen using information from the question.
Tap FV.



Write the answer.

Carina will have $53 141.15 at the end of 5 years.

b Subtract the principal from the value of the final investment to find the interest earned.

$I = A - P$
$= 53\,141.15 - 36\,000$
$= 17\,141.15$

Write the answer.

Carina earned $17 141.15 interest.

9780170250283

Example 12

How long will it take for an investment of $12 000 to grow to $24 000 if it is invested at 8% p.a. compounded monthly?

Solution

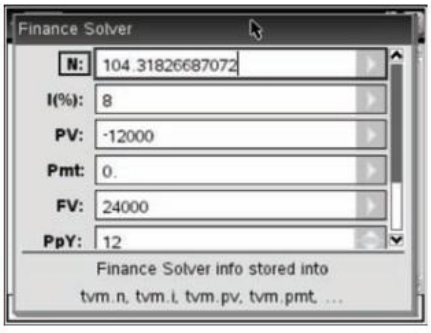
Write down the values of the known quantities.
FV will be positive as that is the amount paid back so FV is 24 000.
The interest rate per annum is 8% so I% is 8.
Interest is compounded monthly so it is paid twelve times in a year. PpY(P/Y) and CpY(C/Y) are both 12.
$12 000 is being invested so PV is −12000.

N = ?
I% = 8
PV = −12000
PMT = 0
FV = 24000
PpY or P/Y = 12
CpY or C/Y = 12

Enter the known values.

TI-Nspire CAS

Move the cursor to the data entry line for N and press enter.



Interest is compounded monthly so N is the number of months.

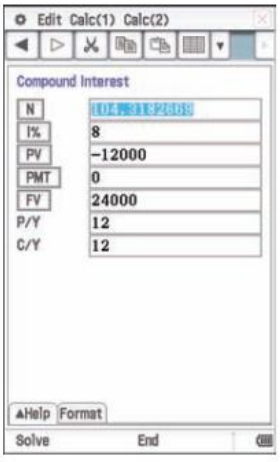
N = 104.318… months

The value of the investment needs to reach $24 000, so investing for 104 months will not be quite long enough. Therefore, round the answer up to the nearest month.

It will take 105 months or 8 years and 9 months for the value of the investment to reach $24 000.

ClassPad

In the **Financial** application, select the **Compound Interest** function as before. Enter the information of the question into the relevant fields. Tap N.



EXERCISE 7.04 The compound interest formula

Concepts and techniques

Using spreadsheets to calculate interest

1 Example 9 Use the formula $A = P\left(1+\frac{r}{100}\right)^n$ to find the future value of each of the following investments.

- **a** $17 000 invested at 5% p.a. compounded yearly for 8 years.
- **b** $125 000 invested at 7.6% p.a. compounded yearly for 4 years.
- **c** $53 500 invested at $6\frac{1}{4}$% p.a. compounded yearly for 12 years.
- **d** $10 000 invested at 4.5% p.a. compounded monthly for 21 years.
- **e** $62 000 invested at 6.3% p.a. compounded six-monthly for $3\frac{1}{2}$ years.
- **f** $8500 invested at $8\frac{1}{2}$% p.a. compounded quarterly for 7 years.
- **g** $147 000 invested at 5.9% p.a. compounded monthly for 4 years.
- **h** $21 500 invested at 4% p.a. compounded daily for 2 years. ← Assume 365 days in a year

2 Find the compound interest earned on each of the investments in question **1**.

3 Example 10 Use the formula $A = P\left(1+\frac{r}{100}\right)^n$ to find the unknown quantity for each of the following investments.

- **a** $57 000 is invested at 6.3% p.a. compounded yearly and grows to $68 500. How long does it take? Round up to the nearest year.
- **b** $9000 is invested for 6 years and grows to $12 500. If it was compounded monthly, what was the interest rate per annum? Answer as a percentage correct to one decimal place.
- **c** $14 500 is invested at $5\frac{1}{2}$% p.a. compounded quarterly and grows to $20 700. How long does it take?
- **d** An amount of money was invested at 5.8% p.a. compounded monthly. It grows to $43 250 in 4 years. How much was invested, correct to the nearest dollar?

4 Example 11 Evan invested $80 000 in an account earning 6.5% p.a. compounded yearly.

- **a** What is the value of his investment at the end of 4 years?
- **b** How much interest did he earn?

5 Sam invested $46 000 in an account earning 5.3% p.a. compounded quarterly.

- **a** Calculate the value of the investment at the end of 4 years.
- **b** How much interest has he earned?

6 The compound interest earned if $16 500 is invested at 8% p.a. compounded six-monthly for 5 years is

A $1346.40 **B** $3574.77 **C** $7743.91
D $7924.03 **E** $19 122.26

7 Example 12 How long will it take for an investment of $8 000 to grow to $14 000 if it is invested at 6.5% p.a. compounded monthly?

8 An amount of money was invested at 7.4% p.a. compounded quarterly. It grows to $37 200 in 3 years. The amount invested, correct to the nearest dollar, is

A $29 854.55 **B** $30 028.26 **C** $35 196.15
D $35 209.50 **E** $36 688.59

9780170250283

9 Which one of the following will give the amount an investment of $125 000 at 6% p.a. compounded monthly will grow to in 7 years?

A $125\,000\left(1+\frac{6}{100}\right)^{7}$ B $125\,000\left(1+\frac{0.5}{100}\right)^{7}$ C $125\,000\left(1+\frac{0.5}{100}\right)^{12}$

D $125\,000\left(1+\frac{6}{100}\right)^{84}$ E $125\,000\left(1+\frac{0.5}{100}\right)^{84}$

10 Matthew has $16 000 to invest for 4 years. Which of the following will give him the greatest return?

A 7.5% p.a. simple interest
B 7.4% p.a. compounded yearly
C 7.3% p.a. compounded quarterly
D 7.2% p.a. compounded monthly
E 7.1% p.a. compounded daily

11 Angela wants to save $65 000 for a holiday. How long will it take her to achieve her goal if she invests $50 000 in a term deposit earning 6.2% p.a. compounded monthly?

12 On the day he was born, Joshua's parents invested money in an account earning interest of 5.65% p.a. compounded monthly. On his 18th birthday, they gave him the money from the account. If he received $18 750, how much did they invest? Answer correct to the nearest hundred dollars.

Reasoning and communication

13 $45 000 was invested at 8.3% p.a. compounded weekly. Calculate the interest earned in $1\frac{1}{2}$ years.

14 $79 000 was invested for 4 years and was compounded six-monthly. If it earned interest of $37 500, what was the interest rate? Answer as a percentage correct to two decimal places.

15 How long will it take an investment to double if it is invested at 7.2% p.a. compounded monthly?

16 If an account has interest compounded monthly, what interest rate will be required for the investment to triple in 4 years? Answer as a percentage p.a. correct to one decimal place.

17 Ariel invests $10 000 in bonds that pay interest at a rate of 8.2% p.a. for a period of 5 years.
a She will receive a half-yearly interest cheque. How much will she receive each half year?
b Alternatively, Ariel could invest her $10 000 in a fund that pays 8% p.a. compounded half yearly with interest paid at the end of 5 years. Which is the better investment option?

18 Lucy is investing $45 000 in government bonds that pay 1.9% per quarter, compounded monthly.
a What is the annual interest rate?
b What will the investment be worth in 6 years?
c How much interest will she earn?

19 An amount of $12 000 is invested for 8 years. The interest rate for the investment is 6.8% p.a. and interest is compounded quarterly. The interest earned by the investment in the fifth year is

A $12\,000 \times 1.017^{6} - 12\,000 \times 1.017^{5}$ B $12\,000 \times 1.017^{5} - 12\,000 \times 1.017^{4}$
C $12\,000 \times 1.017^{20} - 12\,000 \times 1.017^{16}$ D $12\,000 \times 1.068^{5} - 12\,000 \times 1.068^{4}$
E $12\,000 \times 1.068^{20} - 12\,000 \times 1.068^{16}$



20 Shane inherits $75 000 and decides to invest it until he finishes studying. He is deciding between two different investment accounts.

Account A: 7.9% p.a. compounded six-monthly

Account B: 0.64% per month, compounded monthly

a He anticipates that he will finish his degree in 2 years. Calculate the value of his investment if he decides to use:

i Account A ii Account B.

b Which should Shane choose? Justify your answer.

c Shane is contemplating doing post graduate studies at the end of his degree. This will mean he will complete a further 5 years of study, rather than 2. Would this affect his decision in part **b**? Explain.

21 $10 000 is invested at 8% p.a. for 6 years.

a Calculate the interest earned if it was

i a simple interest investment

ii a compound interest investment with interest compounded annually.

b Why does one type of investment earn more interest than the other? Explain.

22 $45 000 is invested at 7% p.a.

a Calculate the value of the investment after 4 years if the interest is compounded:

i yearly ii six-monthly iii quarterly

iv monthly v daily

b As the frequency of compounding increases, what happens to the amount of interest earned? Why?

7.05 INTEREST GRAPHS

A graph is an easy way to show the amount of interest that is earned over a certain period for an investment. They provide a valuable tool in comparing different investments.

A graph of simple interest against time will always resemble an increasing straight-line graph. The dots are never to be joined because interest is calculated in discrete periods (e.g., yearly or monthly), not calculated continuously.

Example 13

Rachel invests $2000 in an account that earns 6% p.a. simple interest.

a Construct a graph that shows the value of the investment, A, in dollars for n years, for values of n from 0 to 5.

b If this were a straight line graph, what would be its gradient? Explain what this represents.

c Use the graph to estimate:

i the value of the investment after 8 years

ii when $1800 in interest will be earned.

9780170250283

Solution

a Write the values of the known variables.

$P = 2000$, $r = 6$, $n = ?$

Write the simple interest formula and substitute the known values.

$$I = \frac{Prn}{100}$$

$$= \frac{2000 \times 6 \times n}{100}$$

$$= 120n$$

Write the formula for the investment amount and substitute the values.

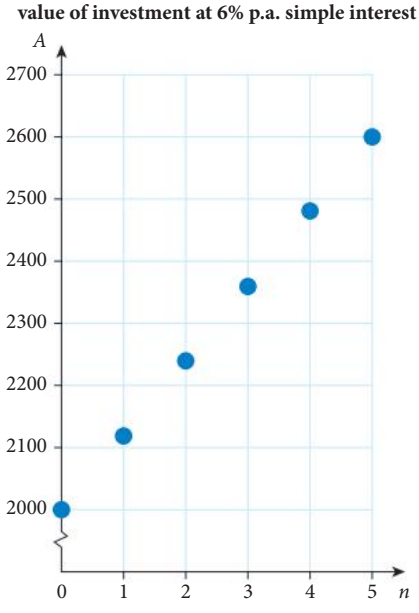
$$A = P + I$$

$$= 2000 + 120n$$

Complete a table of values for $n = 0$ to 5, and complete the row for the value of the investment, A, by substituting n into the formula $A = 2000 + 120n$.

n	0	1	2	3	4	5
A	2000	2120	2240	2360	2480	2600

Construct axes with n on the horizontal axis and I on the vertical axis.
Plot the points from the table of values: (0, 2000), (1, 2120), (2, 2240), (3, 2360), (4, 2480), (5, 2600).



b Rise = 2600 – 2000
Run = 5 – 0

Rise is the vertical distance covered, run is the horizontal distance covered over the same period.

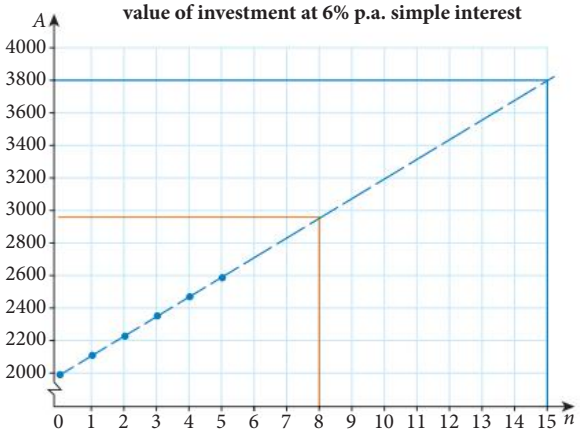
$$\text{Gradient} = \frac{\text{rise}}{\text{run}}$$

$$= \frac{600}{5}$$

$$= 120$$

The gradient would represent the simple interest earned per year.

c i Use a ruler to extend a line through the points to the point where $n = 8$ and read the value of the investment from the vertical axis. (See the red line on the graph).



The value of the investment should be approximately \$2950.

ii When the simple interest earned is \$1800, this means that the total investment would be \$2000 + \$1800 = \$3800. Use a ruler to extend the line through the points to the point where $A = \$3800$ and read off the horizontal axis to find what n is equal to. (See the blue line on the graph).

It will take approximately 15 years for Rachel to earn \$1800 in interest.

A graph of the value that an investment earning compound interest will grow to against time will always resemble an increasing **exponential** graph. The dots should never be joined because interest is calculated in discrete periods (e.g. yearly or monthly), not calculated continuously.

Example 14

Ross invests \$2000 in an account that earns 8% p.a. interest, compounded yearly.

a Construct a graph that shows the amount, A, of the investment in dollars for n years, for values of n from 0 to 10.

b Use the graph to estimate:

i the value of the investment after 15 years

ii when the value of the investment reaches \$5000.

Solution

a Write the values of the known variables. $P = 2000, r = 8$

Write the compound interest formula. $A = P\left(1 + \frac{r}{100}\right)^n$

Substitute into the formula. $= 2000\left(1 + \frac{8}{100}\right)^n$

$= 2000 \times 1.08^n$

9780170250283

Construct a table of values for $n = 0$ to 10, and complete the row for the Amount of the investment, A, by substituting n into the formula $A = 2000 \times 1.08^n$.

For $n = 2$,

$A = 2000 \times 1.08^2$

$= 2332.8$

≈ 2333

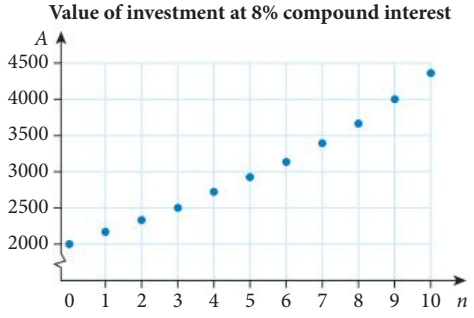
n	0	1	2	3	4	5
A	2000	2160	2333	2519	2721	2939

n	6	7	8	9	10
A	3174	3428	3702	3998	4318

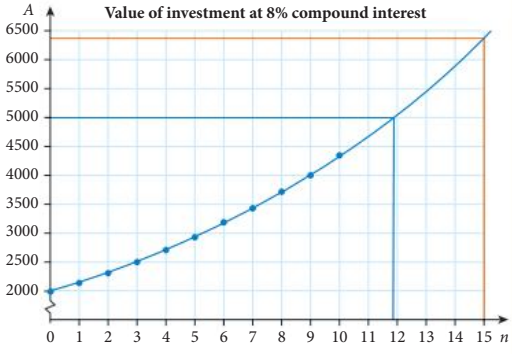
Graph as accurately as possible but it is important to remember that on a hand drawn graph it is not possible to graph with this level of accuracy.

Draw up an axes with n on the horizontal axes and A on the vertical axes.

Plot the points from the table of values: (0, 2000), (1, 2160), …, (10, 4318).



b i Use a pencil to lightly sketch the exponential graph that is forming to the point where $n = 15$. Read off the vertical axis to find what A should be equal to (see the red line on the graph).



The value of the investment should be approximately \$6350.

ii Use a pencil to lightly sketch the exponential graph through the existing points to the point where $A = \$5000$. Read off the horizontal axis to find the value of n (see the blue line on the graph).

It will take approximately 12 years for the investment to reach \$5000.

A CAS calculator can also be used to create graphs.

Example 15

Emily has $6000 to invest. She is comparing two different accounts. They offer:
Account 1: simple interest at 6.5% p.a.
Account 2: compound interest at 6% p.a. compounded annually.

a Create a graph that shows the value of each investment over a 10 year period on the same number plane.

b What would your advice to Emily be?

Solution

a Find an expression for the value of each investment after n years.

Account 1
Use the simple interest formula with $P = 6000$ and $r = 6.5$.

$$I = \frac{Prn}{100}$$
$$= \frac{6000 \times 6.5 \times n}{100}$$
$$= 390n$$

The value of the investment after n years is equal to the sum of the principal and the interest.

$$A = P + I$$
$$= 6000 + 390n$$

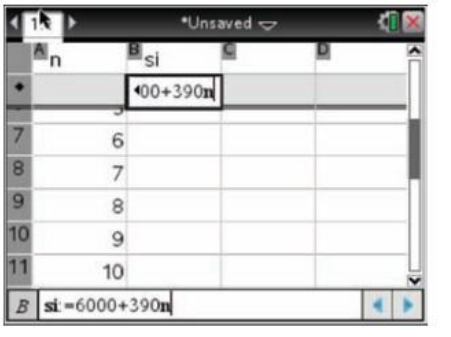
Account 2
Use the compound interest formula with $P = 6000$ and $r = 6$.

$$A = P\left(1 + \frac{r}{100}\right)^n$$
$$= 6000\left(1 + \frac{6}{100}\right)^n$$
$$= 6000 \times 1.06^n$$

Construct a graph of each investment over 10 years using a CAS calculator.

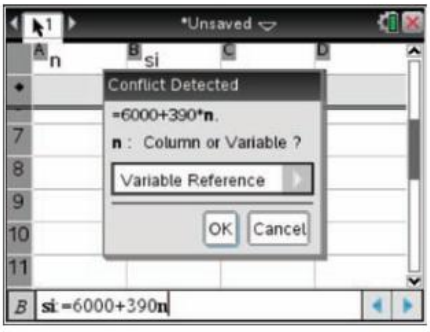
TI-Nspire CAS

Open a New Document with a Lists & Spreadsheet page.
Give column A the label n.
Type in the values 0 to 10 inclusive.
Give column B the label si.
In the shaded entry line, type the formula
= 6000 + 390n
Then press enter.

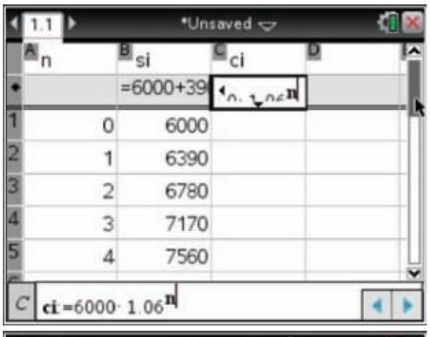


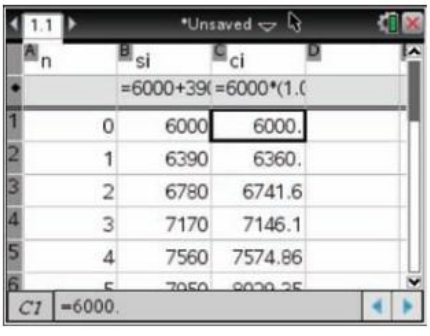
9780170250283

In the pop-up window that appears, select Variable Reference then click OK.

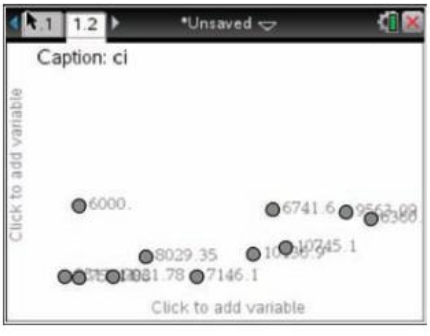


Give column C the label ci.
In the shaded entry line type the formula $= 6000 \times 1.06^n$
Then press enter.
In the pop up window that appears select variable reference then click OK.



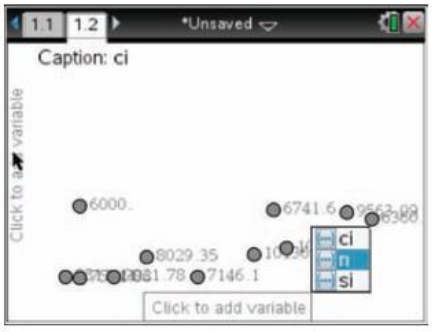


Add a Data & statistics page.

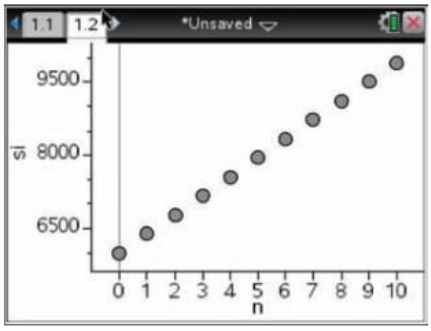




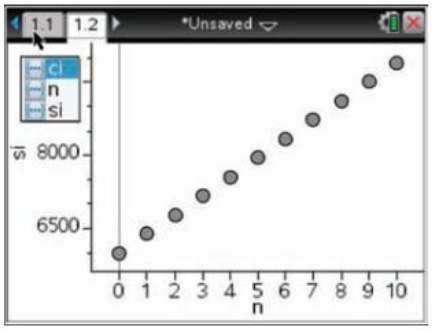
Move the cursor to the 'Click to add variable' box at the bottom of the screen and select n.
Move the cursor to the 'Click to add variable' box at the left of the screen and select si.

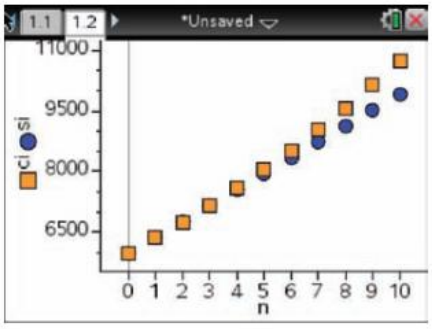


This gives a graph showing the value of the simple interest investment over 10 years.



Press menu then 2: Plot properties then 8: Add Y variable.
Select ci.
This graph now displays the value of each investment over 10 years.

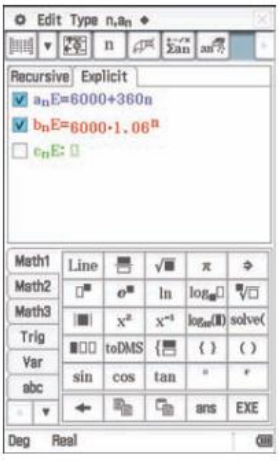




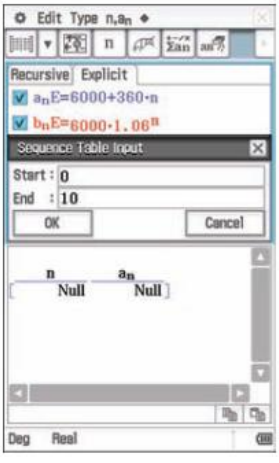
9780170250283

ClassPad

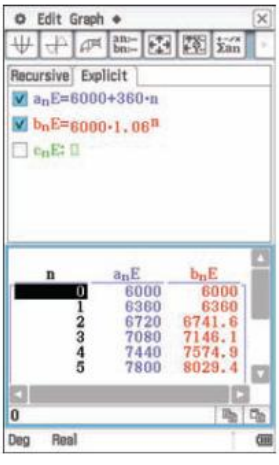
Use the **Sequence** application.
Tap on the **Explicit** tab.
Type 6000 + 360n next to anE:
Type 6000×1.06^n next to bnE:
Make sure that both anE and bnE are checked.



Tap and set Start to 0 and End to 10.
Tap OK.

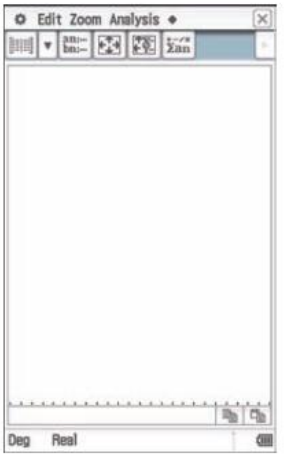


Tap to create a table showing the value of each investment.

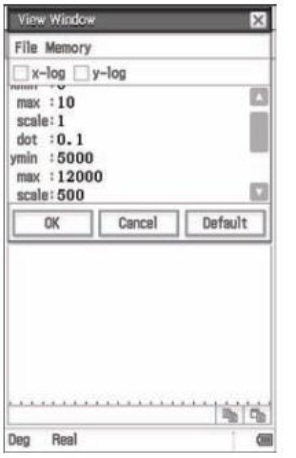




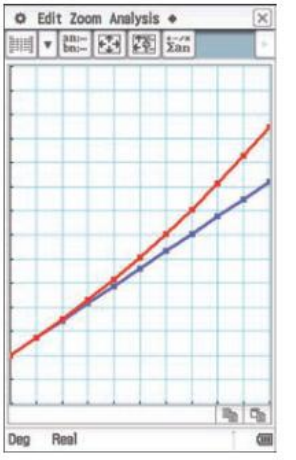
Tap , then Resize .
You may not see a graph yet.



Tap set **xmin:** 0 and continue to complete the screen as shown.



Tap OK.
This graph displays the value of each investment over 10 years.



b Looking at the graphs, the simple interest graph lies above the compound interest graph until approximately $n = 4$ years. Then the compound interest graph lies above the simple interest graph.

If Emily is going to invest for less than 4 years, then she would be better off with Account 1 but if she is going to invest for longer than that, then Account 2 is the better option.

9780170250283

EXERCISE 7.05 Interest graphs

Concepts and techniques

Graphing by hand

Graphing on the CAS

1 Example 13 Use the simple interest graph from Example **13** to estimate the value of the investment when $2000 is invested at 6% p.a. for:
a 9 years **b** 12 years **c** 15 years.

2 Use the simple interest graph from Example **13** to estimate when a principal of $2000 invested at 6% p.a. will have earned simple interest to the value of:
a $1200 **b** $1560 **c** $2040.

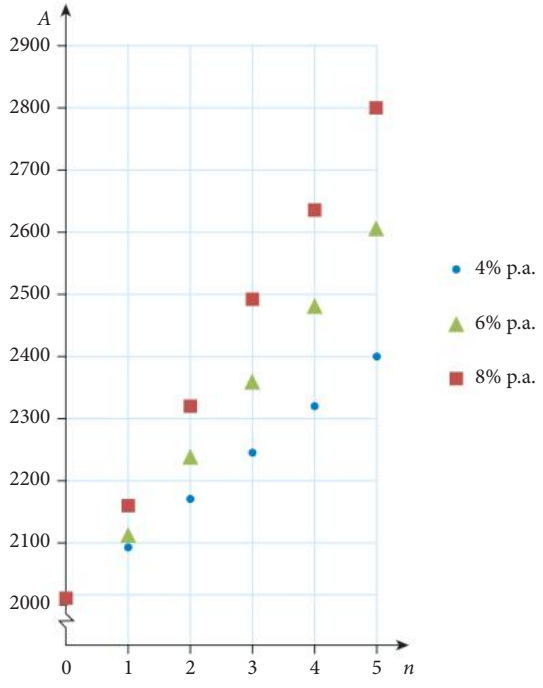
3 Graph on the same axes the value that the investment, A, grow to in n years, for values of n from 0 to 6, when $4000 is invested in an account earning:
a 4% p.a. **b** 12% p.a.

4 **a** Use your graphs from question **3** to estimate the value of an investment with a principal of $4000 earning simple interest
i over 10 years at 4% p.a. **ii** over 8 years at 12% p.a.
b How do the two different interest rates appear on the graphs?
c What would be the gradient of the graph you drew in question **3a**?
d What would be the gradient of the graph you drew in question **3b**?
e What does the gradient of each of the graphs represent?

5 The graph on the right shows the value of an investment with a principal of $2000, invested for up to 5 years at three different simple interest rates per annum: 4%, 6% and 8%.

Use the graph to estimate:
a the value of the investment earning 6% p.a. after 7 years
b the interest rate when the investment reaches $2960 after 6 years
c the number of years it takes for the investment to reach $2720 at 4% p.a.
d the time it takes to earn $1920 in interest at 8% p.a.
e the interest earned after 15 years at 4% p.a.
f the interest rate if $1560 interest is earned after 13 years.

Value of investment at different simple interest rates



6 Example 14 Use the compound interest graph from Example **14** to estimate the value of a $2000 investment earning 8% p.a. interest compounded yearly for:
a 13 years **b** 15 years

7 Use the compound interest graph from Example **14** to estimate when each of the following amounts will be reached if $2000 is invested at 8% p.a. interest compounded yearly.
   a $5800 b $4000

8 A $5000 investment is earning interest compounded yearly for n years at a rate of:
   i 4% p.a. ii 12% p.a.
   a For each rate, complete the table of values below, showing the amount that the investment will grow to (A) for values of n from 0 to 10.

n	0	1	2	3	4	5	6	7	8	9	10
A											

   b Graph on the same axes the amount the investment will grow to (A) for values of n from 0 to 10.

9 a Use your graph from question **8** to estimate the value of an investment of $5000
   i after 15 years at 4% p.a. ii after 12 years at 12% p.a.
   b How do the two different interest rates appear on the graphs?

10 Example 15 For the scenarios listed **a**–**d** below,
   i graph the values which an investment of $1000 will grow to for $n = 0$ to 10
   ii find the approximate value of each investment after its end date.
   a Simple interest at 5.8% p.a for 15 years.
   b Compound interest at 7.2% p.a. compounded annually for 15 years.
   c Compound interest at 6.4% p.a. compounded six-monthly for 10 years.
   d Compound interest at 1.1% per month compounded monthly for 2 years.

11 For the scenarios listed **a**–**d** below,
   i graph the values that an investment of $1000 will grow to for $n = 0$ to 10
   ii find the approximate time for each investment to earn the required interest.
   a Sienna invests in an account with simple interest at 9.2% p.a and wants to earn $1500 in interest.
   b Ava invests in an account with compound interest at 9.2% p.a. compounded annually and wants to earn $1800 in interest.
   c Isaac invests in an account with compound interest at 7.2% p.a. compounded monthly, and wants to earn $150 in interest.
   d Caleb invests in an account with compound interest at 2.4% per quarter compounded quarterly, and wants to earn $500 in interest.

Reasoning and communication

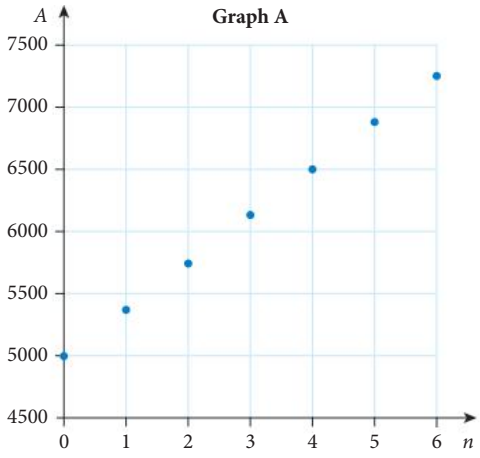
12 Matthew has $15 000 to invest. He is comparing two different options.
Option 1: Compound interest at 7.5% p.a. compounded quarterly.
Option 2: Compound interest at 3.9% per six months compounded six-monthly.
   a On the same number plane, graph the future value of the investment for each option.
   b What would your advice to Matthew be?

13 Maddy has $50 000 to invest. She is comparing three different investments.
Investment 1: Simple interest at 8% p.a.
Investment 2: Compound interest at 8% p.a. compounded quarterly.
Investment 3: Compound interest at 8% p.a. compounded six-monthly.
   a Without doing any graphing, can you tell which investment would be the better option? Why?
   b Graph the future value of the investments on the same number plane.
   c Were you correct in part **a**?
   d What would your advice to Maddy be?

14 **a** Graph A on the right shows an investment with simple interest.

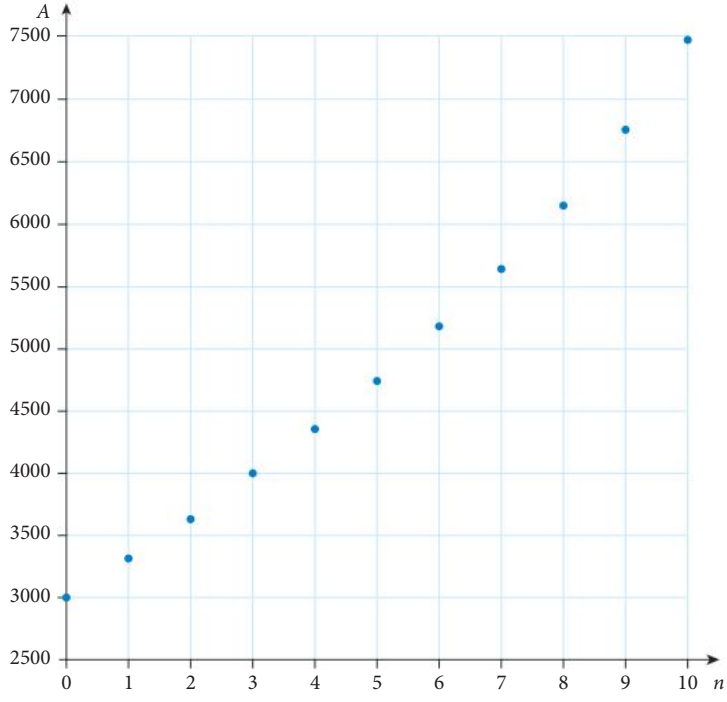
- **i** What is the principal amount invested?
- **ii** Write down the formula used to calculate the total value of an investment using simple interest.
- **iii** If the final point on the graph is (6, 7250), substitute the known values into the formula you wrote in part **ii** and rearrange it to find the interest rate.
- **iv** Find the amount of interest paid annually.

Graph A



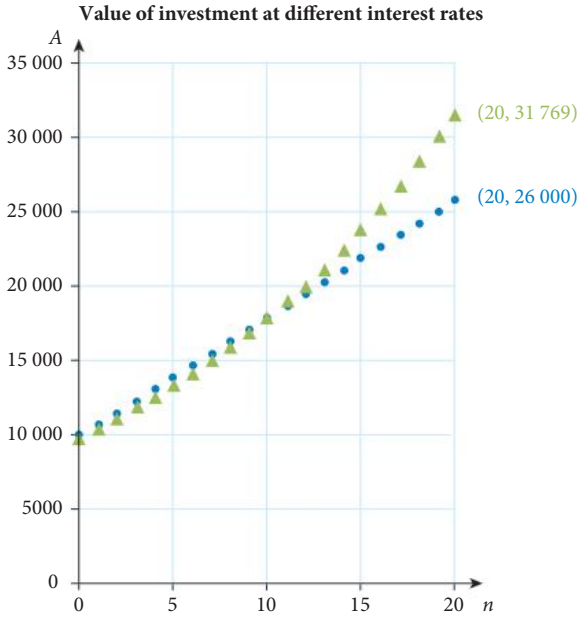
b Graph B shows an investment with compound interest, compounding annually.

Graph B



- **i** What is the principal amount invested?
- **ii** Write down the formula used to calculate the total value of an investment using compounding interest. Answer correct to one decimal place.
- **iii** If the final point on the graph is (10, 7435), substitute the known values into the formula you wrote in part **ii** and find the interest rate, correct to one decimal place.
- **iv** What will be the value of the investment in 20 years?

15 Consider this graph.



a Which set of points represents the balance of an investment earning simple interest and which set of points represents the balance of an investment earning compound interest? How can you tell?

b Which account offers a higher interest rate? How can you tell? (Hint: look at the first set of points.)

c Write the formulas for finding the total value of an investment for each of these set of points.

d Using the points shown on the graph, substitute the values into the formulas.

e Find the interest rates for both sets of points, correct to one decimal place.

f Hence find the value of each investment after 30 years.

INVESTIGATION Comparing interest

$200 000 is invested at 7% p.a.

a If it was a simple interest investment, calculate the interest earned after:

i 1 year ii 2 years iii 5 years

iv 7 years v 10 years

b If it was a compound interest investment with interest compounded annually, calculate the interest earned, correct to the nearest dollar, after:

i 1 year ii 2 years iii 5 years

iv 7 years v 10 years

c On the same axes, graph the interest earned (I) against the number of years (n) for the investments in part **a** and **b**. Use n from 0 to 10.

d Which is the better investment over 10 years?

e Why does one type of interest earn more than the other? Explain.

$45 000 is invested at 7% p.a.

f Calculate, correct to the nearest dollar, the interest (I) earned after 4 years if the interest is compounded:

i yearly ii six-monthly iii quarterly

iv monthly v daily

g Construct a graph showing the number of compounding periods in one year on the horizontal axis and the amount of interest earned on the vertical axis.

9780170250283

h Which is the better investment over 4 years?

i Between what two values would you expect the amount of interest to lie if compounded weekly?

j Calculate the amount of interest if it was compounded weekly. Were you correct?

k As the frequency of compounding increases, what happens to the amount of interest earned? Why?

l Without doing any calculations, which is the better investment option in each pair below?

i 7.5% p.a. simple interest or 7.5% p.a. compounded yearly?

ii 6.9% p.a. compounded monthly or 6.9% p.a. simple interest?

iii 5.85% p.a. compounded weekly or 5.85% p.a. compounded daily?

iv $7\frac{1}{4}$% p.a. compounded yearly or $7\frac{1}{4}$% p.a. compounded half-yearly?

v 11% p.a. compounded quarterly or 11% p.a. compounded monthly?

m Check your answers to part **l** by calculating the interest earned on $10 000 for each of the investments over 2 years.

7.06 EFFECTIVE INTEREST RATE

The interest rate quoted for a loan or investment is also known as the **nominal interest rate.** The amount of interest paid for a compound interest investment depends primarily on the nominal interest rate and the term. However, it is also affected by the frequency of the compounding periods. The **effective interest rate** is used to compare the interest paid on loans (or investments) with the same nominal annual interest rate i but with different compounding periods (daily, monthly, quarterly, annually, other).

As a decimal, it is the interest that would be paid on $1 invested for 1 year. As a percentage, it is the interest that would be paid on $100 invested for 1 year.

IMPORTANT

The effective interest rate can be found using:

$$e = \left(1 + \frac{i}{n}\right)^n - 1$$

where

e = effective yearly interest rate as a decimal

i = nominal yearly interest rate as a decimal

n = number of compounding periods in a year



Example 16

What is the effective rate of interest for an investment earning 7% p.a. compounded monthly?

Solution

Write the values of the known variables. Interest is compounded monthly, so there are 12 compounding periods per year.

$i = 7\% = 0.07,\ n = 12$

Write the effective interest rate formula.

$$e = \left(1 + \frac{i}{n}\right)^n - 1$$

Substitute the values of i and n into the formula.

$$= \left(1 + \frac{0.07}{12}\right)^{12} - 1$$

$$= 0.07229\ldots$$

$$\approx 7.23\%$$

Write the answer.

The effective rate of interest is approximately 7.23% p.a.

Example 17

Liana is looking to invest her money. She has done some research and found the best offers from three different banks.

Bank 1: 8.45 % p.a. compounded daily

Bank 2: 8.6 % p.a. compounded monthly

Bank 3: 8.7 % p.a. compounded six-monthly

Which bank should Liana choose if she wants to earn the most interest?

Solution

Find the effective rate of interest for each bank.

Bank 1

Write the values of the known variables.

$i = 8.45\% = 0.0845,\ n = 365$

Substitute into the formula.

$$e = \left(1 + \frac{0.0845}{365}\right)^{365} - 1$$

$$= 0.08816...$$

$$\approx 8.82\%\ \text{p.a.}$$

9780170250283

Bank 2

Write the values of the known variables.
$i = 8.6\% = 0.086,\ n = 12$

Substitute into the formula.
$$e = \left(1 + \frac{0.086}{12}\right)^{12} - 1$$
$$= 0.08947...$$
$$\approx 8.95\% \text{ p.a.}$$

Bank 3

Write the values of the known variables.
$i = 8.7\% = 0.087,\ n = 2$

Substitute into the formula.
$$e = \left(1 + \frac{0.087}{2}\right)^{2} - 1$$
$$= 0.08889...$$
$$\approx 8.89\% \text{ p.a.}$$

Write the answer.

Liana should choose Bank 2 as it pays the higher effective rate of interest and will therefore pay more interest.

EXERCISE 7.06 Effective interest rate

Concepts and techniques

1 Example 16 Determine the effective interest rate for each of the following investments. Answer as a percentage, correct to two decimal places.

a 9% p.a. compounding monthly
b 11% p.a. compounding weekly
c 12% p.a. compounding half-yearly
d 6% p.a. compounding daily

2 What is the effective interest rate (answer as a percentage, correct to two decimal places) for:

a 7.5% p.a. compounding monthly?
b 13.4% p.a. compounding daily?
c 8.4% p.a. compounding quarterly?
d 8% p.a. compounding fortnightly?

3 Which one of the following has the highest effective interest rate?

A 5.8 % p.a. compounded monthly
B 5.4 % p.a. compounded daily
C 5.3 % p.a. compounded fortnightly
D 5.6% p.a. compounded six-monthly
E 5.85 % p.a. compounded quarterly

Reasoning and communication

4 Example 17 Georgio has the choice of three investments at his bank. He can invest at 7.6% p.a. compounded quarterly, 7.5% p.a. compounded monthly or 7.4% p.a. compounded daily. Calculate the effective rate for each option and decide which is best.

5 Kylie invests $10 000 at 8.2% p.a. compounded weekly. What is the effective rate of interest? Answer correct to two decimal places.



6 Yannick inherits $30 000 from his aunt. He is offered the following investment options.

- A term deposit at 7.5% p.a. compounded daily
- Insurance bonds at 7.7% p.a. compounded monthly
- Government bonds at 7.9% p.a. compounded annually.

Which investment should Yannick choose if he wants to earn the most interest?

7 Find the effective interest rate for each of the following. Answer correct to two decimal places.

a 8.05% p.a. compounding weekly
b 10.13% p.a. compounding monthly
c 14.24% p.a. compounding fortnightly
d 17.11% p.a. compounding daily

8 Which investment would provide the best return: 6.08% p.a. compounding quarterly or 5.94% p.a. compounding daily?

9 **a** Without performing any calculations, which investment option will have the higher effective rate of interest: 6.2 % p.a. compounded weekly or 6.2 % p.a. compounded quarterly? Explain your reasoning.

b Check your answer to part **a**.

9780170250283

CHAPTER SUMMARY
LOANS AND INVESTMENTS

7

- **Simple interest** is interest calculated as a percentage of the **principal**.
- Simple interest can be calculated using the formula: $I = \frac{Prn}{100}$, where I = simple interest, P = principal, r = interest rate per period and n = number of periods.
- $A = I + P$, where A is the final value of an investment.
- If the amount of interest is known, the simple interest formula can also be used to calculate one of the other values, principal, interest rate or time. Substitute in the given values and solve the resulting equation.
- **Hire-purchase** means 'paying off' an expensive item over time, sometimes after a deposit is made. By calculating the total amount paid, the interest rate that has been charged can be calculated.
- A deferred payment plan is when there are no repayments until a later date. If the item is not paid off fully by the agreed date, interest is charged on the amount still owed from the date of purchase. This can be a very expensive option.
- Compound interest is interest that is added to the principal and then the total amount is reinvested. It can be calculated using repeated applications of the simple interest formula but it is much easier to use the compound interest formula.
- The compound interest formula is $A = P\left(1+\frac{r}{100}\right)^n$ where
  A (amount) = future value or final balance
  P = principal or initial investment
  r = interest rate **per compounding period**, expressed as a percentage
  n = number of compounding periods
- The interest earned can be found by subtracting the principal from the total amount of the investment, i.e., $I = A - P$.
- When using the formula, it is important to realise that when interest is not compounded yearly, the values of r and n need to be adjusted accordingly.
- The formula can also be used to calculate the principal, interest rate or time by substituting in the given information and solving the resulting equation.
- A graph of simple interest against time will always give an increasing straight-line graph.
- A graph of the future value of an investment that earns compound interest against time will always give an increasing **exponential graph**.
- The **effective interest rate** is used to compare the interest paid on loans (or investments) with the same nominal annual interest rate i, but with different compounding periods (daily, monthly, quarterly, annually, other).
- The effective interest rate can be found using:
  $e = \left(1+\frac{i}{n}\right)^n - 1$
  where
  e = effective yearly interest rate as a decimal
  i = nominal yearly interest rate as a decimal
  n = number of compounding periods in a year.

7 CHAPTER REVIEW
LOANS AND INVESTMENTS

Multiple choice

1 Example 2 Mark invests \$23 500 at $8\frac{1}{2}$% p.a. simple interest. How much will he have after 9 months?

A \$1498.13 B \$5522.50 C \$17 977.50
D \$24 998.13 E \$41 477.50

2 Example 4 Jayden purchased a new jet ski valued at \$6200 on terms of 10% deposit and monthly payments of \$265.98 for two years. What is the interest rate charged per annum?

A 4.29% B 6.48% C 7.2% D 12.96% E 14.4%

3 Example 7 The difference equation $t_{n+1} = 1.05t_n$ where $t_0 = 27\,000$, represents an investment of:

A \$5000 at 27% p.a. compounded yearly?
B \$27 000 at 10.5% p.a. compounded yearly?
C \$10 500 at 27% p.a. compounded yearly?
D \$27 000 at 5% p.a. compounded yearly?
E \$27 000 at 1.05% p.a. compounded yearly?

4 Example 9 Josephine buys government bonds worth \$98 500. They are paying 9% p.a. compounded monthly. Which calculation will give the value of the bonds in 6 years?

A $98\,500\left(1+\frac{0.75}{100}\right)^{6}$ B $98\,500\left(1+\frac{0.75}{100}\right)^{72}$ C $98\,500\left(1+\frac{9}{100}\right)^{72}$

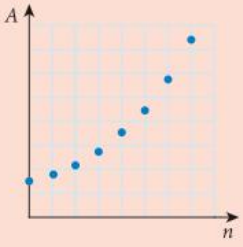
D $98\,500\left(1+\frac{9}{100}\right)^{6}$ E $98\,500\left(1+\frac{9}{12}\right)^{72}$

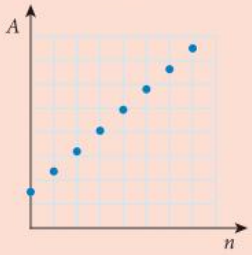
5 Example 11 \$28 000 is invested for 4 years. Which one of the following investment options will give the greatest return?

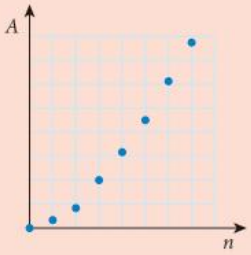
A 6.5% p.a. simple interest B 6.1% p.a. compounded yearly
C 5.95% p.a. compounded monthly D 0.5% per month, compounded quarterly
E 5.9% p.a. compounded daily

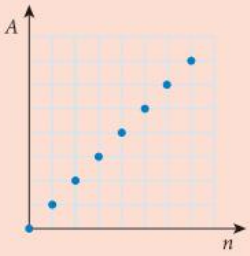
9780170250283

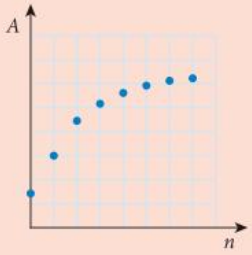
6 Example 14 Jane makes a compound interest investment. Which of the following graphs displays the value of the investment after n compounding periods?

A 

B 

C 

D 

E 

7 Example 17 Which has the lowest effective interest rate?

A 7.9% p.a. compounded quarterly
B 7.7% p.a. compounded fortnightly
C 7.5% p.a. compounded monthly
D 7.45% p.a. compounded daily
E 7.58% p.a. compounded six-monthly

Short answer

8 Example 1 Find the weekly interest paid when \$79 000 is invested at 9.4% p.a. simple interest.

9 Example 3 Isaac invests money in an account earning 7.2% p.a. simple interest. How much did he invest if he earned \$1750 interest in $1\frac{1}{2}$ years? Answer correct to the nearest hundred dollars.

10 Example 3 \$14 000 grows to \$16 550 when it is invested for 3 years in an account earning simple interest. What was the annual interest rate? Answer as a percentage correct to two decimal places.

11 Example 4 Joanna bought a snowboard and gear for \$1570 on terms of 15% deposit and 9.3% p.a. simple interest on the balance over 18 months.

a How much interest did she pay?
b She paid it off in equal fortnightly instalments. Calculate the value of each instalment.

12 Example 5 A company offered a deferred payment plan of 'buy now and pay nothing for 150 days'. If it wasn't paid in full by the end of the 150 days, they charged interest of 0.13% per day simple interest on the outstanding amount from the purchase date.

Sharon bought a lounge suite for \$2949 and paid it off in full after 160 days.

a Calculate how much she had to pay in total.
b How much would she have paid had she paid for it in the 150 days?

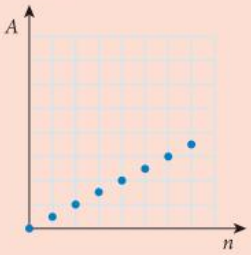
13 Example 8 $15 000 is invested at 6.9% p.a. compounded yearly.
- a Write down a difference equation to describe the value of the investment at the end of each year.
- b What will be the value of the investment after 5 years?

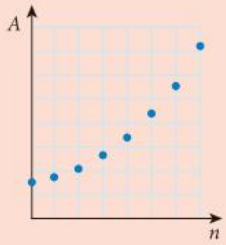
14 Example 11 Jack invests $1700 at 8.1% p.a. compounded six-monthly.
- a Calculate the amount his investment will grow to in 3 years.
- b Calculate the interest earned.

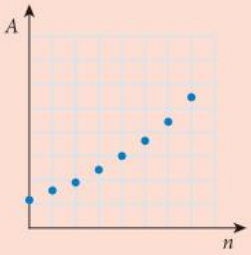
15 Example 12 Peter wants to buy a new car in 3 years. He has an account that earns 8.7% p.a. compounded monthly. How much should he invest now if he estimates that he will need $45 000? Round up to nearest dollar.

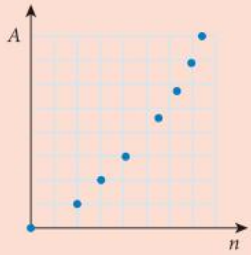
16 Example 12 $72 500 was invested at 7.1% p.a. compounded quarterly. How long would it take to earn $25 000 interest?

17 Example 13, 14 Match each of the following situations to its graph below.
- a Simple interest, I, earned on an investment of $6000 at 3% p.a. over n years.
- b Simple interest, I, earned on an investment of $6000 at 15% p.a. over n years.
- c The amount, A, of an investment after $5000 is invested at 2.2% p.a. compounded yearly over n years.
- d The amount, A, of an investment after $5000 is invested at 10.75 % p.a. compounded yearly over n years.

A


B


C


D


18 Example 15 Chloe has $8000 to invest. She is comparing two different options. They offer:

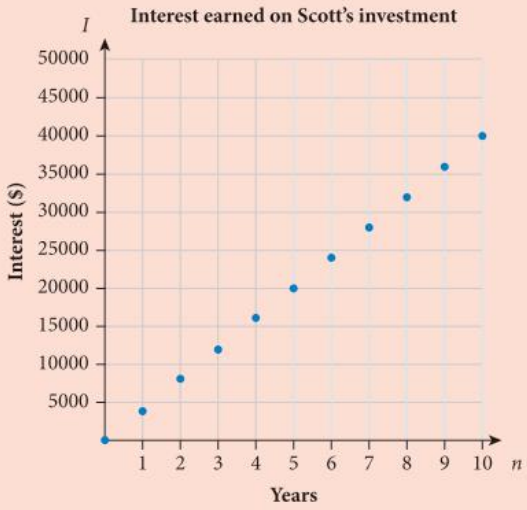
Option 1: Simple interest at 7.5% p.a.
Option 2: Compound interest at 7% p.a. compounded annually.
- a On the same number plane, graph the values that the investments will grow to using each account.
- b What would your advice to Chloe be?

19 Example 16 What is the effective interest rate for an investment earning 9.5% p.a. compounded quarterly? Answer as a percentage, correct to two decimal places.

9780170250283

Application

20 Scott, Grant and Jodie each inherited money from their grandmother. The graph below represents the interest (I) Scott earned on his investment over n years.



a Explain how you know that this graph represents an investment earning simple interest.

b From the graph, estimate:

i the amount of interest earned after 8 years

ii how long it will take the investment to earn $10 000 interest.

c Calculate the gradient if, after 10 years, the interest is $40 000.

d The gradient of the line represents the simple interest earned per year. What is the simple interest earned per year?

e Show that Scott's initial investment was $50 000.

Grant also invested $50 000 for ten years but the investment he chose paid interest of 7.2% p.a. compounded annually.

f Calculate how long will it take him to earn $10 000 interest.

g Calculate:

i the amount his investment will grow to in 10 years

ii his interest earned in 10 years.

h Complete the table of values for the interest (I) Grant earned on his investment over n years. Round I to the nearest hundred dollars.

n	0	1	2	3	4	5	6	7	8	9	10
I											

i On the axes above showing Scott's interest, graph the interest (I) Grant earned on his investment over n years.

j When is Grant's investment the best option?

k Jodie invests her money in an account earning 7.05% p.a. compounding monthly. By calculating the effective interest rate, decide who made the better investment option: Grant or Jodie?

Practice quiz


TE
ROUGHOUT
ERRITORIES
TEN DOLLARS TEN DOLLARS TEN DOLLARS
HE WAS RIGHT AMONG THE HORSES
WILD HORSES RACING YET WITH TH
AND HE RAN THEM SINGLE HANDED
AND DOWN BY KOSCIUSKO WHERE
THE COLD AND FROSTY SKY
WORD TODAY AND THE STOCK
THERE WAS MOVEMENT
SO ALL THE CRACKS
RIDING WHERE THE WILD BUSH HORSES ARE
TEN DOLLARS TEN DOLLARS TEN DOLLARS
HE HAILS FROM SNOWY
MAN THAT HOLDS HIS
SINCE I FIRST COMMENCED
TEN DOLLARS TEN DOLLARS
FIERCELY WE MAY
TEN DOLLARS TEN DOLLARS
WHEN THEY REACHED THE MOUNTAINS SUMMIT
OF WOMBAT HOLES AND ANY SLIP WAS DEATH
MOUNTAIN LIKE A TORRENT DOWN ITS BED
TEN DOLLARS TEN DOLLARS TEN DOLLARS
HE SENT THE FLINT STONES FLYING OUT
MOUNTAIN HORSEMEN RIDE THROUGH
LANDED SAFE AND SOUND AT THE
TEN DOLLARS TEN DOLLARS TEN DOLLARS
HE WAS RIGHT AMONG THE HORSES AS THEY CLIMBED THE
HE RACED ACROSS THE CLEARING IN PURSUIT THEN THEY LOST
WILD HORSES RACING YET WITH THE MAN FROM SNOWY RIVER
TEN DOLLARS TEN DOLLARS TEN DOLLARS TEN DOLLARS
AND HE RAN THEM SINGLE HANDED TILL THEIR SIDES WERE
HOME AND ALONE AND UNASSISTED BROUGHT THEM BACK
STILL UNDAUNTED AND HIS COURAGE FIERY HOT FOR NEW
TEN DOLLARS TEN DOLLARS TEN DOLLARS TEN DOLLARS
AND DOWN BY KOSCIUSKO WHERE THE PINE CLAD RIDGES
THE COLD AND FROSTY SKY AND WHERE AROUND THE
WORD TODAY AND THE STOCKMEN TELL THE STORY OF
TEN DOLLARS TEN DOLLARS TEN DOLLARS TEN DOLLARS
THERE WAS MOVEMENT AT THE STATION FOR THE
SO ALL THE CRACKS HAD GATHERED TO THE
RIDING WHERE THE WILD BUSH HORSES ARE
TEN DOLLARS TEN DOLLARS TEN DOLLARS
AND ONE
rent at the Playhouse
LARS



8

TERMINOLOGY

annuity
annuity investment
compounding period
deposit
future value
interest rate
outstanding balance
perpetuity
present value
reducing balance loan
repayment
withdrawal

Prior learning

LOANS, INVESTMENTS AND ANNUITIES

REDUCING BALANCE LOANS AND ANNUITIES

8.01 Reducing balance loans
8.02 Using finance solvers for reducing balance loan calculations
8.03 Changing the terms of a loan
8.04 Annuities
8.05 Perpetuities
8.06 Compound interest investments with regular deposits
Chapter summary
Chapter review

REDUCING BALANCE LOANS (COMPOUND INTEREST LOANS WITH PERIODIC REPAYMENTS)

- use a recurrence relation to model a reducing balance loan and investigate (numerically or graphically) the effect of the interest rate and repayment amount on the time taken to repay the loan (ACMGM097)
- with the aid of a financial calculator or computer-based financial software, solve problems involving reducing balance loans; for example, determining the monthly repayments required to pay off a housing loan. (ACMGM098)

ANNUITIES AND PERPETUITIES (COMPOUND INTEREST INVESTMENTS WITH PERIODIC PAYMENTS MADE FROM THE INVESTMENT)

- use a recurrence relation to model an annuity, and investigate (numerically or graphically) the effect of the amount invested, the interest rate, and the payment amount on the duration of the annuity (ACMGM099)
- with the aid of a financial calculator or computer-based financial software, solve problems involving annuities (including perpetuities as a special case); for example, determining the amount to be invested in an annuity to provide a regular monthly income of a certain amount. (ACMGM100) AC

8.01 REDUCING BALANCE LOANS

A **reducing balance loan** is one where interest is calculated on the amount still owing after each **repayment** is made. Over the life of the loan, the amount still owing reduces with each repayment and thus the interest charges also reduce.

Example 1

Carlos borrows $40 000, on 15 March 2014, to landscape his garden. The bank lends him this amount with interest calculated monthly at a rate of 11.2% per annum. To repay this loan, Carlos is required to make repayments of $873.69 each month for five years.

a Determine the amount still owing on 15 July 2014.

b How much interest is charged during this period?

Solution

a Each month, interest is charged at a rate of $\frac{11.2}{12}\%$ and then the monthly repayment of $873.69 is deducted from the balance of the loan.

Date	Amount owing
15 March 2014	$40 000
15 April 2014	$\left(1+\frac{11.2}{1200}\right)\times 40\,000 - 873.69 = \$39\,499.64$
15 May 2014	$\left(1+\frac{11.2}{1200}\right)\times 39\,499.64 - 873.69 = \$38\,994.62$
15 June 2014	$\left(1+\frac{11.2}{1200}\right)\times 38\,994.62 - 873.69 = \$38\,484.88$
15 July 2014	$\left(1+\frac{11.2}{1200}\right)\times 38\,484.88 - 873.69 = \$37970.38$

The amount owing on 15 July 2014 is $37 970.38.

9780170250283

b The interest paid is the difference between the total of the repayments made and the reduction in the amount owing on the loan.

Total of repayments = 4 × \$873.69
= \$3494.76
Reduction in amount owing = \$40 000 – \$37 970.38
= \$2029.62
Interest = \$3494.76 – \$2029.62
= \$1465.14

In Example **1**, it can be seen that the balance of the loan in any month is calculated using the balance for the previous month. The recurrence relation or difference equation

$$A_n = \left(1 + \frac{11.2}{1200}\right)A_{n-1} - 873.69$$

where $A_0 = 40\,000$, can be used to find A_n, the amount still owing after the nth repayment.

IMPORTANT

In general, the amount still owing on a reducing balance loan, A_n, after the nth repayment is given by the recurrence relation

$$A_n = \left(1 + \frac{r}{100}\right)A_{n-1} - \text{monthly repayment}$$

where A_0 = amount borrowed and r is the **interest rate** per repayment period.

Example 2

Erica borrowed \$25 000 to cover her expenses whilst studying at University. She makes a monthly repayment on the loan of \$500 and interest is calculated at a rate of 9.5% per annum adjusted monthly.

a Write a recurrence relation that gives the amount still owing on the loan after n months.

b How much does Erica owe after two years?

c If Erica could have obtained a loan at 8.5% instead, how much would she have left to pay after 2 years?

Solution

a Define the variable in the recurrence relation.

Let A_n be the amount still owing after the n months.

The recurrence relation takes the form $A_n = \left(1 + \frac{r}{100}\right)A_{n-1} - \text{monthly repayment}$ where r is the interest rate per month. The interest rate per annum is 9.5% and the monthly repayment is \$500.

$$A_n = \left(1 + \frac{9.5}{12 \times 100}\right)A_{n-1} - 500$$
$$= \left(1 + \frac{9.5}{1200}\right)A_{n-1} - 500$$

State the initial loan balance to fully define the recurrence relation.

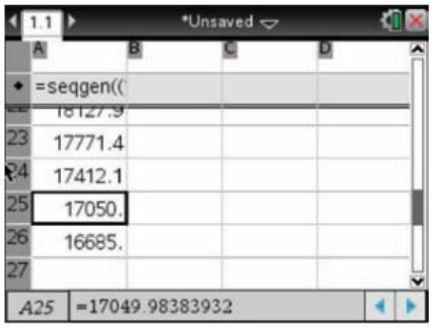
$$A_n = \left(1 + \frac{9.5}{1200}\right)A_{n-1} - 500, \text{ where } A_0 = 25\,000$$



b The amount left owing after 2 years or 24 months is A_{24}.

TI-Nspire CAS

$n = 1$ corresponds to A_0 and therefore $n = 25$ will correspond to A_{24}.

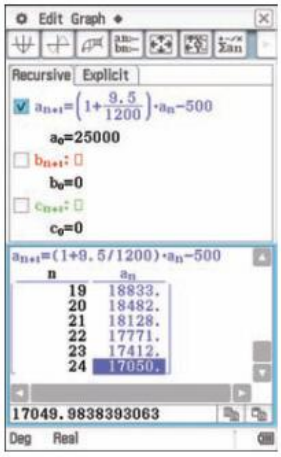


ClassPad

Tap Main $\sqrt{\alpha}$ and select the **Sequence** application. Tap the **Recursive** tab. Select the $\frac{n+1}{a0}$ template from the second drop-down menu on the top toolbar and enter the recurrence relation. Ensure that the box on the left of the expression is checked.

Tap and in the pop-up box set **Start** to 0 and End to 24.

Calculate the values of the sequence by tapping . A_{24} will then correspond to $n = 24$.



The value of A_{24}, correct to two decimal places, is the amount still owing after 2 years.

Erica still owes $17 049.98 after two years.

c Write the recurrence relation.

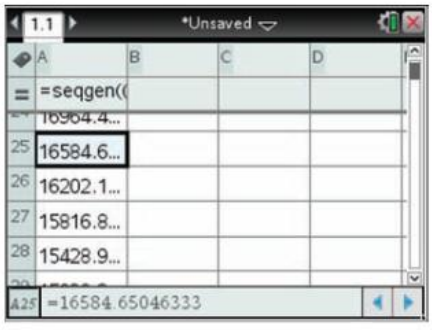
$$A_n = \left(1 + \frac{8.5}{12 \times 100}\right) A_{n-1} - 500$$

Simplify.

$$= \left(1 + \frac{8.5}{1200}\right) A_{n-1} - 500, \text{ where } A_0 = 25\,000$$

TI-Nspire CAS

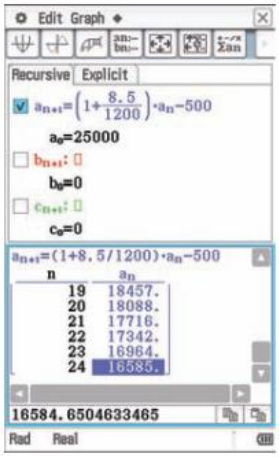
$n = 1$ corresponds to A_0 and therefore $n = 25$ will correspond to A_{24}.



9780170250283

ClassPad

$n = 0$ corresponds to A_0 and therefore $n = 24$ will correspond to A_{24}.



The value of A_{24}, correct to two decimal places, is the amount still owing after 2 years.

Erica would owe $16 584.65 after 2 years, so she would owe $465.33 less.

You can continue to use a recurrence relation to find out how long it takes to pay off a loan, and find the effects of changing interest rates or payments on the period of the loan. In Example 2 part c, the spreadsheet shows the following further down.

61	961.296…		
62	468.105…		
63	−28.578…		
64	−528.781…		

This means that the payment at the end of 62 months would be $28.58 too much, so the last payment would be $471.42 at the end of the 62nd month. This means that it would take 5 years and two months to pay off the loan.



INVESTIGATION Payments and interest rates

Consider a loan of $10 000 with an interest rate of 7%.

- Write a recurrence relation for payments of $50 a month
- What happens to the loan with this payment?
- What happens to the loan with a payment of $60 a month?
- What happens with a payment of $70 a month?
- How long does the loan take to pay off with payments of $60 a month?
- How long does it take with payments of $70 a month?
- How long does it take with payments of $100 a month?
- What conclusions can you make about interest rates and payments on reducing balance loans?

WS

Recurrence relations and compound interest loans

EXERCISE 8.01 Reducing balance loans

Concepts and techniques

1 Example 1 Gaetano borrows $70 000 for a new swimming pool on 1 June 2014. The terms of the loan are that interest is charged at 8% per annum compounded monthly and Gaetano makes regular monthly repayments of $849.30.
   a Determine the amount still owing on 1 October 2014.
   b How much interest is charged during this period?

2 Anh borrows $60 000 from a bank at 7.65% per annum adjusted monthly. She is required to make a regular monthly repayment of $720. How much does she still owe on the loan after six repayments?

3 Example 2 Paul borrowed $45 000 for a holiday. He makes a monthly repayment on the loan of $850 and interest is calculated at a rate of 11% per annum adjusted monthly.
   a Write a recurrence relation that gives the amount still owing on the loan after n months.
   b How much does Paul still owe after four years?

4 For a loan of $50 000, interest is calculated at a rate of 12% p.a. on the reducing monthly balance. The loan is repaid with regular monthly payments of $1120.
   a Find the amount still owing on the loan after two years.
   b How much interest is charged during this period?

9780170250283

5 Marina borrows $200 000 at a rate of 6% per annum calculated monthly on the reducing balance. Each month, Marina repays $1700.

If A_n is the amount owing after the nth repayment, a recurrence relation that describes the way in which the amount owing on this loan changes is

A $A_n = 0.06A_{n-1} - 1700$, where $A_0 = 200\,000$
B $A_n = 0.06A_{n-1} - 200\,000$, where $A_0 = 1700$
C $A_n = 0.005A_{n-1} - 1700$, where $A_0 = 200\,000$
D $A_n = 1.005A_{n-1} - 1700$, where $A_0 = 200\,000$
E $A_n = 1.005A_{n-1} - 200\,000$, where $A_0 = 1700$

Reasoning and communication

6 Hannah borrows $50 000 from the bank as a bridging loan to use as the deposit for a new house. Interest is charged on this loan at a rate of 11.5% per annum, calculated monthly.

a Determine the amount owing after 6 months if Hannah makes monthly repayments of
i $750 ii $900 iii $1000
b For each scenario in **a**, determine the amount of interest charged by the bank over the six-month period.
c How does increasing the size of the repayment affect the amount of interest paid?
d i Calculate the amount owing after six months if Hannah makes monthly repayments of $450.
ii Interpret your answer to **d i**.

7 Adeline wishes to take out a business loan for $40 000. She has two options available to her.

Option 1: Interest charged at 10% per annum adjusted monthly and an additional management fee of $50 per month.
Option 2: Interest charged at 10.5% per annum adjusted monthly with no additional fees.

Adeline plans to make regular monthly repayments of $800.

a Determine the amount owed on each loan after one year.
b Which loan should Adeline choose?

8 Madeleine borrows $10 000 for one year. Interest is charged at 10% per annum adjusted monthly. The debt is to be repaid in monthly instalments of $879.16.

a Plot a graph of the amount owing in the nth month for $n = 0$ to $n = 12$.
b Describe how the rate at which the loan is paid off changes over time.

9 Dominique borrowed $22 000 at 7.5% reducing balance interest, and has to make payments of $180 a month.

a Write a recurrence relation for the balance of the loan.
b Find the time taken to pay off the loan and the amount of the last payment.

10 Carly borrowed $16 500 at 8% reducing balance interest, and must choose to make payments of at least $135 a month.

a Write a recurrence relation for the balance of the loan with payments of $150 a month.
b Find the time taken to pay off the loan and the last repayment at $150 a month.
c What is the total amount paid with payments of $150 a month?
d Write a recurrence relation for the balance of the loan with payments of $155 a month.
e Find the time taken to pay off the loan and the last repayment at $155 a month.
f What is the total amount paid with payments of $155 a month?



8.02 USING FINANCE SOLVERS FOR REDUCING BALANCE LOAN CALCULATIONS

The **outstanding balance** on a reducing balance loan, the value of an annuity or the value of an investment with regular **deposits** can all be found using the annuities formula

$$A = PR^n - \frac{Q\left(R^n - 1\right)}{R - 1} \text{ where } R = 1 + \frac{r}{100}$$

A is the amount owing or invested after n repayments

P is the amount borrowed or invested

r is the interest rate per **compounding period**

Q is the value of each payment

n is the number of payments

For calculations requiring the use of this formula, the 'Finance Solver' on the TI-Nspire CAS or 'Financial Application' on the ClassPad can be used.

Annuities and investments with regular deposits are covered later in this chapter.

IMPORTANT

For a calculation relating to a reducing balance loan, the fields for the Financial Solver or Financial Application (Compound interest) are as follows.

N is the number of time periods.

I% is the interest rate as a percentage per annum.

PV is the **present value** (for a loan, this is the amount borrowed and is positive since this amount is being given by the bank).

Pmt or **PMT** is the value of the regular payments being made (for a loan, this is negative because this is being paid to the bank).

FV is the **future value** (this will be 0 if the loan is to be fully repaid, otherwise it is negative, as it is money that needs to be paid to the bank).

PpY or **P/Y** is the number of payments per year.

CpY or **C/Y** is the number of times in a year interest is compounded.

PpY or P/Y and CpY or C/Y usually take the same value for reducing balance loan calculations.

9780170250283

Example 3

Massimo borrows $140 000 from a building society at 7.5% p.a. compound interest, adjusted monthly. Find the

a monthly repayment if the loan is to be repaid after 15 years
b total amount repaid on the loan
c amount of interest paid on the loan.

Solution

a Use a finance solver to find the monthly repayment.
N is the number of months in 15 years.
I% is the annual interest rate.
PV is positive since the bank has given Massimo the $140 000.
Pmt or PMT is what needs to be found, so leave this field blank.
FV is 0 since the loan will be fully repaid in 15 years.
PpY or P/Y and CpY or C/Y are both 12 since repayments are made monthly and interest compounds monthly.

N = 180
I% = 7.5
PV = 140000
Pmt or PMT =
FV = 0
PpY or P/Y = 12
CpY or C/Y = 12

TI-Nspire CAS

Move the cursor into the Pmt field and press enter.

ClassPad

Tap PMT.
Note that the result is negative since this is money being paid to the bank.

Pmt or PMT = –1297.817304

Write the answer correct to the nearest cent.

Massimo's monthly repayment is $1297.82

b Massimo makes 180 monthly payments of $1297.82.

Amount repaid = 180 × 1297.82
= $233 608

c The amount of interest paid is the difference between the total amount Massimo repaid and the amount he borrowed.

Interest = Amount repaid – amount borrowed
= 233 608 – 140 000
= $93 608



Example 4

Lily borrows $30 000 at a rate of 8.5% per annum, adjusted monthly. She makes repayments of $650 per month.

a Find how much is still owing on this loan after two years.

b How many repayments are required in order for her to fully repay the loan?

Solution

a Use a finance solver to find the amount still owing after 2 years.
N is the number of months in 2 years.
I% is the annual interest rate.
PV is positive, since the bank has given Lily the $30 000.
Pmt or PMT is –650, since this amount is being given to the bank each month.
FV is 0 since the loan will be fully repaid.
PpY or P/Y and CpY or C/Y are both 12 since repayments are made monthly and interest compounds monthly.

N = 24
I% = 8.5
PV = 30000
Pmt or PMT = -650
FV =
PpY or P/Y = 12
CpY or C/Y = 12

TI-Nspire CAS

Move the cursor into the FV field and press enter.

ClassPad

Tap FV.
Note that the result is negative, since this is money to be paid back to the bank.

FV = –18 598.558 69

Write the answer correct to the nearest cent.

After two years, Lily still owes $18 598.56 on this loan.

b Use a finance solver to find the number of payments required to fully repay the loan.
N is what needs to be found so leave this blank.
I% is the annual interest rate.
PV is positive, since the bank has given Lily the $30 000.
Pmt or PMT is –650, since this amount is being given to the bank each month.
FV is what needs to be found, so leave this field blank.
PpY or P/Y and CpY or C/Y are both 12, since repayments are made monthly and interest compounds monthly.

N =
I% = 8.5
PV = 30000
Pmt or PMT = -650
FV = 0
PpY or P/Y = 12
CpY or C/Y = 12

TI-Nspire CAS

Move the cursor into the N field and press enter.

ClassPad

Tap N.

N = 56.088866

Round up to the nearest whole number, since she needs to make more than 56 repayments, the last repayment will be less than $650.

It takes 57 payments in order to fully repay the loan.

9780170250283

EXERCISE 8.02 Using finance solvers for reducing balance loan calculations

Using finance solvers for compound interest loans

Concepts and techniques

1 Find the repayment required to fully repay each of the following loans.
- a $5000 borrowed at 10% per annum for 6 years with monthly repayments and interest compounded monthly.
- b $75 000 borrowed at 9.25% per annum for 10 years with fortnightly repayments and interest compounded fortnightly.
- c $300 000 borrowed at 7.5% per annum for 30 years with monthly repayments and interest compounded monthly.
- d $35 000 borrowed at 12% per annum for 10 years with quarterly repayments and interest compounded quarterly.

2 Example 3 Tayla borrows $250 000 from a bank at 8.25% p.a. compound interest, adjusted monthly. Find the
- a monthly repayment if the loan is to be repaid after 25 years
- b total amount repaid on the loan
- c amount of interest paid on the loan.

3 Find the number of repayments required to pay off each of the following loans.
- a $56 000 borrowed at 9.5% p.a. with monthly repayments of $835 and interest compounded monthly.
- b $28 000 borrowed at 12.5% p.a. with quarterly repayments of $2250 and interest compounded quarterly.
- c $260 000 borrowed at 8.25% p.a. with monthly repayments of $2050 and interest compounded monthly.
- d $45 000 borrowed at 10.25% p.a. with fortnightly repayments of $1031 and interest compounded fortnightly.

4 Example 4 Michael borrows $50 000 at a rate of 8.0% per annum, adjusted monthly. He makes repayments of $620 per month.
- a Find how much is still owing on this loan after three years.
- b How many repayments are required in order for him to fully repay the loan?

5 Deana borrowed $10 000 to pay for a holiday. She agreed to repay the money at $2000 per quarter. Interest was charged at 10.5% p.a. compounded quarterly.
- a How much was still owing on this loan after one year?
- b How long did it take Deana to fully repay the loan?

6 Jonah borrows $75 000 at 8.5% p.a. compound interest, adjusted monthly.
- a Determine the monthly repayment if the loan is to be fully repaid in 15 years.
- b If Jonah decides to make repayments of $1200 per month, how long, in years and months, does it take to fully repay the loan.



7 Remi borrows $15 000 at 7.5% p.a. interest, compounding monthly. The loan is to be fully repaid over five years in equal monthly repayments.
   a Determine the monthly repayment for this loan, correct to the nearest cent.
   b How much interest is paid on this loan over five years?
   c After ten equal payments have been made, how much does Remi still owe on this loan?

8 Marco borrows $60 000 from the bank to renovate his café. Interest is charged at a rate of 8.6% p.a. on the reducing monthly balance. The bank requires him to reduce the balance of the loan to $20 000 in five years. Determine the monthly repayment he will need to make.

9 Lenny borrowed $26 000 for a holiday. The loan will be fully repaid in 5 years. Interest is charged at a rate of 10.3% per annum, calculated monthly on the reducing balance. How much has Lenny paid off the principal immediately following his 12th repayment?

10 Rosa borrowed $385 000 to buy an apartment. The interest rate is 7.5% per annum, calculated on the reducing balance over the life of the loan. She pays equal monthly instalments and will fully repay the loan in 20 years. Find the total amount of interest that she will pay on the loan.

11 Poh borrowed $25 000 and will fully repay the loan in five years. Interest is charged at a rate of 8.9% per annum, calculated monthly on the reducing balance. The amount that Poh has paid off the principal immediately following her eighth repayment is closest to
   A $1413 **B** $2379 **C** $2729 **D** $3081 **E** $4142

Reasoning and communication

12 An oven is advertised for sale at Crazy Carl's for $4500. The sale terms are 10% deposit and the balance repaid in monthly repayments over two years with interest charged at 9.5% per annum, calculated monthly on the reducing balance.
   a What is the monthly repayment for this loan?
   b What is the total cost of the oven from Crazy Carl's if it is purchased on these terms?
   The same oven is also advertised at Mad Marika's for $4800. The sale terms are no deposit and interest free payments over two years.
   c Which is a better option, purchasing the oven from Crazy Carl's or Mad Marika's? Explain.

13 Luisa wants to borrow $25 000 for five years. Interest is calculated at a rate of 8.5% per annum, calculated quarterly. She can afford to pay $1700 per quarter. Will this amount enable her to fully repay the loan in five years? Explain.

14 Giancarlo takes out a loan for $120 000. Interest is charged at a rate of 9.5% p.a. with interest adjusted monthly. Find
   i the monthly repayment, and
   ii the interest charged over the life of the loan if the loan is fully repaid in each of
      a 10 years b 15 years c 20 years d 25 years.
   e What happens to the repayment value as the time taken to fully repay the loan increases?
   f What happens to the total amount of interest paid as the time taken to fully repay the loan increases?

INVESTIGATION Comparing flat rate and reducing balance loans

Flat-rate loans charge interest at a fixed rate on the amount borrowed and do not take into account any reduction in the amount owing due to repayments.

a i Determine the total amount to be repaid on a flat-rate loan of $20 000 at 11% per annum simple interest for 5 years.

ii Determine the total amount to be repaid on a reducing balance loan of $20 000 at 11% per annum compounded monthly with monthly repayments for 5 years.

iii How much is saved in interest by taking out a reducing balance loan rather than a flat-rate loan?

b An amount of $25 000 is borrowed at a rate of 9% per annum to be repaid over 10 years.

i What is the monthly repayment if interest is calculated on the amount borrowed?

ii What is the monthly repayment if interest is calculated monthly on the reducing balance?

iii Compare the total amounts repaid with each type of loan.

iv What would the flat rate of interest need to be in order for the total amount repaid with the flat-rate loan to be the same as that for the reducing balance loan? Answer correct to 1 decimal place.

c Two different lenders offer the following terms for a loan of $30 000.

Lender 1: A reducing balance loan with monthly repayments of 622.75 for 5 years.

Lender 2: A flat-rate loan with interest charged at 9% p.a. over 4 years.

i Which loan has the best terms?

ii How much interest is saved with this loan?

8.03 CHANGING THE TERMS OF A LOAN

It is common for loans to have flexible terms. Interest rates are usually variable rather than fixed, meaning that they change over the life of the loan. A borrower may choose to change the frequency of their repayments or the size of the repayment. Therefore, the life of a loan can be extended or reduced as required.

Example 5

Christian borrows $100 000 at 8% per annum compounding monthly, to be repaid over 10 years with repayments of $1213.28 each month. After 5 years, the interest rate is reduced to 7.5% per annum.

a Find the new repayment required in order for Christian to pay off the loan in the same amount of time.

b How much does Christian save due to the interest rate cut?



Solution

a Firstly find the amount still owing after 5 years.
N is the number of months in 5 years.
I% is the annual interest rate.
PV is positive, since the bank has given Christian $100 000.
Pmt or PMT is negative, since this is money given to the bank.
PpY or P/Y and CpY or C/Y are both 12 since repayments are made monthly and interest compounds monthly.

N = 5 × 12 = 60
I% = 8
PV = 100000
Pmt or PMT = -1213.28
FV =
PpY or P/Y = 12
CpY or C/Y = 12

TI-Nspire CAS

Move the cursor into the FV field and press enter.

ClassPad

Tap FV.

FV = –59 836.57
After five years, the amount still owing is $59 836.57.

Now find the new repayment.
There are 5 years remaining to repay the loan so N is the number of months in 5 years.
I% is the new annual interest rate.
The amount still owing after the first five years is the present value for the second stage of the loan, so PV = 59 836.57.
The loan needs to be fully repaid, so FV = 0.
PpY or P/Y and CpY or C/Y are both 12 since repayments are made monthly and interest compounds monthly.

N = 5 × 12 = 60
I% = 7.5
PV = 59836.57
Pmt or PMT =
FV = 0
PpY or P/Y = 12
CpY or C/Y = 12

Find the monthly repayment.

Pmt or PMT = –1199.00
So Christian's monthly repayment reduces to $1199.00 as a result of the interest rate cut.

b The interest rate cut applies to the last five years of the loan, so it is only necessary to find the difference between the total amounts that would have been repaid at the different rates.

Total repaid with old repayment amount
= 60 × 1213.28
= $72 796.80
Total repaid with new repayment amount
= 60 × 1199.00
= $71 940.00

Interest saved = 72796.80 – 71940.00
= $856.80

9780170250283

Example 6

Andrew took out a reducing balance loan for $60 000 with interest calculated monthly at a rate of 10.25% per annum. He makes monthly repayments of $918.41.

a How long will it take Andrew to repay the loan?
After 3 years, he increases his monthly repayment to $1200.

b By how many months is the length of the loan reduced?

c Find the amount saved in interest by increasing the size of the monthly repayment.

Solution

a Find the number of months needed to fully repay the loan.
I% is the annual interest rate.
PV is positive, since this is money given to Andrew by the bank.
Pmt or PMT is negative, since this is money given to the bank by Andrew.
FV is 0, since the loan is to be fully repaid.
PpY or P/Y and CpY or C/Y are both 12, since repayments are made monthly and interest compounds monthly.

N =
I% = 10.25
PV = 60000
Pmt or PMT = –918.41
FV = 0
PpY or P/Y = 12
CpY or C/Y = 12

Find the number of payments, N.

N = 95.999…

Write the answer.

It takes Andrew 96 months, or 8 years, to repay the loan.

b First find the amount still owing on the loan after 3 years.
N is the number of months in 3 years.
I% is the annual interest rate.
PV is positive, since this is money given to Andrew by the bank.
Pmt or PMT is negative, since this is money given to the bank by Andrew.
PpY or P/Y and CpY or C/Y are both 12, since repayments are made monthly and interest compounds monthly.

N = 3 × 12 = 36
I% = 10.25
PV = 60000
Pmt or PMT = –918.41
FV =
PpY or P/Y = 12
CpY or C/Y = 12

Find the future value, which is the amount still owing after 3 years.

FV = –42 975.73
Andrew still owes $42 975.73 after 3 years.



	Now find the length of time it will take to repay the remainder of the loan with the new repayment amount. I% is the new annual interest rate. The amount still owing after the first three years is the present value for the second stage of the loan, so PV = 42 975.73. Pmt or PMT is –1200, since this is the new repayment amount and it is money paid to the bank. The loan needs to be fully repaid, so FV is 0. PpY or P/Y and CpY or C/Y are both 12, since repayments are made monthly and interest compounds monthly.	N = I% = 10.25 PV = 42975.73 Pmt or PMT = –1200 FV = 0 PpY or P/Y = 12 CpY or C/Y = 12
	Find the number of payments, N.	N = 42.93 It takes Andrew 43 months to pay the remainder of the loan.
	Find the total number of months required to pay the loan with the new repayment after 3 years.	Total time to repay loan = 3 × 12 + 43 = 79 months
	With the original repayment, the loan takes 96 months to repay. Find the difference between the lengths of time taken to pay the loan with the original repayment amounts and the new ones.	Time difference = 96 – 79 = 17 months
	Write the answer.	The length of the loan reduces by 17 months when the repayment amount is increased after 3 years.
c	Find the total amount repaid with the repayments of $918.41 for 96 months.	Total with original repayments = 96 × $918.41 = $88 167.36
	Find the total amount repaid with repayments of $918.41 for 36 months, then $1200 for 43 months.	Total with new repayments = 36 × $918.41 + 43 × $1200 = $84 662.76
	Find the difference in the amounts repaid to determine the amount saved.	Amount saved = $88 167.36 – $84 662.76 = $3504.60
	Write the answer.	By increasing the size of his regular repayment after 3 years, Andrew saves $3504.60 in interest.

Increasing the frequency of repayments reduces the life of the loan but only by a small amount if the overall yearly repayment amount remains the same.

Example 7

a Find the monthly repayment on a loan of $120 000, reducing monthly over 20 years at 8.25% per annum with interest compounded monthly.

b Find the equivalent fortnightly repayment for this loan.

c How much earlier is the loan paid off by making fortnightly repayments, with interest compounded fortnightly, rather than monthly repayments?

Solution

a To find the monthly repayment, list the information at hand.
N is the number of months in 20 years.
I% is the annual percentage interest rate.
PV is positive, since this money is given by the bank.
FV is 0, since the loan is to be fully repaid.
PpY or P/Y and CpY or C/Y are both 12, since repayments are made monthly and interest compounds monthly.

$N = 20 \times 12 = 240$
$I\% = 8.25$
$PV = 120000$
Pmt or PMT =
$FV = 0$
PpY or P/Y = 12
CpY or C/Y = 12

Find the monthly repayment, Pmt or PMT.

Pmt or PMT = –1022.48
The monthly repayment is $1022.48.

b Firstly, convert the monthly repayment amount found in **a** to a yearly repayment amount.

Yearly repayment = $12 \times \$1022.48$
$= \$12\,269.76$

Convert the yearly repayment amount to a fortnightly repayment amount.

Fortnightly repayment = $\$12\,669.76 \div 26$
$= \$471.91$

c Find the length of time it will take to repay the the loan with the fortnightly repayments.
I% is the new annual interest rate.
PV is positive, since this is money given by the bank.
Pmt or PMT is –471.91, since this is the fortnightly repayment amount and it is money paid to the bank.
The loan needs to be fully repaid, so FV is 0.
PpY or P/Y and CpY or C/Y are both 26, since repayments are made monthly and interest compounds monthly.

N =
$I\% = 8.25$
$PV = 120000$
Pmt or PMT = –471.91
$FV = 0$
PpY or P/Y = 26
CpY or C/Y = 26

Find the number of repayments, N.

N = 519
It takes 519 fortnights to repay the loan.

The time saved is the difference in the number of repayments.

The loan with monthly repayments takes $20 \times 26 = 520$ fortnights to repay.
The loan with fortnightly repayments takes 519 months to repay.
The time saved is one fortnight or two weeks.



EXERCISE 8.03 Changing the terms of a loan

Concepts and techniques

Reasoning and communication

1 Example 5 Tom borrows $200 000 at 7.5% per annum compounding monthly, to be repaid over 15 years with repayments of $1854.02 each month. After 5 years, the interest rate is reduced to 7.2% per annum.
   a Find the new repayment required in order for Tom to pay off the loan in the same amount of time.
   b How much does Tom save due to the interest rate cut?

2 Mina takes out a loan of $50 000 for renovations on her flat. The loan was for 10 years at an interest rate of 7.2% p.a., calculated quarterly on the reducing balance.
   a What is Mina's quarterly repayment on this loan?
   b How much will she repay altogether?
   c How much interest will she pay?

   After 4 years, the interest rate is changed to 6.8% p.a., calculated quarterly on the reducing balance.
   d What will Mina's new quarterly repayment be after this interest rate cut if she still intends to repay the loan in 10 years?
   e How much does she save in interest due to the rate cut?

3 Tarnya borrows $80 000 for 20 years at an interest rate of 9% p.a. calculated monthly on the reducing balance. She makes monthly repayments of $719.78. After 8 years, the interest rate increases to 10.5% p.a. and remains at this rate for the rest of the loan.
   a If Tarnya continues to make repayments of $719.78 per month, find the new life of the loan. Answer correct to the nearest month.
   b How much does she repay in total? Answer correct to the nearest hundred dollars.
   c How much extra interest does she pay as a result of the interest rate increase? Answer correct to the nearest hundred dollars.

4 Example 6 Molly took out a reducing balance loan for $75 000 with interest calculated monthly at a rate of 9.5% per annum. She makes monthly repayments of $874.78.
   a How long will it take for Molly to repay the loan?

   After 5 years, she increases her monthly repayment to $1000.
   b By how many months is the length of the loan reduced?
   c Find the amount saved in interest by increasing the size of the monthly repayment.

5 Jinghan borrows $160 000 from a bank at a rate of 9.8% p.a. adjusted monthly.
   She can afford to make regular monthly repayments of $1440.
   a Find the amount still owing after 10 years.
   b How long will it take her to fully repay the loan with repayments of $1440 per month? Answer in years and months correct to the nearest month.
   c What monthly repayment is necessary for her to fully repay the loan in 10 years?
   d How much would she save by repaying the loan in 10 years? Answer correct to the nearest hundred dollars.

9780170250283

6 Marti takes out a business loan for $130 000 to be fully repaid in 20 years with interest charged at a rate of 10.25% p.a. compounded monthly.
   a What is the monthly payment for this loan?
   b How much is repaid in total over the 20 year period?
   c How much interest is paid over the 20 year period?

   After 10 years, Marti increases the monthly repayment by $300.
   d What is the new monthly payment for this loan?
   e How much does he save in interest by increasing the repayment amount? Answer correct to the nearest ten dollars.

7 For a reducing balance loan of $250 000 at 8.2% p.a. adjusted monthly,
   a determine the monthly repayment if the loan is to be fully repaid over
      i 15 years ii 20 years iii 25 years
   b determine the total amount paid if the loan is to be fully repaid over
      i 15 years ii 20 years iii 25 years
   c find the total amount of interest charged for each loan
   d how does extending the life of the loan affect
      i the monthly repayment? ii the total amount of interest charged?

8 An amount of $300 000 is borrowed at an interest rate of 6.9% p.a. compounding monthly. This loan is fully repaid over 20 years with equal monthly repayments. Which one of the following statements is **incorrect**?
   A The total number of repayments is 240.
   B A monthly repayment of $2000 will reduce the life of the loan.
   C After 10 years, there is more than $150 000 still owing on this loan.
   D Extending the life of the loan to 25 years will increase the total amount of interest paid.
   E A reduction in the interest rate after 10 years will reduce the life of the loan if the same monthly repayments are maintained.

9 Example 7 a Find the monthly repayment on a loan of $480 000 reducing monthly over 20 years at 7.8% per annum with interest compounded monthly.
   b Find the equivalent fortnightly repayment for this loan.
   c How much earlier is the loan paid off by making fortnightly repayments, with interest compounded fortnightly, rather than monthly repayments?

10 Larni wishes to borrow $80 000 over 15 years. She has two loan options available.

   Option A: Interest charged at 7.8% p.a. compounded monthly with repayments of $755.31 month.

   Option B: Interest charged at 7.8% p.a. compounded quarterly with quarterly repayments.
   a Determine the quarterly repayment for loan Option B.
   b Is the quarterly repayment for loan Option B equivalent to $755.31 per month?
   c Is there a better loan option? Give a reason for your answer.



8.04 ANNUITIES

An **annuity** is a type of investment. A sum is invested and interest is compounded at some fixed rate. **Withdrawals** are made at regular intervals, usually until the value of the investment is $0.

Example 8

Michel receives an inheritance of $150 000. He invests the money at 7.5% per annum. Interest is paid at the end of each year and after the interest is paid, Michel withdraws $12 000. The amount remaining gets invested for another year.

a Find a recurrence relation of the form $A_{n+1} = rA_n + d$, where $A_1 = a$, that gives the value of the investment, A_n, at the start of the nth year.

b What is the value of the investment at the start of the 10th year?

c How long will the annuity last?

d What amount could be withdrawn at the end of each year so that the value of the investment remains at $150 000?

e If Michel was only able to invest the money at 7%, but still withdrew $12 000 a year, how long would it last?

f At the interest rate of 7.5%, how long would it last if Michel withdrew $15 000 a year?

Solution

a Write the general form of the recurrence relation. $A_{n+1} = rA_n + d$, where $A_1 = a$

A_1 is the initial amount invested. $A_1 = 150\,000$

Each year, the investment earns 7.5% interest so 7.5% of its value is added on. $r = 1 + \frac{7.5}{100}$
$= 1.075$

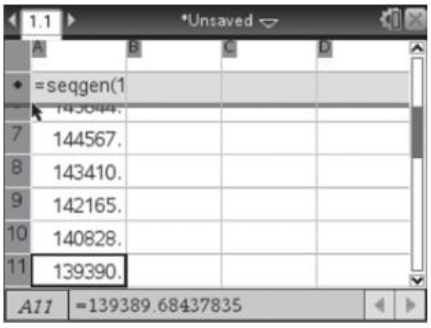
After the interest is added, $12 000 is withdrawn from the investment. $d = -12000$

Write the recurrence relation by substituting in the values of the constants. $A_{n+1} = 1.075A_n - 12\,000$, where $A_1 = 150\,000$

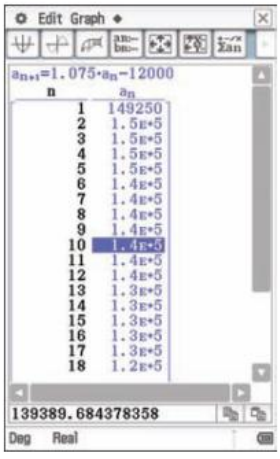
9780170250283

b At the start of the 10th year, $n = 10$. So find A_{10}.

TI-Nspire CAS



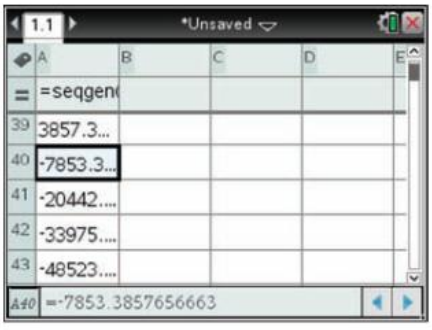
ClassPad



The value of the investment at the start of the 10th year is $139 389.68.

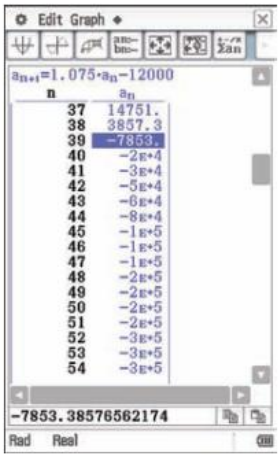
c Find where the balance changes to a negative amount.

TI-Nspire CAS



The balance is negative for $n = 40$ and the last payment will be $12 000 – $7853.39.

ClassPad



The annuity will last 39 years and the last payment will be $4146.61

d If the value of the investment is to remain unchanged, then only the interest can be withdrawn at the end of each year.

$$\text{Interest} = \frac{7.5}{100} \times 150\,000$$
$$= 11\,250$$

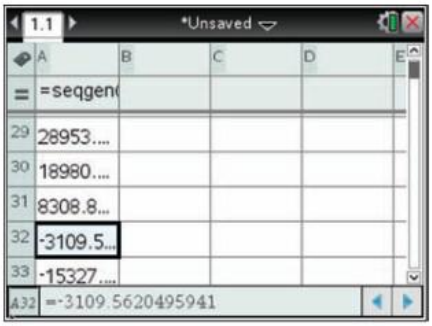
The amount that can be withdrawn each year is $11 250.



e State the new recurrence relation.

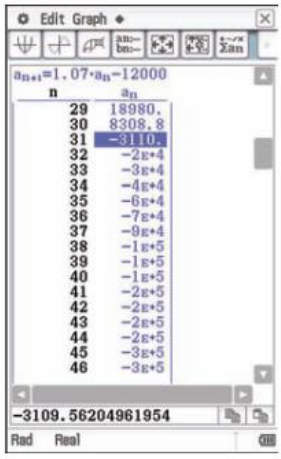
$A_{n+1} = 1.07A_n - 12\,000$

TI-Nspire CAS



The balance is negative for $n = 32$ and the last payment will be \$12000 – \$3109.56.

ClassPad

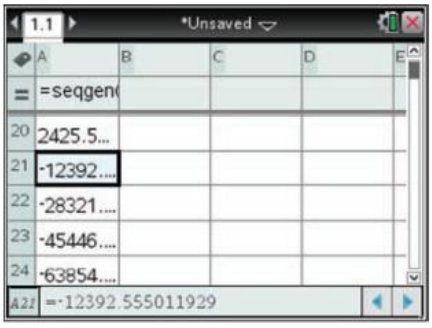


The annuity will last 31 years and the last payment will be \$8890.44

f State the new recurrence relation.

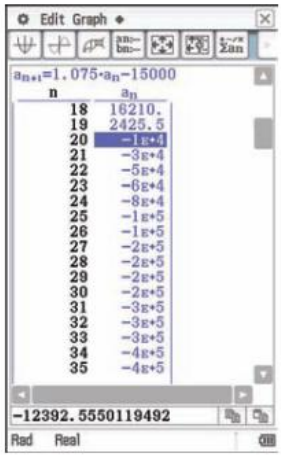
$A_{n+1} = 1.07A_n - 15\,000$

TI-Nspire CAS



The balance is negative for $n = 21$ and the last payment is \$15 000 – \$12 392.56.

ClassPad



The annuity will last 20 years and the last payment will be \$2607.44

The annuities formula that was introduced in Section 8.02 can be used for calculations relating to annuities.

The 'Finance Solver' on the TI-Nspire CAS or 'Financial Application' on the ClassPad can be used.

9780170250283

IMPORTANT

For a calculation relating to an annuity, the fields for the Financial Solver or Financial Application (Compound Interest) are as follows.

N is the number of time periods.

I% is the interest rate as a percentage per annum.

PV is the present value (for an annuity, this is the amount invested and is negative since this amount is being given to the bank).

Pmt or **PMT** is the value of the regular payments being made (for an annuity, this is positive because this is being paid by the bank).

FV is the future value (this is 0 at the end of the annuity. Otherwise, it is positive as it is money that will be given back by the bank at the end of the investment period).

PpY or **P/Y** is the number of payments per year.

CpY or **C/Y** is the number of times in a year interest is compounded.

PpY or P/Y and CpY or C/Y usually take the same value for annuities calculations.

Example 9

Simeon purchases a $400 000 annuity. Interest is paid at 8% per annum compounded monthly. If he receives monthly payments of $3500, how long will the annuity last? Answer correct to the nearest month.

Solution

Use a finance solver to find how long the annuity will last.	N =
I% is the annual interest rate.	I% = 8
PV is negative, since Simeon has effectively given the bank the $400 000.	PV = –400000
Pmt or PMT is positive since this amount is being given by the bank each month.	Pmt or PMT = 3500
FV is 0, since the annuity will last until it no longer has any value.	FV = 0
PpY or P/Y and CpY or C/Y are both 12, since payments are made monthly and interest compounds monthly.	PpY or P/Y = 12 CpY or C/Y = 12

TI-Nspire CAS

Move the cursor into the N field and press enter.

ClassPad

Tap N.

	N = 215.98
Write the answer correct to the nearest month.	Simeon's annuity will last for 216 months.



Example 10

Christina invests $350 000 with interest paid at 7.5% per annum compounded monthly. She receives monthly payments from this investment. What monthly payment will she receive, if she wishes to receive payments for 15 years?

Solution

Use a finance solver to find the monthly payment.	N = 15 × 12 = 180
N is the number of months in 15 years.	I% = 7.5
I% is the annual interest rate.	PV = –350000
PV is negative, since Christina has given the bank the $350 000.	Pmt or PMT =
Pmt or PMT is positive, since this amount is being given by the bank each month.	FV = 0
FV is 0, since the annuity is to last for 15 years.	PpY or P/Y = 12
PpY or P/Y and CpY or C/Y are both 12, since payments are made monthly and interest compounds monthly.	CpY or C/Y = 12

TI-Nspire CAS

Move the cursor into the Pmt field and press enter.

ClassPad

Tap PMT.

Pmt or PMT = 3244.54

Write the answer.

Christina will receive monthly payments of $3244.54 for 15 years.

WS Annuities

EXERCISE 8.04 Annuities

Concepts and techniques

1 Example 8 Annika wins a sum of $80 000. She invests the money at 6.6% per annum. Interest is paid at the end of each year and after the interest is paid, Annika withdraws $6000. The amount remaining gets invested for another year.

- **a** Find a recurrence relation of the form $A_{n+1} = rA_n + d$, where $A_1 = a$, that gives the value of the investment, A_n, at the start of the nth year.
- **b** What is the value of the investment at the start of the 6th year?
- **c** What amount could be withdrawn at the end of each year so that the value of the investment remains at $150 000?
- **d** How long will the investment last for withdrawals of $6000 a year?
- **e** How long will the investment last for withdrawals of $10 000 a year?
- **f** How long would it last if the interest rate was only 6% for withdrawals of $6000?

2 At the start of 2015, Felicity invests $20 000 at 7.5% per annum. At the end of each year, after interest has been paid, she withdraws $1600. No other withdrawals or deposits are made. If A_n represents the value of her investment at the start of year n and $A_1 = \$20\,000$, which one of the following recurrence relation gives the value of her investment at the start of year n?

A $A_{n+1} = 1.075A_n - 1600$
B $A_{n+1} = 0.075A_n - 1600$
C $A_{n+1} = 1.075(A_n - 1600)$
D $A_{n+1} = 0.075(A_n - 1600)$
E $A_{n+1} = 1.075^n A_n - 1600$

3 Ani invests $200 000 at a rate of 9% p.a. compounded monthly. Each month, after interest is paid, she withdraws $2000 and the remaining amount is reinvested. The value of her investment A_n at the start of month n is given by a recurrence relation of the form $A_{n+1} = rA_n + d$, where $A_1 = a$.

a Find the values of r, d and a.
b For how long, correct to the nearest month, will Ani receive a monthly income from this investment?
c How long would it last if she withdrew $3000 a month?
d How long would it last at an interest rate of 8% with withdrawals of $2000?

4 Example 9 Ashton purchases a $500 000 annuity. Interest is paid at 7% per annum compounded monthly. If he receives monthly payments of $3200, for how long will the annuity last? Answer correct to the nearest month.

5 Prasanti invests $250 000 at a rate of 6% p.a. compounding monthly. Each month she withdraws $2000 from her investment after interest is paid and then she reinvests the balance. For how long will her investment last? Answer correct to the nearest month.

Reasoning and communication

6 $1000 is deposited into an account with a fixed interest rate of 7% per annum compounding monthly. Each month, after interest has been paid, a sum of $100 is withdrawn from the account. What is the value of this investment after 6 months?

7 Example 10 Anton invests $450 000 with interest paid at 6.5% per annum compounded monthly. He receives monthly payments from this investment. What monthly payment will Anton receive, if he wishes to receive payments for 20 years?

8 Melanie purchases a $250 000 annuity. Interest is paid at 7.2% per annum compounding twice yearly. If she wishes to receive a six-monthly payment of $30 000, for how long will her annuity last? Answer correct to the nearest half year.

9 Asa invests $160 000 in an annuity at a fixed interest rate of 8.2% p.a. compound interest. Each year, immediately after interest is paid, she receives a payment for $15 000 which is used to provide an Art scholarship for a Year 12 student at the local school.

One year after the money is invested, the first scholarship is awarded to a student.

a After this scholarship is awarded, what sum remains invested?
b How much is left in Asa's account after 8 scholarships have been awarded?
c If the scholarships are to be awarded forever, what is the maximum amount that could be awarded for each scholarship?



10 Enrico invests $300 000 at a fixed rate of 8% p.a. compounding quarterly. Every three months, he receives a regular payment from his investment. What is the maximum payment that he can receive if payments are to continue for an indefinite period?

8.05 PERPETUITIES

A **perpetuity** is a type of annuity where regular payments are paid from the investment that continue forever. Thus, the value of the payment per time period is equal to the interest earned on the investment over that period. The value of the investment remains unchanged over time.

A perpetuity is also referred to as a perpetual annuity.

IMPORTANT

For a perpetuity

$$Q = \frac{Pr}{100}$$

where Q is the value of the regular payment per period

P is the amount invested

r is the percentage interest rate per period

Example 11

Jacob has $500 000 to set up a perpetuity for his granddaughter Sophie. He invests the money in bonds that return 3.5% per annum compounded annually. How much will Sophie receive each year from this investment?

Solution

Write down the formula for a perpetuity investment.

$$Q = \frac{Pr}{100}$$

Write down the known values of any variables.
The amount invested is $500 000.
The interest rate per annum is 3.5%.

$P = 500\,000$
$r = 3.5$

Substitute the values of these variables into the formula and evaluate.

$$Q = \frac{Pr}{100}$$
$$= \frac{500\,000 \times 3.5}{100}$$
$$= 17\,500$$

Write the answer.

Sophie will receive $17 500 annually from this investment.

Scholarships are most commonly set up in perpetuity, which allows the scholarship to exist for all time.

9780170250283

Example 12

Lucy wishes to set up a scholarship fund in her name so that each year an amount of $5000 is awarded to a promising cross-country runner at the school she attended. If interest on the investment averages to 4% per annum simple interest, compounded yearly, how much should she invest?

Solution

Write down the formula for a perpetuity investment.

$$Q = \frac{Pr}{100}$$

Write down the known values of any variables.
The regular payment per year is $5000.
The interest rate per annum is 4%.

$Q = 5000$
$r = 4$

Substitute the values of these variables into the formula.

$$Q = \frac{Pr}{100}$$

Solve for P.

$$5000 = \frac{P \times 4}{100}$$

$$\frac{5000 \times 100}{4} = P$$

$$P = 125\,000$$

Write the answer.

Lucy needs to invest $125 000.

EXERCISE 8.05 Perpetuities

Concepts and techniques

1 Example 11 Mila invests $650 000 in a perpetuity that returns 4.2% per annum compounded annually. How much will Mila receive each year from this investment?

2 Naoto receives a sum of $200 000 and invests the money in a perpetuity that pays him a fixed amount each month. The interest rate is 4% p.a. The amount that he receives each month is

A $153.85 **B** $307.69 **C** $615.40 **D** $666.67 **E** $8000

Use the monthly interest rate when determining the monthly payment from a perpetuity.

3 Shin places his retirement savings of $750 000 in a perpetuity that pays 5.25% per annum compounded monthly. What monthly payment does he receive?

4 Example 12 Emile sets up a scholarship fund in her name so that each year an amount of $3000 is awarded to the top mathematics student in their final year at the school he attended. If interest on the investment averages 4.5% per annum simple interest, compounded yearly, how much should he invest?

5 Mica has $600 000 to invest in a perpetuity with interest compounded monthly. She wants this investment to provide a monthly donation of $2700 to her favourite charity. What annual interest rate is required for this to be achieved?

6 Angela wishes to provide an ongoing payment of $5000 per year to her granddaughter Elina. If the interest on the investment averages 4.8% per annum, how much does she need to invest?



Reasoning and communication

7 Daniel wins $800 000 in a lottery and invests this amount in a perpetuity that pays 5.1% per annum interest. He receives regular monthly payments from his investment. How much does he receive in total over 10 years?

8 Roula invests $140 000 in a perpetuity at 4.5% per annum compounded monthly.
- a What monthly payment is provided by this perpetuity?
- b How much remains invested after 5 years of monthly payments?

8.06 COMPOUND INTEREST INVESTMENTS WITH REGULAR DEPOSITS

An **annuity investment** is an investment that involves making an initial deposit followed by additional regular payments into an account earning a fixed rate of compound interest.

People often use this type of investment to save for retirement, or the deposit for a house, or large purchases such as a car.

Example 13

At the start of each year, Vinh deposits $1200 into an account earning 8% p.a. interest compounded annually. Interest is paid at the end of the year before his next deposit. The table below shows the progress of the account for the first three years.

Year	Annual payment ($)	Balance at start of year ($)	Interest ($)	Balance at end of year ($)
1	1200	1200	96	1296
2	1200	2496	199.68	2695.68
3	1200	3895.68	311.65	4207.33
4	1200			
5	1200			

a Complete the table for years 4 and 5.

b Write a recurrence relation that gives the balance at the end of year n, B_n.

Solution

a The annual payment for year 4 is $1200. The balance at the start of year 4 is the sum of the previous balance and the annual payment, so it is 1200 + 4207.33. The interest is 8% of the balance at the start of the year, so it is 0.08 × 5407.33. The balance at the end of the year is the sum of the balance at the start of the year and the interest, so it is 5407.33 + 432.59.

Year	Annual payment ($)	Balance at start of year ($)	Interest ($)	Balance at end of year ($)
1	1200	1200	96	1296
2	1200	2496	199.68	2695.68
3	1200	3895.68	311.65	4207.33
4	1200	5407.33	432.59	5839.92

9780170250283

The annual payment for year 5 is \$1200. The balance at the start of year 5 is the sum of the previous balance and the annual payment, so it is 1200 + 5839.92. The interest is 8% of the balance at the start of the year, so it is 0.08 × 7039.92. The balance at the end of the year is the sum of the balance at the start of the year and the interest, so it is 7039.92 + 563.19.

Year	Annual payment (\$)	Balance at start of year (\$)	Interest (\$)	Balance at end of year (\$)
1	1200	1200	96	1296
2	1200	2496	199.68	2695.68
3	1200	3895.68	311.65	4207.33
4	1200	5407.33	432.59	5839.92
5	1200	7039.92	563.19	7603.11

b B_n is the balance at the end of the nth year, so $B_1 = 1296$.

$B_1 = 1296$.

Each year, the additional deposit is added onto the balance from the previous year.

Balance at start of year = $B_n + 1200$

Interest is calculated at a rate of 8% on this new amount. So the balance at the end of the year is 1.08 × the balance at the start of the year.

$B_{n+1} = 1.08(B_n + 1200)$

Write the recurrence relation.

The recurrence relation is $B_{n+1} = 1.08\ (B_n + 1200)$, where $B_1 = 1296$.

The annuities formula that was introduced in Section 8.02 can be used for calculations relating to savings accounts with regular fixed contributions.

The 'Finance Solver' on the TI-Nspire CAS or 'Financial Application' on the ClassPad can be used.

IMPORTANT

For a calculation relating to an annuity investment, the fields for the Financial Solver or Financial Application (Compound Interest) are as follows.

N is the number of time periods.

I% is the interest rate as a percentage per annum.

PV is the present value (for an annuity investment, this is the amount invested and it is negative since this amount is being given to the bank).

Pmt or **PMT** is the value of the regular payments being made (for an annuity investment, this is negative because this is being paid to the bank).

FV is the future value (this is positive as it is money that will be given back by the bank at the end of the investment period).

PpY or **P/Y** is the number of payments per year.

CpY or **C/Y** is the number of times in a year interest is compounded.

PpY or P/Y and CpY or C/Y usually take the same value for annuities calculations.



Example 14

Kim opens a savings account with an initial investment of $3000 at 6.5% p.a. compounded monthly. At the end of each month, she makes a deposit of $800. Find

a the amount of her investment after 5 years

b the interest earned.

Solution

a Use a finance solver to find the amount of the investment after 5 years.
N is the number of months in 5 years.
I% is the annual interest rate.
PV is negative, since Kim has given the bank the $3000.
Pmt or PMT is negative, since this amount is being given to the bank each month.
PpY or P/Y and CpY or C/Y are both 12, since payments are made monthly and interest compounds monthly.

$N = 5 \times 12 = 60$
$I\% = 6.5$
$PV = -3000$
$Pmt \text{ or } PMT = -800$
$FV =$
$PpY \text{ or } P/Y = 12$
$CpY \text{ or } C/Y = 12$

TI-Nspire CAS
Move the cursor into the FV field and press enter.

ClassPad
Tap FV.

Note that the result is positive since this is money to be paid to Kim by the bank.

$FV = 60687.63$
The amount of Kim's investment after 5 years is $60 687.63.

b Determine the total amount invested by Kim. She makes an initial deposit of $3000 and then 60 payments of $800.

$\text{Total invested} = \$3000 + 60 \times \$800$
$= \$51\,000$

The interest earned is the difference between the final value of the investment and the total invested by Kim.

$\text{Interest} = \text{Final value of investment} - \text{Total invested}$
$= \$60\,687.63 - \$51\,000$
$= \$9687.63$

9780170250283

Example 15

The owners of a restaurant estimate that $40 000 will be needed in 4 years time for renovations. They open an investment account with an initial deposit of $2000 and interest of 6% p.a. compounded quarterly. How much must be deposited into the account each quarter in order for them to have enough for the renovations in 4 years time?



Shutterstock.com/Pavel L Photo and Video

Solution

Use a finance solver to find the value of the regular deposit.
N is the number of quarters in 4 years.
I% is the annual interest rate.
PV is negative, since the owners have given the bank the $2000.
FV is positive, since the bank will return $40 000 in 4 years time.
PpY or P/Y and CpY or C/Y are both 4, since payments are made quarterly and interest compounds quarterly.

TI-Nspire CAS

Move the cursor into the Pmt field and press enter.

ClassPad

Tap PMT.

N = 4 × 4 = 16
I% = 6
PV = –2000
Pmt or PMT =
FV = 40000
PpY or P/Y = 4
CpY or C/Y = 4

Pmt or PMT = 2809.07
The regular quarterly deposit must be $2089.07.

EXERCISE 8.06 Compound interest investments with regular deposits

Concepts and techniques

1 Example 13 At the start of each year, Megan deposits $500 into an account earning 8% p.a. interest compounded annually. The table below shows the progress of the account for the first three years.

Year	Annual payment ($)	Balance at start of year ($)	Interest ($)	Balance at end of year ($)
1	500	500	40	540
2	500	1040	83.20	1123.20
3	500	1623.20	129.86	1753.06
4				
5				

a Copy and complete the table for years 4 and 5.
b Write a difference equation that gives the balance at the of year n, B_n.

2 At the start of each year, Eric deposits $8000 into an account earning 7.2% p.a. interest compounded annually. Interest is paid at the end of the year before his next deposit. If A_n is the amount in his account at the end of year n, a recurrence relation for determining A_n is
A $A_{n+1} = 1.072A_n + 8000$, where $A_1 = 8000$
B $A_{n+1} = 1.072A_n + 8000$, where $A_1 = 8576$
C $A_{n+1} = 0.072A_n + 8000$, where $A_1 = 8000$
D $A_{n+1} = 1.072(A_n + 8000)$, where $A_1 = 8000$
E $A_{n+1} = 1.072(A_n + 8000)$, where $A_1 = 8576$

3 At the start of 2015, Jason makes a deposit into a savings account earning r% per annum compounded monthly. Interest is paid at the end of the month, and at the start of the next month he makes a deposit equal to his initial deposit. A recurrence relation that gives the amount in his account A_n at the end of the nth month is

$$A_{n+1} = 1.007(A_n + 400), \text{ where } A_1 = 402.80$$

a How much does Jason deposit into the account each month?
b Find the value of r.
c What is the value of Jason's investment at the end of 2015?

4 Example 14 Hannah opens a savings account with an initial investment of $5000 at 6.2% p.a. compounded monthly. At the end of each month, she makes a deposit of $500. Find
a the amount of her investment after 8 years
b the interest earned.

5 Lana invests $1000 in a savings account earning 8% per annum compounded annually. She makes regular deposits of $1000 per year into the account. How much is in the account after 5 years?

6 Each month for 4 years, Tom deposits $165 into an account earning 6.2% p.a. compounded monthly. At the end of the 4 years, he withdraws all of his money to buy a boat for $9300. Is there enough money in the account to pay for the boat?

9780170250283

7 Jacinta deposits $100 each week for 50 weeks into an account earning 8.4% p.a. compounded weekly. The value of her investment at the end of the 50 weeks is

A $5000 **B** $5008 **C** $5203.10 **D** $5403.85 **E** $5420

Reasoning and communication

8 Effie receives an account containing $2000 for her first birthday. The account earns 8.6% p.a. compounded annually. Her aunt and uncle decide to add to this account by contributing $2000 each year on her birthday.

a How much will she have in the account after the last contribution is made on her 12th birthday?

b Calculate the total interest earned.

9 Example 15 John and Nancy start a trust fund for their newborn granddaughter, Emma, to finance her university education. They open a savings account earning 7.1% p.a. compounded annually. What amount should they deposit each year so that she will have $50 000 after 18 years?

10 Madeleine and Daniel are planning a trip in 2 years' time and open an account which they refer to as the 'trip fund'. How much should they deposit into the 'trip fund' each quarter if the account earns 5% p.a. compounded quarterly and they need $8000 for their trip?



Shutterstock.com/Webspark

11 Fadia starts an investment with a fixed interest rate of 9% per annum compounding monthly and an initial deposit of $5000. Every month, she makes an additional deposit of $500.

a What is the value of Fadia's investment after 5 years?

b Fadia needs to save $75 000 as a deposit for a house. How long, correct to the nearest month, will it take her to save this deposit?

12 Massimo deposits $200 each month for 2 years into an account earning 9.5% p.a. compounded monthly. What single sum of money could he invest at the start to achieve the same financial result at the end of the two years?



8 CHAPTER SUMMARY

REDUCING BALANCE LOANS AND ANNUITIES

- A **reducing balance loan** is one where interest is calculated on the amount still owing after each repayment is made.
- In general, the amount still owing on a reducing balance loan, A_n, after the nth repayment is given by the recurrence relation

  $A_n = \left(1 + \frac{r}{100}\right) A_{n-1} -$ monthly repayment

  where A_0 = amount borrowed and r is the interest rate per repayment period.
- For a calculation relating to a reducing balance loan, the fields for the Financial Solver or Financial Application (Compound Interest) are as follows.

  N is the number of time periods.

  I% is the interest rate as a percentage per annum.

  PV is the present value (for a loan, this is the amount borrowed and is positive since this amount is being given by the bank).

  Pmt or **PMT** is the value of the regular payments being made (for a loan, this is negative because this is being paid to the bank).

  FV is the future value (this will be 0 if the loan is to be fully repaid, otherwise it is negative as it is money that needs to be paid to the bank).

  PpY or **P/Y** is the number of payments per year.

  CpY or **C/Y** is the number of times in a year interest is compounded.

  PpY or P/Y and CpY or C/Y usually take the same value for reducing balance loan calculations.
- With a reducing balance loan, when changes occur to the interest rate, repayment or frequency of repayments, then changes also occur in the amount of interest paid and the life of the loan.
- An **annuity** is a type of investment. A sum is invested and interest is compounded at some fixed rate. Withdrawals are made at regular intervals, usually until the value of the investment is $0.
- For a calculation relating to an annuity, the fields for the Financial Solver or Financial Application (Compound Interest) are as follows.

  N is the number of time periods.

  I% is the interest rate as a percentage per annum.

  PV is the present value (for an annuity, this is the amount invested and is negative since this amount is being given to the bank).

  Pmt or **PMT** is the value of the regular payments being made (for an annuity, this is positive because this is being paid by the bank).

  FV is the future value (this is 0 at the end of the annuity. At other times, it is positive as it is money that will be given back by the bank at the end of the investment period).

  PpY or **P/Y** is the number of payments per year.

  PpY or P/Y and CpY or C/Y usually take the same value for annuities calculations.
- A **perpetuity** is a type of annuity where regular payments are paid from the investment. This continues forever.

9780170250283

- For a perpetuity,

$$Q = \frac{Pr}{100}$$

where Q is the value of the regular payment made per period

P is the amount invested

r is the percentage interest rate per period

- An **annuity investment** is an investment that involves making an initial deposit followed by additional regular repayments into an account earning a fixed rate of compound interest.
- For a calculation relating to an annuity investment, the fields for the Financial Solver or Financial Application (Compound Interest) are as follows.

N is the number of time periods.

I% is the interest rate as a percentage per annum.

PV is the present value (for an annuity investment, this is the amount invested and is negative since it is being given to the bank).

Pmt or **PMT** is the value of the regular payments being made (for an investment, this is negative because this is being paid to the bank).

FV is the future value (this is positive as it is money that will be given back by the bank at the end of the investment period).

PpY or **P/Y** is the number of payments per year.

CpY or **C/Y** is the number of times in a year interest is compounded.

PpY or P/Y and CpY or C/Y usually take the same value for annuities calculations.

8 CHAPTER REVIEW

REDUCING BALANCE LOANS AND ANNUITIES

Multiple choice

1 Example 2 Andrew borrowed $30 000 in order to buy a new car. He makes a monthly repayment of $800, and interest is calculated at a rate of 12.5% per annum adjusted monthly. How much does Andrew owe after three years?

A $1200 B $8476.62 C $8838.35
D $9538.99 E $14 765.16

2 Example 4 Sienna borrows $50 000 for an overseas trip. Interest is charged at 9% per annum, adjusted monthly. She makes repayments of $1000 per month. How many repayments are required in order for her to fully repay the loan?

A 6 B 50 C 55 D 62 E 63

3 Example 5 Zoe borrows $45 000 at 7.5% per annum compounding monthly, to be repaid over 5 years with repayments of $901.71 each month. After 2 years, the interest rate is reduced to 6.5% per annum. If Zoe chooses to reduce her monthly repayments and keep the loan term the same, how much will she save due to the interest rate cut?

A $13.26 B $477.36 C $795.60 D $888.45 E $28 988.02

4 Example 9 Isaac purchases a $250 000 annuity. Interest is paid at 5.5% per annum compounded monthly. If he receives monthly payments of $2000, how long will the annuity last for?

A 15.5 months B 125 months C 132 months
D 186 months E 187 months

5 Example 12 Caleb wishes to set up a perpetuity for his daughter so that each year she receives $10 000. If interest on the investment averages to 5% per annum simple interest, compounded yearly, how much should he invest?

A $500 B $200 000 C $2 400 000
D $20 000 000 E $953 380.95

6 Example 15 A couple estimate that they will need about $50 000 to cover a wedding and honeymoon in two years time. They open an investment account with an initial deposit of $5000 and interest of 4.5% compounded quarterly. How much must be deposited into the account each quarter for them to have enough for the wedding in two years time?

A $1587.47 B $1776.65 C $2065.24
D $5350.98 E $6665.09

9780170250283

Short answer

7 Example 1 Ava borrowed $25 000 on 23 April 2012. The bank lends her this amount with interest calculated monthly at a rate of 9.75% per annum. To repay this loan, Ava is required to make repayments of $631.07 each month for four years.
- a Determine the amount still owing on 23 April 2013.
- b How much interest is charged during this period?

8 Example 2 Ryan borrowed $10 000 to cover his expenses whilst moving house. He makes a monthly repayment on the loan of $300 and interest is calculated at a rate of 8.25% per annum adjusted monthly.
- a Write a recurrence relation that gives the amount still owing on the loan after n months.
- b How much does Ryan owe after 3 years?

9 Example 3 Chelsea borrowed $450 000 from a bank at 5.65% per annum compound interest, adjusted monthly. Find the:
- a monthly repayment if the loan is to be repaid after 20 years
- b total amount repaid on the loan
- c amount of interest paid on the loan.

10 Example 6 Dane took out a reducing balance loan for $80 000, with interest calculated monthly at a rate of 4.85% per annum. He makes monthly repayments of $842.67.
- a How long will it take for Dane to repay the loan?

After 4 years, he increases his monthly repayment to $950.
- b By how many months is the length of the loan reduced?
- c Find the amount saved in interest by increasing the size of the monthly repayment.

11 Examples 7, 11 A loan is taken out for $500 000 reducing monthly over 25 years at 6.75% per annum with interest compounded monthly.
- a Find the monthly repayment on this loan.
- b Find the equivalent fortnightly repayment for this loan.
- c How much earlier is the loan paid off by making fortnightly repayments, with interest compounded fortnightly, rather than monthly repayments?

12 Example 8 Lily receives an inheritance of $65 000. She invests the money at 4.2% per annum. Interest is paid at the end of each year and after the interest is paid, Lily withdraws $5000. The amount remaining gets invested for another year.
- a Find a recurrence relation of the form $A_{n+1} = rA_n + d$, where $A_1 = a$, that gives the value of the investment, A_n, at the start of the nth year.
- b What is the value of the investment at the start of the 15th year?
- c What amount could be withdrawn at the end of each year so that the value of the investment remains at $65 000?

13 Example 10 Luisa invests $100 000 with interest paid at 5.25% per annum compounded monthly. She receives monthly payments from this investment. What monthly payment will she receive, if she wishes to receive payments for 20 years?

14 Example 11 Jack has $1 000 000 to set up a perpetuity for his family. He invests the money in bonds that return 4.25% per annum compounded annually. How much will his family receive each year from this investment?

15 Example 13 At the start of each year, Oliver deposits $5000 into an account earning 4.8% p.a. interest compounded annually. Interest is paid at the end of the year before his next deposit. The table below shows the progress of the account for the first three years.

Year	Annual payment ($)	Balance at start of year ($)	Interest ($)	Balance at end of year ($)
1	5000	5000	240	5240
2	5000	10 240	491.52	10 731.52
3	5000	15 731.52	755.11	16 486.63
4	5000			
5	5000			
6	5000			

a Complete the table for years 4, 5 and 6.
b Write a recurrence relation that gives the balance at the end of year n, B_n.

16 Example 14 Lisa opens a savings account with an initial investment of $8000 at 5.9% p.a. compounded monthly. At the end of each month, she makes a deposit of $1000. Find
a the amount of her investment after 4 years.
b the interest earned.

Application

17 Sarah has $5000 in savings that she wants to put into the bank. She aims to deposit at least $200 a month into the account, with the goal of reaching $20 000 so that she can travel overseas. She has the option of putting the money in an account A with 7.0% interest per annum, compounding annually, or another account B with 6.5% interest, compounding monthly.
a Which account will help Sarah reach her goal more quickly? Assume that she would deposit $200 per month into account B or deposit $2400 a year into account A.
b How much time will she save with this option?

Sarah deposits $200 a month into this account for the first 12 months, but then gets a promotion with a good pay rise, and is then able to deposit $350 a month.
c If she continues to deposit at this rate, how long would it take her in total to save $20 000?
d The bank drops its interest rates to 5.75%. How much money would Sarah have to deposit each month now in order to reach her goal of $20 000 in the same time frame that you calculated in part **c**?

To finance her overseas trip, Sarah plans to withdraw $3000 per month.
e For how long could Sarah travel if she did this?

Sarah manages to only withdraw $2000 per month for the first two months.
f If Sarah stuck to this budget, for how much longer could she travel than the timeframe you calculated in part **e**?
g If Sarah wanted to travel for a full year, how much would she be able to withdraw each month after the first two months?
h Assuming that Sarah spent $3000 per month for the time following the first two months, find how much money Sarah would have remaining if she returned home six months after she left.

9780170250283

18 Cameron receives an inheritance of $750 000. He wants to set up a perpetuity for his three children. He invests the money in an account at 4.25% per annum, compounding annually.

a How much would each child get per year?

b Cameron decides that he wants each child to receive $20 000 per year instead. For how long will the annuity now last?

One of Cameron's children, Hunter, decides to invest the money she gets each year. She opens an account that returns 4.5% per annum, compounding annually.

c If she invests all the money she receives, how much would she have after

i five years ii ten years?

Hunter decides to use half the money to supplement her income, and invest the other half.

d How much less money would she have after

i five years ii ten years?

e Assuming that Hunter deposits $10 000 a year and does not touch it, how much money would she have invested when the original annuity runs out?

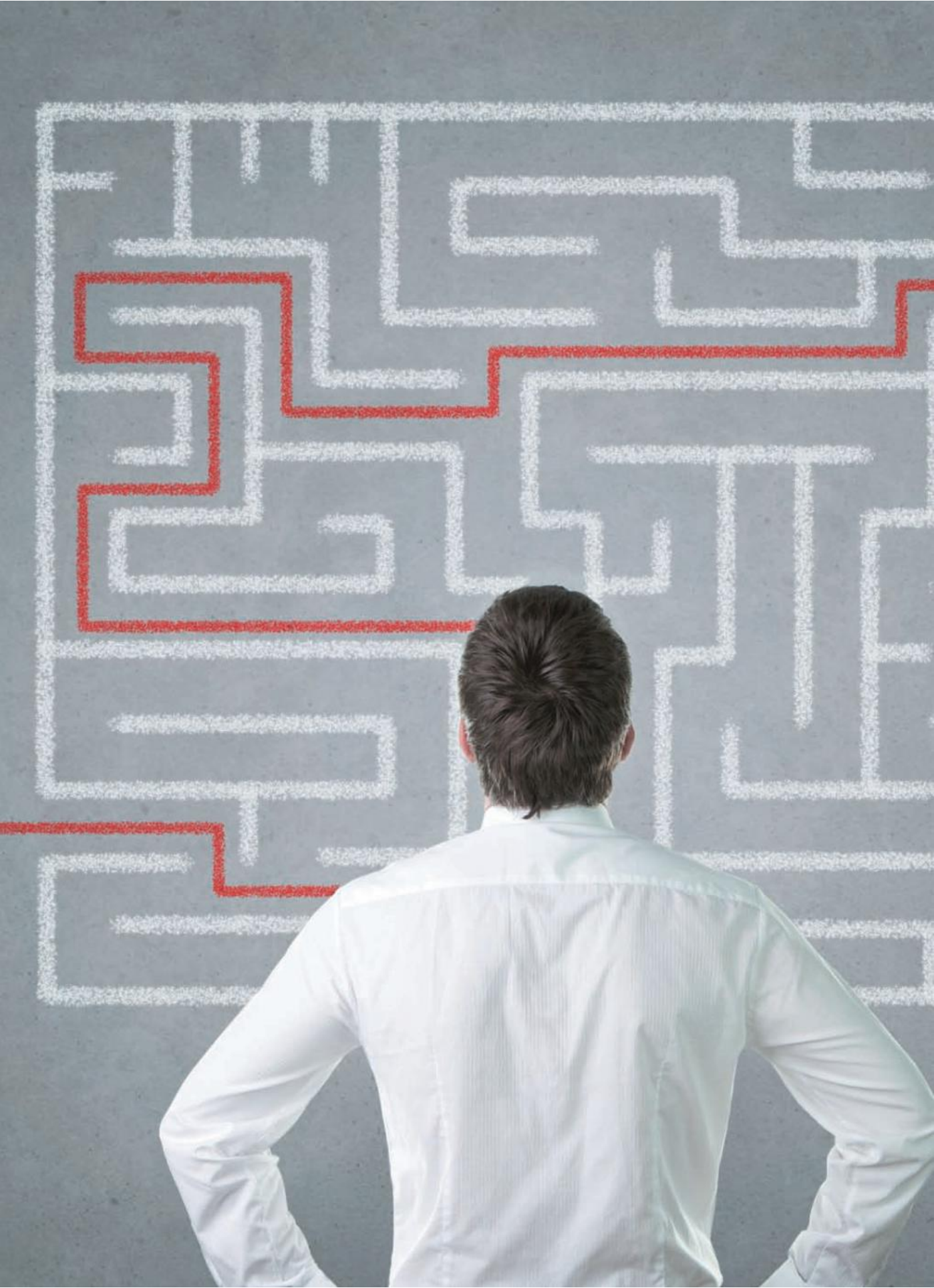
When the annuity runs out, Hunter decides to put her savings towards a new home worth $550 000. She uses the entire savings as a deposit and takes out a loan from the bank for the rest.

f If the loan had an interest rate of 5.2% per annum compounding monthly, how much would Hunter need to pay each month to pay the loan off in

i 10 years? ii 20 years?

g Hunter decides to put $3000 towards the loan each month for the first three years, then $3500 towards the loan thereafter. How long will it be before she completely owns the house?

Practice quiz





9

TERMINOLOGY

activity
adjacency matrix
allocation
assignment
backward scan
bipartite graph
capacity
crashing
critical activity
critical path
cut
digraph
dominance
dummy activity
earliest starting time
edge
event
float
flow
forward scan
Hungarian algorithm
inflow
latest starting time
matrix
maximum
minimum
network
node
non-critical activity
one-stage pathway
optimum
outflow
predecessor
reachability
reduction
resultant matrix
sink
source
two-stage pathway
valid
vertex

Prior learning

NETWORKS AND DECISION MATHEMATICS

DIRECTED GRAPHS AND NETWORKS

9.01 Directed graphs, reachability and dominance
9.02 Drawing network diagrams
9.03 Forward and backward scanning
9.04 Locating the critical path
9.05 Network flow problems
9.06 Bipartite graphs and the assignment problem
Chapter summary
Chapter review

Shutterstock.com / ranjuzzz

PROJECT PLANNING AND SCHEDULING USING CRITICAL PATH ANALYSIS (CPA)

- construct a network to represent the durations and interdependencies of activities that must be completed during the project; for example, preparing a meal (ACMGM104)
- use forward and backward scanning to determine the earliest starting time (EST) and latest starting times (LST) for each activity in the project (ACMGM105)
- use ESTs and LSTs to locate the critical path(s) for the project (ACMGM106)
- use the critical path to determine the minimum time for a project to be completed (ACMGM107)
- calculate float times for non-critical activities. (ACMGM108)

FLOW NETWORKS

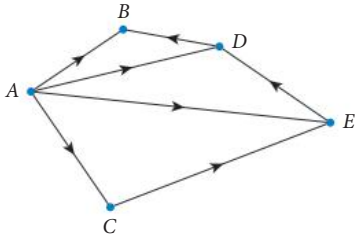
- solve small-scale network flow problems including the use of the 'maximum-flow minimum- cut' theorem; for example, determining the maximum volume of oil that can flow through a network of pipes from an oil storage tank (the source) to a terminal (the sink). (ACMGM109)

ASSIGNMENT PROBLEMS

- use a bipartite graph and/or its tabular or matrix form to represent an assignment/ allocation problem; for example, assigning four swimmers to the four places in a medley relay team to maximise the team's chances of winning (ACMGM110)
- determine the optimum assignment(s), by inspection for small-scale problems, or by use of the Hungarian algorithm for larger problems. (ACMGM111) AC

9.01 DIRECTED GRAPHS, REACHABILITY AND DOMINANCE

A graph where direction is indicated along each edge is called a directed graph or **digraph**.



Reachability

Reachability is the ability to get from one vertex to another vertex in a directed graph.

In the diagram above,

- point B is reachable from point A along one edge on path AB
- point E is reachable from point A along one edge on path AE
- point E is also reachable from point A along two edges on path ACE
- point D is not reachable from point B.

9780170250283

In the directed graph, there are **one-stage pathways** that contain only one edge.

There are **two-stage pathways** that contain two edges.

There are also three-stage pathways that contain three edges.

Matrix representation of a network

In order to fully describe the **network**, we need to show all the paths between all the vertices.

The first matrix records all the one-stage pathways in the network. It is called the **adjacency matrix** and is denoted by R^1.

The rows indicate the vertex we are coming from and the columns indicate the vertex we are going to.

The matrix shows there are one-stage pathways from A to B, A to C, A to D and A to E.

There are no one-stage pathways from B to any other vertex.

There is a one-stage pathway from C to E.

There is a one-stage pathway from D to B.

There is a one-stage pathway from E to D.

$$\begin{array}{cc} & \text{to} \\ & \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \text{from} \begin{array}{c} A \\ B \\ C \\ D \\ E \end{array} & \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

The second matrix records all the two-stage pathways in the network and is denoted by R^2.

The two-stage pathways are shown in the matrix on the right.

This matrix shows that there are two-stage pathways from

A to B along path ADB

A to D along path AED

A to E along path ACE

C to D along path CED and

E to B along path EDB

$$\begin{array}{cc} & \text{to} \\ & \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \text{from} \begin{array}{c} A \\ B \\ C \\ D \\ E \end{array} & \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

The third matrix records all the three-stage pathways in the network and is denoted by R^3.

The three-stage pathways are shown in the matrix on the right.

This matrix shows that there are three-stage pathways from

A to B along path $AEDB$

A to D along path $ACED$

C to B along path $CEDB$

$$\begin{array}{cc} & \text{to} \\ & \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \text{from} \begin{array}{c} A \\ B \\ C \\ D \\ E \end{array} & \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

There is one four-stage pathway from A to B along path $ACEDB$, which is shown by the matrix R^4 on the right.

Finally, we can combine all three matrices to find a single reachability matrix R,

where $R = R^1 + R^2 + R^3 + R^4$.

$$\begin{array}{cc} & \text{to} \\ & \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \text{from} \begin{array}{c} A \\ B \\ C \\ D \\ E \end{array} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

$$R^1 = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad R^2 = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad R^3 = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad R^4 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$R = R^1 + R^2 + R^3 + R^4 = \begin{bmatrix} 0 & 4 & 1 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

This final matrix indicates the reachability of each of the vertices in the matrix. The final reachability of each vertex can be found by adding the numbers in each column.

Vertex	Reachability
A	0
B	7
C	1
D	5
E	3

The table shows the total number of ways of reaching each of the vertices.

In this network, the vertex with the greatest reachability is vertex B. The reachability value of 7 indicates that there are 7 different paths to vertex B.

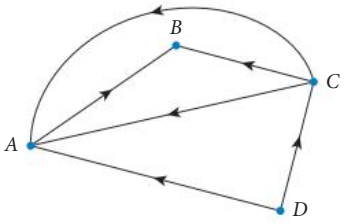
9780170250283

Example 1

The bus routes between landmarks A, B, C and D are shown in the directed graph.

a Which locations are reachable from D?

b List all possible paths from C.



Solution

a Start at D and travel along the edges in the direction indicated on the edge.
There is a one stage path from D to C along DC.
There is a one stage path from D to A along DA.
There are 2 two stage paths from D to B along DAB and DCB.

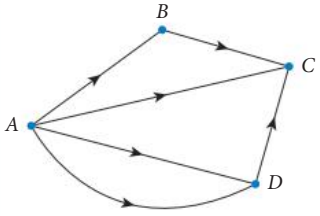
The locations reachable from D are C, A and B.

b There are two one stage paths from C to A along CA.
There is a one stage path from C to B along CB.
There is a two stage path from C to B along CAB.

The paths from C are CA (along two edges), CB and CAB.

Example 2

Represent the one and two-stage pathways in the directed network in matrix form.



Solution

Construct a matrix with four rows and four columns and label the rows and columns A, B, C and D.

$$R^1 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} & & & \\ & & & \\ & & & \\ & & & \end{bmatrix} \end{array}$$

The one-stage pathway matrix is labelled R^1. Start with row A and enter 1 in any column representing a vertex that can be reached from vertex A along one edge and 0 in the columns that cannot be reached from A along one edge.
This shows that there is no path from A to A.
There is a one stage path from A to B and from A to C.
There are two one stage paths from A to D.

$$R^1 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 2 \\ & & & \\ & & & \\ & & & \end{bmatrix} \end{array}$$

Repeat this process for row B, recording all the one stage paths from vertex B and then complete rows C and D.

$$R^1 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

Find matrix R^2, which shows all the two-stage pathways from each vertex in the network.
There are three two-stage pathways all from A to C. These must start at A, end at C and travel along two edges. The paths are ADC (2 paths) and ABC.

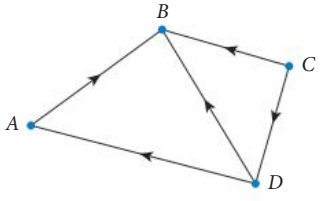
$$R^2 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Dominance

Four basketball players, Abbie, Bess, Cath and Denise, play off against each other in a series of one-on-one basketball games. In these games,

- Abbie beats Bess
- Bess has no wins
- Cath beats Bess and Denise
- Denise beats Abbie and Bess

We can represent this graphically using a **dominance** graph.



9780170250283

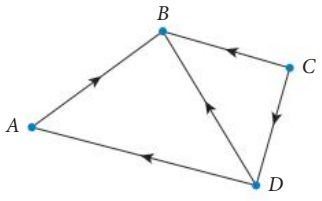
The arrow on the edge points to the individual who is defeated.



In this competition, two players, Cath and Denise, are drawn on two wins each. A dominance matrix will enable us to identify the dominant player.

Example 3

The results of a basketball competition between Abbie, Bess, Cath and Denise is shown in the dominance graph below, where each player is represented by their initial.



a Find the first dominance matrix D^1, showing the one step dominance between the players.

b Find the second dominance matrix D^2, showing the two step dominance between the players.

c Find the dominance matrix by summing D^1 and D^2.

d Find the dominant player by summing the rows of the dominance matrix.

Solution

Construct a matrix with four rows and four columns and label the rows and columns *A*, *B*, *C* and *D*.

$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} & & & \\ & & & \\ & & & \\ & & & \end{bmatrix} \end{array}$$

The one stage dominance matrix is labelled D^1.

Start with row *A* and enter 1 in any column representing a vertex that can be reached from vertex *A* along one edge and 0 in the columns that cannot be reached from *A* along one edge.

Player *A* does not compete against player *A* so a zero is placed in this cell.

Player *A*'s only win is against player *B*, so 1 is placed in row *A* column *B* and 0 is placed in rows *C* and *D*.

$$D^1 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ & & & \\ & & & \\ & & & \end{bmatrix} \end{array}$$

Repeat this process for rows *B*, *C* and *D*.

$$D^1 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

The second dominance matrix D^2 shows vertices that can be reached along two edges. Starting with row A, there are no vertices that can be reached from vertex A in two stages so 0 is placed in each of columns A, B, C and D.

$$D^2 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 \\ & & & \\ & & & \\ & & & \end{bmatrix} \end{array}$$

Repeat this process for rows B, C and D.
Vertex B has no vertices that can be reached in two stages either, so 0 is placed in each column.
Vertex C can reach vertices A and B in two stages, so 1 is recorded in these cells. The "1" in row C column A indicates that player C defeated a player who defeated player A.

$$D^2 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

Add D^1 and D^2 to find the dominance matrix D.

$$D = D^1 + D^2 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & 2 & 0 & 0 \end{bmatrix} \end{array}$$

Sum the rows of the matrix to determine player dominance.

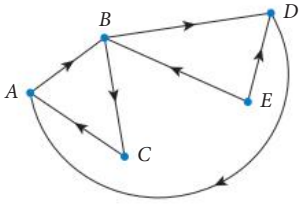
Player	Dominance
A	1
B	0
C	4
D	3

Player C (Cath) is the dominant player.

EXERCISE 9.01 Directed graphs, reachability and dominance

Concepts and techniques

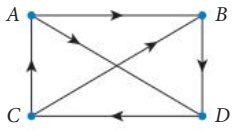
1 Example 1 In the network shown, the vertex that is not reachable from vertex B is



A A **B** C **C** D **D** E
E All vertices are reachable.

9780170250283

2 The network shows the one-stage pathways between vertices A, B, C and D.



A two-stage pathway is possible from vertex A to

A vertex B **B** vertices D and C **C** vertices B and C

D vertices B and D **E** vertices B, C and D

3 The adjacency matrix for five vertices A, B, C, D and E is shown.

A two-stage pathway exists between

A from A to C

B from B to C

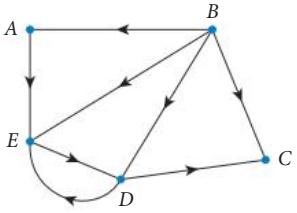
C from C to A

D from D to E

E from E to D

$$R^1 = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

4 The digraph shows a series of rivers that flow between towns A, B, C, D and E.

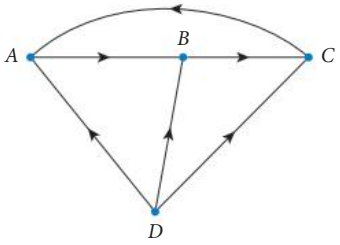


a How many paths exist from B to D?

b Name the paths from B to D.

c Find the one stage adjacency matrix for this network.

5 Example 2 The directed graph shows one way roads between 4 landmarks A, B, C and D.



a Draw the matrix R^1 that shows all the one-stage pathways in the network.

b Draw the matrix R^2 that shows all the two-stage pathways in the network.

6 The adjacency matrix for four vertices A, B, C and D is shown.

$$R^1 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

a Draw the directed graph.

b Find the matrix R^2 for the two-stage pathways between the vertices.

c How many pathways exist from vertex A to vertex B?

7 A matrix showing the combined one and two step dominance between four soccer teams W, X, Y and Z is shown.

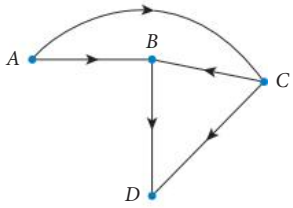
$$D = \begin{array}{c} \\ W \\ X \\ Y \\ Z \end{array} \begin{array}{c} \begin{array}{cccc} W & X & Y & Z \end{array} \\ \begin{bmatrix} 0 & 2 & 1 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 2 & 0 & 2 \\ 1 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

a Find the dominance scores for each team.

b Which was the most dominant team in the soccer competition?

Reasoning and communication

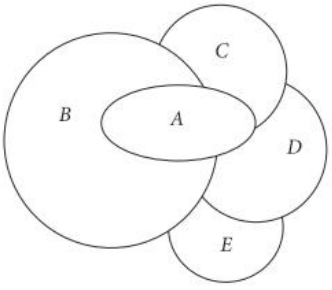
8 Example 3 Four players compete in a round robin darts competition. The results of the competition are shown in the dominance graph. The directed edge from A to B indicates that A defeated B.



- **a** Find the first dominance matrix D^1 showing the one step dominance between the players.
- **b** Find the second dominance matrix D^2 showing the two step dominance between the players.
- **c** Find the dominance matrix by summing D^1 and D^2.
- **d** Find the dominant player by summing the rows of the dominance matrix.

9 A landscape garden is planted in five tiers A, B, C, D and E. Tier A is the highest, B is the second highest, followed by C and D, with E being the lowest tier.

Only level A is watered and flows down to the other levels. The adjacency matrix shows the possible paths the water can take.



$$R^1 = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ & & & & \\ & & & & \\ & & & & \end{bmatrix} \end{array}$$

- **a** Complete the matrix.
- **b** Draw a directed graph where the vertices represent the tiers and the edges represent the paths the water can take.
- **c** Find R^2 for this network.
- **d** Name the possible paths from A to E.

10 There are four different brands of cereal that are labelled W, X, Y and Z. The cereals are rated at a blind tasting and it is found that

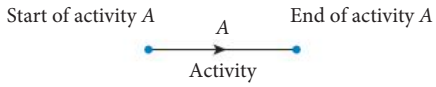
- W is rated higher than Y and Z
- X is rated higher than W and Y
- Y is rated higher than Z
- Z is rated higher than X

- **a** Draw a dominance digraph where a cereal has dominance if it is rated higher.
- **b** Determine the one step dominance matrix.
- **c** Determine the two step dominance matrix.
- **d** Find the dominance matrix $D^1 + D^2$ and use this result to find the dominant cereal.

9780170250283

9.02 DRAWING NETWORK DIAGRAMS

When a large number of activities are involved in the completion of a task or project, it is often useful to draw a network to show the sequence of activities and the time required for each **activity**. The diagram used is a directed graph or digraph and is made up of vertices and edges. The edges (lines) represent the activities and the vertices (circles) represent the **events**.



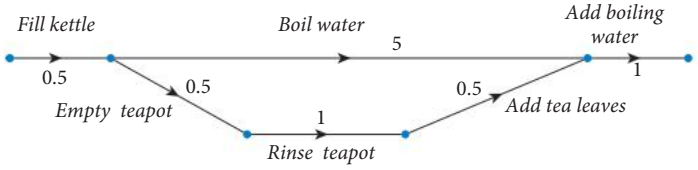
Project networks

IMPORTANT

In a network representing a project, each activity is represented by an **edge**, shown by a line with an arrow. The activity is labelled using a letter and the time is marked on the line. It is only possible to move along the network from the start to the end in the direction of the arrows.

Each **vertex** or **node** is shown by a circle. The vertices represent the events for the start and end of a particular activity. This is shown in the diagram above.

In a network, some activities can be performed at the same time, while in other cases an activity cannot commence until another is complete. This can be shown using the network for making a pot of tea. It is possible to boil the water at the same time as the tea pot is emptied. However, the water cannot be added to the pot until the tea pot is emptied, rinsed and fresh tea leaves added.



Drawing a network from an activity table

An activity table shows the order and estimated time for each activity.

An **immediate predecessor** to an activity is any activity that must be completed before this activity can commence.

Types of connections

There are three different types of connections.

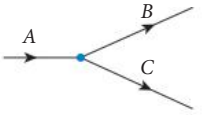
One preceding event

Event *B* is preceded by *A*



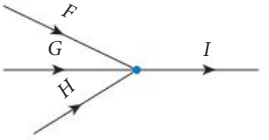
More than one activity with the same preceding event

Events *B* and *C* are preceded by *A*



An activity with more than one preceding event

Event *I* is preceded by *F*, *G* and *H*



Guidelines to follow when drawing a network diagram

IMPORTANT

1. Use a vertex to represent the start of the network.
2. Look for any activities that do not have any predecessors. These will be your starting activities.
3. Activities that have multiple predecessors will all end at the same vertex.
4. An activity should not be represented by more than one edge in the network.
5. Two vertices can be connected by one edge only.
6. A vertex indicating the completion of the project needs to be included in the network.

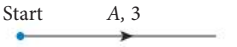
Example 4

Draw a network for the activity table shown.

Activity	Time (hours)	Immediate predecessor
A	3	none
B	4	*A*
C	3	*A*
D	1	*B*
E	5	*B*
F	7	*C*
G	3	*D*
H	1	*E*
I	3	*F*, *G*, *H*

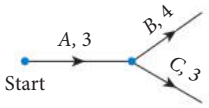
Solution

Activity *A* is the only activity with no predecessors. Draw the first vertex and from this a directed edge labelled *A*, 3. Label the first vertex as the start.

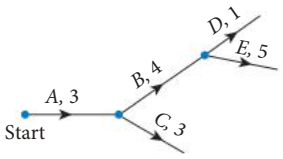


9780170250283

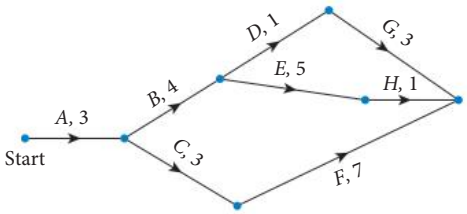
Activities *B* and *C* are preceded by *A*. End activity *A* with a vertex, then draw two directed edges starting at this vertex labelled *B*, 4 and *C*, 3.



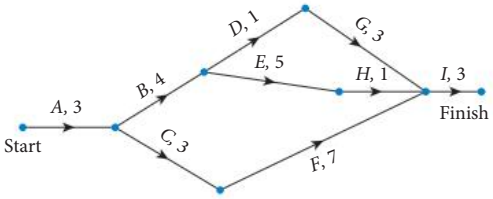
Activities *D* and *E* are preceded by *B*. End activity *B* with a vertex, then draw directed edges starting at this vertex labelled *D*, 1 and *E*, 5.



Activities *F*, *G* and *H* all precede *I*, so each directed edge representing one these activities must end at the same vertex.
Activity *F* is preceded by *C*. Draw a directed edge starting from the vertex that finished *C*, 3 and label it *F*, 7. End with a vertex.
Activity *G* is preceded by *D*. Draw a directed edge from the vertex that finished *D*, 1 and label it *G*, 3. End at the same vertex as *F*, 7.
Activity *H* is preceded by *E*. Draw a directed edge from the vertex that finished *E*, 5 and label it *H*, 1. End at the same vertex as *F*,7 and *G*,3.



Activity *I* is an example of an activity with more than one preceding event.
Activity *I* is preceded by *F*, *G* and *H*. Draw a directed edge from the vertex that finishes *G*, 3, *H*,1 and *F*, 7 and label it *I*, 3. End with a vertex and label this vertex as the finish.



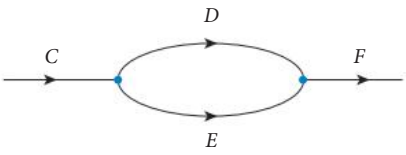
Dummy activities

Sometimes a dummy edge needs to be added to a network to maintain precedence structure whilst ensuring that no two vertices are connected by multiple edges. A **dummy activity** has zero time and is shown as a directed edge with a broken line.

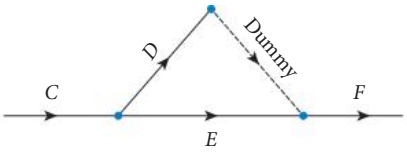
Multiple edges

A network cannot be drawn with multiple edges connecting the same vertices. In the example below, the directed graph cannot be drawn with two vertices both connected by activities *D* and *E*. A dummy activity needs to be included. It can occur before or after activities *D* or *E* to overcome this problem.

Activity	Immediate predecessor
D	*C*
E	*C*
F	*D*, *E*



Incorrect

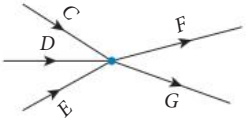


Correct

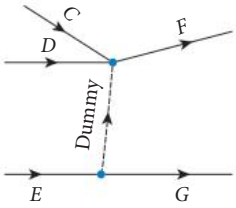
Logic difficulties in the precedence structure

The second function of dummy activities is to ensure that the logic of the network is maintained. The first network has activity *G* preceded by activities *C*, *D* and *E*, which does not follow the logic of the activity table. A second vertex and a dummy activity needs to be included so activities *F* and *G* do not both start from the same vertex and activities *C*, *D* and *E* do not finish at the same vertex.

Activity	Immediate predecessor
F	*C, D, E*
G	*E*



Incorrect



Correct

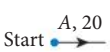
Example 5

Draw a network to show the task of publishing a book, using the activities table below.
Publishing a book

Code	Activity	Time (weeks)	Predecessors
A	Write manuscript	20	None
B	Edit manuscript	6	*A*
C	Typeset manuscript	4	*B*
D	Perform rough pagination	3	*C*
E	Perform final pagination	3	*D*
F	Do artwork	10	*D*
G	Mount artwork in final pages	5	*E, F*
H	Promote book	4	*A*
I	Print book	6	*G*
J	Distribute book for sale	2	*H, I*

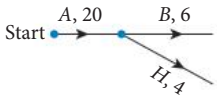
Solution

Only activity *A* has no predecessors.

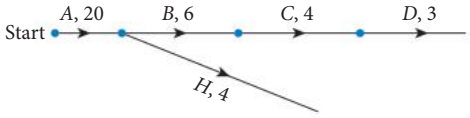


9780170250283

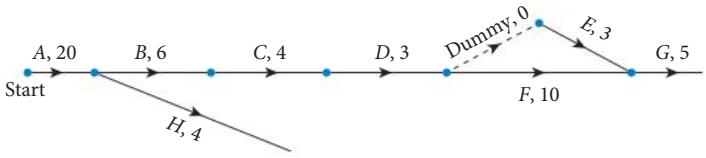
Activities *B* and *H* are both preceded by *A*.
Draw directed edges to represent each activity but do not add the end vertex.



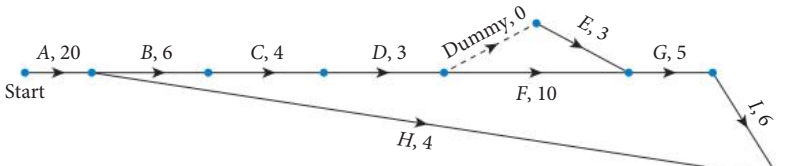
Activity *C* is preceded by *B* and activity *D* is preceded by *C*.



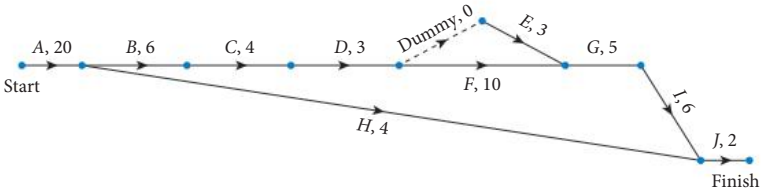
Activities *E* and *F* are both preceded by activity *D*. Activity *G* is preceded by activities *E* and *F*. This means *E* and *F* both end at the same vertex. Two edges cannot be connected by multiple edges, so a dummy activity will need to be included.



Activity *I* is preceded by activity *G* but activities *H* and *I* both precede activity *J*. This means *H* and *I* both end at the same vertex.



The final activity is *J*, which is preceded by activities *H* and *I*.



INVESTIGATION Drawing your own network.

Design your own network for a project. Some possible projects are

- building a tree house
- planting a vegetable garden
- baking a cake
- organising a children's party

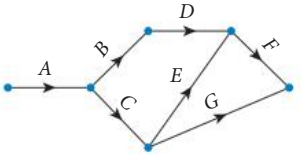
Write your own activity table and estimated activity times. In your planning, you will need to decide on the sequence of activities, the activities that can be done at the same time as other activities, and the activities that depend on another activity being completed before they can start. Once this is done, the information can be added to the predecessor column in the activity table. You can now draw and label the network.

EXERCISE 9.02 Drawing network diagrams

WS Drawing directed graphs

Concepts and techniques

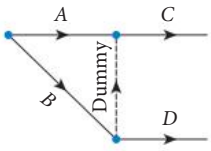
1 Which one of the following is **incorrect** about the directed graph on the right?

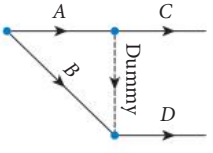


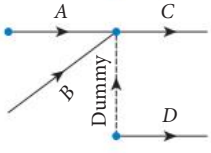
A *B* is preceded by *A*
B *C* is preceded by *F*
C *E* and *G* are preceded by *C*
D *F* is preceded by *D* and *E*
E *D* is preceded by *B*.

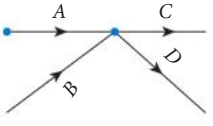
2 The directed graph for the network in the table below is

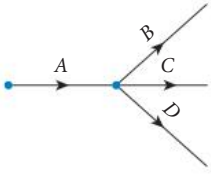
Activity	Immediate predecessor
A	None
B	none
C	*A*
D	*A*, *B*

A


B


C


D


E


9780170250283

3 Draw the networks for each activity table, including the dummy activities where required.

a

Activity	Time (days)	Immediate Predecessor/s
A	4	none
B	3	none
C	5	A, B

b

Activity	Time (hours)	Immediate Predecessor/s
A	3	none
B	5	A
C	8	A
D	6	B, C

c

Activity	Time (days)	Immediate Predecessor/s
A	3	none
B	6	none
C	5	A
D	7	A, B

d

Activity	Time (days)	Immediate Predecessor/s
A	8	none
B	3	A
C	6	A
D	7	A
E	5	B
F	9	C, E
G	4	C, E, D

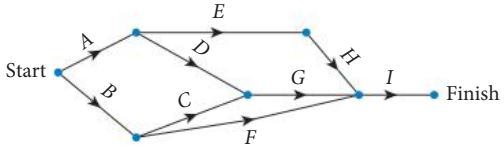
4 For the network shown, write the activity or activities that precede

a activity E

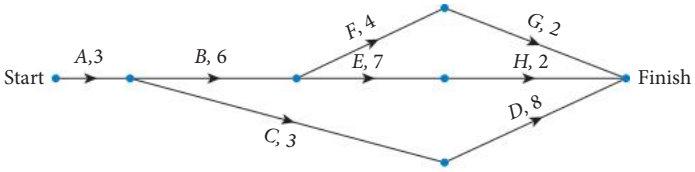
b activity B

c activity G

d activity I



5 Complete the activity table for the following network.



Activity	Time (days)	Predecessor/s
A		
B		
C		
D		
E		
F		
G		
H		

6 Draw the network for the following activity tables, including dummy activities where required.

a

Activity	Time (hours)	Predecessor/s
A	4	none
B	6	*A*
C	2	*A*

b

Activity	Time (hours)	Predecessor/s
A	2	none
B	3	none
C	2	none
D	4	*A*
E	5	*B*
F	7	*C*

c

Activity	Time (hours)	Predecessor/s
A	6	none
B	3	*A*
C	5	*A*
D	4	none
E	3	*B*
F	3	*C*
G	7	*D, E, F*

d

Activity	Time (days)	Predecessor/s
A	7	none
B	5	*A*
C	5	none
D	8	*B, C*
E	8	none
F	4	*D, E*

9780170250283

7 Example 4 Draw a project network for the task of making a dressing table.

Code	Activity	Time (days)	Predecessor/s
A	Order mirror	10	None
B	Order drawer pull handles	10	None
C	Make frame	5	None
D	Make top	5	*C*
E	Make drawers	7	*C*
F	Fit top	1	*D*
G	Fit drawers	1	*E*
H	Finish and polish surfaces	3	*F, G*
I	Fit drawer pull handles	1	*B, H*
J	Fit mirror	1	*A, H*

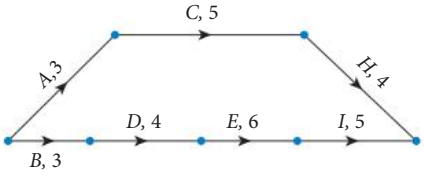
8 Example 5 Draw a project network for the task of making a raised garden bed using sleepers.

Code	Activity	Time (hours)	Prerequisites
A	Take the trailer to the dump	2	*F*
B	Put the new plants into the bed	3	*D, H*
C	Mark out the site of the garden bed	1	None
D	Put the new soil into the bed	3	*G, M*
E	Water the plants in	$\frac{1}{2}$	*B*
F	Remove the existing vegetation and put in trailer	4	*C*
G	Put the side sleepers in and bolt into place	3	*J*
H	Buy the plants to put in the bed	2	*L*
I	Concrete corner posts into place	$1\frac{1}{2}$	*K*
J	Allow concrete to dry	10	*I*
K	Dig holes for the corner posts	$3\frac{1}{2}$	*C*
L	Plan the layout of the plants to be placed in the bed	2	None
M	Use the trailer to get the new soil	$2\frac{1}{2}$	*A*

Reasoning and communication

9 The following activity table is partially illustrated in the network below.

Activity	Time (days)	Predecessor/s
A	3	none
B	3	none
C	5	*A*
D	4	*B*
E	6	*D*
F	7	*B*
G	2	none
H	4	*C, G*
I	5	*E, F*



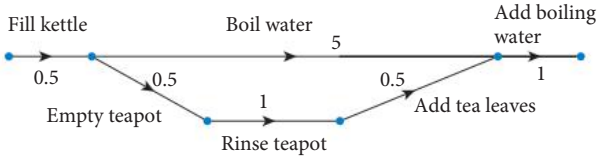
a Add activities *F* and *G* to the network diagram.

b Determine the total time for the paths

i *BDEI* ii *ACH* iii *BFI* iv *GH*

9.03 FORWARD AND BACKWARD SCANNING

Any project ranging from making a cup of tea to building a house needs to be completed within a certain time. Activities in a network often have preceding activities that must be completed before they can commence. This can be seen by looking at the network for making a pot of tea. Consider the activity of adding the boiling water. It is not possible to add the boiling water until the water has boiled and the tea has been added to the pot.





Alamy/Image Source/IE347

Forward scanning

Forward scanning through a network enables us to determine the **earliest starting time** (EST) for every activity in the network. This is the earliest time that it is possible to start the activity.

IMPORTANT

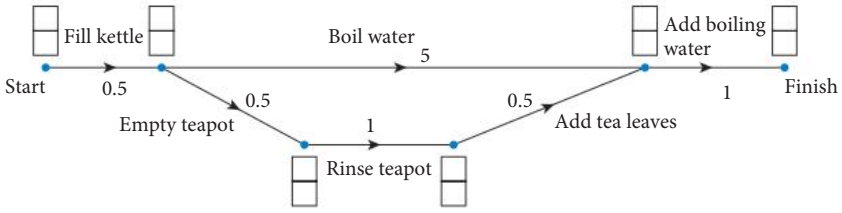
Follow the steps below to determine the earliest starting times.

1. Begin with the first vertex, which has an earliest starting time of 0. The earliest starting times are written in the top box next to each vertex.
2. The EST for any activity is the maximum time taken to reach that activity. So, wherever two or more paths join at a vertex, choose the highest total time as the EST.
3. Continue until the end is reached.

9780170250283

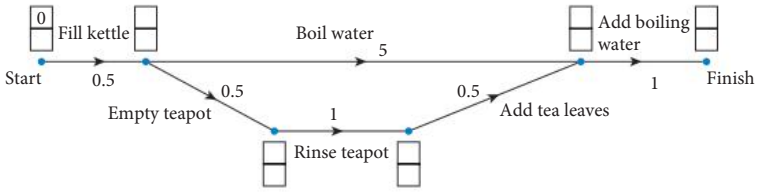
Example 6

Complete the earliest starting times for the following network.

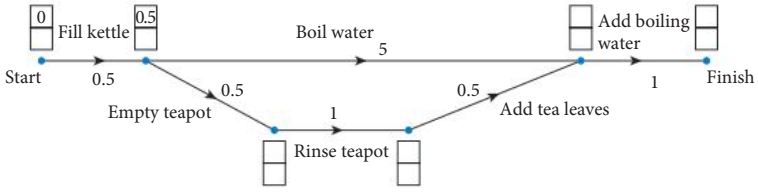


Solution

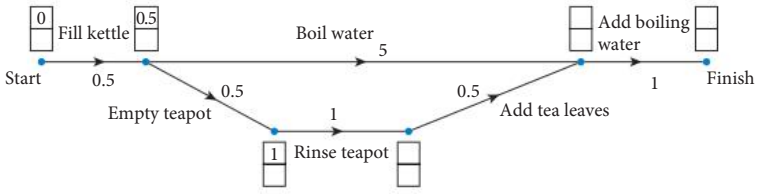
Start at the first vertex with an EST of 0.



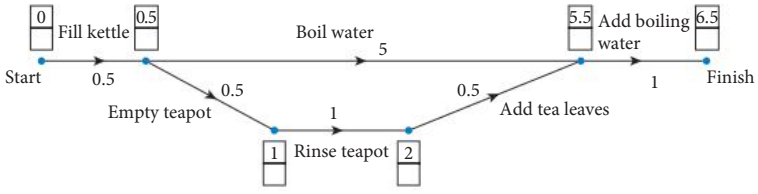
Add the activity time to the EST for the previous vertex.
The EST at the end of "fill the kettle" is $0 + 0.5 = 0.5$



Continue this process.
The EST for the vertex at the end of "empty the pot"
$= 0.5 + 0.5 = 1$

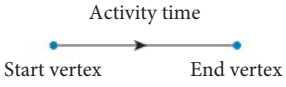


The EST at the start of "add the boiling water" produces two totals.
Activity "boil water"
$= 0.5 + 5$
$= 5.5$
Activity "add tea leaves"
$= 2 + 0.5$
$= 2.5$
The EST is 5.5, which is the largest total.



Backward scanning

Backward scanning is the process used to find the **latest starting times** (LST) for activities. The latest starting time is the latest time you can start an activity without affecting the project completion time.



The LST at the start vertex = the LST at the end vertex – the activity time

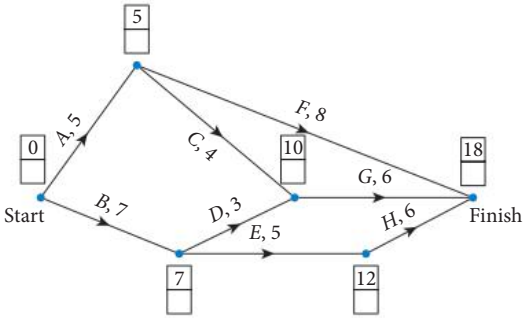
IMPORTANT

Follow the steps below to determine the latest starting times (LST).

1. Start at the final vertex. The latest starting time at the final vertex is the same as the earliest starting time.
2. LSTs are written in the bottom box next to each vertex.
3. To find the LST at the start vertex, work backwards, subtracting the activity time from the LST at the end vertex.
4. Where there is more than one path to the previous vertex, take the smallest result as the LST.
5. The LST at the first vertex must be zero.

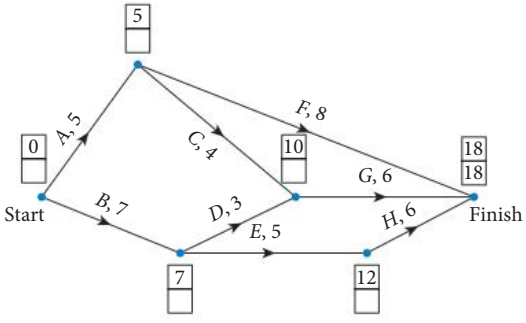
Example 7

The earliest starting times are shown in the network below. Complete the latest starting times for this network.



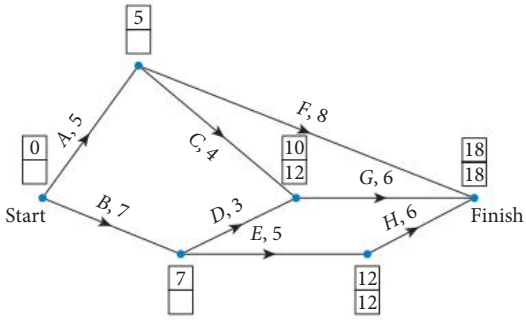
Solution

Start at the last vertex in the network. This vertex has a latest starting time (LST) that is the same as the earliest starting time (EST).

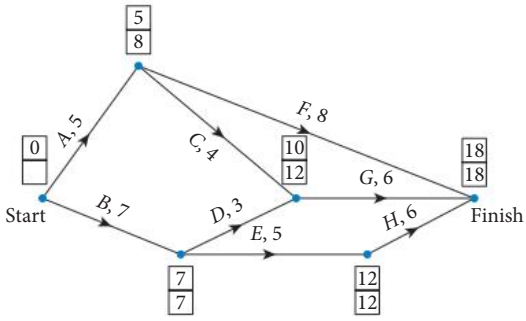


9780170250283

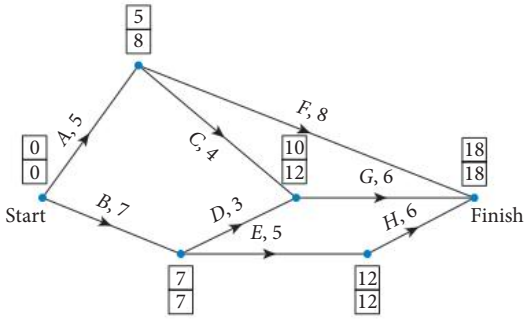
There is only one path backwards to the vertex at the start of activity *H*. The LST is 18 − 6 = 12
There is only one path backwards to the vertex and the start of activity *G*. The LST is 18 − 6 = 12



There are two paths backwards to the vertex at the start activities *D* and *E*. The LST working backwards along
Activity *E* = 12 − 5 = 7
Activity *D* = 12 − 3 = 9
Take the smallest result so the LST is the 7.
For the vertex at the start of activities *C* and *F*, the LST backwards along
Activity *F* = 18 − 8 = 10
Activity *C* = 12 − 4 = 8
Take the smallest result so the LST is 8.



The process is repeated for the start vertex.
The LST backwards along
Activity *A* = 8 − 5 = 3
Activity *B* = 7− 7 = 0
This confirms that the LST for the start of the project is 0.

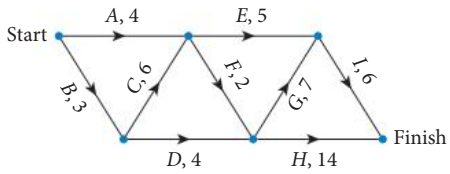


EXERCISE 9.03 Forward and backward scanning

Concepts and techniques

Questions **1**, **2** and **3** refer to the following network.

Activities *A*, *B*, *C*, *D*, *E*, *F*, *G*, *H* and *I* must be completed for the project shown in the network. Activity times are in hours.



Calculating EST and LST

1 The earliest starting time for activity *I* is

A 5 hours **B** 9 hours **C** 14 hours **D** 18 hours **E** 25 hours

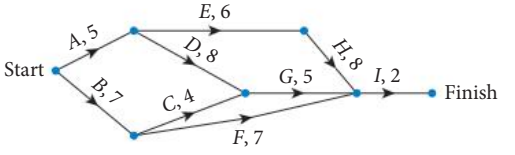
2 The latest starting time for activity *I* is

A 9 hours **B** 18 hours **C** 19 hours **D** 25 hours **E** 51 hours

3 If each of the activities can be reduced by a maximum of one hour, the number of activities that need to be altered for the project to be completed in 22 hours is

A 1 **B** 2 **C** 3 **D** 4 **E** 9

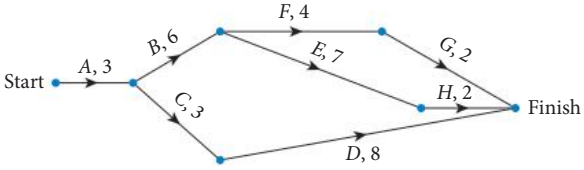
4 Example 6 Use the network diagram below to find the earliest starting time for



a the node at the start **b** the node at the end of activity *A*

c the node at the start of activity *G* **d** the node at the end of activities *G* and *H*.

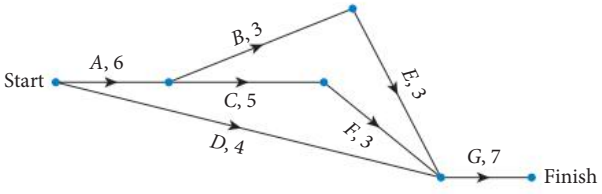
5 Use the network diagram below to find the earliest starting time for



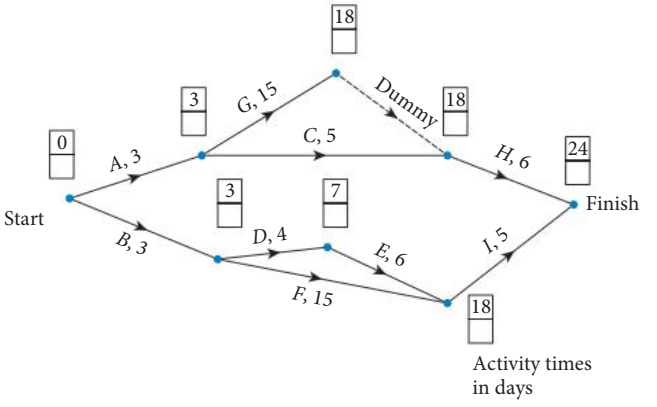
a the vertex at the end of activity *B* **b** the vertex at the start of activity *C*

c the vertex at the end of activity *E* **d** the vertex at the end of the network.

6 Example 6 Forward scan to find the earliest starting times for the activities in the network below.



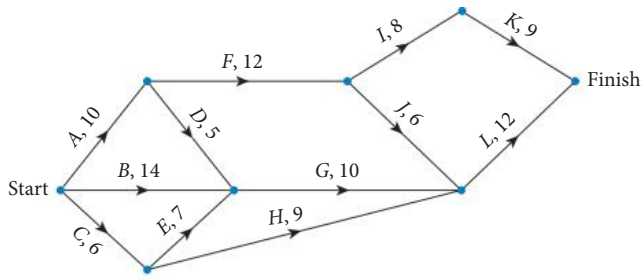
7 Example 7 The ESTs have been completed in the network below. Backward scan to find the latest starting times for each activity in the network.



9780170250283

8 In the network shown in question **4**, the time for activity *F* is changed to 12 days and the time for activity *G* is changed to 10 days. Calculate the EST and LST for each vertex in the network.

9 Calculate the earliest starting times and latest starting times for each activity in the network below. Activity times are in days.



10 In the network in question **6** it is found that activity *B* can now be completed in 4 days. What impact does this have on the EST and LST for the final vertex?

11 The activity table of a project is shown below.

Activity	Time (days)	Predecesor/s
A	8	None
B	9	None
C	2	None
D	5	*A*
E	6	*B*
F	7	*B*
G	10	*C*
H	6	*D, E*
I	3	*F, G*
J	4	*F, G*
K	4	*F, G*
L	5	*H, I*
M	5	*K*

a Draw the network for this project.
b Use forward scanning to determine the EST for each activity.
c Use backward scanning to find the LST for each activity.

Reasoning and communication

12 A rock-fill dam is being constructed. The activity table for the project is shown below.

Code	Activity	Time (days)	Prerequisites
A	Preliminary works	85	None
B	River diversion	150	A
C	Site excavation	200	B
D	Bank construction	600	C
E	Permanent roads	50	B
F	Bank facing	200	D
G	Spillway construction	250	D
H	Inlet tower construction	100	D
I	Outlet valve house	100	D
J	Outlet pipeline	60	I
K	Remove river diversion	50	E, F, G, H, J
L	Clean up and move out	40	K

a Draw a project network for the construction of the dam.
b Use forward scanning to find the earliest starting times for each activity.
c Use backward scanning to find latest starting times for each activity.
d What is the effect of a delay of 60 days in permanent road construction on the final EST?
e What is the effect of a delay of 60 days in facing the bank on the final EST?
f What is the effect of a delay of 60 days in spillway construction on the final EST?

13 A local publisher decides to print calendars that advertise local businesses. The publisher will arrange the printing and distribution of the calendars and will charge businesses to advertise on the calendar. The activity table for production of the calendar is given below.
a Draw a project network for the production of the calendar.
b Use forward scanning to find the earliest starting times for each activity.
c Use backward scanning to find the latest starting times for each activity.
d What is the effect of a delay of 20 days in advertising the calendar on the final EST?
e What is the effect of a delay of 20 days in selling advertising space on the final EST?
f What is the effect of a delay of 20 days in designing the calendar layout on the final EST?

Code	Activity	Time (days)	Prerequisites
A	Advertise calendar	30	None
B	Sell advertising space	40	None
C	Design calendar layout	30	None
D	Typeset calendar with advertisements	10	C
E	Let contract for delivery of calendars	10	None
F	Print calendars	30	B, D
G	Send calendars to distributors	5	F
H	Send 1 copy to each advertiser	1	F

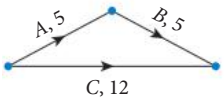
9780170250283

9.04 LOCATING THE CRITICAL PATH

The critical path

There are often multiple paths between the start and end vertices of a network. The **critical path** is the longest time path between the start and end and determines the project completion time.

Consider the simple network.



Activities *A*, *B* and *C* must all be done for the project to be completed. There are two paths through the network. *AB*, which takes 10 days to complete and *C*, which takes 12 days to complete. The project will take 12 days to complete, so this makes *C* the critical path.

In more complex networks, the EST and LST values can be used to find the critical path or paths.

IMPORTANT

The critical path is the longest time path in the network.

On the critical paths, the EST and LST values will be equal at each vertex.

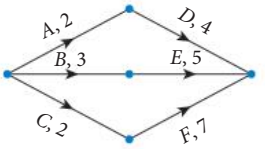
IMPORTANT

Determining the critical path

1. Use forward scanning to determine the earliest starting times for each activity.
2. Identify the path or paths that produce the final EST.
3. Use backward scanning to determine the latest starting times for each activity.
4. Activities for which EST = LST are on the critical path.

Example 8

A project has an activity network where the completion times of activities are in hours.



a List all the paths from start to end in the network and find the total of the activity times for each path.
b Hence find the critical path and the shortest time for project completion.

Solution

a Write all the paths starting at the first vertex and ending at the last vertex. There are three paths.

AD
BE
CF

Add the durations of the activities for each path.

$AD = 2 + 4 = 6$ h
$BE = 3 + 5 = 8$ h
$CF = 2 + 7 = 9$ h

b The critical path is the longest time path, and the shortest time for project completion is the total time for the activities along the path.

Critical path is *CF*, and the shortest time for project completion is 9 hours.

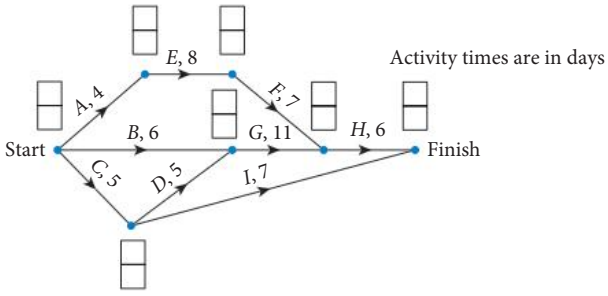
Finding the critical path using EST and LST values.

In more complex networks, it is not practical to list all the paths to find the critical path. A more efficient method is to find the EST and LST values at each vertex in the network and use these to determine the critical path and shortest time for project completion.

9780170250283

Example 9

Consider the network shown below.

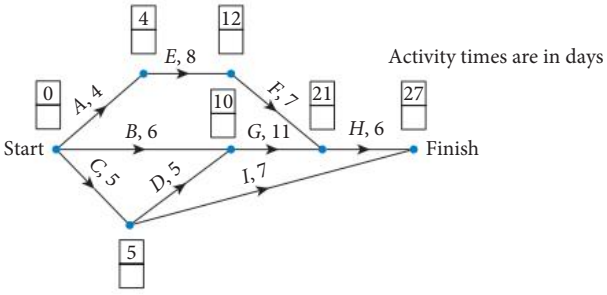


a Calculate the earliest starting times and latest starting times for each activity in the network.

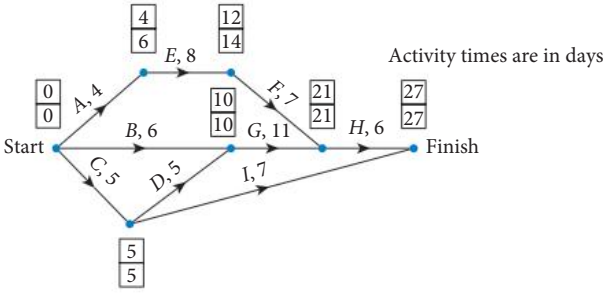
b Hence determine the critical path and the project completion time.

Solution

Calculate earliest starting times (EST) for each activity.
It is evident that the path *CDGH* is the longest time path. The sum of the activity times along this path, $5 + 5 + 11 + 6 = 27$, produces the final EST.



Calculate the latest starting times (LST) for each activity.

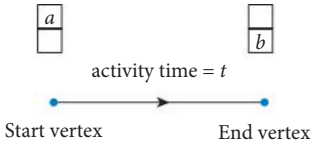


On the critical path, EST = LST at each vertex. The project completion time is the shortest time in which the project can be completed and is calculated by adding the times for each activity on this path.

The critical path is *CDGH* and the project completion time is 27 days.

Activity float times

The float time for an activity is the **maximum** time that an activity can be delayed without affecting the project completion time. Activities on the critical path all have **float** times of zero.



Activities on the critical path are called **critical activities** and those not on the critical path are called **non-critical activities**.

In the diagram

a = the EST at the start of an activity
b = the LST at the end of an activity
t = the activity time.

Float = $b - a - t$

Consider activity E in Example **2**.

IMPORTANT

$\text{Float} = \text{LST}_{\text{end}} - \text{EST}_{\text{start}} - \text{activity time}$

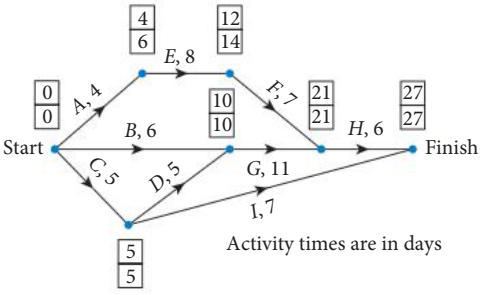


The earliest time activity E can start is after 4 days and the latest that activity E can finish is after 14 days. This allows 10 days for the activity. However, activity E requires 8 days, which means that the longest delay possible is 2 days. The delay of 2 days is the float for activity E. When the formula is used, float = 14 − 4 − 8 = 2.

Example 10

The EST and LST values have been completed for the network

Calculate the float times for activity



a A **b** C **c** G **d** F

9780170250283

Solution

a Float = $LST_{end} - EST_{start}$ – activity time

For activity A
$LST_{end} = 6$
$EST_{start} = 0$
Activity time = 4
Activity A float = $6 - 0 - 4$
= 2 days

b All activities along the critical path have a float time of zero. The critical path is $CDGH$, which means that activity C has a float time of zero.

Activity C float = 0 days

c Activity G is on the critical path.

Activity G float = 0 days

d Float = $LST_{end} - EST_{start}$ – activity time

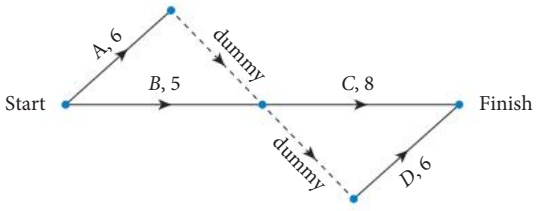
For activity F
$LST_{end} = 21$
$EST_{start} = 12$
Activity time = 7
Activity F float = $21 - 12 - 7$
= 2 days

Project crashing

It is possible to speed up a project by employing more resources and this enables some activities to be completed more quickly. When this occurs with critical activities, the process is called **crashing**.

Example 11

The network for a project is shown below.



(Activity times are in days)

a Determine the critical path.

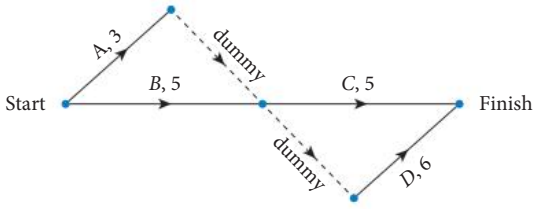
b Extra resources are employed so that the time required to complete each of activities A and C is reduced by three days. Determine the critical path and project completion time.

Solution

a The critical path is the longest time path from the start to the end of the network. This will be AC, which will take $6 + 8 = 14$ days to complete. The dummy activity is not included when describing the path.

Critical path = AC

b The times for activities *A* and *C* will change to 3 days and 5 days. The new network is shown.



The longest time path is now *BD*, which will take $5 + 6 = 11$ days to complete.

The new critical path is *BD*.
The project completion time is 11 days

EXERCISE 9.04 Locating the critical path

Concepts and techniques

Critical paths, critical times and activity float times

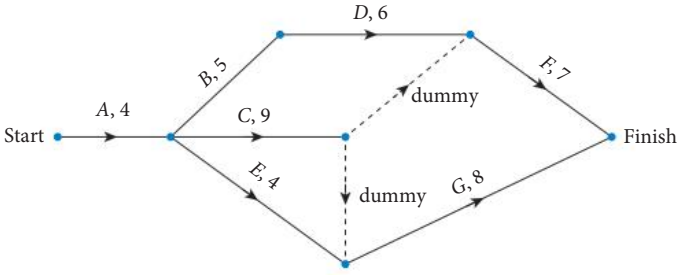
1 Bella, Grace, Lily and Cassie all work on an assembly line where activities *C*, *D*, *E* and *F* must be completed. The activity table for their project is shown below.

Person	Activity	Time (minutes)	Immediate predecessor
Bella	*C*	15	none
Grace	*D*	18	none
Lily	*E*	10	*C*
Cassie	*F*	11	*D*

Lily's activity has a delay of 5 minutes. The time taken to complete the project is

A 18 minutes **B** 29 minutes **C** 30 minutes
D 33 minutes **E** 59 minutes

Questions **2** and **3** refer to the following network where activity times are in days.



2 If activity *C* is delayed by 6 days, the project would be delayed by

A 1 day **B** 3 days **C** 5 days **D** 6 days **E** 22 days

3 Extra resources are used to speed up the original project, resulting in every activity being reduced by 2 days. The new project completion time is

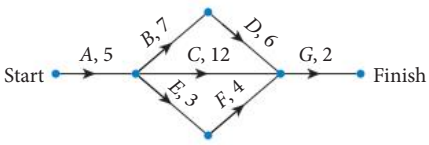
A 14 days **B** 15 days **C** 18 days **D** 20 days **E** 24 days

9780170250283

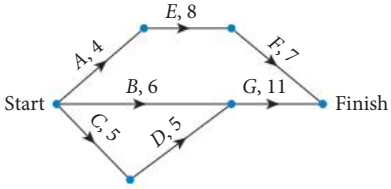
4 **a** Example 8 List all the possible paths for the networks below and find the total of the activity times for each path (activity times are in hours).

b Hence find the critical path and the critical time for the network.

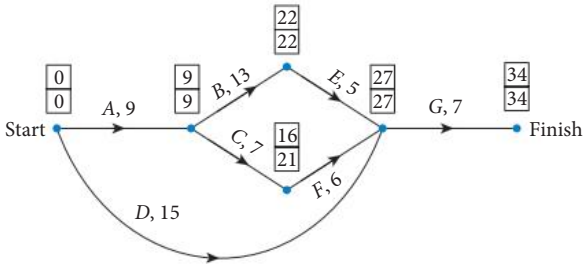
i



ii

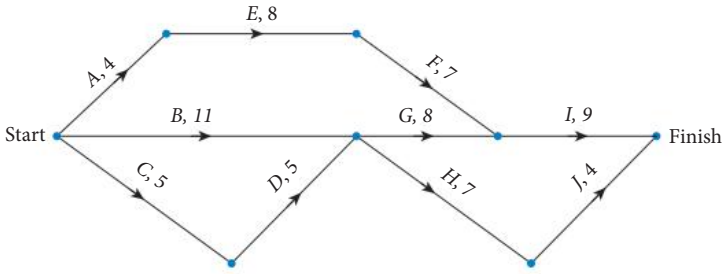


5 Example 9 The EST and LST values have been calculated for the network below. Find the critical path and the minimum time required to complete the project.



Activity times in hours

6 The activities times shown in the following network are in hours.



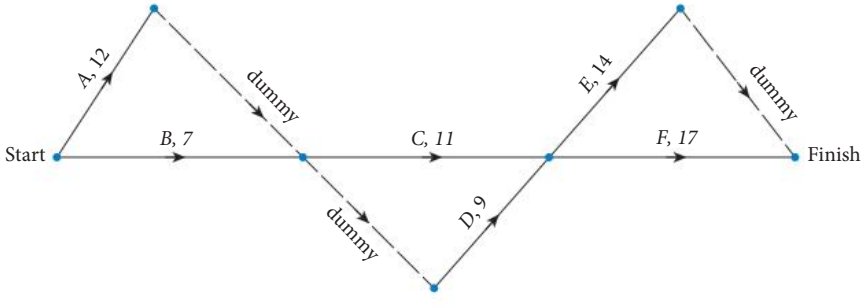
a Determine the EST and LST values for each activity.

b Find the critical path.

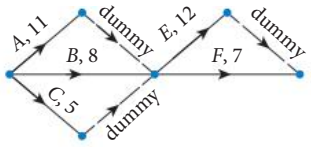
c Find the minimum time in which the project can be completed.

d Extra resources are employed so that activities *B* and *E* can be completed in 7 days and 3 days respectively. Find the critical path and the critical time.

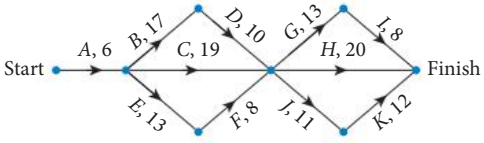
7 Example 11 The network for a project is shown below and activity times are given in hours. Extra resources are employed so that activities *A*, *C* and *E* can be completed in 9 hours, 7 hours and 12 hours respectively. Determine the time saved on the project.



8 The network for a project is shown below and activity times are given in days. Extra resources are employed so that the times for activities *A*, *E* and *F* are reduced by 6 days, 7 days and 2 days respectively. Determine the time saved on the project.



9 The activities times shown in the following network are in hours.
- a Determine the EST and LST values for each activity.
- b Find the critical path.
- c Find the minimum time in which the project can be completed.



10 Example 10 From the network diagram in question **6**, determine the float times for all the non-critical activities in the network.

11 From the network diagram in question **9**, determine the float times for all the non-critical activities in the network.

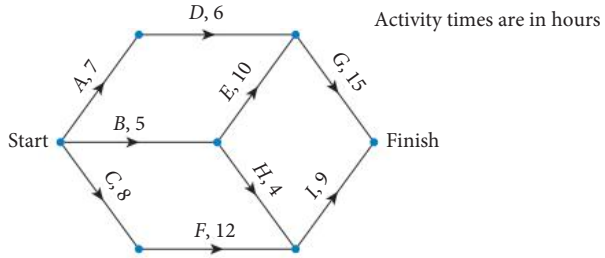
12 From the following activity table
- a construct a network diagram
- b forward scan to determine the earliest starting times for each activity
- c backward scan to determine the latest starting times for each activity
- d determine the critical path or paths
- e determine the earliest completion time for the project
- f determine the float times for all non-critical activities.

Activity	Time (days	Predecessor/s
A	3	none
B	4	none
C	7	*A*
D	6	none
E	8	*B*
F	5	*B*
G	12	*D*
H	10	*F*
I	11	*C*, *E*
J	4	*H*
K	4	*G*

9780170250283

Reasoning and communication

13 The network for a project is shown below.



a Forward scan to determine the earliest starting times for each activity.
b Backward scan to determine the latest starting times for each activity.
c Determine the critical path.
d Determine the earliest completion time for the project.
e Determine the float times for all non-critical activities.
f Determine the effect on the project completion time and the critical path if activity D is delayed by 10 days.
g Determine the effect on the project completion time and the critical path if activity I is delayed by 5 days.

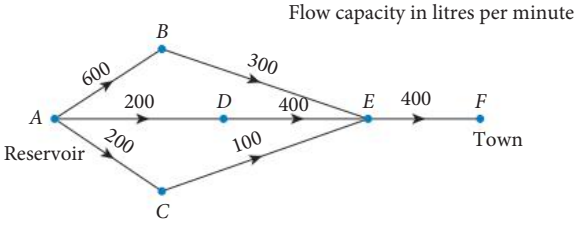
14 The activity table for a network shows the 9 activities required to complete a project.

Activity	Time (days)	Predecessor/s
A	2	None
B	4	None
C	5	A, B
D	3	C
E	2	D
F	7	E
G	3	E
H	2	E
I	6	F, G, H

a Draw the network.
b Calculate the EST and LST values for each activity.
c Determine the critical path and the project completion time.
d Determine the float times in all non-critical activities.
e Which activities can be delayed without affecting the project?
f Activities F and G are combined to make one activity with an activity time of 10 days. What impact does this have on the project completion time?

9.05 NETWORK FLOW PROBLEMS

The applications of flow networks range from transporting water through a network of water pipes to moving people or product by road or a rail network.



Consider the situation where water is transported from a reservoir at A to a town at F via a series of pipes. The network is a directed graph where the edges show the direction and **capacity** of **flow** between the start (**source**) and the end (**sink**).

Flow capacity and maximum flow

The start of the network is called the source which in this example is the reservoir. The end of this network is the town and it is called the sink.

Inflow capacity. This is the total flow capacity entering a vertex.

At vertex B there is a capacity for 600 litres per minute to enter.

Outflow capacity. This is the total flow capacity leaving the vertex.

There is only a capacity of 300 litres per minute to leave vertex B.

Maximum flow. This is the smaller of total inflow capacity and total outflow capacity.

The maximum flow out of vertex B is 300 litres per minute.

At vertex D there is only an inflow capacity for 200 litres per minute however there is an outflow capacity of 400 litres per minute to leave. The maximum flow out of vertex B is 200 litres per minute.

IMPORTANT

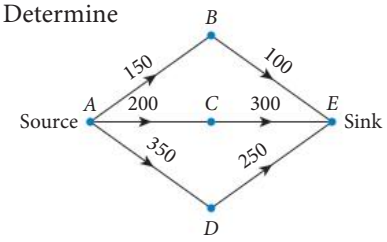
The maximum flow out of a vertex is the smaller of total inflow capacity and total outflow capacity.

9780170250283

Example 12

The network shows the flow capacities in litres per minute. Determine

a the inflow capacity of vertex B
b the outflow capacity of vertex B
c the maximum flow out of vertex B
d the maximum flow out of vertex C.
e the maximum flow out of vertex D.



Solution

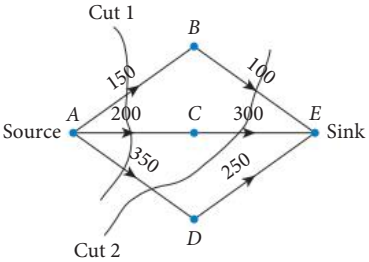
a The edge entering vertex B is the inflow and the value on the edge indicates the capacity.	The inflow capacity of vertex B is 150 litres per minute.
b The edge leaving vertex B is the outflow and the value on the edge indicates the capacity.	The outflow capacity of vertex B is 100 litres per minute.
c Maximum flow is the smaller of inflow and outflow capacity.	The maximum flow out of vertex B is 100 litres per minute.
d For vertex C, the maximum flow out is the smaller of inflow = 200 and outflow = 300.	The maximum flow out of vertex C is 200 litres per minute.
e For vertex D, the maximum flow out is the smaller of inflow = 350 and outflow = 250.	The maximum flow out of vertex D is 250 litres per minute.

Maximum flow – minimum cut

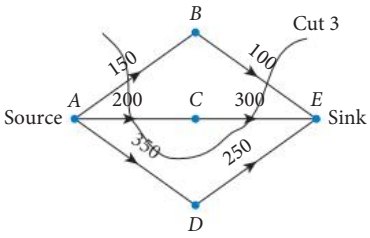
One method of finding the **maximum** flow through a network is to use the **minimum** capacity of the cuts.

A **cut** through a network must stop all flow from the start (*source*) to the end (*sink*).

In the network shown, cut 1 and cut 2 are both **valid**. The cuts sever the pipes in the network in such a way that there is no flow path from A to E. Cut 3 is not a valid cut because it does not stop all the flow between A and E. It is still possible for flow to occur from A to D to E.



Valid cuts



Invalid cut

The capacity of a cut

The capacity of a cut can be determined by adding all the flow capacities that pass across the cut in the direction from source to sink. In our example, this is from the left to the right-hand side of the cut.

Capacity of cut 1 = 150 + 200 + 350 = 700

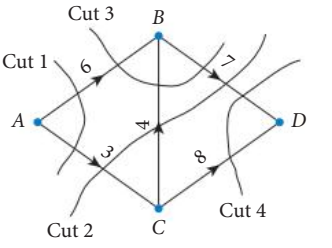
Capacity of cut 2 = 100 + 300 + 350 = 650

Example 13

The network shows the flow from the source at A to the sink at D.

a Determine which of the cuts in the network are valid.

b Determine the capacity of the valid cuts.



Solution

a The cuts that are valid will cut the supply from A to D.

Cuts 1, 2 and 4 are valid.

Cut 3 is not valid because it is still possible for flow to occur between A, C and D.

Cut 3 is not valid.

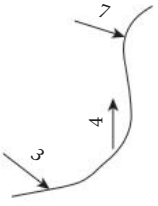
b The capacity of the cut is determined by adding the capacities of the edges that cross the cut from left to right.

Capacity of cut 1 = 6 + 3 = 9
Capacity of cut 4 = 7 + 8 = 15

For cut 2, the edge from C to B is not included because the flow does not travel towards the cut from left to right. The flow of 4 would only be included if the direction of flow was reversed.

Capacity of cut 2 = 3 + 7 = 10

The vertex C is on the sink side of the cut and the vertex B is on the source side of the cut and hence the direction of flow does not travel towards the cut.



The direction of the flow must be towards the cut for the flow capacity to be included.

9780170250283

IMPORTANT

Determining the maximum flow for a network

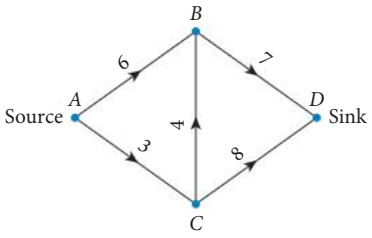
1. Identify cuts through the network.
2. Find the capacity of each cut.
3. Maximum flow = capacity of the minimum cut.

When determining the capacity of a cut, any edges that do not contribute to the flow of the network in the required direction are not included.

The starting vertex of an edge must be on the source side of the cut and the end vertex must be on the sink side of the cut for the edge to be included in the capacity of a cut.

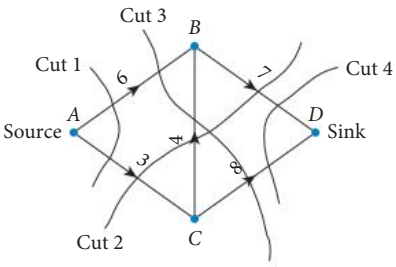
Example 14

Determine the maximum flow for this network where flow occurs between the source A and the sink D.



Solution

Identify the possible cuts that can be made to the network. The capacity of cuts 1, 2 and 4 was determined in Example **13**. A new cut 3 has also been added.



Capacity of cut 1 = 6 + 3 = 9
Capacity of cut 4 = 7 + 8 = 15
Capacity of cut 2 = 3 + 7 = 10
Capacity of cut 3 = 6 + 4 + 8 = 18

Maximum flow capacity = capacity of the minimum cut
Identify the smallest of the cut capacities, which is 9.

The maximum flow is 9.

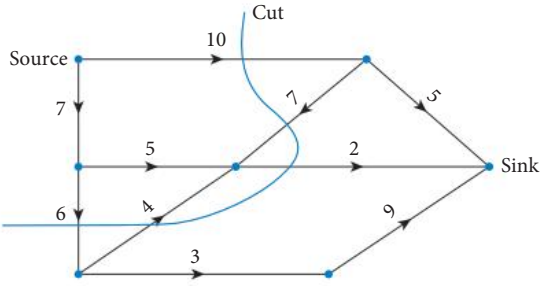
EXERCISE 9.05 Network flow problems

Network flow capacity

Concepts and techniques

The following information refers to questions **1** and **2**.

In the network below, the numbers on the edges show the maximum possible flow between the vertices. The direction of the arrow indicates the direction of the flow. A cut separating the sink from the source is also shown.



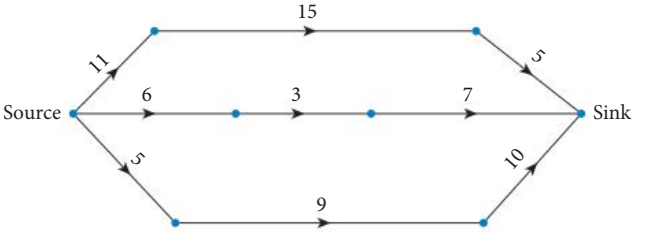
1 The capacity of the cut is

A 16 **B** 18 **C** 22 **D** 27 **E** 29

2 The maximum flow from the source to the sink is

A 13 **B** 15 **C** 16 **D** 17 **E** 28

3 The maximum flow from source to sink in the network is



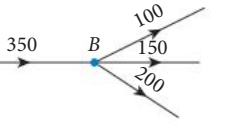
A 13 **B** 14 **C** 20 **D** 22 **E** 32

4 Example 12 For vertex *B* in the network, determine

a the inflow capacity of vertex *B*

b the outflow capacity of vertex *B*

c the maximum flow out of vertex *B*.

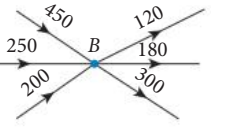


5 For vertex *B* in the network, determine

a the inflow capacity of vertex *B*

b the outflow capacity of vertex *B*

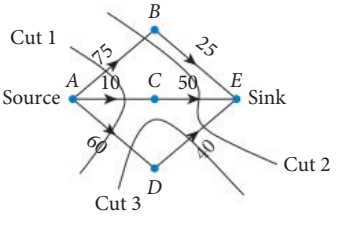
c the maximum flow out of vertex *B*.



6 Example 13 For the network shown, determine

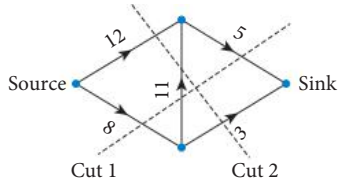
a which of the cuts are valid

b the capacity of the valid cuts.

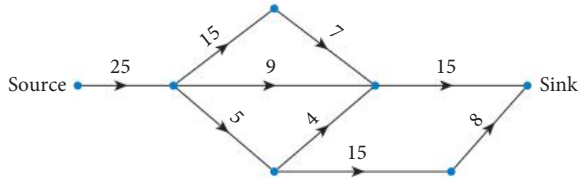


9780170250283

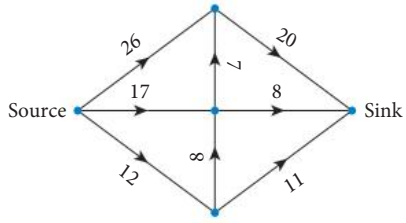
7 Find the capacity of the cuts in the network below.



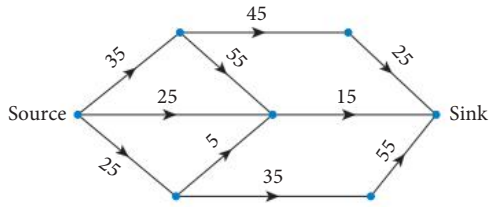
8 Example 14 Find the maximum flow capacity of the network.



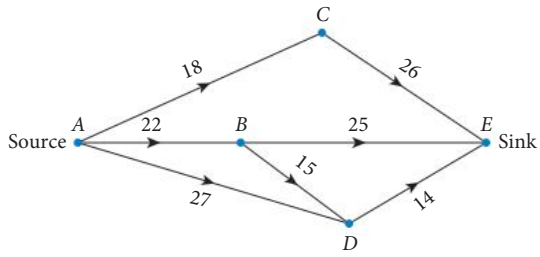
9 Find the maximum flow for the network.



10 Find the maximum flow for the network.



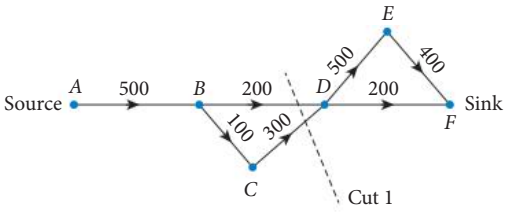
11 **a** Determine the number of cuts, and the capacity of each cut for the network shown.



b Hence determine the maximum flow for the network.

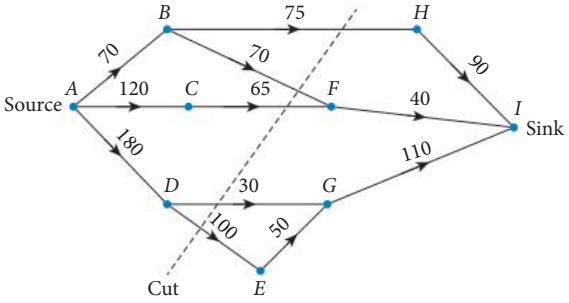
Reasoning and communication

12 The network shown is for a series of pipes that carry storm water from the source at A to an ocean outlet at B. The capacity of each pipe is shown in kilolitres per minute.



- **a** Determine the capacity of cut 1.
- **b** Determine the number of possible cuts and the capacity of each cut.
- **c** Find the maximum flow from A to F.
- **d** The pipe from B to C is to upgraded so the network can carry more water. Find the minimum capacity of the upgraded pipe so that maximum flow through the network can be achieved.

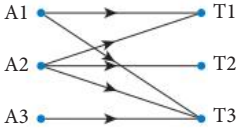
13 The network for a small airline running flights connecting cities A, B, C, D, E, F, G, H and I is shown. The maximum number of passengers in each flight is shown on each edge.



- **a** Find the capacity of the cut.
- **b** Find the maximum number of passengers that can be transported from A to I.
- **c** The flight from B to H is provided with a larger plane. What is the minimum number of passengers it must carry for the maximum capacity of the network to be achieved?
- **d** A company books 60 of their employees on the flight from A to D. Find the maximum number of these people that will get to I without leaving the group.

9.06 BIPARTITE GRAPHS AND THE ASSIGNMENT PROBLEM

The head chef in a kitchen has three tasks that need to be performed $\{T1, T2, T3\}$ and three apprentice chefs $\{A1, A2, A3\}$ who will each be assigned a task. The apprentices are not all experienced enough to complete the three tasks so the head chef draws a bipartite graph to help him decide how to allocate the tasks.



9780170250283

The graphs shows

- Apprentice 1 can complete tasks 1 and 3
- Apprentice 2 can complete tasks 1, 2 and 3
- Apprentice 3 can complete task 3

$A2$ is the only qualified apprentice for task 2. This now makes $A1$ the only qualified apprentice left for task 1.

Based on this information, the head chef allocates $A2$ to $T2$, $A1$ to $T1$ and $A3$ to $T3$.

Bipartite graphs

A **bipartite graph** has its vertices in two distinct sets and the edges join elements in the first set to elements in the second set. In the example above, the two sets are the apprentices and the tasks and the edges indicate which apprentices have the experience to complete the particular tasks.

The assignment or allocation problem

The **assignment** problem arises in a number of different situations.

- Assigning people to tasks
- Assigning jobs to machines
- Assigning deliveries from warehouses to stores
- Assigning taxis to customers

In each case, one agent is assigned only one task.

In an assignment problem, we are finding the best way of matching the elements in the two groups that optimises a stated objective such as minimising cost, distance or time.

Example 15

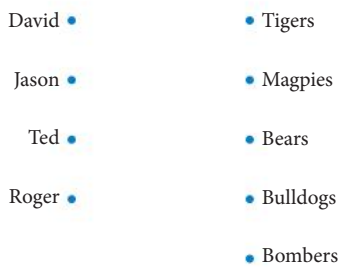
The clubs that four footballers have played for are shown in the table below.

Players	Clubs
David	Tigers, Magpies
Jason	Bears, Bulldogs
Ted	Bulldogs
Roger	Bombers, Bears

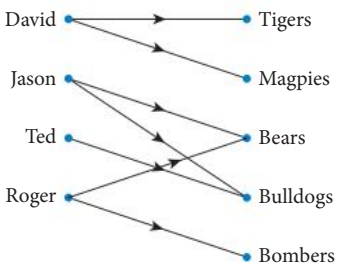
Represent this as a bipartite graph.

Solution

On the left-hand side the players are represented as 4 vertices, and on the right-hand side the teams are represented as 5 vertices.



Draw in edges from the players connected to the teams they have played for.



In many **allocation** problems data is supplied as a table or **matrix**. This could be the times workers take to complete particular tasks. The solution to the problem shows the best way to allocate workers to tasks so the tasks are completed in the shortest time possible. In very simple problems this can be done by inspection.

Example 16

A company has to allocate three tasks, Task 1, Task 2 and Task 3 that have to be matched to three employees *A*, *B* and *C*. Each of the employees has a time recorded for each task and this is shown in the table. Times are in minutes.

a Match the tasks to the employees so as to minimise the total time.

b Find the time required to complete all three tasks.

	Task 1	Task 2	Task 3
A	17	12	5
B	6	7	6
C	9	4	8

Solution

a Find the employee that completes each task in the shortest time. Employee *B* has the shortest time.

Task 1 is completed by *B*.

Tasks 2 and 3 remain for *A* and *C*. For the shortest time, Task 2 is best completed by employee *C*.

Task 2 is completed by *C*.

This leaves *A* for Task 3.

Task 3 is completed by *A*.

b Add the times for all three tasks.

A does Task 3 5 minutes
B does Task 1 6 minutes
C does Task 2 4 minutes
Total time = 15 minutes

9780170250283

Row and column reduction

More complex allocation problems cannot be solved by inspection and for these problems we need to use row and column **reduction**.

IMPORTANT

Row and column reduction method

1. Choose the smallest number in each row and subtract it from every element in the same row.
2. For every column that does not have a zero value, choose the smallest number in the column and subtract it from every element in the same column.
3. Cover all the zeros with the smallest number of lines (horizontal or vertical). The process is complete if the number of lines is equal to the number of rows.
4. The allocations are determined by using the zeros. Draw a bipartite graph and choose a connection from each row to a different column that produces the smallest total.

Example 17

A taxi company has 3 taxis *A*, *B* and *C* available and there are three customers 1, 2 and 3 that require a taxi. The distance (km) that each taxi must travel to get to the customers is shown in the table.

	1	2	3
A	26	17	9
B	15	13	17
C	13	16	11

a Perform row and column reductions.
b Find the optimum allocation and represent it as a bipartite graph.
c Find the minimum total distance travelled by the taxis.

Solution

a **Row reductions**
Row 1: Minimum is 9.
Subtract 9 from the numbers in the *A* row to give 17, 8, 0
Row 2: Minimum is 13.
Subtract 13 from the numbers in the *B* row to give 2, 0, 4
Row 3: Minimum is 11.
Subtract 11 from the numbers in the *A* row to give 2, 5, 0

Row reduction

	1	2	3
A	17	8	0
B	2	0	4
C	2	5	0

Column reduction

Column 1: Minimum is 2

Subtract 2 from the numbers in column 1 to give 15, 0, 5

Columns 2 and 3 do not require reduction as they contain zeros.

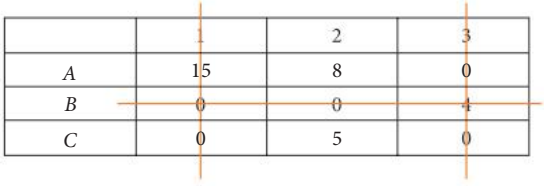
Column reduction

	1	2	3
A	15	8	0
B	0	0	4
C	0	5	0

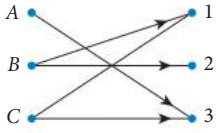
b Cover the zeros with the smallest number of lines

The number of lines is equal to the number of rows so the reduction process is complete.

	1	2	3
A	15	8	0
B	0	0	4
C	0	5	0



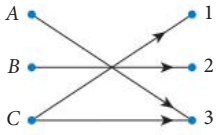
The zeros in the final reduced table indicate the allocation of taxis to customers. Draw a bipartite graph showing all connections.



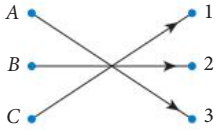
Look at the customers first.

Only Taxi *B* is connected to customer 2.

Eliminate the other connection for Taxi *B*.



This leaves only Taxi C for customer 1 and Taxi *A* for customer 3.



Taxi *A* Customer 3

Taxi *B* Customer 2

Taxi *C* Customer 1

c The distance travelled is obtained by matching the allocations with the distances in the original table.

Distance = 9 + 13 + 13

= 35 km

9780170250283

The Hungarian Algorithm

In cases where a row and column reduction does not result in the same number of lines as there are rows, you use the **Hungarian algorithm** to continue the process.

An algorithm is a set of step-by-step procedures for a calculation.

IMPORTANT

The Hungarian algorithm

1. After performing a row and column reduction and covering the zeros, find the smallest uncovered number and add it to every covered number. Numbers that are covered twice have this number added twice.
2. Subtract the smallest number in the matrix or table from all of the numbers in the table.
3. Cover all the zeros with the smallest number of horizontal or vertical lines and if the number of lines is equal to the number of rows, the process is complete. If this is not the case, repeat steps 1, 2 and 3 until a solution is found.
4. Draw a bipartite graph and use it to determine the **optimum** allocation.

Example 18

Four workers *A*, *B*, *C* and *D* all provide quotes for jobs 1, 2, 3 and 4. Their quotes are shown in the matrix, where the rows represent the workers *A*, *B*, *C* and *D* and the columns represent the jobs 1, 2, 3 and 4.

$$\begin{bmatrix} 15 & 24 & 40 & 21 \\ 24 & 18 & 19 & 20 \\ 28 & 13 & 30 & 22 \\ 10 & 18 & 17 & 13 \end{bmatrix}$$

a Determine the best allocation of workers that minimises the total quote.

b Find the cost of the four jobs.

Solution

a Perform a row reduction.

Row *A* subtract 15

Row *B* subtract 18

Row *C* subtract 13

Row *D* subtract 10

$$\begin{bmatrix} 0 & 9 & 25 & 6 \\ 6 & 0 & 1 & 2 \\ 15 & 0 & 17 & 9 \\ 0 & 8 & 7 & 3 \end{bmatrix}$$

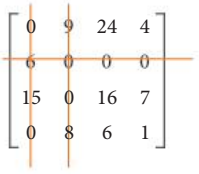
Perform a column reduction.

Column 3 subtract 1

Column 4 subtract 2

$$\begin{bmatrix} 0 & 9 & 24 & 4 \\ 6 & 0 & 0 & 0 \\ 15 & 0 & 16 & 7 \\ 0 & 8 & 6 & 1 \end{bmatrix}$$

Cover the zeros
There are three lines but four rows so the Hungarian algorithm must be used.

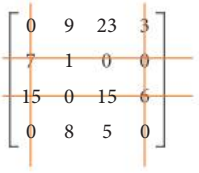


$$\begin{bmatrix} 0 & 9 & 24 & 4 \\ 6 & 0 & 0 & 0 \\ 15 & 0 & 16 & 7 \\ 0 & 8 & 6 & 1 \end{bmatrix}$$

The smallest uncovered number is 1. Add this to every covered number but add 2 to the numbers covered twice.

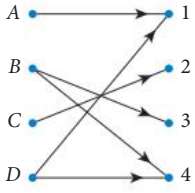
$$\begin{bmatrix} 1 & 10 & 24 & 4 \\ 8 & 2 & 1 & 1 \\ 16 & 1 & 16 & 7 \\ 1 & 9 & 6 & 1 \end{bmatrix}$$

The smallest value is 1, so subtract this from every element in the matrix.
Now cover the zeros. There are now four lines and 4 rows so the allocation is complete.

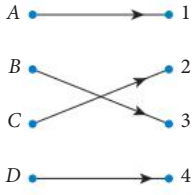


$$\begin{bmatrix} 0 & 9 & 23 & 3 \\ 7 & 1 & 0 & 0 \\ 15 & 0 & 15 & 6 \\ 0 & 8 & 5 & 0 \end{bmatrix}$$

Use the zeros to draw a bipartite graph.



Determine the allocation.
Two jobs have a single worker connected.
Worker *C* for job 2
Worker *B* for job 3
This leaves
Worker *D* for job 4 and
Worker *A* for job 1.



Allocation
*A*1, *B*3, *C*2, *D*4

b Look at the original matrix to determine the cost.
*A*1 = 15, *B*3 = 19, *C*2 = 13, *D*4 = 13

Cost = \$15 + \$19 + \$13 + \$13
= \$60

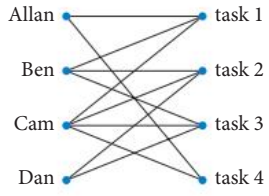
9780170250283

EXERCISE 9.06 Bipartite graphs and the assignment problem

Concepts and techniques

Allocation and the Hungarian algorithm

1 The bipartite graph shows the tasks that each of the four people is able to undertake. If each person must complete one task, a valid allocation is



A

Allan	Task 2
Ben	Task 3
Cam	Task 1
Dan	Task 4

B

Allan	Task 1
Ben	Task 2
Cam	Task 3
Dan	Task 4

C

Allan	Task 1
Ben	Task 3
Cam	Task 4
Dan	Task 2

D

Allan	Task 2
Ben	Task 3
Cam	Task 4
Dan	Task 1

E

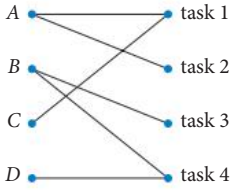
Allan	Task 4
Ben	Task 3
Cam	Task 2
Dan	Task 1

2 Four workers A, B, C and D must complete four tasks 1, 2, 3 and 4. After row and column reduction, the following matrix is obtained.

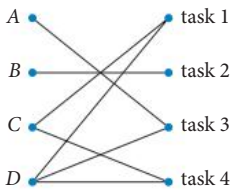
	Task 1	Task 2	Task 3	Task 4
A	2	1	0	4
B	3	0	4	5
C	0	2	9	0
D	0	4	0	0

The bipartite graph for this network is

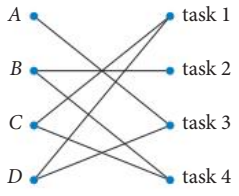
A



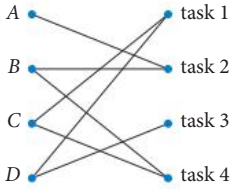
B



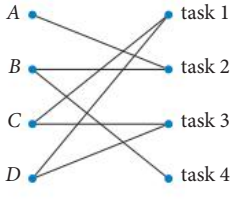
C



D



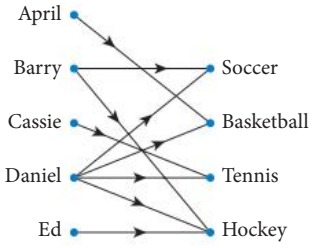
E



3 Example 15 The subjects studied by four students are shown in the table. Represent this using a bipartite graph

Student	Subjects
Amy	Maths, Chemistry, Media
Beth	Physics, Chemistry
Charles	Art, Media
Darren	Maths, Art

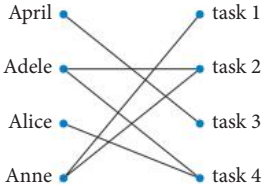
4 The bipartite graph shows the sports played by four students. Indicate if the following statements are true or false.



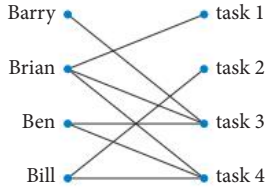
- **a** Only one person plays all four sports.
- **b** April plays soccer.
- **c** Three people play hockey.
- **d** April and Cassie play basketball.
- **e** All the boys play hockey.

5 Determine the allocations of tasks to people from the bipartite graphs below.

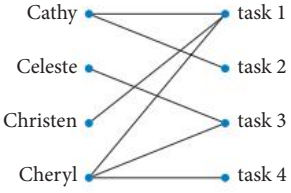
a



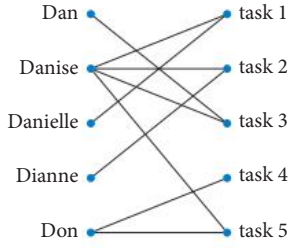
b



c



d



6 Example 17 Perform row and column reductions on the matrices, which represent the cost ($) for workers A, B and C to complete tasks 1, 2 and 3 and determine the optimum allocation and the minimum cost.

a $\begin{bmatrix} 2 & 6 & 9 \\ 2 & 3 & 8 \\ 7 & 9 & 5 \end{bmatrix}$

b $\begin{bmatrix} 4 & 3 & 7 \\ 9 & 4 & 6 \\ 5 & 6 & 7 \end{bmatrix}$

7 Perform row and column reductions on the matrices which represent the times taken (hours) for workers A, B, C and D to complete tasks 1, 2, 3 and 4 and determine the optimum allocation and the minimum time.

a $\begin{bmatrix} 25 & 14 & 7 & 15 \\ 6 & 10 & 8 & 9 \\ 13 & 5 & 9 & 5 \\ 8 & 5 & 10 & 11 \end{bmatrix}$

b $\begin{bmatrix} 8 & 5 & 9 & 3 \\ 3 & 11 & 5 & 7 \\ 10 & 4 & 8 & 2 \\ 4 & 9 & 7 & 12 \end{bmatrix}$

9780170250283

8 A garage has four mechanics Rob, Ben, Claire and Dianne and four tasks1, 2, 3 and 4 that must be performed. The mechanics have been timed on each task and the results in minutes are shown in the table below. Determine the optimum allocation to minimise the time taken.

	Task 1	Task 2	Task 3	Task 4
Rob	8	26	17	11
Ben	13	28	4	26
Claire	38	19	18	15
Dianne	19	26	24	10

9 Example 18 A taxi company has 3 taxis *A*, *B* and *C* available and there are three customers 1, 2 and 3 that require a taxi. The distance (km) that each taxi must travel to get to the customers is shown in the table.

	1	2	3
A	11	19	17
B	21	15	13
C	15	18	21

a Use the Hungarian algorithm find the optimum allocation that minimises the distance travelled.

b Find the total distance travelled.

10 The matrix shown shows the costs (in $1000s) for 4 builders on 4 small projects. Use the Hungarian algorithm to find the optimum allocation and hence determine the minimum cost.

$$\begin{bmatrix} 26 & 19 & 23 & 44 \\ 20 & 27 & 22 & 21 \\ 16 & 31 & 24 & 33 \\ 22 & 13 & 16 & 19 \end{bmatrix}$$

Reasoning and communication

11 A factory has 4 workers Andrew, Michael, Chris and Steph who will each be allocated one of 4 tasks *W*, *X*, *Y* and *Z*. The times that each person takes to complete the tasks (in minutes) is shown in the table below.

	W	*X*	*Y*	*Z*
Andrew	22	18	23	19
Michael	35	23	28	18
Chris	24	23	26	17
Steph	32	23	35	23

a Find the table that results from a row and column reduction.

b Explain why an optimum allocation cannot be made at this stage.

c Use the Hungarian method to produce a table from which an optimum allocation can be made.

d Write the tasks that each worker will complete so that the tasks can be completed in the least amount of time.

12 Stock must be delivered from four warehouses *A*, *B*, *C* and *D* to four supermarkets *W*, *X*, *Y* and *Z*. The distance from each warehouse to each supermarket (in kms) is shown in the matrix below. The rows represent the warehouses and the columns represent the supermarkets.

	W	*X*	*Y*	*Z*
A	20	28	30	26
B	19	29	35	24
C	25	30	28	23
D	20	31	26	25

a Find the table that results from a row and column reduction.
b Explain why an optimum allocation cannot be made at this stage.
c Use the Hungarian method to produce a table from which an optimum allocation can be made.
d Write the optimum method of transporting stock from the warehouses to the supermarkets.
e Determine the minimum total number of kilometres travelled.

13 Darcie, Shannon, Chanelle and Morgan enter their times (in seconds) with their local athletics club for the 100 m, 200 m, 400 m and 800 m track events. The club wishes to select the best competitor for each event. Use the Hungarian algorithm to solve this problem.

	100 m	**200 m**	**400 m**	**800 m**
Darcie	15	34	65	156
Shannon	13	39	67	150
Chanelle	15	30	73	162
Morgan	14	36	66	169

14 The RACV has four service vehicles *A*, *B*, *C* and *D* and receives calls from four motorists *W*, *X*, *Y* and *Z* who have broken down. The table shows the distance from each of the service vehicles to the four motorists. Determine the assignments of service vehicles to motorists that minimises the total distance travelled by using the Hungarian algorithm..

	W	X	Y	Z
A	20	23	20	21
B	22	14	16	18
C	24	30	23	26
D	27	12	24	28

9780170250283

CHAPTER SUMMARY
DIRECTED GRAPHS AND NETWORKS

9

- A directed graph or digraph and is made up of vertices and directed edges.
- In a network, the edges represent the activities and the vertices represent the events.
- Reachability is the ability to get from one vertex to another vertex in a directed graph.
- *A* has one stage dominance over *B* if there is a directed edge from *A* to *B*.
- If *A* dominates *B* and *B* dominates *C*, then *A* has two stage dominance over *C*.
- The dominance value of a particular vertex is given by the total number of one step and two step dominances for that vertex.
- An activity table shows the order and estimated time for each activity.
- An **immediate predecessor** is any activity that must be completed before another activity can commence.
- Steps to follow when drawing your network diagram.
  1. Look for any activities that do not have any predecessor. These will be your starting activities.
  2. Look for activities that have multiple predecessors. These will all end with the same vertex (circle).
  3. Draw activity lines first and do not draw the end vertex until you know what vertex it connects to.
  4. As you look through the activities list, look for other activities that have the same predecessor.
- The critical path is the longest time path in the network.
- On critical paths, the EST and LST values will be equal at each vertex.
- Steps to follow to determine the earliest starting times (EST).
  1. Begin with the first vertex, which has an earliest starting time of 0. The earliest starting times are written in the left-hand section of each vertex.
  2. Work along each path from the start. Add the activity time to the EST at the start of the activity.
  3. Wherever two or more paths join, choose the highest total as the EST.
  4. Continue until the end is reached.
- Steps to follow to determine the latest starting times (LST).
  1. Start at the final vertex. The latest starting time at the final vertex is the same as the earliest starting time.
  2. LSTs are written in the right-hand section of the vertex.
  3. To find the LST at the start vertex, work backwards, subtracting the activity time from the LST at the end vertex.
  4. When there is more than one path to the previous vertex, take the smallest result.
  5. The LST at the first vertex must be zero.
- A dummy edge is added to a network to maintain precedence structure whilst ensuring that no two vertices are connected by multiple edges. A dummy activity has zero time and is shown as a directed edge with a broken line.

- The float time for an activity is the maximum time that an activity can be delayed without affectinging the project completion time.
- float = $LST_{end} - EST_{start}$ – activity time
- It is possible to speed up a project by employing more resources and this enables some activities to be completed more quickly. When this occurs with critical activities, the process is called *crashing*.
- The start of the network is called the source. The end of the network is called the sink.
- Inflow capacity:This is the total flow capacity entering a vertex.

  Outflow capacity: This is the total flow capacity leaving the vertex.
- Maximum flow: This is the smaller of total inflow capacity and total outflow capacity.
- The capacity of a cut can be determined by adding all the flow capacities that pass across the cut in the direction from source to sink.
- Maximum flow = capacity of the minimum cut.
- A bipartite graph has its vertices in two distinct sets and the edges join elements in the first set to elements in the second set.
- Row and column reduction method
  1. Choose the smallest number in each row and subtract it from every element in the same row.
  2. For every column that does not have a zero value, choose the smallest number in the column and subtract it from every element in the same column.
  3. Cover all the zeros with the smallest number of lines (horizontal or vertical). The process is complete if the number of lines is equal to the number of rows.
  4. The allocations are determined by looking at the zeros. Draw a bipartite graph and choose a connection from each row to a different column that produces the smallest total.
- In cases where a row and column reduction does not result in an equal number of lines and rows, you use the Hungarian algorithm to continue the process.
- The Hungarian algorithm
  1. After performing a row and column reduction and covering the zeros, find the smallest uncovered number and add it to every covered number. Numbers that are covered twice have this number added twice.
  2. Subtract the smallest number from all the numbers in the table.
  3. Cover all the zeros and if the number of lines is equal to the number of rows, the process is complete. If this is not the case, repeat steps 1, 2 and 3 until a solution is found.
  4. Draw a bipartite graph and use it to determine the optimum allocation.

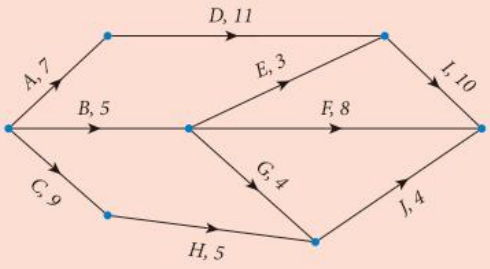
9780170250283

CHAPTER REVIEW
DIRECTED GRAPHS AND NETWORKS

9

Multiple choice

Questions **1**, **2**, **3**, **4** and **5** relate to the directed graph below, where activity times are in days.



1 Which of the statements below is true?
- A Activity *A* is preceded by activity *D*.
- B Activity *B* is preceded by activities *E*, *F* and *G*.
- C Activity *C* is preceded by activity *J*.
- D Activity *J* is preceded by activities *G* and *H*.
- E Activity *A* is preceded by activity *B*.

2 The earliest starting time for activity *J* is

A 5 days B 9 days C 14 days D 24 days E 28 days

3 The critical path is

A *ADI* B *CHJ* C *BF* D *ABEI* E *BGJ*

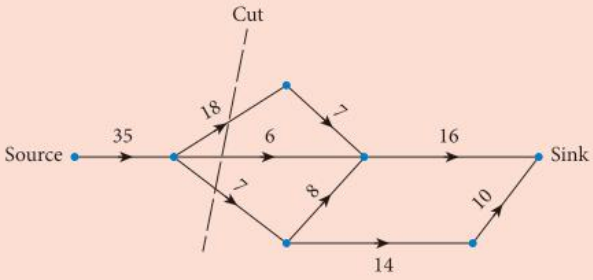
4 The latest starting time for activity *E* is

A 5 days B 8 days C 13 days D 15 days E 20 days

5 The float for activity *F* is

A 0 days B 6 days C 8 days D 10 days E 15 days

Questions **6** and **7** refer to the flow network below, where capacities are in kL / min.



6 The capacity of the cut shown in the flow network is

A 13 kL/min **B** 18 kL/min **C** 30 kL/min
D 31 kL/min **E** 35 kL/min

7 The maximum flow for the network is

A 20 kL/min **B** 26 kL/min **C** 31 kL/min
D 35 kL/min **E** 46 kL/min

8 When a row reduction is completed on the matrix $\begin{bmatrix} 3 & 6 & 9 \\ 5 & 2 & 8 \\ 1 & 9 & 12 \end{bmatrix}$, the second row becomes

A 7, 4, 8 **B** 5, 2, 8 **C** 6, 11, 20 **D** 5, 0, 8 **E** 3, 0, 6

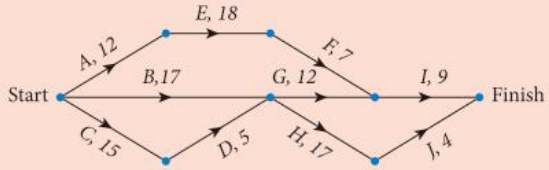
Short answer

9 **a** Draw the network from the activity table below.
b Determine the critical path.

Activity	Time (hours)	Predecessor/s
A	16	none
B	13	none
C	15	*A*
D	14	*B*
E	13	*A*
F	13	*C*, *D*
G	17	*B*
H	15	*G*

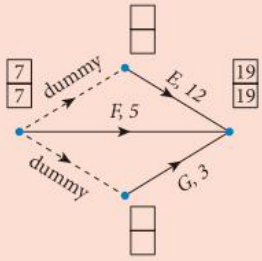
9780170250283

10 In the network below, activity times are in hours.

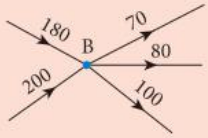


a Forward scan to determine the earliest starting times for each activity.
b Backward scan to determine the latest starting times for each activity.
c Determine the critical path or paths.
d Determine the earliest completion time for the project.
e Determine the float times for all non-critical activities .

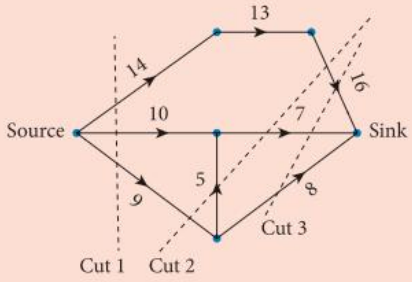
11 Determine the float in days for
a activity *E*
b activity *F*
c activity *G*.



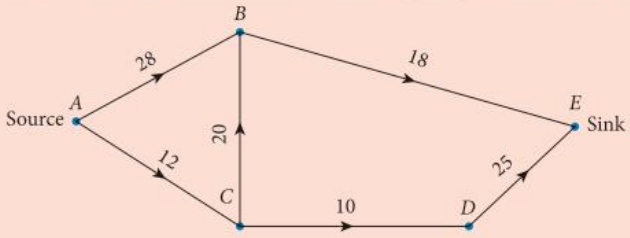
12 For the vertex *B* find
a the inflow capacity
b the outflow capacity
c the maximum flow out of vertex *B*.



13 **a** Determine the capacity of the cuts.
b Find the maximum flow of the network.



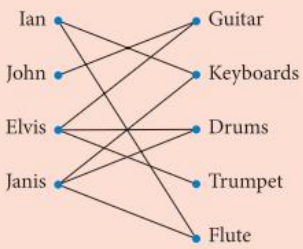
14 For the flow network below, the flow capacity is shown in litres/second. Find



a the maximum number of valid cuts
b the flow capacity of each cut
c the maximum flow from the source at A to the sink at E.

15 The instruments played by 4 people in a band are shown in the bipartite graph on the right.

a Which people play guitar?
b How many people play more than two instruments?
c What instruments does Elvis play?
d Who plays the drums?



16 The time taken in hours for workers A, B, C and D to perform tasks 1, 2, 3 and 4 is given in the table. Determine the optimum allocation that minimises the time taken to complete these tasks.

	1	2	3	4
A	8	9	10	11
B	12	7	9	8
C	7	11	6	8
D	7	6	10	6

17 Four tankers A, B, C and D must deliver petrol to four petrol stations 1, 2, 3 and 4. The distance (km) from each tanker to each petrol station is given in the table.

	1	2	3	4
A	48	36	45	36
B	39	36	37	30
C	41	36	48	31
D	36	31	35	32

a Determine the optimum allocation that minimises the distance travelled.
b Find the total distance travelled by the tankers.

9780170250283

Application

18 Four taxis $\{A, B, C, D\}$ are required to collect 4 passengers $\{W, X, Y, Z\}$. The distance (km) from each taxi to each customer is shown in the matrix, where the rows represent the taxis and the columns represent the customers. Determine the optimum allocation and the total distance travelled.

$$\begin{bmatrix} 29 & 25 & 11 & 10 \\ 35 & 21 & 17 & 20 \\ 23 & 20 & 27 & 19 \\ 33 & 28 & 26 & 14 \end{bmatrix}$$

a Find the table that results from a row and column reduction.
b Explain why an optimum allocation cannot be made at this stage.
c Use the Hungarian method to produce a table from which an optimum allocation can be made.
d Write the optimum method of allocating taxis to customers.
e Determine the minimum total number of kilometres travelled.

19 The network for a project is shown in the activity table below.

Activity	Time (days)	Predecessor/s
A	10	none
B	5	A
C	4	A
D	5	A
E	9	D
F	6	B
G	4	B
H	7	C, E, G
I	3	D
J	5	I
K	10	F, H, J

a Draw the network
b Forward scan to determine the earliest starting times for each activity.
c Backward scan to determine the latest starting times for each activity.
d Determine the critical path or paths.
e Determine the earliest completion time for the project.
f Determine the float times for all non-critical activities.
g Determine the impact on the project if activity B is delayed by 4 days.
h Determine the impact on the project if activity F is delayed by 12 days.

Practice quiz

3 MIXED REVISION
CHAPTERS 7 • 8 • 9

Multiple choice

1 Paul invests $15 000 at 6% p.a. compounding half-yearly. The value of his investment after 3 years is given by

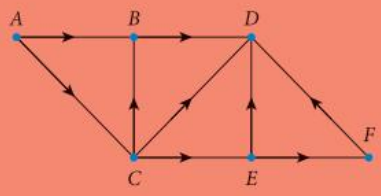
A $15\,000 \times 0.06 \times 3$
B $15\,000 \times 0.03 \times 6$
C $15\,000 \times 0.03^6$
D $15\,000 \times 1.03^6$
E $15\,000 \times 0.06^3$

2 Rowena borrows $150 000 at 9.75% p.a. compound interest, adjusted monthly. The monthly payment required to repay the loan over 20 years is closest to

A $625 B $677.37 C $1422.78
D $1443.27 E $18 423.75

3 In the digraph, it would be impossible to reach vertex F from vertex A if we removed which edge?

A A to B B C to E C B to D
D D to E E E to D



4 Andrew invests $30 000 at 8.5% per annum. At the end of each year, after interest has been paid, he withdraws $1800. No other withdrawals or deposits are made. If A_n represents the value of his investment at the start of year n and $A_1 = \$30\,000$, which one of the following recurrence relations gives the value of his investment at the start of year n?

A $A_{n+1} = 0.085A_n - 1800$
B $A_{n+1} = 1.085A_n - 1800$
C $A_{n+1} = 1.085(A_n - 1800)$
D $A_{n+1} = 0.085(A_n - 1800)$
E $A_{n+1} = 1.085^n A_n - 1800$

5 $140 000 is invested for 1 year. The interest rate is 7.2% p.a. compounding monthly for the first four months and then increases to 7.4% p.a. compounding monthly. The total interest earned for the year is closest to:

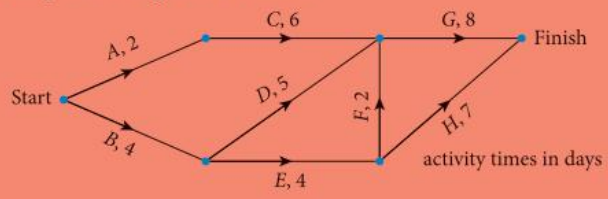
A $10 080.00
B $10 220.00
C $10 360.00
D $10 447.95
E $10 618.86

9780170250283

6 On the day he was born, Peter's parents invested money at 6.1% p.a. simple interest. On his 18th birthday, they gave him the interest on the money as a gift. If they gave him $4200, how much was their initial investment?

A $3825.14
B $4611.60
C $8025.14
D $8811.60
E $12 393.44

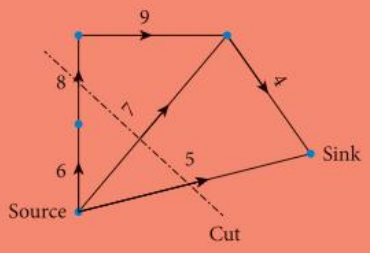
7 In the network below, eight activities must be completed for a project. Seven workers are available for the project, however because of the skills required for a task, activities *C* and *E* must both be completed by one particular worker. The minimum time (in days) required to complete all eight activities is



A 18 B 20 C 21 D 22 E 24

8 In the network below, the values on the edges show the maximum flow possible between the vertices. The direction of the flow is indicated by the arrow on the edge. The capacity of the cut shown is

A 9 B 12 C 13
D 15 E 20



9 Michelle invests $900 000 in a perpetuity to provide her with regular repayments in her retirement. The interest rate is 6.2% per annum and she will receive fortnightly payments. The amount that she will receive each fortnight is closest to

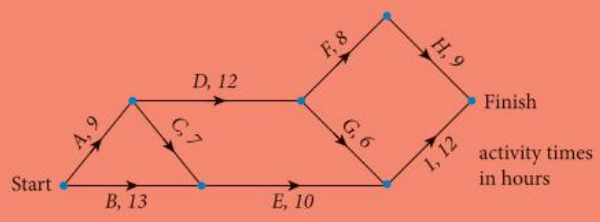
A $1073 B $1163 C $2146 D $2325 E $4650

Short answer

1 Hugh opens a savings account with an initial investment of $10 000 at 6.5% p.a. compounded monthly. At the end of each month, he makes a deposit of $1000. Find

a the amount of his investment after 10 years
b the interest earned.

2 Calculate earliest starting times and latest starting times for the network below.

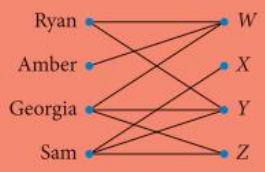


3 How long would it take for $24 000 to earn $6500 interest if it was invested at 7.9% p.a. compounded quarterly?

4 Angela has $500 000 to invest in a perpetuity with interest compounded yearly. She wants this investment to provide annual art scholarships of total value $29 000 for her former school. What annual interest rate is required for this to be achieved?

5 What is the effective interest rate for an investment earning 7.8% p.a. compounded monthly? Answer as a percentage, correct to one decimal place.

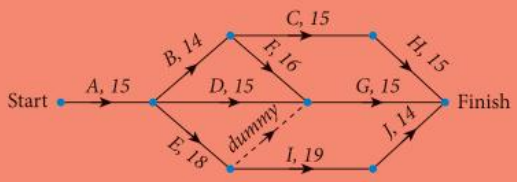
6 Tasks *W*, *X*, *Y* and *Z* are to be completed by Ryan, Amber, Georgia and Sam. Each worker is to be allocated one task. Determine the allocation of workers to tasks indicated by the bipartite graph.



Application

1 Hannah gets a 10-year, $40 000 loan at 8.5% p.a. compounding monthly.
- a What will be the monthly repayments that she makes?
  After 4 years the interest rate drops to 7.75%.
- b If Hannah wanted to keep up her current repayments, how much time would she save on the loan?
- c If Hannah kept the loan term at 10 years, what would be the new repayments?
- d What would be the interest saved if Hannah
  - i kept up her current repayments?
  - ii kept the loan term at 10 years?

2 A commercial kitchen draws a network for the production of their hearty beef pies. The pies require ten separate activities that are shown in the digraph. All activity times are in minutes.



- a Determine the critical path.
- b Find the earliest starting time for activity *H*.
- c Find the latest starting time for activity *C*.
- d What activities must be completed before activity *G* can commence?
- e Extra resources are employed so that activities *D*, *E* and *I* can be completed in 5 less minutes. Determine the new project completion time.

3 Nicole bought a new kitchen valued at $15 500 on terms of 15% deposit and 8.2% p.a. on the balance over 2 years.
- a Calculate her deposit.
- b What amount did she have to pay interest on?
- c How much interest did she have to pay?
- d What was the total amount she had to repay?
- e If she paid it in equal fortnightly instalments, what was the amount of each instalment? Answer correct to the nearest cent.

9780170250283

4 Desi receives an inheritance of $70 000. She invests the money at 5% per annum. Interest is paid at the end of each year and after the interest is paid, Desi withdraws $6000. The amount remaining gets invested for another year.

a Find a recurrence relation of the form $A_{n+1} = rA_n + d$, where $A_1 = a$, that gives the value of the investment, A_n, at the start of the nth year.

b What is the value of the investment at the start of the 10th year?

c What amount could be withdrawn at the end of each year so that the value of the investment remains at $70 000?

5 Louis was comparing accounts for investing $10 000. He had narrowed it down to two different options.

Account 1: 6.5% p.a. simple interest

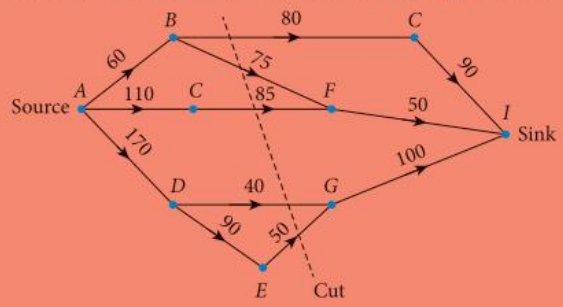
Account 2: 6.2% p.a. compounding yearly

a Draw a graph of the value ($A) that his investment will grow to in n years if invested in Account 1. Use n from 0 to 10 years.

b On the same axes, graph the value ($A) that his investment will grow to in n years if it is invested in Account 2. Use n from 0 to 10 years.

c What advice would you give Louis? Explain.

6 A small airline runs flights connecting cities A, B, C, D, E, F, G, H and I as shown in the network. The maximum number of passengers for each flight is shown on each edge.



a Find the capacity of the cut.

b Find the maximum number of passengers that can be transported from A to I.

c The flight from G to I must use a smaller plane. What is the maximum number of passengers it must carry for the maximum capacity of the network to be unaffected?

d A company books 60 of their employees on the flight from A to D. Find the maximum number of these people that will get to I without leaving the group.

ANSWERS

1.01

1 a Ordinal requires an order to be present (e.g. finishing position in a race), nominal does not, (e.g. favourite pizza topping).
b Discrete represents whole numbers only (e.g. number of children), continuous does not (e.g. height).

2 a categorical and nominal
b numerical and discrete
c numerical and continuous
d categorical and nominal
e categorical and ordinal
f numerical and continuous
g numerical and continuous
h categorical and ordinal
i numerical and discrete

3 C

4 a Not appropriate as there is a large range of answers, provide groups of numbers to select from 0–9 runs, 10–19 runs, etc.
b Appropriate, as there is a small range of answers.
c Not appropriate, as there are a large number of responses; provide a list to select from or change to favourite type of movie.
d Not appropriate, as there is a large range of possible answers; provide a list to select from.
e Not appropriate, as the first sentence could affect a person's answer; remove the first sentence.

5 a B b A c A
d C e C

6 B

7 Answers will vary.

8 Answers will vary.

9 a Yearly income: numerical continuous; age: numerical discrete; educational level: categorical ordinal; hours worked per week: numerical continuous and gender: categorical nominal.
b too large
c To reduce the amount to investigate and to eliminate bias; randomly assign people numbers and use a random number generator to select numbers.
d Answers will vary, calculator step *randint(1,10000,100)* or *randList(100, 1, 10000)*

1.02

1 a $425.56, $400, $300
b 172.375 kg, 171.5 kg, no mode
c 22.8, 23, 24
d 120.93 secs, 121 secs, 123 and 137 secs

2 a $450, $310, $165.92
b 55 kg, 39.5 kg, 21.48 kg
c 6, 3, 1.80
d 60 secs, 28 secs, 16.67 secs

3 a

Mean	23.57
Median	23.5
Mode	27
Range	8
IQR	5
Standard deviation	2.71

b

Mean	7.89
Median	6
Mode	6
Range	24
IQR	3
Standard deviation	6.38

c

Mean	7.25
Median	7
Mode	7
Range	9
IQR	3.5
Standard deviation	2.4

d

Mean	26.29
Median	26
Mode	26
Range	10
IQR	3
Standard deviation	2.42

e

Mean	18.03
Median class	15–19
Modal class	15–19
Range	24
IQR	10
Standard deviation	5.73

f

Mean	277.71
Median	200–299
Modal class	200–299
Range	599
IQR	200
Standard deviation	125.55

9780170250283

4 a The cost of two concert tickets has a centre around $400. The mean is slightly higher than the median, and there are reasonably large measures of spread, so the costs are quite varied. As the middle 50% covers a spread of $300, this indicates that there is unlikely to be any outliers.

b The masses of large rocks have an average of around 170 kg, with the mean and median very similar. The masses vary somewhat with a range of 55 kg and the middle 50% covering a bit over 21 kg.

c The class sizes for grade 1 at a primary school have an average of around 23 students with the mean and median almost identical, indicating that there are no outliers. The spread is quite small, only a range of 6 and the middle 50% of the data is within a spread of 3.

d The time in seconds of competitors in a go-kart race have an average of around 121 seconds with the mean and median almost identical, indicating that there are no outliers. The times are quite varied, with a range of one minute, although the middle 50% covers a spread of almost 30 seconds.

5 E

6 D

7 a 24.5 min b 25 min

c 2 min d 1.8 min

e Any, as there are no outliers and the data is not skewed

8 a mean = 167.78 cm, median = 168 cm, mode = 175 cm; range = 27 cm, IQR = 15 cm, s = 8.21 cm

b The mean and median are very similar, which indicate that there are no outliers within the data set. 50% of the heights lie below 168 cm and 50% lie above it, the most common height was 175cm; although this is different to the other two measures of centre, the mode is a less accurate measure. The data set is reasonably varied, having a range of 27 cm separating the shortest and tallest students. The middle 50% of the class covers a spread of 15 cm. Knowing there are no outliers, the standard deviation of 8.21 cm shows that the data set is reasonably varied.

9 a 15 b Holden

c Mean = $14 333, median = $13 000

d The mean is $1333 larger, there is an outlier of $40 000, which raises the mean.

e Median due to presence of outlier

10 a

Class interval	Frequency
150 – <300	4
300 – <450	9
450 – <600	9
600 – <750	9
750 – <900	11
900 – <1050	5
1050 – <1200	0
1200 – <1350	2
4950 – <5100	1

b mean = $722.18, median = $644.50, modes = $346 and $579.

c range = $4882, IQR = $414, s = $672.74.

d The data is best described as having a centre of around $650; the middle 50% lies within a spread of $414. It is best to use the median and IQR due to the presence of an outlier.

11 Mean = 1.72, median = 2, mode = 2, range = 19, IQR = 1, s = 2.08

The average household can be said to own two pets, give or take a pet. (It is best to use the median and IQR to describe these results, due to the presence of an outlier.)

12

Class interval	Frequency
100 – <150	1
150 – <200	2
200 – <250	8
250 – <300	14
300 – <350	14
350 – <400	7
400 – <450	4

Mean = 302.04, median = 303, modes = 202, 267, 278, 312, 336, 346 and 384; range = 290, IQR = 90, s = 69.01

The average time taken to complete a marathon is just over 300 minutes (302 or 303 minutes). There was a large variation in the finish times (290 minutes), and the middle 50% finished within a spread of 90 minutes. Although the times are quite varied, the large IQR and the similarity between the mean and median indicate that there are no outliers.

13 a black and red are 10

b Blonde

c The data is categorical and the mean and median can only be calculated for numerical data.

1.03

1 a

Stem	Leaf
4	3 5 9
5	0 2 7 8
6	1 2 4 5 7 8
7	0 2 3 9
8	2 4 9

Key: 4|5 = 45

b 20 matches **c** 89 points **d** 25%

2 a

Stem	Leaf
16	2 3 4 4
16*	5 5 5 8 8
17	0 2 3 3 4
17*	5 6 6 8 8
18	0 2
18*	5 5 6 7 9 9

Key: 16|2 = 162

b 12 applicants **c** 33.3%

3 a i

Minimum	Q_1	Median	Q_3	Maximum
1	3	5	8	11

ii 3 to 8

b i

Minimum	Q_1	Median	Q_3	Maximum
7	10	12	13	16

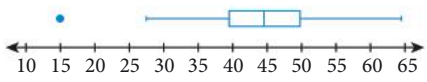
ii 10 to 13

c i

Minimum	Q_1	Median	Q_3	Maximum
1	2	6	8	10

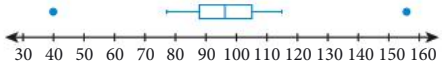
ii 2 to 8

4 a



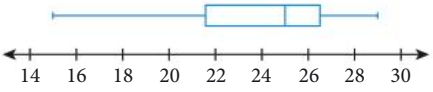
15 is an outlier as it is outside the range $22 \leq x \leq 66$

b



40 and 156 are outliers as they are outside the range $62.5 \leq x \leq 130.5$

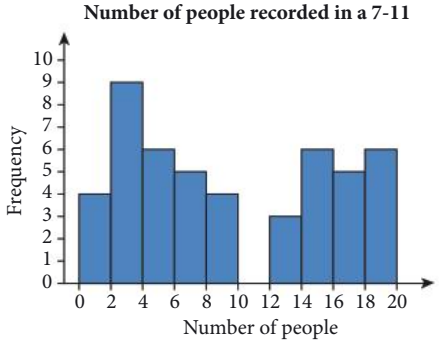
c



5 C

6 B

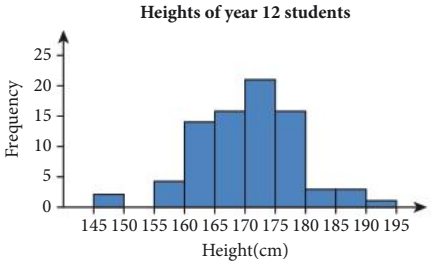
7 a



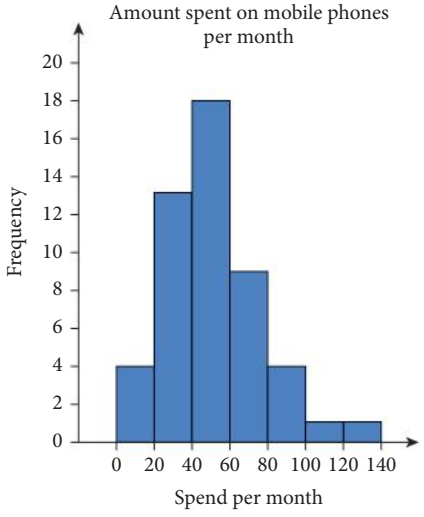
b



c

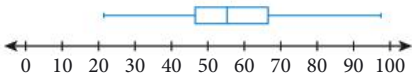


8 a



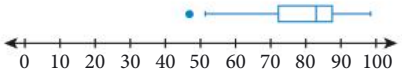
b 30% **c** 62% **d** $\$30 \leq x \leq \$69$

9 a i



ii Approximately symmetrical **iii** 67

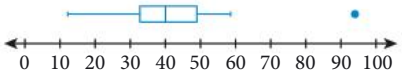
b i



ii Negatively skewed with an outlier

iii 87

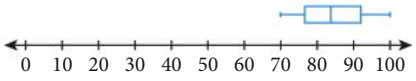
c i



ii Approximately symmetrical with an outlier

iii 48

d i



ii Approximately symmetrical **iii** 92

9780170250283

10 B

11 a 26 students b positively skewed

c

Score	Frequency
5	8
6	6
7	5
8	4
9	2
10	1
Total	26

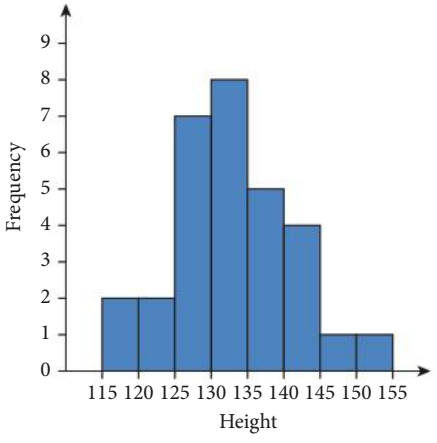
d 11.5%

12 a Numerical, discrete.

b

Height	Frequency
115 – <120	2
120 – <125	2
125 – <130	7
130 – <135	8
135 – <140	5
140 – <145	4
145 – <150	1
150 – <155	1
Total	30

c



d 116 119 121 122 125 125 126
127 127 127 128 130 131 131
131 132 132 132 134 135 135
136 138 139 140 141 143 147
147 154

e

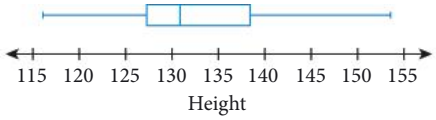
Stem	Leaf
11	6 9
12	1 2 5 5 6 7 7 7 8
13	0 1 1 1 2 2 2 4 5 5 6 8 9
14	0 1 3 7 7
15	4

Key: 11|9 = 119

f Approximately symmetrical.

g 30%

h



There are no outliers as all values are within the range $110.5 \leq x \leq 154.5$.

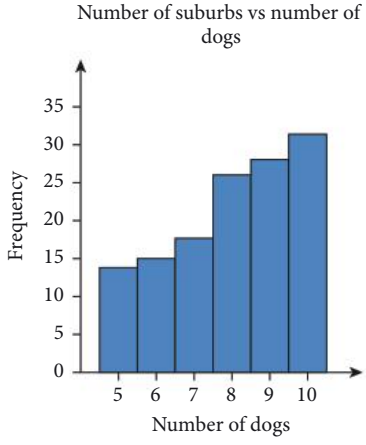
i 116 cm to 127 cm, so the range is 11 cm.

j Mean = 132.37 cm, median = 131.5 cm, modes = 127 cm, 131 cm and 132 cm, range = 38 cm, IQR = 11 cm, $s = 8.70$.

k The average height in this class of 30 is around 131–132 cm. The mean and the median are fairly similar, which indicates that there are no outliers. This is supported by the boxplot, which shows an approximately symmetrical distribution with no outliers. The spread of data is reasonably large (38 cm), though the middle 50% of the class ranges from 127 cm to 138 cm.

13 a 131

b

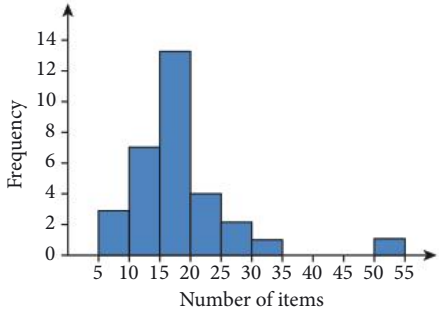


c negatively skewed d 8 dogs e 1.65

f No, as the data is skewed. Median and IQR should be used to describe the data.

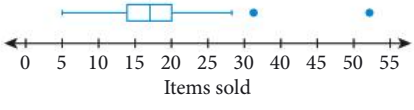
14 a Mean = 18.03, median = 17, mode = 19, range = 47, IQR = 6, $s = 8.38$

b



c The values 31 and 52 are outliers as they fall outside the range of $5 \leq x \leq 29$.

d



e Approximately symmetrical with two outliers.
f 25%
g 159
h It seems that on average, 17 or 18 items are sold per day at this eBay store. It has a large range (47 items) but a small IQR (6), which indicates the presence of an outlier. If we look at the displays, the bulk of the data is approximately symmetrical, although there are two outliers (one in particular is very far outside the IQR, hence the large range). The worst sales day (5 items) was within the IQR, which suggests that this is more likely to occur again than the maximum (52).

1.04

1 a RV = amount of rainfall (mm), EV = temperature (°C)
b RV = arm length, EV = height
c RV = distance covered (m), EV = time (mins)
d RV = weight loss (kg), EV = days spent dieting
e RV = wage ($), EV = hours of work
f RV = yield of vegetables (kg), EV = size of a vegetable patch (m^2)
g RV = height (cm), EV = age (years)
h RV = weekly budget ($), EV = weekly wage ($)
i RV = heart rate, EV = caffeine consumption
j RV = number of skiers, EV = amount of snow

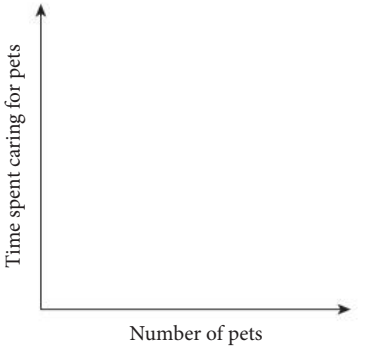
2 Because in most circumstances, nothing can cause a change in time.

3 a Height = EV, Basketball ability = RV
b No association
c Size of house = EV, time spent cleaning per week = RV
d Number of people in a house = EV, number of rooms in a house = RV
e No association
f Smoking status = EV and health = RV

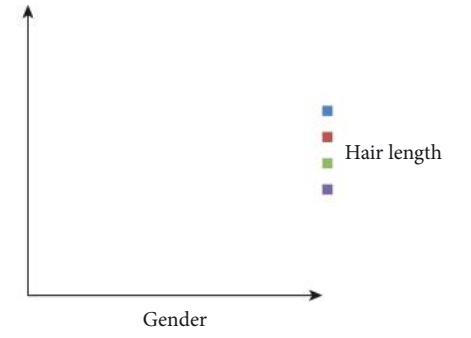
4 D

5 a Yes b No c No
d Yes e Yes f No

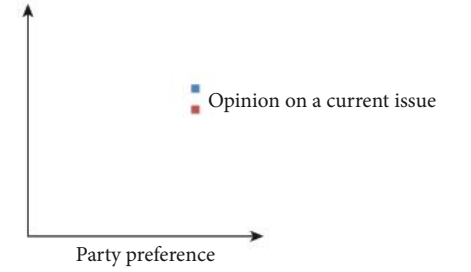
6 a No association
b



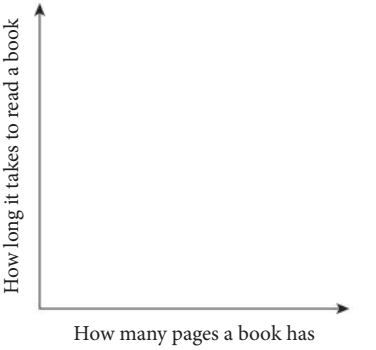
c



d



e No association
f



7 E

1.05

1 Because a segmented column graph gives you the variables in percentages, allowing for a more accurate comparison, rather than just straight numbers.

2 a EV = Gender, RV = Library member
b

	Gender		
Library membership	Male	Female	
Library member	20	32	52
Not a library member	8	15	23
	28	47	75

c 23 d 47 e 38.5% f 31.9%

3

	Gender		
Suncreen use	Men	Women	
Use sunscreen	35	30	65
Do not use sunscreen	5	10	15
	40	40	80

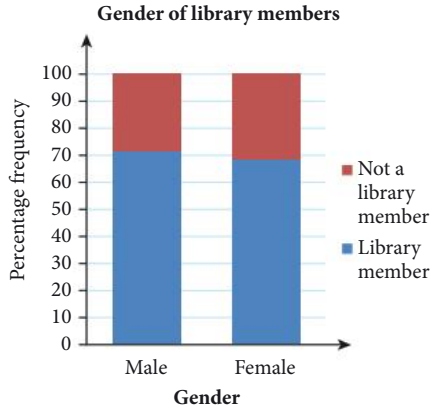
9780170250283

4 E

5 a

	Gender	
Library membership	Male	Female
Library member	71%	68%
Not a library member	29%	32%
	100%	100%

b

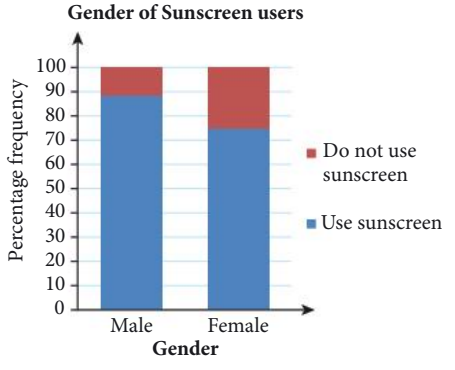


c The graph shows that there are slightly more male library members than females. As the column splits are nearly the same, library membership is probably not associated with gender.

6 a

	Gender	
Suncreen use	Men	Women
Use sunscreen	88%	75%
Do not use sunscreen	12%	25%
	100%	100%

b



c The graph shows that a higher percentage of men use sunscreen than women, therefore the use of sunscreen may be associated with gender.

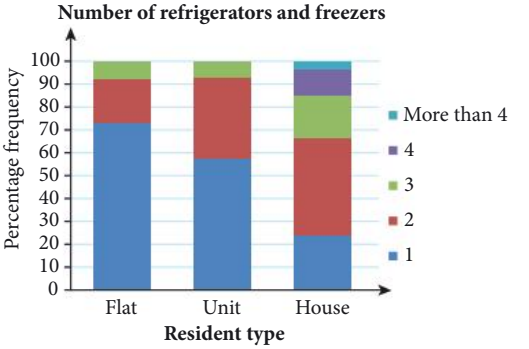
7 a

	Type of residence			
Number of refrigerators/freezers	Flat	Unit	House	Total
1	40	35	40	115
2	10	20	70	100
3	5	5	30	40
4	0	0	20	20
More than 4	0	0	5	5
Total	55	60	165	280

b

	Type of residence		
Number of refrigerators/freezers	Flat	Unit	House
1	72.7%	58.3%	24.2%
2	18.2%	33.3%	42.4%
3	9.1%	8.3%	18.2%
4	0	0	12.1%
More than 4	0	0	3%
Total	100%	100%	100%

*Note that rounding may mean some totals do not appear to add to 100%.



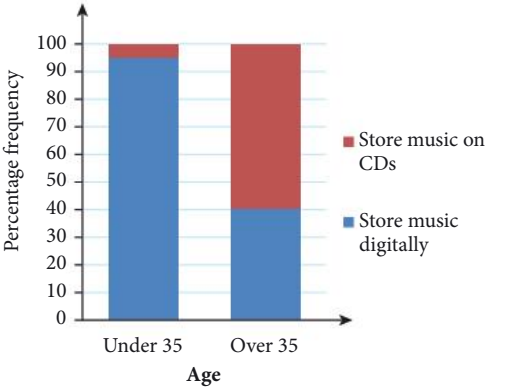
c People living in houses seem to own more refrigerators and freezers than those who live in flats and units. This graphical display suggests that the number of refrigerators and freezers may be associated with the type of residence you live in.

8 a

	Under 35	Over 35	
Store music digitally	38	14	52
Store music on CDs	2	21	23
	40	35	75

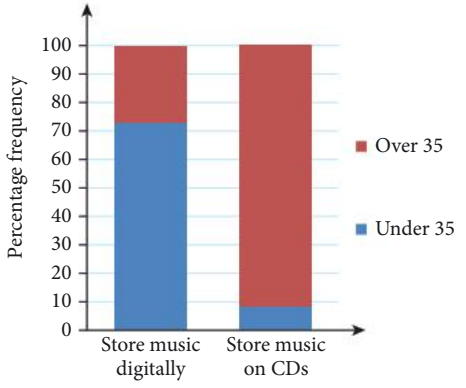
b

	Under 35	Over 35
Store music digitally	95%	40%
Store music on CDs	5%	60%
	100%	100%



c

	Under 35	Over 35	
Store music digitally	73.1%	26.9%	100%
Store music on CDs	8.7%	91.3%	100%



d The segmented column graph in part **b** looks at the how people store their music according to their age (e.g. what percentage of those over 35 choose to store music digitally); whereas the segmented column graph in part **c** looks at the age of people according to how they store their music (e.g. what percentage of those who store music digitally are over 35).

e i 5% ii 40% iii 91.3% iv 26.9%

f Part **ii** looks at the percentage of digital users out of the total of those over 35, whereas part **iv** looks at the percentage of those over 35 out of the total of digital users.

9

	City	Country	
Manual licence	7	36	43
Automatic licence	39	13	52
	46	49	95

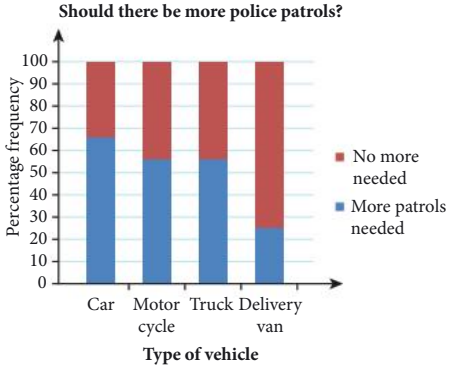
a 26.5% b 15.2% c 75% d 83.7%

10 a

	Type of vehicle				
Opinion	Car	Motorcycle	Truck	Delivery van	Total
More patrols needed	30	8	20	6	64
No more needed	15	6	15	18	54
Total	45	14	35	24	118

b

	Type of vehicle			
Opinion	Car	Motorcycle	Truck	Delivery van
More patrols needed	67%	57%	57%	25%
No more needed	33%	43%	43%	75%
Total	100%	100%	100%	100%



c In general, those people who drive cars have the highest percentage that think more patrols are needed with motorcycles and trucks only slightly behind, whereas only a small percentage of delivery van drivers agree. Due to the differences, it appears that a person's opinion on police patrols may be associated with the type of vehicle they drive.

11 a Type of death (RV) and smoking status with (EV)

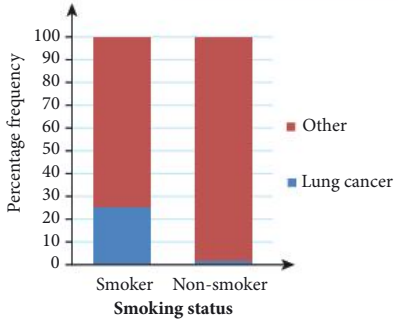
b

	Smoking status		
Type of death	Smoker	Non-smoker	
Lung cancer	630	85	715
Other	1870	7415	9285
	2500	7500	10 000

c

	Smoking status	
Type of death	Smoker	Non-smoker
Lung cancer	25%	1%
Other	75%	99%
	100%	100%

d



9780170250283

e The graph clearly shows that if you are a smoker, then you have a greater chance of dying from lung cancer. It also shows that fewer non-smokers die of lung cancer than other types of death. In general, the way you die may be associated with your smoking status.

1.06

1 a

Camera 1	Stem	Camera 2
9 9 9 7 5 3	6	
8 8 6 5 4 4 3 3 2 1 1	7	
9 3 3 3 0	8	
0	9	
2	10	
	11	2 6 6 7 9
	12	0 0 3 3 5 7
	13	0 0 0 1 1 4 5 6 8 9
	14	0 8
	15	
	16	9

Key: 6|3 = 63

b

	Mean	Standard deviation
Camera 1	76.13	8.90
Camera 2	129.54	12.34

c Camera 2 has recorded a much higher average speed than Camera 1. Camera 2 has a higher standard deviation, indicating that the speeds recorded have a greater spread about the mean.

2 a

	Range	IQR	Median
Test A	90	58.5	56.5
Test B	91	18	53

b

	Mean	Standard deviation
Test A	50.5	33.3
Test B	50.6	25.4

c Test B due to the smaller IQR and standard deviation.

3 a

Before	Stem	6 weeks later
9	0	6 8
8 6 4 0	1	3 6 6 8 9
1 1	2	1 4 8
6 4 2	3	1 2
5 2	4	

Key: 1|3 = 13

b Before: 20, 6 weeks later: 11.5

c Before: no specific shape, 6 weeks later: approximately symmetrical

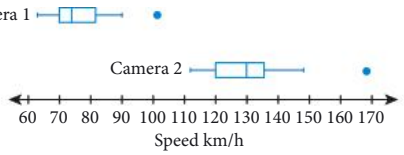
d

	Mean	Standard deviation
Before	25	12
6 weeks later	19	8

e The mean is lower for 6 weeks after the program and the standard deviation is smaller for 6 weeks after the program.

f Yes, as the average number of cigarettes has decreased by 5 and the standard deviation is smaller, indicating that the data is less variable.

4 a

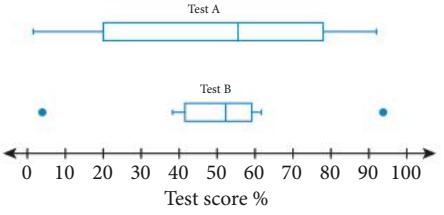


b Camera 1: 63 to 70 km/h, Camera 2: 112 to 120 km/h

c Camera 2

d Yes, there are significant differences in the box plots. Camera 2 appears to be positioned in a higher speed limit zone.

5 a

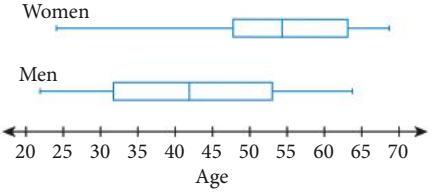


b Test B c Test B

d 3 and 94 are outliers as they are outside the range $14.5 \leq x \leq 86.5$

e Yes, there are significant differences in the boxplots. Test B appears to have been easier as most marks are over a much smaller range than in Test A.

6 a



b 42

c Men's data is approximately symmetrical, whereas the women's data looks more negatively skewed.

d Men, due to lower median and smaller range. Their mean age is also approximately 10 years less than women.

e There appears to be an association between gender and yoga class attendance. In general, the women attending tend to be older than the men attending.

7 D

8 C

9 a i 12 ii 26

b The estimated lengths has a far greater range than the actual lengths.

c The estimated lengths has a far greater IQR than the actual lengths.

d Estimates: Positively skewed; Actual: Approximately symmetrical

e Agree as estimates has a far greater range and lower median although some people did also over estimate.

10 a i Darwin ii Brisbane
iii Darwin iv Canberra

b Yes, even though the range is the same the median price in Sydney is much lower

c There appears to be an association between city and petrol costs as there are differences in each state's boxplot, in particular the median value is different for every state.

11 a

Melbourne	Stem	Sydney
	0	
	0	
	0	
	0	7
	0	8 9
1 0 0 0	1	1
3 2 2 2 2 2 2 2	1	2 2
	1	4 4 4 5 5
	1	6
	1	

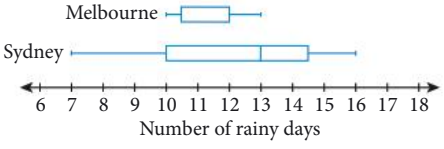
Key: 0|7 = 7

b Melbourne: Approximately symmetrical (though all values are very close together so this is a loose description); Sydney: slightly negatively skewed. *Note that the descriptions depend on how many leaves you choose to use for the plot.*

c

	Melbourne	Sydney
Minimum	10	7
Q_1	10.5	10
Median	12	13
Q_3	12	14.5
Maximum	13	16

d

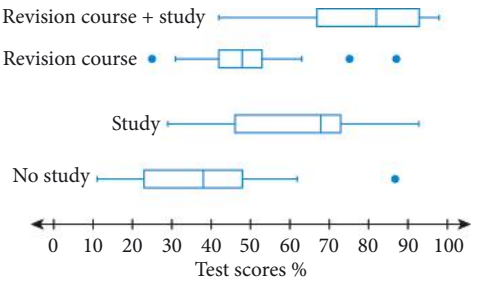


e Sydney has a wider range of rainy days.

f There appears to be an association between the city and the number of rainy days per month as Sydney covers a larger range and has a higher median number of rainy days.

12 a 25% b 5 people c After exercise
d 68 to 90 e 22
f 133 and 140 are outliers as they are outside the range $35 \leq x \leq 123$
g 6 beats/min

13



The boxplots appear to show a difference in results depending on what students did before a test. They were more likely to perform well if they studied, especially if they also attended a revision course. Attending a revision course was better than nothing at all, but studying provided better results (an IQR of 46 – 73 for studying compared to an IQR of 42 to 53 for attending a revision course and 23 – 48 for not studying at all). Although overall ranges were reasonably similar (in size, not in spread), the only students that performed well without studying were classified as outliers. These results suggest that studying appears to improve results.

1.07

1 a positive b negative c no relationship
d positive e negative f no relationship
g positive h positive

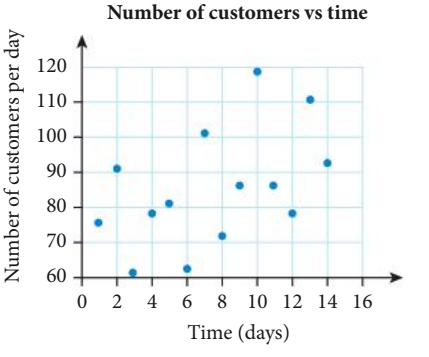
2 a Linear, negative and strong.
b No relationship.
c Linear, negative and weak.
d Linear, positive and moderate.

3 E

4 C

5 D

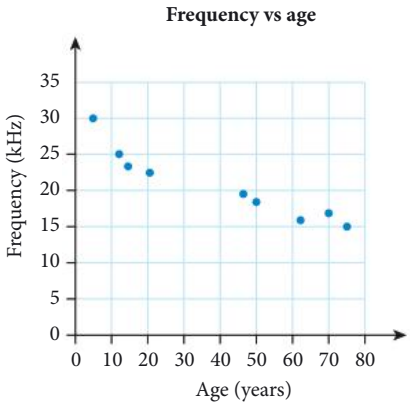
6 a i



ii The relationship between the time and number of customers per day can be described as a weak positive relationship.

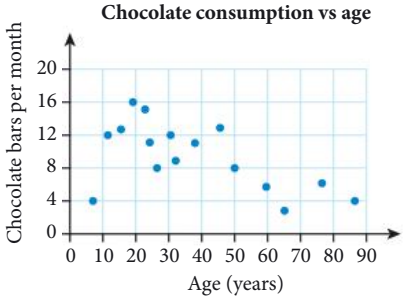
9780170250283

b i



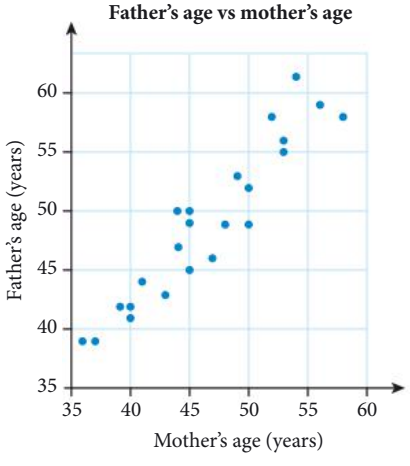
ii The relationship between age and frequency can be described as linear, negative and strong.

c i



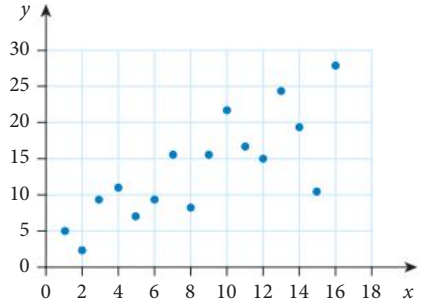
ii The relationship between age and chocolate consumption can be described as linear, negative and moderate.

d i



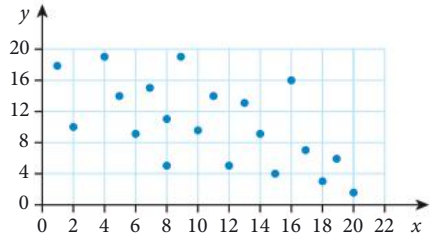
ii The relationship between mother's age and father's age can be described as linear, positive and strong.

7 a



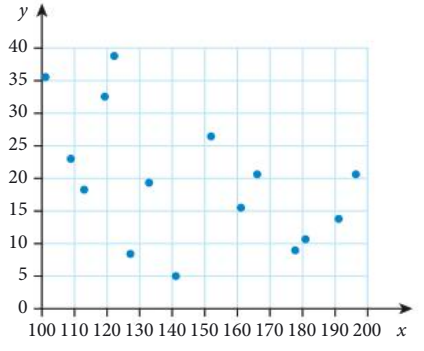
Moderate, positive association

b



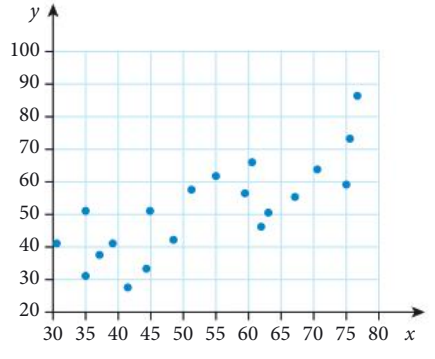
Weak, negative association

c



Weak, negative association

d



Moderate, positive association

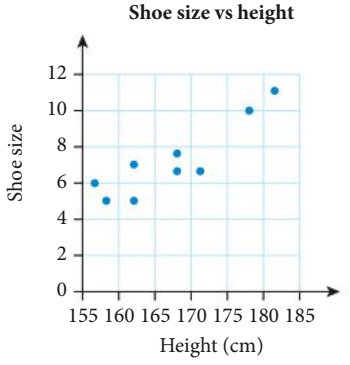
1.08

1 a –0.4 b 0.8 c 0.6
d –0.9 e –0.5 f 0.2

2 a moderate and positive b no relationship
c moderate and negative d weak and positive
e strong and positive f strong and negative

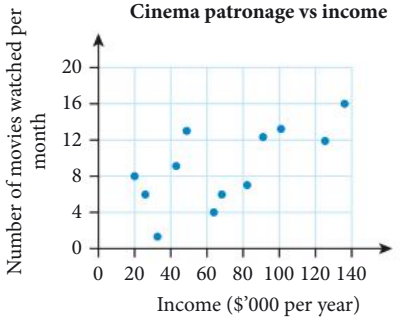
3 a A correlation coefficient of 0.2946 indicates a very weak positive relationship. There is limited evidence that the amount of exercise that a person does increases as their height increases.
b A correlation coefficient of 0.5283 indicates a somewhat moderate positive relationship. There is some evidence to support that a person's weight should increase as the number of hours a person spends sitting down increases.
c A correlation coefficient of –0.9108 indicates a strong negative relationship. It can be concluded that the percentage of good peaches should decrease as the time spent in storage increases.
d A correlation coefficient of 0.7241 indicates a moderate positive relationship. There is some evidence to support that sales at a local store should increase as temperature increases.
e A correlation coefficient of –0.0628 indicates no relationship between the number of pets owned and temperature.
f A correlation coefficient of –0.4637 indicates a weak negative relationship. There is limited evidence to suggest that the number of home cooked meals will decrease as income increases.

4 a



b $r = 0.9218$
c An r value of 0.9218 indicates that there is a very strong, positive linear relationship between height and shoe size. It can be concluded that shoe size should increase as height increases.

5 a



b 0.6484
c An r value of 0.6484 indicates that there is a moderate, positive linear relationship between height and shoe size. There is some evidence to support that cinema patronage should increase as monthly income increases.

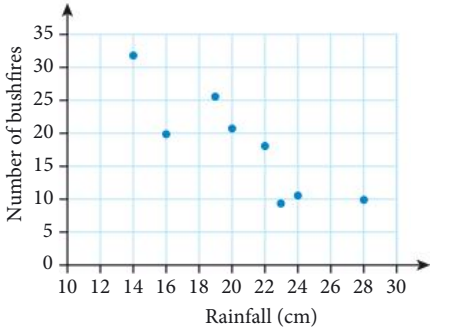
6 a –0.7325
b An r value of –0.7325 indicates that there is a moderate, negative linear relationship between shoe size and height. There is some evidence to support that hot dog sales should decrease as temperature increases.

7 $r = 0.4464$. This indicates that there is a weak, positive relationship between monthly rainfall and monthly electricity bill. There is limited evidence to suggest that the monthly electricity bill will increase as monthly rainfall increases.

8 B

9 D

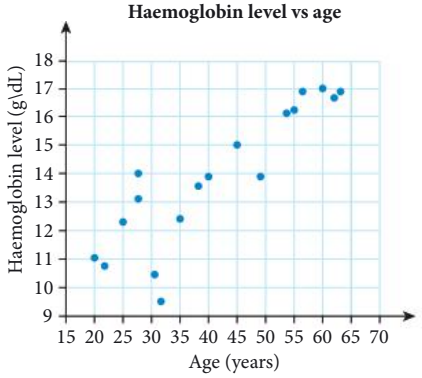
10 a



b The relationship appears to be linear, negative and strong.
c $r = –0.86$
d When $r = –0.86$, it indicates a strong negative relationship, therefore it can be concluded that the number of bushfires should decrease as the amount of rainfall increases.

9780170250283

11 a



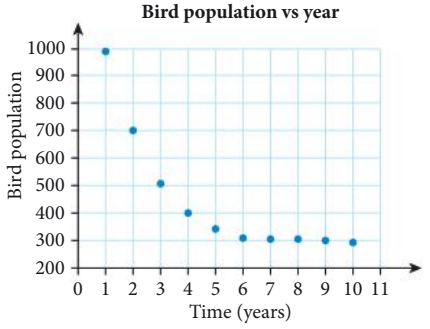
b Linear, positive and strong relationship.

c $r = 0.88$

d When $r = 0.88$, it indicates a strong positive relationship, therefore it can be concluded that haemoglobin levels should increase as age increases.

e The data supports the theory as there is evidence to suggest that haemoglobin level increase as age increases.

12 a

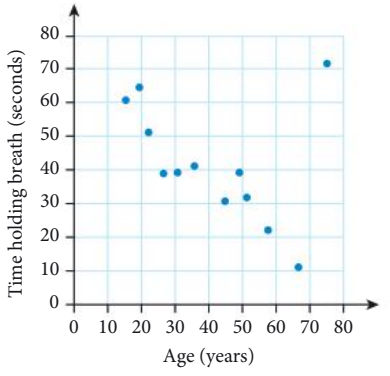


$r = -0.8318$

b When $r = -0.86$ it indicates a reasonably strong negative relationship.

c This data is very clearly non-linear, and therefore the r value is not useful in describing the data. Using a different method, the data should describe a very strong, non-linear relationship as is shown in the scatterplot.

13 a



$r = -0.3969$

b The r value indicates that the data shows a weak, negative relationship. There is limited evidence to suggest that the length of time a person can hold their breath for should decrease as age increases.

c The data set has an outlier.

d $r = -0.9552$. The r value is much higher, indicative of a strong, negative relationship. This allows us to conclude that the length of time a person can hold their breath for should decrease as age increases.

INVESTIGATION: HOW CAN YOU PREDICT A PERSON'S HEIGHT?

Teacher to check.

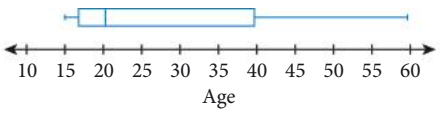
CHAPTER 1 REVISION

1 D
2 D
3 C
4 E
5 C
6 D
7 B
8 E
9 A
10 B
11 E
12 a Numerical, continuous
b Numerical, discrete
c Numerical, continuous
d Categorical, nominal
13 a i mean = 4.78, median = 5, mode = 5
ii range = 9, IQR = 5, standard deviation = 2.91
b i mean = 15.08, median = 13, modes = 6, 17, 18
ii range = 71, IQR = 11.5, standard deviation = 13.49
c i mean = 11.30, median = 11, mode = 10
ii range = 3, IQR = 3, standard deviation = 1.2
d i mean = 4, median = 4, no mode
ii range = 4, IQR = 2, standard deviation = 1.49
e i mean = 5.04, median = 5.4, mode = 5.5
ii range = 4.2, IQR = 2.8, standard deviation = 1.52
14 a 24.5 mins b 25 mins c 2 mins
d 1.8 mins e no outliers in data, so all of them
15 a negatively skewed
b positively skewed
c negatively skewed with an outlier
16 a

Stem	Leaf
1	5 5 5 6 6 7 8 9 9
2	0 1 2 2
3	5 6
4	3 9
5	0 5 9

Key: 1|5 = 15

b



c If all values are within the range $0 \le x \le 74$, then there are no outliers.

d positively skewed

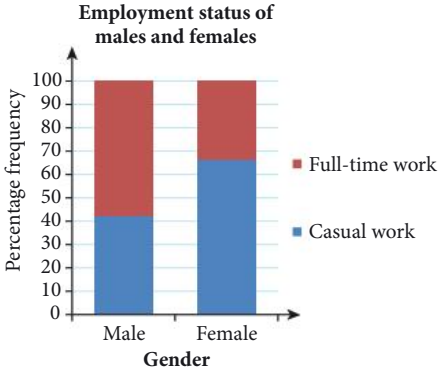
17 a

	Gender		
Employment status	Male	Female	
Casual work	438	542	980
Full-time work	602	269	871
	1040	811	1851

b

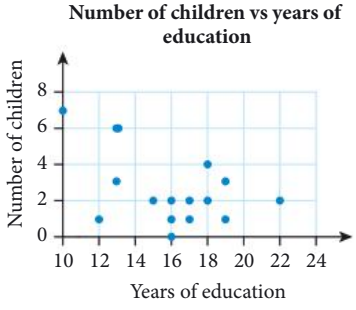
	Gender	
Employment status	Male	Female
Casual work	42%	67%
Full-time work	58%	33%
	100%	100%

c



d From the graph, it is clear that a higher percentage of women are casually employed and a higher percentage of men are employed full-time. This information suggests that employment status **may be associated** with gender.

18 a i

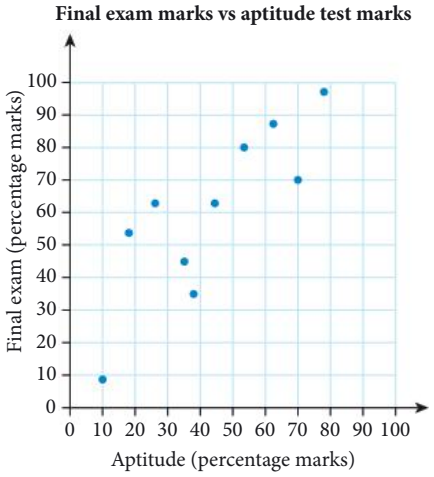


ii $r = -0.4130$

iii There appears to be a linear, weak and negative relationship between the years of education and the number of children a person has.

iv There is limited evidence to suggest that the number of children a person will have should decrease as the number of years of their education increases.

b

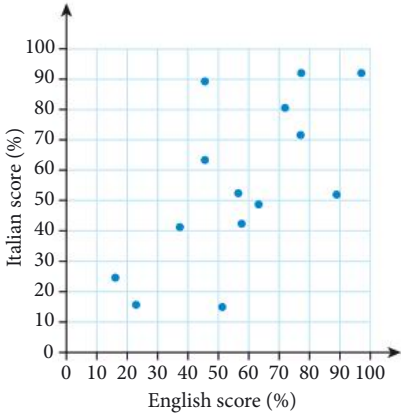


ii $r = 0.8307$

iii There appears to be a linear, strong and positive relationship between aptitude test results and final exam results.

iv We can conclude that the final exam percentage should increase as the aptitude test percentage increases.

c i



ii $r = 0.6589$

iii There appears to be a linear, moderate, positive relationship between a student's English score and their Italian score.

iv There is some evidence to support that a student's Italian score should increase as their English score increases.

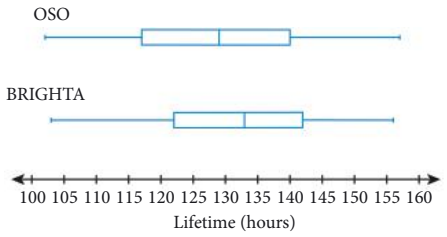
19 a OSO: 130, 129, 117
BRIGHTA: 131.5, 133, 129 and 134
Different Brighta are slightly higher for all.

b OSO: 14.4; BRIGHTA: 13.9

c BRIGHTA

9780170250283

d



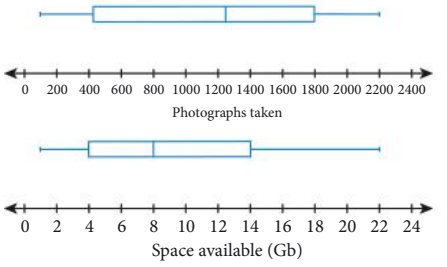
e OSO: 117 to 140; BRIGHTA: 122 to 142

f Distributions are very similar, although BRIGHTA has a slightly higher median, therefore the brand of globe may be associated with the lifetime of the globe.

20 a Number of photos taken: mean = 1183, median = 1250, mode = 2000, range = 2100, IQR = 1300, standard deviation = 695.56.

Amount of free space: mean = 8.93, median = 8, modes = 2 and 6, range = 21, IQR = 10, standard deviation = 6.06.

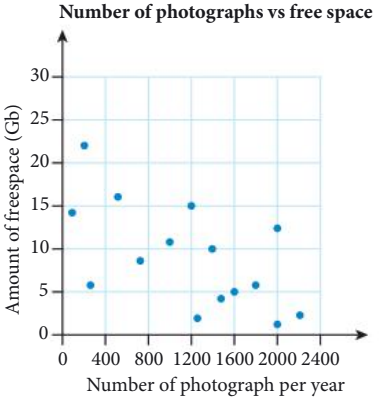
b



c There is an average of around 1200 photos taken per year by Year 12 students on their phone. As the mean and median are reasonably similar (mean = 1183 and median = 1250) and the IQR is a little less than half the range (1300 compared to 2100), this indicates that there are no outliers in the set and the data is roughly symmetrical (this is confirmed by the corresponding boxplot).

There is an average of around 8–9 Gb of free space on phones owned by Year 12 students. The mean is slightly higher than the median (8.93 compared to 8), and the IQR is roughly half the range (10 compared to 21), which implies that there are no outliers but the data may be positively skewed (this is confirmed by the corresponding boxplot).

d

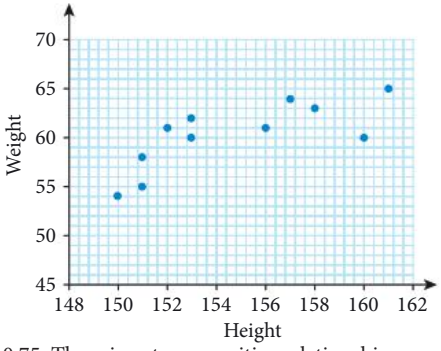


e $r = -0.6192$, which indicates that the relationship between the number of photographs taken in a year and the free space available on a phone is linear, moderate and negative.

f Year 12 students take an average of around 1200 photos a year on their phone, with a symmetric distribution ranging from 100 to 2200. In terms of free space, they have an average 8–9 Gb free on their phones, although this has a slightly positively skewed distribution ranging from 1 to 22 Gb free. The relationship between the number of photographs taken and free space on the phone is a moderate, negative one. This means that there is some evidence to support that the number of Gb free on a phone should decrease as the number of photographs taken increases.

2.01

1 a

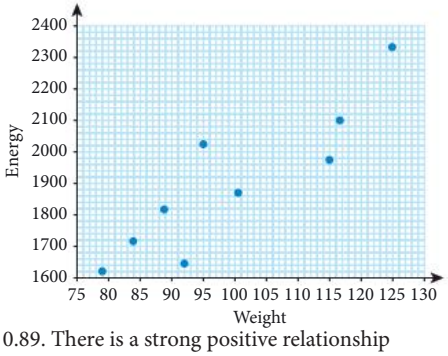


b 0.75. There is a strong positive relationship between the variables weight and height.

c weight = −44 + 0.7 × height

2 a weight

b



c 0.89. There is a strong positive relationship between the energy used and weight.

d energy = 609 + 13 × weight

e 13 kJ

3 The gradient is 0.7, which tells us that weight increases by 0.7 kg for every 1 cm increase in height.

4 a The gradient is 0.8, which tells us that weight loss increases by 0.8 kg for every 1 hour increase in exercise.

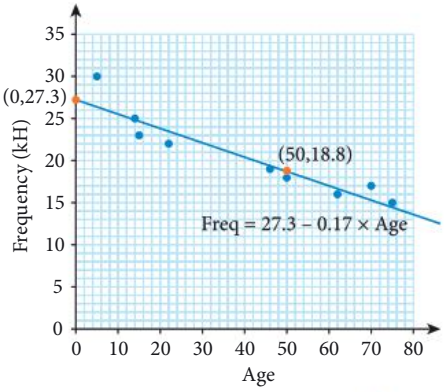
b The gradient is 8.2, which tells us that height increases by 8.2 cm for every year of increase in age.

c The gradient is −1235, which tells us that car value decreases by $1235 for every 1 year of increase in car age.

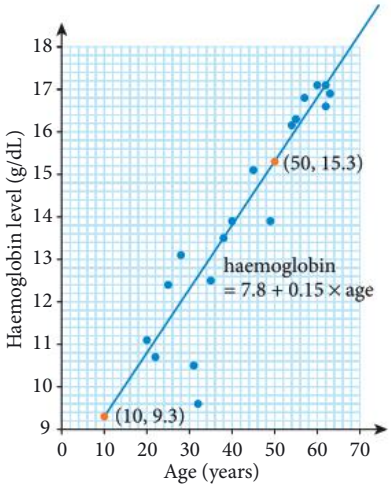
5 number of bushfires = 41.00 − 1.14 × amount of rainfall

6 number of umbrellas sold = 11.07 + 0.11 × amount of rain

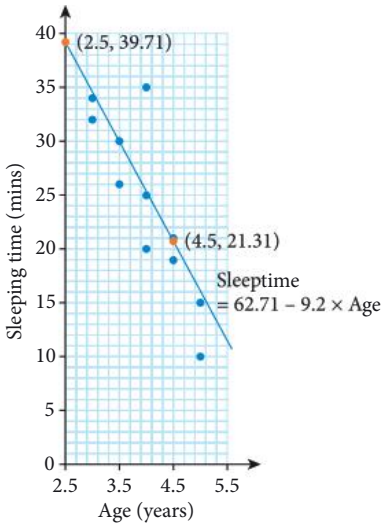
7 a



b



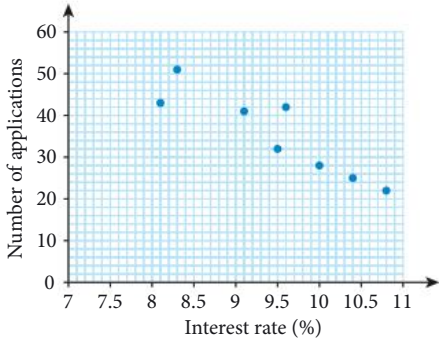
c



8 a 9.49, 0.95 b 35.5, 10.21

c EV = Interest rate and RV = number of applicants

d

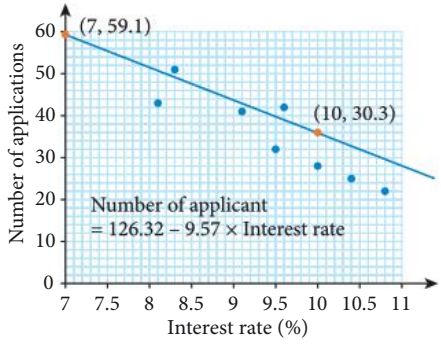


e −0.89, strong negative relationship, therefore we can conclude that the number of applicants decreases as the interest rate increases.

f number of applicants = 126.3 − 9.6 × interest rate

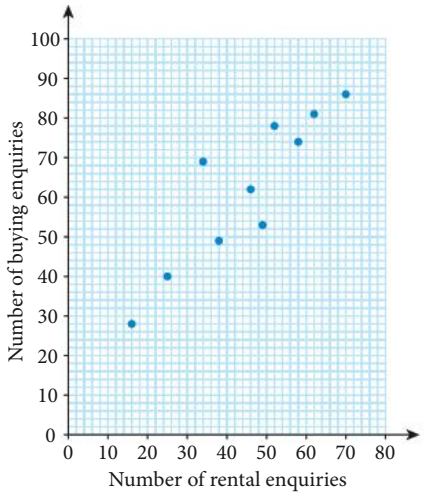
g The *y*-intercept is 126.3. So if the interest rate were 0%, 126.3 or approximately 126 applications would be expected.
The gradient is −9.6 so for every 1% increase in interest rate , the number of applications would be expected to decrease by 9.6.

h



9780170250283

9 a

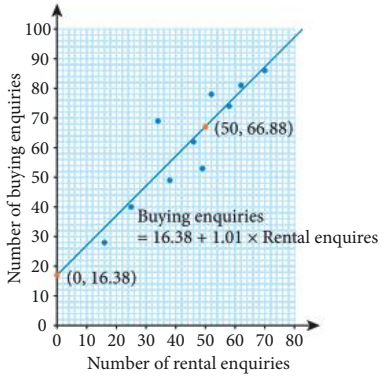


b 0.90, strong positive relationship, therefore it can be concluded that the number of buying enquiries increases as rental enquiries increase.

c buying enquiries = 16.38 + 1.01 × rental enquires

d The y-intercept is 16.38, which tells us that the number of buying enquiries is 16.38 when the number of rental enquiries is 0.
The gradient is 1.01, which tells us that number of buying enquiries increases by 1.01 for every 1 increase in the number of rental enquiries.

e



10 a Explanatory variable is the area of the planting field.

b Number of pumpkins = 101.59 + 0.39 × area of field

c The gradient is 0.39 so for every 1 m^2 increase in area, the number of pumpkins harvested increases on average by 0.39.

2.02

1 a 0.687

b 68.7% of the variation in shoe size can be explained by the variation in height.

c shoe size = −21.83 + 0.17 × height

d The r^2 value by itself suggests it may be an appropriate model.

2 a 0.855

b 85.5% of the variation of the father's age can be explained by the variation of the mother's age.

c father's age = 0.58 + 1.04 × mother's age

d Yes, this is an appropriate model due to the high r^2 value, therefore only 14.5% of the variation between the variables is unexplained.

3 a 0.868

b 86.8% of the variation in the number of laps can be explained by the variation in age.

c number of laps = 11.99 − 0.14 × age

d Yes, this is an appropriate model due to the high r^2 value, therefore only 13.2% of the variation between the variables is unexplained.

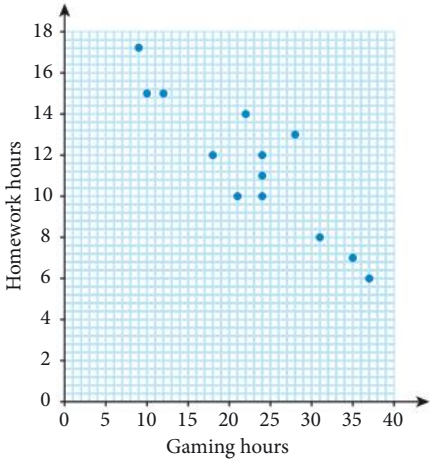
4 a 0.20 b 0.57 c 0.11
d 0.02 e 1

5 a 0.94 b −0.68 c −0.85

6 D

7 D

8 a



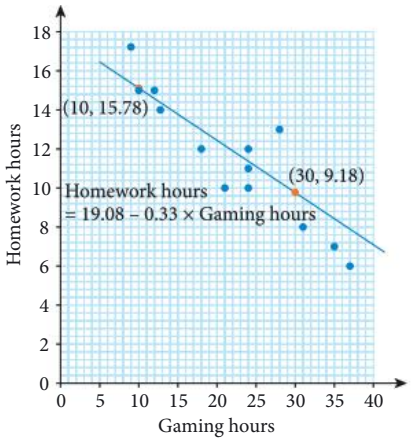
b −0.89, strong negative relationship, therefore it can be concluded that homework hours will decrease as gaming hours increase.

c 0.79, 79% of the variation in the amount of homework hours can be explained by the variation in the number of gaming hours.

d homework hours = 19.08 − 0.33 × gaming hours

e The y-intercept is 19.08, which tells us that the number of homework hours is 19.08 when the number of gaming hours is 0.
The gradient is −0.33, which tells us that number of homework hours decreases by 0.33 hours for every 1 hour of increase in the number of gaming hours.

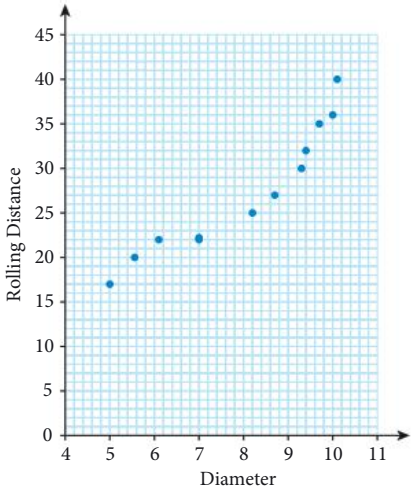
f



g Yes, as the interpretation of the y-intercept is logical and the r^2 value indicates that only 21% of the variation in the variables is unexplained.

2.03

1 a 8.95 laps b 5.43 laps c 2.07 laps
2 a 3.28 b 6.64 c 12.8
3 a 2°C b 8°C c 18°C
4 a 3 levels b 12 levels
5 a 145 grams, interpolation
b 45 grams, extrapolation
c 370 grams, extrapolation
d The prediction for 25 chocolates is the most reliable as it is inside the data range of 15 to 50, the other 2 predictions are extrapolation.
6 a 43 applicants, interpolation
b 89 applicants, extrapolation
c 9.02%, interpolation
d The predictions made in part **a** and **c** are the most reliable as they are examples of interpolation.
7 B
8 D
9 C
10 a Diameter
b



c 0.95, strong positive relationship, therefore it can be concluded that the rolling distance increases as the diameter of the wheel increases.
d 0.89, 89% of the variation in the rolling distance can be explained by the variation in the diameter of the wheel.
e Rolling distance = $-1.63 + 3.66 \times$ diameter
f The y-intercept is -1.63, which tells us that the rolling distance is -1.63 metres when the diameter of the wheel is 0. This is an inappropriate interpretation since if there is no wheel, it cannot roll.
The gradient is 3.66, which tells us that the rolling distance increases by 3.66 metres for every 1 cm of increase in the diameter of the wheel.
g i 28.75 m ii 37.53 m
h predicition in i as it is interpolation
i i 3.18 cm ii 6.18 cm
j prediction in ii as it is interpolation

2.04

1 a 0.46 metres b -0.65 metres
2 a -3.8 hours b 2.02 hours
c 0 hours
3

Height (cm)	Femur length (cm)	Residual
178	50.2	–0.3
173	48.4	0.4
165	45.1	1.1
164	44.6	1.1
168	45	–0.5
165	42.6	–1.4
155	39.9	0.8
155	38	–1.1

4

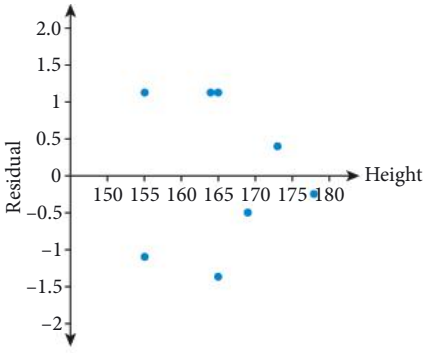
Max temperature (°C)	16	18	19	20	22	23	25	27	29	31
No. of meat pies sold	54	56	52	49	43	44	39	25	20	15
Residual	–6	1.8	0.7	0.6	0.4	4.3	5.1	–3.1	–2.3	–1.5

5

Restaurant bill ($)	45	52.80	61.55	70.20	75.50	79	84.50	91
Tip ($)	1.20	1.60	1.80	2	2.40	3	3.50	4.20
Residual	0.3	0.2	–0.1	–0.5	–0.4	0	0.1	0.4

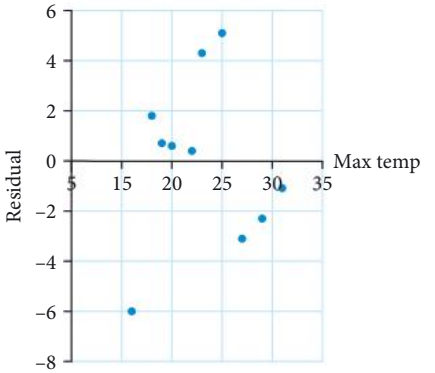
9780170250283

6 a



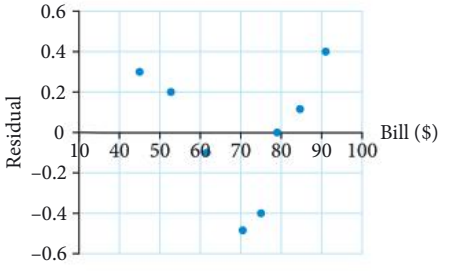
b The data is probably linear as the points in the residual plot are scattered above and below the x-axis and there is no clear curved pattern.

7 a



b The data is probably linear as the points in the residual plot are scattered above and below the x-axis and there is no clear curved pattern.

8 a



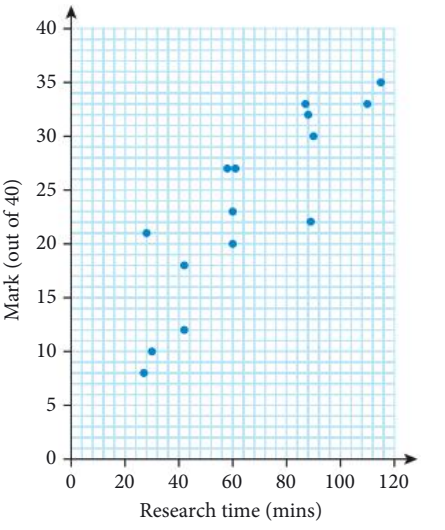
b Data is probably non-linear as the points in the residual plot make a clear curved pattern.

9 A

10 C

11 a The final mark will depend on the amount of research a student does.

b



c mark = 6.53 + 0.26 × research

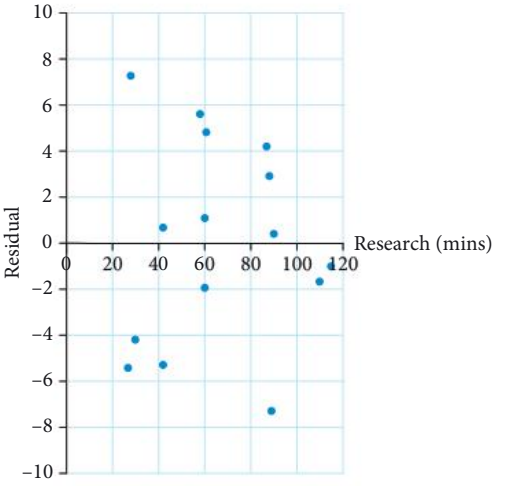
d $r = 0.86$ and $r^2 = 0.74$. A regression model is appropriate as the data has a strong positive relationship and 74% of the variation in the mark is explained by the variation in research time.

e 21.6 out of 40

f 27 − 21.6 = 5.4

g

Research	27	28	30	42	42	58	60	60	61	89	90	88	87	110	115
Mark	8	21	10	12	18	27	20	23	27	22	30	32	33	33	35
Residual	−5.5	7.3	−4.2	−5.3	0.7	5.6	−1.9	1.1	4.8	−7.3	0.4	2.9	4.2	−1.7	−1



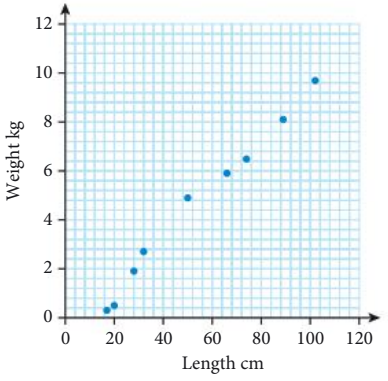
h Yes, the regression model is appropriate as the data is probably linear since the points in the residual plot are randomly scattered above and below the x-axis

INVESTIGATION: PREDICTING HOURS OF SLEEP PER NIGHT FOR YEAR 12 STUDENTS

Answers will vary.

CHAPTER 2 REVISION

1 B
2 A
3 D
4 C
5 E
6 E
7 A
8 D
9 A
10 E
11 **a**

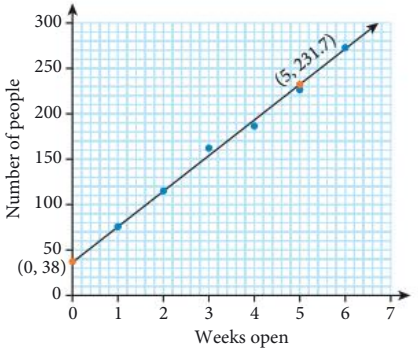


b 0.99, strong positive relationship, therefore it can be concluded that the weight of the mammal increases as its length increases.
c Weight = −1.2 + 0.1 × length

12 number of soft drink cans = 39.80 + 1.30 × temperature

13 **a** The heart rate increases by 3.8 beats per minute for every 1 minute of increase in the amount of exercise.
b Height increases by 2.1 centimetres for every 1 centimetre increase in arm length.
c The value of a car decreases by $1112 for every one year of increase in age.

14



15 **a** 0.97, 97% of the variation in the y variable can be explained by the variation in the x variable.
b 0.99, 99% of the variation in the y variable can be explained by the variation in the x variable.

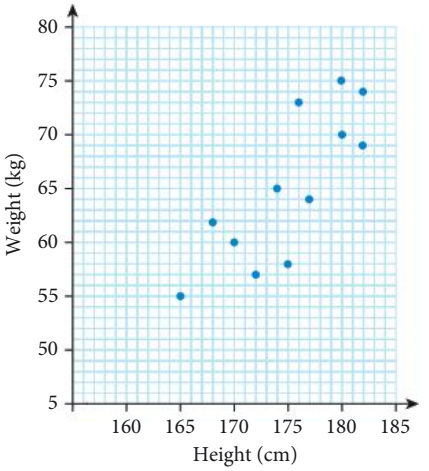
16 **a** 0.38 **b** 0.02 **c** 0.28 **d** 0.19

17 **a** −0.81 **b** 0.933

18 **a** 179 units, extrapolation
b 129 units, interpolation
c 1806 kg, interpolation
d The part **a** prediction, as it was extrapolation.

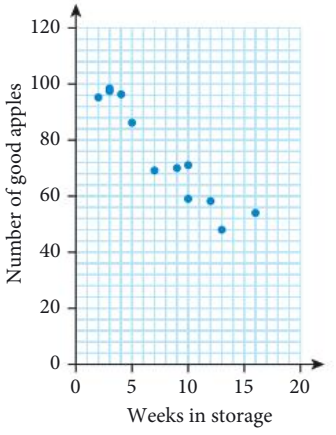
19 **a** 10.2 cm **b** −3.3 cm

20 **a**



b 0.81, strong positive relationship, therefore it can be concluded that weight increases as height increases.
c weight = −112.47 + 1.01 × height
d 0.65, probably not as 35% of the variation in weight is not explained by the variation in height.
e **i** 62.26 kg **ii** 40.04 kg **iii** 201.46 cm
f The prediction in part **ei** as it is interpolation; the other 2 are extrapolation.

21 **a** Number of good apples.
b



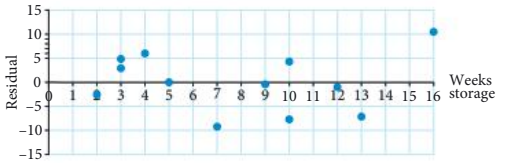
c Number of good apples = 105.3 − 3.9 × weeks in storage

9780170250283

d

Weeks in storage	13	3	4	5	3	7	16	10	9	10	2	12
Residual	−7.1	3.2	6.1	0	4.2	−9.3	10.5	4.3	−0.6	−7.7	−2.6	−1

e



f Yes, as the residual plot has points randomly scattered above and below the x-axis. The coefficient of determination supports this with 89% of the variation in the number of good apples being explained by the variation in weeks in storage.

3.01

1 **a** 1, 7, 13, 19, 25 **b** 10, 4, −2, −8, −14
c 9, 12, 15, 18, 21 **d** −2, −6, −10, −14, −18

2 **a** $a = 11, d = 8$ **b** $a = -10, d = 6$
c $a = 31, d = -2$ **d** $a = -9, d = -4$

3 **a** 43 **b** 55 **c** 37.5 **d** −43

4 **a** $a = 10$ **b** $d = 3$
c $t_{10} = 37$ **d** $t_{20} = 67$

5 **a** $a = 1$ **b** $d = -4$
c $t_{11} = -39$ **d** $t_{100} = -395$

6 $a = 3, d = 4, t_n = 4n - 1, t_{25} = 99$

7 **a** $d = 7$
b $t_{10} = 1 + (10 - 1) \times 7 = 64$

8 **a** $t_{10} = -22, t_{20} = -102$ **b** $t_{10} = -97, t_{20} = -187$
c $t_{10} = 103, t_{20} = 213$ **d** $t_{10} = 43, t_{20} = 113$

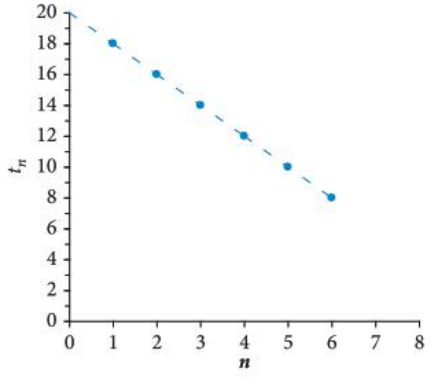
9 A

10 A

11 $d = 32 - 21 = 43 - 32 = 54 - 43 = 11$

12 **a** 18, 16, 14, 12, 10, 8

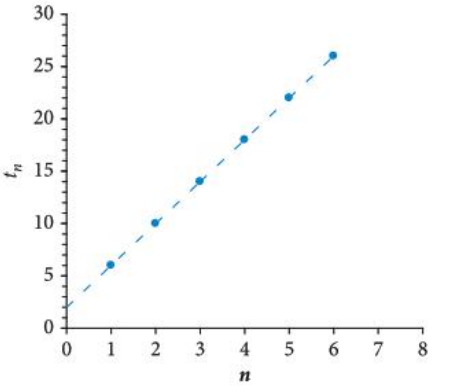
b



c It is a straight line with slope −2.

13 **a** 6, 10, 14, 18, 22, 26

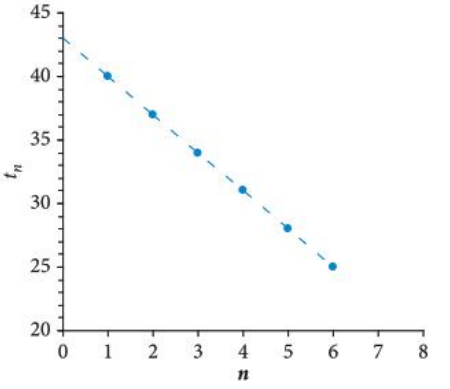
b



c It is a straight line with slope 4.

14 **a** 40, 37, 34, 31, 28, 25

b



c It is a straight line with slope −3.

15 $t_{19} = 100$

16 **a** $t_{22} = 0$ **b** $t_{12} = -100$
c $t_{47} = 1000$ **d** $t_{17} = 250$

17 $a = \frac{90}{7}, d = \frac{-37}{7}$, sequence is $\frac{90}{7}, \frac{53}{7}, \frac{16}{7}, \frac{-21}{7}, \ldots$

18 t_{15}

19 t_{49}

20 $t_{20} = 180$

21 $w = 0.75$

22 $x = 3$

3.02

1 **a** 280 **b** −170 **c** 225 **d** −200

2 **a** 91 **b** 408 **c** 88.5 **d** −910

3 **a** $S_{20} = 1640$ **b** $S_{25} = -1175$
c $S_{100} = 30\,000$ **d** $S_{10} = -125$

4 −306

5 **a** $S_{17} = 1717$ **b** $S_{27} = -1755$
c $S_{11} = 1067$ **d** $S_{25} = 525$

6 $S_{44} = 5126$

7 D

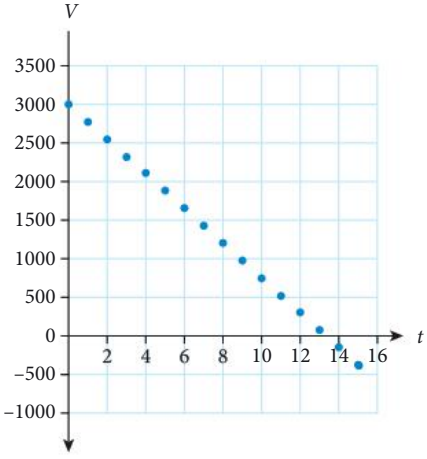
8 A

9 $\frac{5n(n+1)}{2}$

10 $n = 12$
11 $S_{20} = 820$
12 $n = 21$
13 $n = 8$
14 $n = 11$
15 $a = 14, d = 4$

3.03

1 \$1120
2 1.3 L
3 **a** 14 900 **b** 17 900
4 **a** \$24 050 **b** \$220 250
5 280 laps
6 **a** 23 **b** 100 **c** 610
7 **a** \$1875
b



c 13.3 years
8 **a** 150 **b** 14 months
9 **a** 252 **b** 29 days
10 **a** 39 words per minute **b** 15 weeks
11 **a** 210 **b** 13th **c** 57 rows
12 **a** $3k$
b Sequence is 6, 12, 18, ..., where $a = 6$, $d = 6$. $3k^2$
c 9
13 **a** $d = 3$ **b** $S_{20} = 670$

3.04

1 **a** 1, 6, 36, 216, 1296
b 10, −30, 90, − 270, 810
c 9, 27, 81, 243, 729
d −2, 8, −32, 128, −512
2 **a** $a = 7, r = 3$
b $a = 1000, r = \frac{1}{5}$
c $a = -3, r = 4$
d $a = -400, r = -\frac{1}{2}$
3 **a** −1250 **b** 1296 **c** $\frac{3}{8}$ **d** 1250
4 **a** 5 **b** 3 **c** 1215 **d** 98 415
5 **a** $\frac{1}{4}$ **b** −3 **c** $-\frac{19683}{4}$
d $-\frac{1162\,261\,467}{4}$
6 $a = 3, r = 3, t_n = 3(3^{n-1}), t_{20} = 3(3^{20-1}) =$ 3 486 784 401
7 **a** $r = \frac{0.2}{1} = \frac{0.04}{0.2} = \frac{0.008}{0.04} = 0.2$
b 0.000 064
8 **a** $t_{10} = 262\,144, t_{20} = 274\,877\,906\,944$
b $t_{10} = \frac{3}{32}, t_{20} = \frac{3}{32768}$
c $t_{10} = -5120, t_{20} = -5\,242\,880$
d $t_{10} = \frac{4}{6561}, t_{20} = \frac{4}{387\,420\,489}$
9 E
10 B
11 Using $r = \frac{t_2}{t_1} = \frac{t_3}{t_2} = \frac{t_4}{t_3} = \cdots, r = \frac{6}{21} = \frac{\frac{12}{7}}{6} = \frac{\frac{24}{49}}{\frac{12}{7}} = \frac{2}{7}$
12 t_8
13 **a** t_{14} **b** t_6 **c** t_{11} **d** t_8
14 $-\frac{1}{16}, -\frac{1}{8}, -\frac{1}{4}, -\frac{1}{2}$, −1, −2, −4, −8, −16, −32, −64, −128, −256, ...
15 4th
16 5th
17 $a = \frac{2}{9}, r = \frac{3}{2}, t_{11} = \frac{6561}{512}$
18 $m = \pm\sqrt{6}$

3.05

1 A
2 D
3 E
4 **a** 12 093 235 **b** −86 380 250
c 265 716 **d** 419 430
5 **a** 2 097 150 **b** 7 324 218
6 **a** 720 600 **b** 26 240
7 **a** 131 068 **b** $\frac{32769}{65536}$
8 **a** D **b** B **c** B **d** C
9 **a** 2 **b** 12 **c** $\frac{27}{4}$ **d** $-\frac{4}{3}$
10 **a** 12 **b** 121.5 **c** $-\frac{32}{3}$ **d** 25
11 **a** $\frac{640}{7}$ **b** 1280 **c** 62.5 **d** 121.5
12 C
13 9
14 10
15 9
16 10
17 **a** 7812 **b** $\frac{2295}{64}$ **c** 8403
d 273 **e** 255

9780170250283

3.06

1 $1 008 363
2 204 800
3 81
4 14 m
5 160 cm
6 80 kg
7 a i $23 200 ii $26 912 iii $31 217.92
b $102 345.29 c 8 years
8 True. Kim will never pay off the credit card.
9 a $\frac{7}{9}$ b $\frac{2}{9}$ c $\frac{16}{99}$
d $\frac{234}{999}=\frac{26}{111}$
10 $15 000
11 13th year
12 a $r=6$ b 48 372 940
13 a $p>-\frac{1}{2}$ and $p<\frac{1}{2}$ b $\frac{5}{7}$ c $\frac{1}{14}$
14 $S_\infty - S_n = \frac{a}{1-r} - \frac{a(1-r^n)}{1-r}$ giving

$$\frac{a}{1-r}\left(1-\frac{(1-r^n)}{1}\right)=\frac{a}{1-r}\left(1-1+r^n\right)$$

$$=\frac{ar^n}{1-r}$$

QED

3.07

1 E
2 a 386 b 36 374
3 a 96 rows b 6960 logs
4 10
5 1275 oranges
6 12 000 L
7 a 26 km b 740 km
8 $122 102
9 $482 000
10 $50 363
11 $6821
12 a 72.9% b 7 layers
13 37
14 a $960 b $10 275 c After 17 years
15 1.6 m
16 a 735 b 4315
17 $152 171
18 a 15 b 39 c After 13 years
19 a $r = 0.8$, or improves by 20%
b 1.92 minutes = 1 minute 55 seconds
c Approximately 23 minutes (22 minutes 55 seconds) after cutting 10.08 minutes from time.
d After 10 weeks
20 203 kg and 5 kg

CHAPTER 3 REVISION

1 B
2 D
3 D
4 E
5 E
6 a $d=4$ and $a=1$ b 37
7 t_7
8 –13, –7, –1, 5, 11, 17, 23, 29, 35, …
9 820
10 832
11 n^2
12 14
13 190 km
14 a $5250 b In its 15th year.
15 a 19 b 25 rows
16 a $r=2$ and $a=3$ b 1536
17 10th term
18 $\frac{2}{3125}, \frac{2}{125}, \frac{2}{5}$, 10, 250, 6250, 156 250, ...
19 –4 782 966
20 468.75
21 $a=192$, $r=-\frac{1}{4}$, $S_\infty = 153.6$
22 Approximately 883
23 $2.8\overline{3}$ metres
24 $\frac{4}{9}$
25 2.4 m
26 13 hours
27 a i 95% ii 90.25% iii 85.74%
b 46.33%
c 27 weeks
28 1554 customers
29 a $110 - 105 = 105 - 100 = 5$
b 125 c 10 d 1530
30 a 33 b 6th year
c 197 d 9th year

MIXED REVISION 1

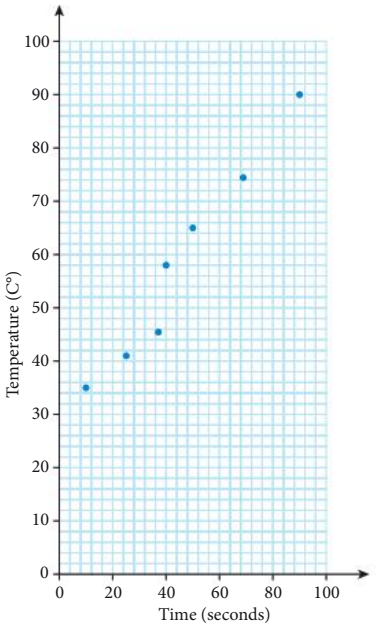
Multiple choice

1 C
2 D
3 E
4 A
5 A
6 D
7 C
8 B
9 C

Short answer

1 22

2 a



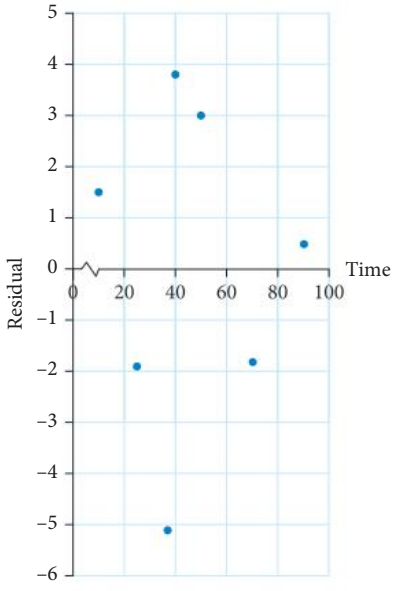
b 0.99, indicates a strong positive relationship therefore it can be concluded that temperature increases as time on the gas increases.

c temperature = 26.7 + 0.7 × time

3 a

Time (seconds)	10	25	37	40	50	70	90
Residual	1.5	−1.9	−5.1	3.8	3	−1.8	0.5

b



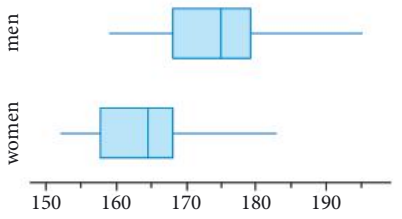
c As the points in the residual plot are randomly scattered above and below the *x*-axis, the original data is probably linear.

4 a 19.72, 20, 18 and 20

b 35 is an outlier as it is outside the range $11.25 \leq x \leq 29.25$

c 21, 20, 18 and 20. Only changed the mean.

5 a



b Men

c Men

d Neither of the sets are skewed.

e Men = 174.91 cm and women = 164.15 cm

f Any, as the data is not skewed and the answers for median and mean are very similar.

6 136

7 28th term

Application

1 a

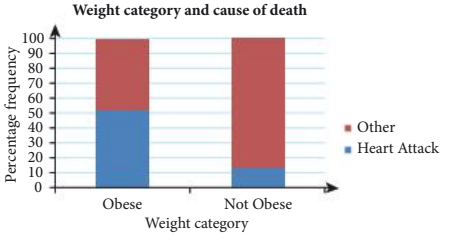
	Weight category		
Cause of death	Obese	Not Obese	
Heart attack	3020	6380	9400
Other	2820	41 780	44 600
	5840	48 160	54 000

b 32%

c

	Weight category	
Cause of death	Obese	Not Obese
Heart attack	52%	13%
Other	48%	87%
	100%	100%

d



e A higher percentage of people with obesity die from heart attacks than those who are not obese. This information suggests that cause of death **may be associated** to weight category.

2 a weight = −113.85 + 1.02 × height

b A student's weight will be −113.85 kg when their height is 0, this interpretation is not sensible.
A student's weight will increase by 1.02 kg for every 1 cm increase in height.

9780170250283

c 0.70

d Yes, as 70% of the variation in weight can be explained by the variation in height.

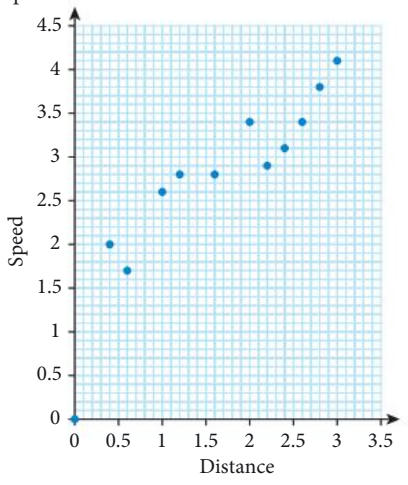
e i 62.61 kg ii 40.17 kg iii 200.83 cm

f The prediction in **ei** is the most reliable as it is interpolation, the other two are extrapolation.

3 21st term

4 a speed

b

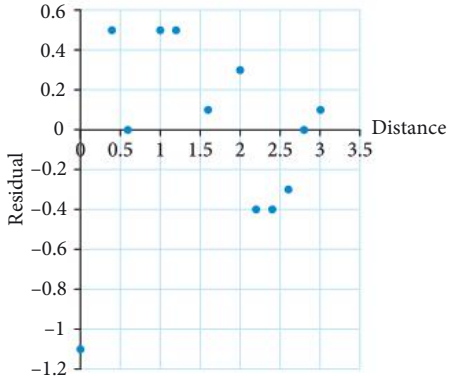


c speed = 1.1 + 1 × distance

d

Distance (m)	0	0.4	0.6	1	1.2	1.6	2	2.2	2.4	2.6	2.8	3
Residual	−1.1	0.5	0	0.5	0.5	0.1	0.3	−0.4	−0.4	−0.3	0	0.1

e



f Yes, the regression model is appropriate as the residual plot is randomly scattered, suggesting that data is probably linear and the correlation coefficient and coefficient of determination are high (0.90 and 0.81).

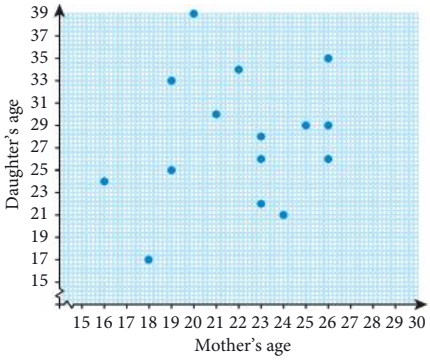
5 a 100, 97, 94

b the 35th term is the first negative term

6 a 22.07 years old b 27.87 years old

c The average suggests that daughters have their first child at an older age than their mothers.

d



e Linear, postive and weak OR no relationship.

f $r = 0.17$

g There is no relationship between the ages of mothers and their daughters on the birth of their first child.

7 a Using $d = t_{n-1} - t_n = t_{n-2} - t_{n-1}$ where

$d = (2x + 4) - (x + 1) = (3x + 7) - (2x + 4)$

LHS $= (2x + 4) - (x + 1)$

$= 2x + 4 - x - 1$

$= x + 3$

RHS $= (3x + 7) - (2x + 4)$

$= 3x + 7 - 2x - 4$

$= x + 3$

So it is an AP.

b $S_{50} = 1275x + 3725$

4.01

1 a 15, 13, 11, 9, 7, 5 b 2, 8, 32, 128, 512, 2048

c −8, −1, 6, 13, 20, 27

d 64, −32, 16, −8, 4, −2

e 50, 45, 40, 35, 30, 25

f 3, 7, 15, 31, 63, 127

g 1, −2, 13, −62, 313, −1562

h 40, 22, 13, 8.5, 6.25, 5.125

2 a 4, 7, 10 b −10, −40, −160

c 5, 0, 10 d 100, 120, 128

e 18, 17, 16.5 f 4, 5, 8

3 51

4 20

5 a 4.75 b 36

6 D

7 a The third term is the sum of the first two terms.

b 3, 7, 10, 17, 27 c 12th term

8 a 21, 50 b 10, 34 c 5, 11

4.02

1 a $t_{n+1} = t_n + 4$ where $t_1 = 4$

b $t_{n+1} = 2t_n$ where $t_1 = 3$

c $t_{n+1} = \frac{1}{2}t_n$ where $t_1 = 100$

d $t_{n+1} = t_n - 7$ where $t_1 = 49$

2 a 3 b 1825

3 **a** -30 **b** -1122

4 $a = 28, r = 2, d = -6$

5 **a** $r = \frac{1}{2}, d = 12$ **b** $a = 56$ **c** 26

6 **a** $t_{n+1} = 2t_n - 1$ where $t_1 = 2$

b $t_{n+1} = -t_n + 4$ where $t_1 = 8$

c $t_{n+1} = \frac{1}{2}t_n + 2$ where $t_1 = 20$

d $t_{n+1} = -0.25t_n + 4$ where $t_1 = 40$

e $t_{n+1} = 4t_n + 3$ where $t_1 = 1$

7 B

8 **a** 6, 8, 14 **b** $t_{n+1} = 3t_n - 10$ where $t_1 = 6$

9 **a** $r = 3, x = 16$ **b** 148

10 **a** $C_{n+1} = 1.15C_n - 50$ where $C_1 = 400$ **b** 422

4.03

1 **a** $t_{n+1} = t_n + 3$ where $t_1 = -4$

b $t_{n+1} = t_n - 9$ where $t_1 = 30$

c $t_{n+1} = t_n + 2.5$ where $t_1 = 1$

d $t_{n+1} = t_n + \frac{1}{2}$ where $t_1 = \frac{1}{4}$

e $t_{n+1} = t_n - 1.1$ where $t_1 = 3.8$

2 a and c

3 $t_n = 40n - 20$

4 **a** $t_n = 8 - 3_n$ **b** $S_n = 4n - 1$

c $u_n = 20 - 5n$ **d** $w_n = 10n - 2$

5 C

6 **a** 41 **b** $t_{n+1} = t_n - 4$ where $t_1 = 41$

c 41, 37, 33

7 **a** -68 **b** $t_{n+1} = t_n - 68$ where $t_1 = 348$

c 348, 280, 212

8 D

9 **a** 2, 6, 10 **b** $t_{n+1} = t_n + 4; t_1 = 2$

10 $t_{n+1} = t_n + 7$ with $t_1 = -49$

11 **a** $a_{n+1} = a_n + 10$ with $a_1 = 5$ **b** 75

4.04

1 **a** $t_{n+1} = 2t_n; t_1 = 2$ **b** $t_{n+1} = 5t_n; t_1 = 3$

c $t_{n+1} = -6t_n; t_1 = 5$ **d** $t_{n+1} = -\frac{1}{2}t_n; t_1 = 48$

e $t_{n+1} = 2.8t_n; t_1 = 2$ **f** $t_{n+1} = -\frac{1}{3}t_n; t_1 = 81$

2 **a** 7.776 **b** $t_n = 4.5(1.2)^{n-1}$

3 **a** geometric **b** arithmetic

c neither **d** neither

e arithmetic **f** geometric

g both arithmetic and geometric

4 $t_{n+1} = 2t_n$ where $t_1 = 3$

5 **a** $t_{n+1} = 3t_n$ where $t_1 = 4$

b $t_{n+1} = 2t_n$ where $t_1 = -2$

c $t_{n+1} = 0.2t_n$ where $t_1 = 3$

d $t_{n+1} = -\frac{1}{2}t_n$ where $t_1 = 8$

e $t_{n+1} = 1.3t_n$ where $t_1 = 5$

6 **a** 5 **b** $t_{n+1} = -3t_n$ where $t_1 = 5$

7 **a** $t_{n+1} = -3t_n$ where $t_1 = 12$

b $t_{n+1} = 3t_n$ where $t_1 = 12$

8 E

9 **a** $h_{n+1} = \frac{3}{4}h_n$ with $h_1 = 8$ **b** sixth bounce

10 **a** $v_{n+1} = 1.08v_n$ where $v_1 = 20\,000$

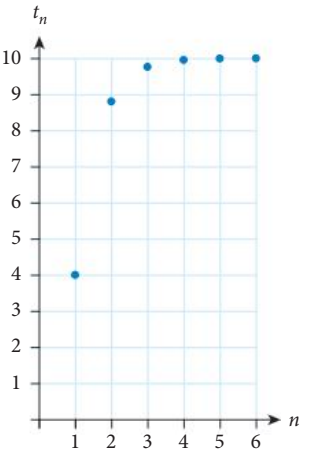
b $v_n = 20\,000(1.08)^{n-1}$ **c** \$43 178.50

4.05

1 $t_{n+1} = t_n + 5$ where $t_1 = 3$

2 D

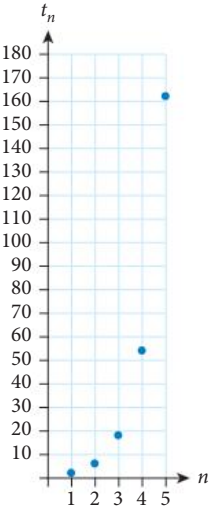
3 **a**



b long term steady state solution

4 **a** 2

b



5 D

6 **a** C **b** E **c** F

d A **e** D **f** B

7 D

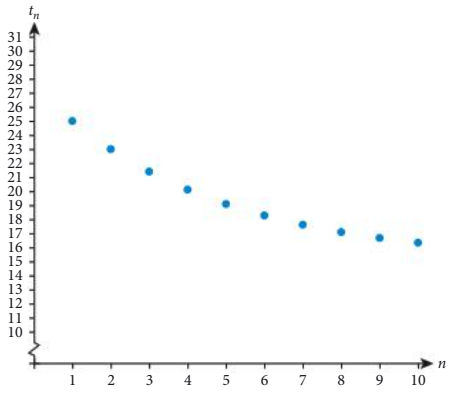
8 C

9 b, c, e, f

10 $-1 < r < 1$

9780170250283

11



In the long run, the amount of pollutant in the lake ceases to change.

4.06

1 **a** 60 **b** −64 **c** 40
d 30 **e** −30
2 **a** 11.65 **b** 7 **c** 40
3 E
4 0.2
5 **a** The difference equation $t_{n+1} = 2t_n - 0.3$ does not have a long term steady state solution since $r > 1$. The difference equation $u_{n+1} = 0.6u_n + 5$ has a long term steady state solution since $-1 < r < 1$.
b 12.5
6 **a** steady state value $= \dfrac{d}{1-r}$
b **i** 20 **ii** $-\dfrac{16}{3}$
7 0.6
8 9
9 250

4.07

1 **a** $t_{n+1} = 1.04t_n$ where $t_1 = 40\,000$
b $t_{n+1} = 0.8t_n$ where $t_1 = 48$
c $t_{n+1} = t_n + 8$ where $t_1 = 12$
d $t_{n+1} = 0.85t_n$ where $t_1 = 20\,000$
e $t_{n+1} = t_n - 10\,000$ where $t_1 = 68\,000$
f $t_{n+1} = 0.88t_n + 200$ where $t_1 = 5000$
2 **a** $d_{n+1} = d_n + 0.2$, $d_1 = 1$
b 4.8 km
3 **a** $V_{n+1} = 0.86V_n$, $V_1 = 80\,000$ **b** 2016
4 **a** $a = 45\,000$, $d = -2500$ **b** 32 500 kg
c 18 days
5 **a** $t_2 - t_1 = -120$, $t_3 - t_2 = -96$ not arithmetic.
$\dfrac{t_2}{t_1} = \dfrac{17}{20}, \dfrac{t_3}{t_2} = \dfrac{73}{85}$ not geometric.
b $t_{n+1} = 0.8t_n + 40$ where $t_1 = 800$
c 396.6 hours **d** 39.3 hours
6 B
7 **a** 2000
b **i** 5552 **ii** 17 432
c 16.84%
8 **a** 16 **b** 268
c Neither: the difference equation is not of the form $t_{n+1} = t_n + d$ or $t_{n+1} = rt_n$
9 **a** D **b** A
10 **a** $t_{n+1} = 1.04t_n + 5000$ where $t_1 = 60\,000$
b \$83 099.84 **c** 2019
11 **a** $t_{n+1} = 1.005t_n - 800$ where $t_1 = 200\,000$
b \$202 467.11
c 15 months
12 **a** $r = 0.9$, $d = 3000$, $k = 50\,000$
b 43 122 litres
c 8th morning
d 30 000 litres
13 **a** $c_{n+1} = 0.8c_n + 32$ with $c_1 = 200$ **b** 160
14 **a** $h_{n+1} = 0.6(h_n + 80)$ $h_1 = 50$
b 104.88 cm
c 120 cm

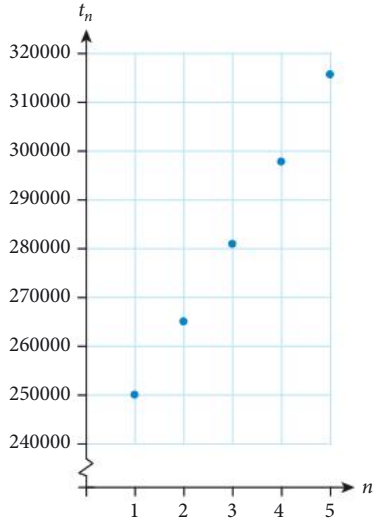
INVESTIGATION: INVESTING MONEY

a $v_{n+1} = 1.06v_n$ where $v_1 = 250\,000$

b

Year	1	2	3	4	5
Value of investment	250 000	265 000	280 900	297 754	315 619.24

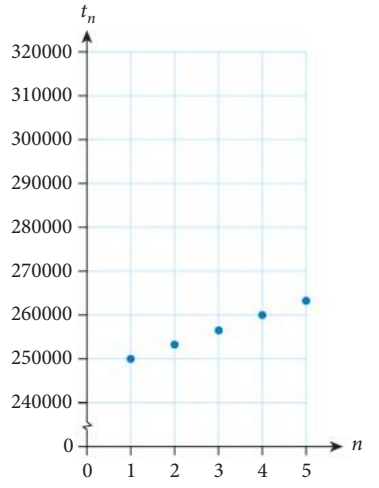
c



d The value of Madeleine's investment increases indefinitely.
e $v_{n+1} = 1.06v_n - 12\,000$ where $v_1 = 250\,000$
f

Year	1	2	3	4	5
Value of investment	250 000	253 000	256 180	259 550.80	263 123.85

g



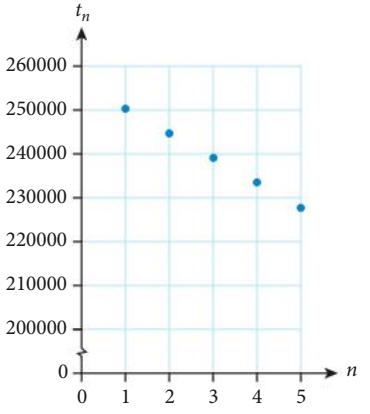
h The value of Madeleine's investment increases indefinitely but had a slower rate of growth than when no withdrawals were being made.

i $v_{n+1} = 1.06v_n - 20000$ where $v_1 = 250\,000$

j

Year	1	2	3	4	5
Value of investment	250 000	245 000	239 700	234 082	228 126.92

k



l The value of her investment is decreasing over time.

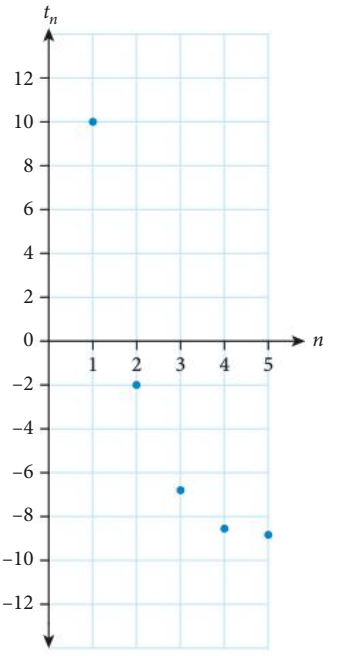
m The value of Madeleine's investment decreases until she has no money left in her account.

n \$15 000

o $v_{n+1} = 1.06v_n - 15000$ where $v_1 = 250\,000$

CHAPTER 4 REVISION

1 C
2 B
3 B
4 A
5 D
6 B
7 B
8 A
9 E
10 B
11 2, –1, –7, –19, –43, –91
12 a 9.8, 12.84 b 7
13 $t_{n+1} = -3t_n + 4$ with $t_1 = 2$.
14 a 40
b $t_{n+1} = t_n - 6$ where $t_1 = 40$
15 $t_{n+1} = -\frac{1}{2}t_n$ where $t_1 = 32$
16 a



b The sequence has a long term steady state solution.

17 17 200
18 a $V_{n+1} = 0.85V_n$ where $V_1 = 45\,000$
b \$32 512.50 c 2019
19 a $k = 1.01$, $m = -1115$, $b = 50\,000$
b i \$49 385 ii \$37 201.30
c 30th month
d \$10 841.47
20 a $t_{n+1} = 1.2t_n - 300$ where $t_1 = 2000$
b 2537 c 2021 d 400

5.01

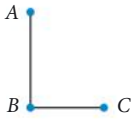
1 a i 3 ii 3
b i 4 ii 3
c i 4 ii 2
d i 4 ii 2
e i 6 ii 10
f i 5 ii 4
g i 3 ii 4
h i 6 ii 6
i i 2 ii 4
2 A

9780170250283

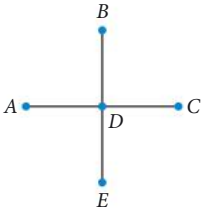
3 **a** *B* **b** *C* **c** *E* **d** *C, D*

4 A

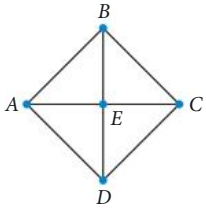
5 **a**



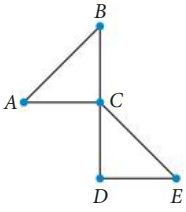
b



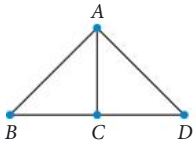
c



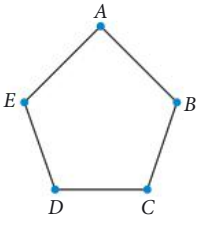
d



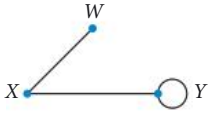
e



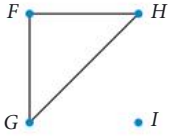
f



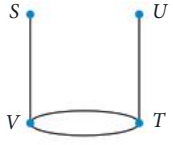
g



h



i



6 E

7 **a** 3 **b** Queensvale and Mullen

c Queensvale

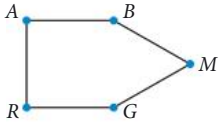
d Long's Landing, Springfield and Walter

8 E

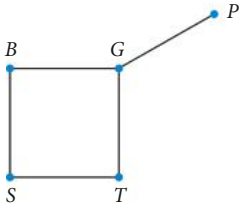
9 **a** $V = \{A, R, B, G, M\}$

b $E = \{AR, AB, GR, GM, MB\}$

c



10 **a**



b Sunvale, Bundoon, Georgetown, Praton; and, Sunvale, Tinnit, Georgetown, Praton

5.02

1 **a** iv **b** v **c** ii

d iii **e** vi **f** i

2 **a** simple, connected **b** connected, complete

c connected, complete **d** simple, bipartite

3 C

4 **a** 3 **b** 66 **c** 276 **d** 4851

5 **a** 4 **b** 20 **c** 44 **d** 1001

6 B

7 **a** 15 **b** 6

8 B

9 **a** Yes; *T–S–P–L–H–C–M*

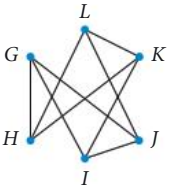
b *C–M–L–H–P* and *C–M–H–P*

c Yes; *H–P–L*

10 A

5.03

1



2 **a** Complete graph

b 10

c



3 C

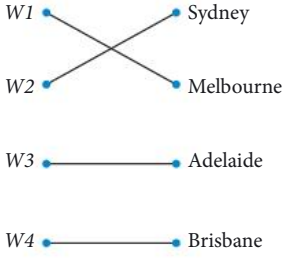
4



5 **a** Melbourne and Adelaide

b *W4* and *W2*

c



6 **a** George **b** Sally and Celeste

c George

7 **a** *A–C–D–E–F* or *A–D–E–F*

b *F–C–D–B–A*

c **i** No **ii** *D*

d Yes; *E–F–C–D–B–A*

8 **a** Black Hole

b Cook's Point and Black Hole

c Car Park–Picnic Ground; Car Park–Dolphin Lookout-Picnic Ground; Car Park–Great Gum Tree–Picnic Ground; or, Car Park–Great Gum Tree–Cook's Point–Picnic Ground

d No. The Picnic Ground is adjacent to every other location (excluding Black Hole).

e Car Park and Great Gum Tree.

9 C

5.04

1 **a** $A = 2, B = 1$ **b** $A = 4, B = 3$

c $A = 3, B = 3$ **d** $A = 0, B = 4$

2 **a** **i** $A = 3, B = 3, C = 3, D = 3$ **ii** 12

b **i** $A = 2, B = 3, C = 3, D = 2, E = 2$ **ii** 12

c **i** $A = 4, B = 2, C = 4, D = 2$ **ii** 12

d **i** $A = 3, B = 2, C = 5$ **ii** 10

3 The sum of all degrees is equal to the degree sum.

4 B

5 **a** **i** 6 **ii** 6

b **i** 10 **ii** 10

c **i** 6 **ii** 6

d **i** 8 **ii** 8

6 The degree sum is the same as the number of entries in the adjacency matrix.

7 **a**

	A	B	C
A	0	2	1
B	2	0	1
C	1	1	0

b

	A	B	C
A	0	1	2
B	1	0	2
C	2	2	0

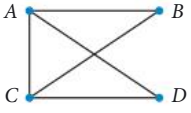
c

	A	B	C	D	E
A	0	1	1	0	0
B	1	0	1	0	0
C	1	1	0	1	1
D	0	0	1	0	1
E	0	0	1	1	0

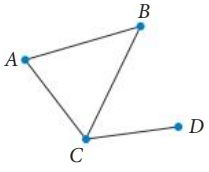
d

	A	B	C	D
A	1	1	0	0
B	1	0	1	0
C	0	1	0	0
D	0	0	0	0

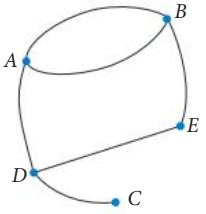
8 **a**



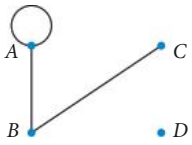
b



c

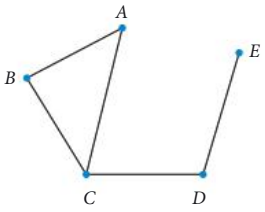


9 **a** **i**



ii 6 **iii** 5 **iv** No

b **i**



ii 10 **iii** 10 **iv** Yes

c When there are no loops.

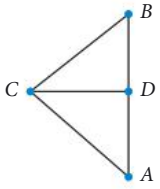
10 **a** D is isolated.

b **i** Another edge BC is missing.

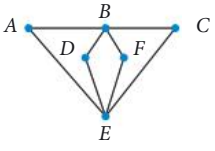
ii Vertex D is missing.

5.05

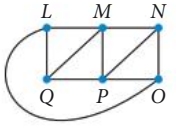
1 **a**



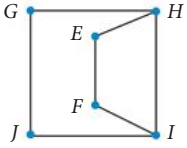
b



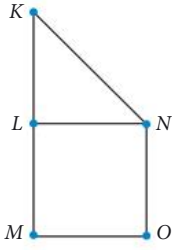
c



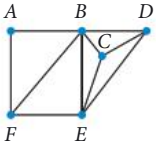
d



e



f

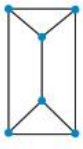


2 **a** Planar **b** Planar **c** Planar

d Planar **e** Planar **f** Planar

3 C

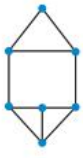
4



5 Square based pyramid

6 D

7 **a**

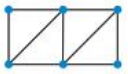


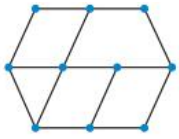
9780170250283

b 5 faces, 10 edges, and 7 vertices (5 = 10 − 7 + 2)

8 B

9 **a** Planar **b** Planar **c** Non-planar

10 **a** 

b 

11 C

5.06

1 **a** Cycle **b** Different **c** Repeated
d Bridge **e** Every

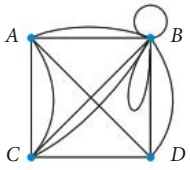
2 **a** Open walk **b** Closed walk
c Open walk **d** Open walk

3 D

4 **a** (*E, B, G, S, P, F, E*) **b** (*C, G, S, P, F, B, E*)

5 **a** **i** (Start, *A, B, C, D, E, F, G, H, I,* Start)
ii (Start, *A, D, I,* Start)
iii (Start, *A, B, C, E, F, G, H, I, D*)
b **i** Cycle **ii** Cycle **iii** Path

6 **a** Two
b Yes; there is a loop at *B*.
c



d Yes; *A* is adjacent to every other vertex.
e *A* and *C* are connected by multiple edges.

7 **a** (*A, B, E, C, D, F, A*)
b (*A, F, G, H, I, J, K, G, F, A*)
c A closed path cannot contain repeated vertices whereas a closed walk can. In part b, Patricia must travel through vertices *F* and *G* twice to cross from North to South and then back again.

5.07

1 **a** Eulerian **b** Not Eulerian
c Eulerian **d** Eulerian

2 **a** Semi-Eulerian **b** Semi-Eulerian
c Not semi-Eulerian **d** Semi-Eulerian

3 **a** *A* and *D* **b** *I* and *K*
c *M* and *P* **d** *V* and *T*

4 **a** **i** Eulerian **ii** (*A, B, E, D, F, C, A*)
b **i** Semi-Eulerian
ii (*B, A, D, G, H, E, G, B, C, F, E*)
c **i** Semi-Eulerian **ii** (*H, L, I, K, G, J, H, K*)
d **i** Eulerian
ii (*P, Q, N, M, O, R, S, O, P, S, T, P*)
e **i** Semi-Eulerian **ii** (*H, I, L, J, M, K, H, J*)
f **i** Semi-Eulerian
ii (*X, V, U, W, T, X, Z, U, Y, X, V, W, Z*)
g **i** Semi-Eulerian
ii (*B, D, A, C, D, E, G, H, C, F, D, G, B, E, F*)
h **i** Semi-Eulerian
ii (*I, P, J, L, N, P, O, N, M, Q, I, J, K, L, M*)

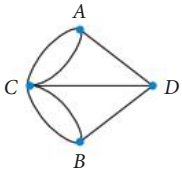
5 C

6 C

7 **a** (Salon, *A, B, C, A, G, C, D, E, F, G,* Salon)
b Eulerian

8 **a** *B* and *I* **b** (*B, D, H, B, A, C, F, I, E, G, H, I*)

9 **a**



b There are more than two vertices of odd degree.
c Any one of *AC* or *BC* or *AD* or *BD* or *CD*

10 A

5.08

1 **a** (*A, D, E, F, C, B*)
b (*G, K, L, I, M, J, H*)
c (*N, O, P, T, S, R, Q*)
d (*A, C, B, D, H, G, E, F*)

2 **a** (*F, G, H, E, B, D, C, A*)
b (*E, D, G, H, F, I, E*)
c (*R, M, P, Q, L, N, D, R*)
d (*A, B, C, F, E, G, D, A*)

3 D

4 C

5 **a** (*A, B, E, C, D, A*) **b** *DE*

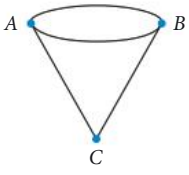
6 **a** 59 km
b 8 ways
N, P, R; N, P, Q, R; N, O, R; N, O, M, Q, R; N, O, M, Q, P, R; N, M, O, R; N, M, Q, R; N, M, Q, P, R
c 159 km

7 (Start, 1, 2, 4, 8, 11, 12, 10, 9, 14, 15, 13, 5, 6, 7, 3, Start)

5.09

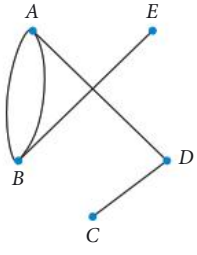
1 **a** Yes **b** No **c** Yes
d No **e** Yes **f** No

2 **a** **i**



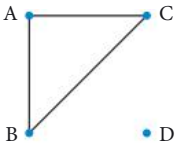
ii Contains a cycle and multiple edges.

b **i**



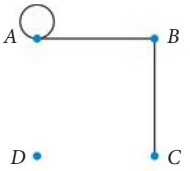
ii Contains a cycle and multiple edges.

c **i**



ii Contains a cycle and an isolated vertex.

d **i**

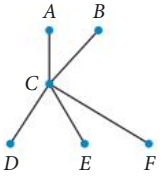


ii Contains a loop and an isolated vertex.

3 **a** False **b** True **c** True
d True **e** True

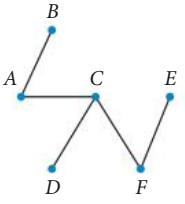
4 C

5 **a**

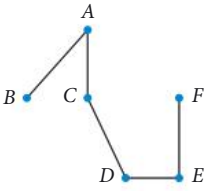


b *DE*, *EF* and *AB*.

6



7



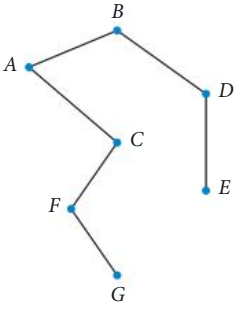
5.10

1 **a** 5.6 **b** 28 **c** 2.15 **d** 45

2 B

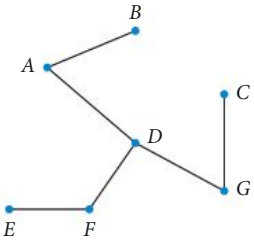
3 293 km

4 **a** **i**



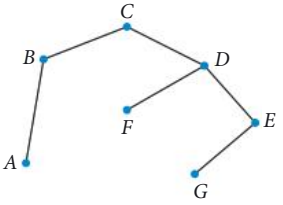
ii 32

b **i**



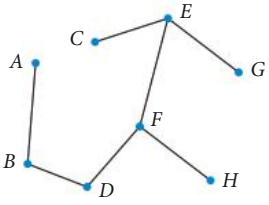
ii 52

c **i**



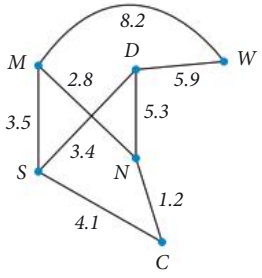
ii 527

d **i**



ii 218

5 **a**



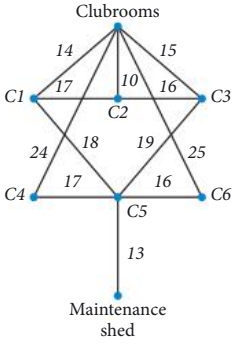
b Moonville–Sunnvale–Dayton = 6.9km

6 3740 m

7 2540

INVESTIGATION: TENNIS CLUB LIGHTING CABLE NETWORK

1

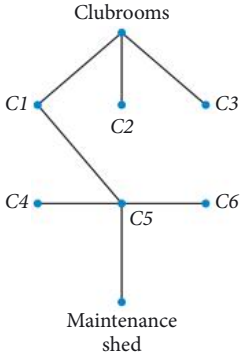


2 The cable from the clubrooms to *C*4 crosses two other cables, and so does the cable from the clubrooms to *C*6. Removing these two cables will allow no cables to be crossed over.

9780170250283

Therefore, remove the cables from the clubrooms to *C*4 and *C*6.

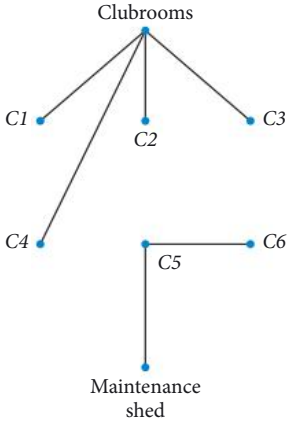
3 a



b 14 + 10 + 15 + 18 + 17 + 16 + 13 = 103 m

4 a There is only one court connected to the maintenance shed and that is *C*5. *C*6 is closer to *C*5 than *C*4, and *C*4 is closer to the clubrooms than *C*6. Therefore, courts *C*5 and *C*6 are to be connected to the generator.

b



c Length of the new cables = 14 + 10 + 15 + 24 + 16 + 13 = 92 m.

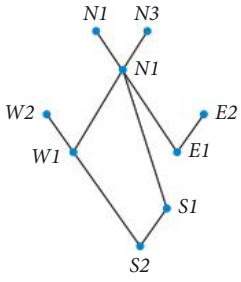
Therefore, the amount saved = 103 − 92 = 11 m.

CHAPTER 5 REVISION

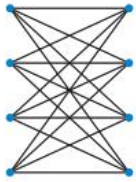
1 D
2 B
3 C
4 D
5 C
6 E
7 A
8 C
9 B
10 D
11 B
12 A
13 C
14 D
15 C
16 a *E* **b** *B* **c** *D* **d** *C, D*
17 a 3 **b** *T* and *V* **c** *T* **d** *W*
18 a *V* = {*N*1, *N*2, *N*3, *E*1, *E*2, *S*1, *S*2, *W*1, *W*2}

b *E* = {*N*1–*N*2, *N*1–*N*3, *N*1–*E*1, *N*1–*S*1, *N*1–*W*1, *E*1–*E*2, *S*1–*S*2, *S*2–*W*1, *W*1–*W*2}

c



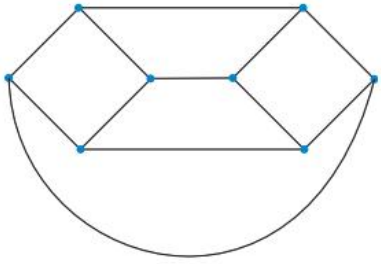
19 Team 1 Team 2



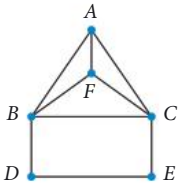
20 a deg *A* = 3, deg *B* = 3, deg *C* = 3, deg *D* = 4, deg *E* = 3, deg *F* = 4

b 20

21



22 a



b 6 = 9 − 5 + 2

23 a (*C, A, E, G, F, E, D, F, B, D, A, B, C*)

b All vertices are of even degree.

c No. This graph is not semi-Eulerian and therefore the trail cannot start and finish at different points.

24 a (*D, E, G, F, H, K, L, J, I, D*)

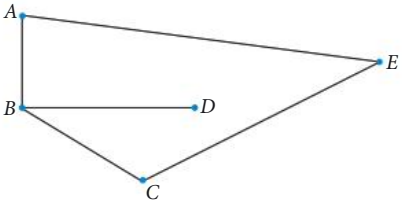
b Hamiltonian cycle.

25 a

	A	*B*	*C*	*D*	*E*
A	0	1	0	0	1
B	1	0	1	1	0
C	0	1	0	0	1
D	0	1	0	0	0
E	1	0	1	0	0

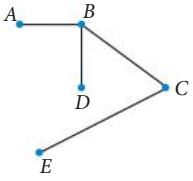
A 1 in the adjacency matrix indicates an edge joining two vertices.

One possible diagram is:



b The graph is not a tree because it contains a circuit *A-B-C-E-A*.

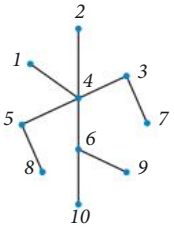
c A possible spanning tree is



d The adjacency matrix for the graph in **c** is

$$\begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

26 a



b 1–5, 6–8, 6–7, 7–9, 9–10.

27 21.6 m

28 a 4 **b** 28 km

c i Hamiltonian cycle

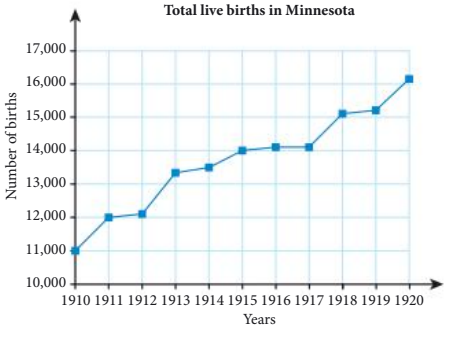
ii (*W, B, A, F, E, D, C, W*)

d i Two vertices now have odd degrees

ii Open trail.

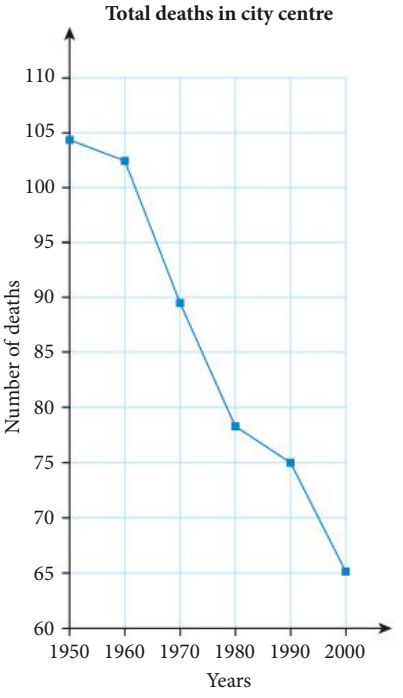
6.01

1 a



b positive secular trend

2 a

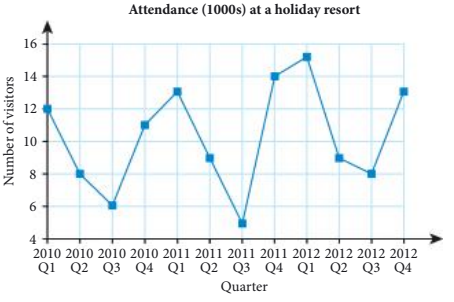


b negative secular trend

3 a C **b** C **c** C **d** E

e A **f** E **g** C **h** A

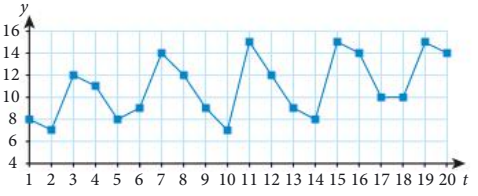
4 a



b The points reach a peak every 4 quarters, matching the summer season. Seasonal pattern.

5 a C **b** E **c** B **d** A

6



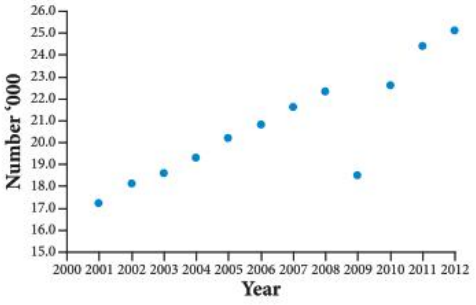
Shows a seasonal pattern with a positive trend.

7 a Neither positive nor negative secular trend.

b Cyclic pattern

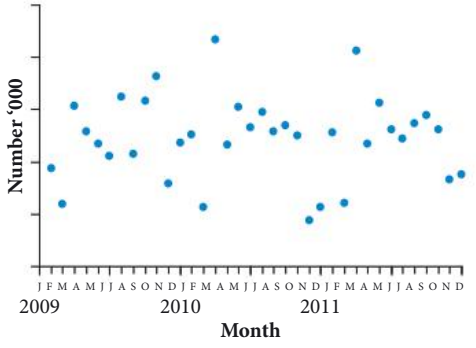
9780170250283

8



There is a positive secular trend, but the year 2009 (and possible 2010) is an outlier.

9



Births appear to have a seasonal pattern.

6.02

1 A

2 a

Year, t	Number of female students, y	Three-point moving average
1	2	
2	1	$\frac{2+1+3}{3}=2$
3	3	$\frac{1+3+5}{3}=3$
4	5	$\frac{3+5+9}{3}=5.7$
5	9	$\frac{5+9+10}{3}=8$
6	10	$\frac{9+10+4}{3}=7.7$
7	4	$\frac{10+4+6}{3}=6.7$
8	6	$\frac{4+6+9}{3}=6.3$
9	9	$\frac{6+9+12}{3}=9$
10	12	$\frac{9+12+15}{3}=12$
11	15	

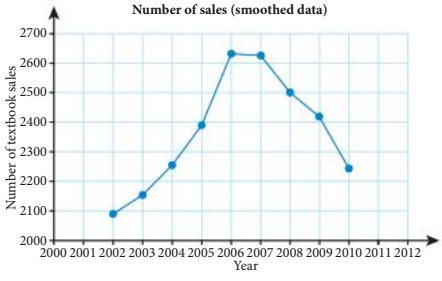
b positive secular trend.

3 a 61 b 61.33 c 74.67 d 66.4
e 46.25 f 78 g 60.25 h 62.125

4 a

Year, t	Sales of textbooks, y	Four-point moving average	4-moving average with centring
2000	2250		
2001	2230		
		$\frac{2250+2230+2000+2010}{4}=2122.5$	
2002	2000		2090
		$\frac{2230+2000+2010+1990}{4}=2057.5$	
2003	2010		2153.75
		$\frac{2000+2010+1990+3000}{4}=2250$	
2004	1990		2255.625
		$\frac{2010+1990+3000+2045}{4}=2261.25$	
2005	3000		2383.625
		$\frac{1990+3000+2045+2989}{4}=2506$	
2006	2045		2632.25
		$\frac{3000+2045+2989+3000}{4}=2758.5$	
2007	2989		2627.25
		$\frac{2045+2989+3000+1950}{4}=2496$	
2008	3000		2505.375
		$\frac{2989+3000+1950+2120}{4}=2514.75$	
2009	1950		2423
		$\frac{3000+1950+2120+2255}{4}=2331.25$	
2010	2120		2243.375
		$\frac{1950+2120+2255+2297}{4}=2155.5$	
2011	2255		
2012	2297		

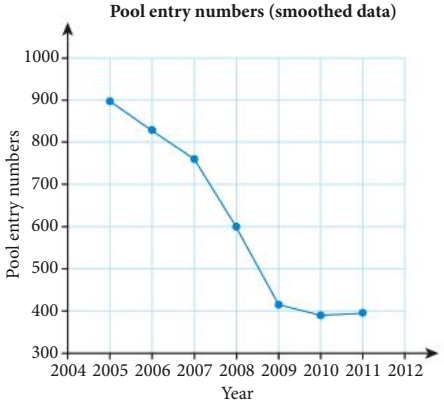
b



5 a

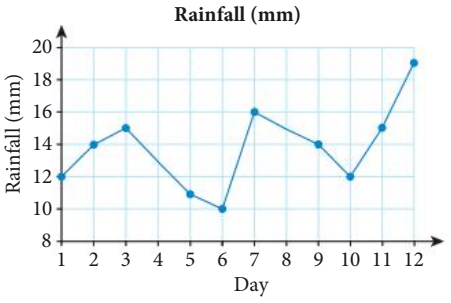
Year, t	Entries at pool, y	Two-point moving average	2-moving average with centring
2004	1000		
		$\frac{1000+889}{2}=944.5$	
2005	889		886.5
		$\frac{889+768}{2}=828.5$	
2006	768		825.25
		$\frac{768+876}{2}=822$	
2007	876		769
		$\frac{876+556}{2}=716$	
2008	556		605
		$\frac{556+432}{2}=494$	
2009	432		430
		$\frac{432+300}{2}=366$	
2010	300		382.5
		$\frac{300+498}{2}=399$	
2011	498		381
		$\frac{498+228}{2}=363$	
2012	228		

b



c Negative secular trend and the councillors should be worried about attendance numbers.

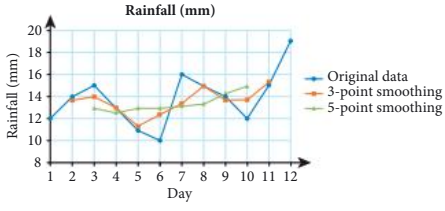
6



b

Day, t	Rainfall (mm), y	Three-point moving average	Five-point moving average
1	12		
2	14	$\frac{12+14+15}{3}=13.67$	
3	15	$\frac{14+15+13}{3}=14$	$\frac{12+14+15+13+11}{5}=13$
4	13	$\frac{15+13+11}{3}=13$	$\frac{14+15+13+11+10}{5}=12.6$
5	11	$\frac{13+11+10}{3}=11.33$	$\frac{15+13+11+10+16}{5}=13$
6	10	$\frac{11+10+16}{3}=12.33$	$\frac{13+11+10+16+15}{5}=13$
7	16	$\frac{10+16+15}{3}=13.67$	$\frac{11+10+16+15+14}{5}=13.2$
8	15	$\frac{16+15+14}{3}=15$	$\frac{10+16+15+14+12}{5}=13.4$
9	14	$\frac{15+14+12}{3}=13.67$	$\frac{16+15+14+12+15}{5}=14.4$
10	12	$\frac{14+12+15}{3}=13.67$	$\frac{15+14+12+15+19}{5}=15$
11	15	$\frac{12+15+19}{3}=15.33$	
12	19		

c



d 13.67 e 13.67 f 13 g 14.4

h The graph shows that the five point moving average has fewer peaks and troughs than the three point moving average.

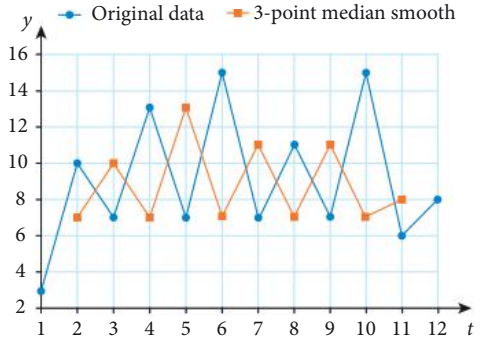
6.03

1 a E b E c D
d A e D f C

2 (1, 3), (2, 10), (3, 7), (4, 13), (5, 7), (6, 15), (7, 7), (8, 11), (9, 7), (10, 15), (11, 6), (12, 8)

9780170250283

3



4

Day, t	Interest rate (%)	Four-point moving median	Centred
1	3.2		
2	4.1		
		median of (3.2, 4.1, 4.6, 3.25) = 3.675	
3	4.6		3.9875
		median of (4.1, 4.6, 3.25, 4.5) = 4.3	
4	3.25		
5	4.5		

Answer = 3.9875

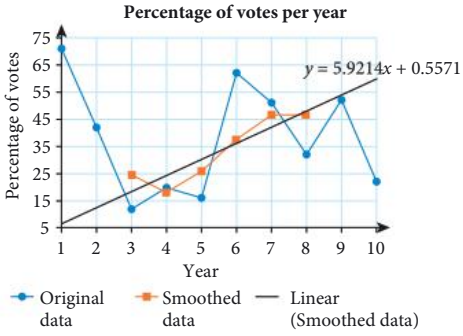
5

Day, t	Entry to museum	Three-point moving median
1	320	
2	420	median of (320, 420, 460) = 420
3	460	median of (420, 460, 325) = 420
4	325	median of (460, 325, 450) = 450
5	450	median of (325, 450, 421) = 421
6	421	

a 420 b 450 c 421

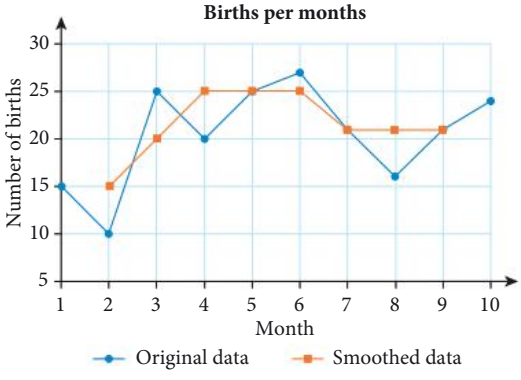
6

Year, t	Votes, y (% per pop'n)	Four-point moving median	Centred
1	71		
2	42		
		median of (71, 42, 12, 20) = 31	
3	12		24.5
		median of (42, 12, 20, 16) = 18	
4	20		18
		median of (12, 20,16, 62) = 18	
5	16		25.75
		median of (20, 16, 62, 51) = 33.5	
6	62		37.5
		median of (16, 62, 51, 32) = 41.5	
7	51		46.5
		median of (62, 51, 32, 52) = 51.5	
8	32		46.5
		median of (51, 32, 52, 22) = 41.5	
9	52		
10	22		



Shows secular positive trend. Equation found is
$y = 5.92t + 0.56$
Substitute $t = 11$ gives $y = 65.69$.

7 a b

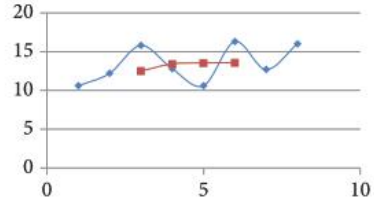


c The data is levelling about at about 21 births per month.

8 a

	med	ctr
	12.5	
t_3	12.5	12.5
t_4	14.3	13.4
t_5	12.75	13.53
t_6	14.35	13.55

b



c sales = 11.8 + 0.32 × year number

d t_7 = 1 468 000

6.04

1

Year	Q1	Q2	Q3	Q4
2002	0.71	0.86	1.08	1.35
2003	0.66	0.85	1.23	1.26
2004	0.55	0.91	1.45	1.09
SI	0.64	0.87	1.25	1.23

a E b C c D d A e B

2

Year	Q1	Q2	Q3	Q4	Yearly average
2011	11	8	1	3	$\frac{11+8+1+3}{4}=5.75$
2011 Step 2	$\frac{11}{5.75}=1.913$	$\frac{8}{5.75}=1.391$	$\frac{1}{5.75}=0.174$	$\frac{3}{5.75}=0.522$	
2012	9	9	3	1	$\frac{9+9+3+1}{4}=5.50$
2012 Step 2	$\frac{9}{5.50}=1.636$	$\frac{9}{5.50}=1.636$	$\frac{3}{5.50}=0.545$	$\frac{1}{5.50}=0.182$	
2013	4	9	1	7	$\frac{4+9+1+7}{4}=5.25$
2013 Step 2	$\frac{4}{5.25}=0.762$	$\frac{9}{5.25}=1.714$	$\frac{1}{5.25}=0.190$	$\frac{7}{5.25}=1.333$	
SI	1.437	1.580	0.303	0.679	

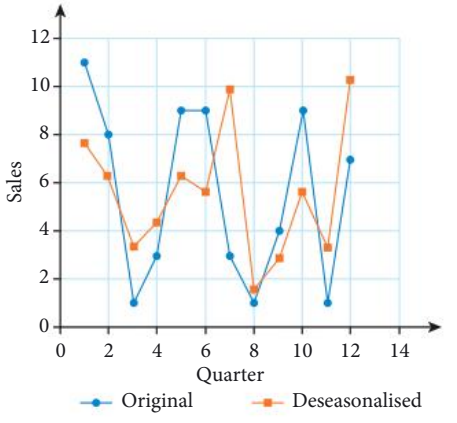
a

Year	Q1	Q2	Q3	Q4
SI	1.437	1.580	0.303	0.679

b

Year	Q1	Q2	Q3	Q4
2011	7.655	6.263	3.300	4.418
2012	6.263	5.696	9.901	1.473
2013	2.784	5.696	3.300	10.309

c



3

Year	Q1	Q2	Q3	Q4
SI	0.643	1.12	1.583	0.653

4 summer $2560
autumn $10 000
winter $10 400
spring $9 040

5 873.68

6 1508

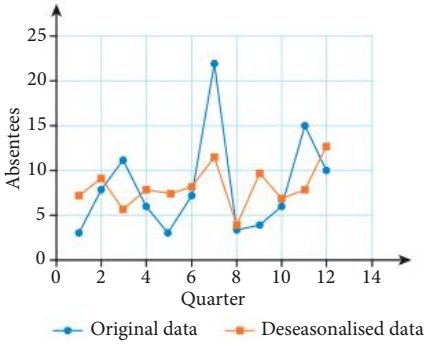
7 a

Year	Q1	Q2	Q3	Q4	Yearly average
2009	3	8	11	6	$\frac{3+8+11+6}{4}=7$
2009 Step 2	$\frac{3}{7}=0.429$	$\frac{8}{7}=1.143$	$\frac{11}{7}=1.571$	$\frac{6}{7}=0.857$	
2010	3	7	22	3	$\frac{3+7+22+3}{4}=8.75$
2010 Step 2	$\frac{3}{8.75}=0.343$	$\frac{7}{8.75}=0.800$	$\frac{22}{8.75}=2.514$	$\frac{3}{8.75}=0.343$	
2011	4	6	15	10	$\frac{4+6+15+10}{4}=8.75$
2011 Step 2	$\frac{4}{8.75}=0.457$	$\frac{6}{8.75}=0.686$	$\frac{15}{8.75}=1.714$	$\frac{10}{8.75}=1.143$	
SI	0.410	0.876	1.933	0.781	

b Deseasonalise the time series.

Year	Q1	Q2	Q3	Q4
2009	7.317	9.132	5.691	7.682
2010	7.317	7.991	11.381	3.841
2011	9.756	6.849	7.760	12.804

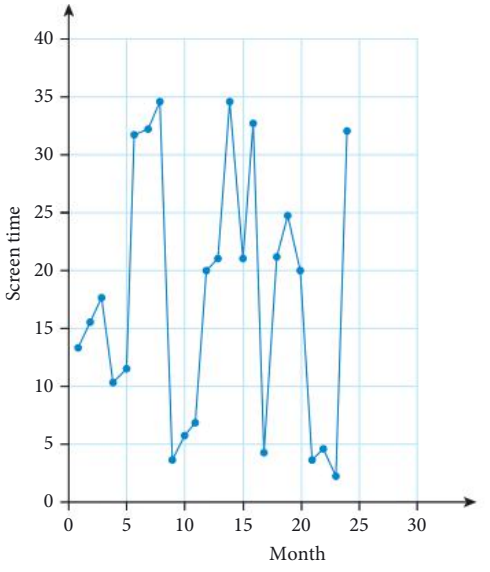
c



d There is no clear pattern in absenteeism. Peaks and troughs occur randomly.

9780170250283

8 a



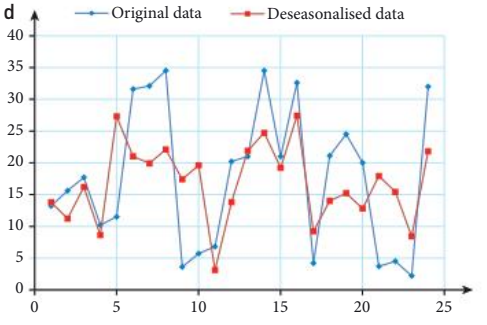
b

Year	J	F	M	A	M	J	J	A	S	O	N	D
SI	0.96	1.398	1.094	1.186	0.455	1.508	1.615	1.564	0.207	0.291	0.261	1.466

c

Year	J	F	M	A	M	J	J	A	S	O	N	D
2010	13.2 / 0.96 =13.8	~~15.6 =11.2~~ 15.6 / 1.398	17.7 / 1.094 =16.2	10.2 / 1.186 =8.6	11.5 / 0.455 =27.3	31.6 / 1.508 =21.0	32.1 / 1.615 =19.9	34.5 / 1.564 =22.1	3.6 / 0.207 =17.4	5.7 / 0.291 =19.6	6.8 / 0.261 =3.1	20.2 / 1.466 =13.8
2011	21 / 0.96 =21.9	34.5 / 1.398 =24.7	21 / 1.094 =19.2	32.6 / 1.186 =27.4	4.2 / 0.455 = 9.2	21.1 / 1.508 =14.0	24.5 / 1.615 =15.2	20 / 1.564 =12.8	3.7 / 0.207 =17.9	4.5 / 0.291 =15.4	2.2 / 0.261 =8.4	32 / 1.466 =21.8

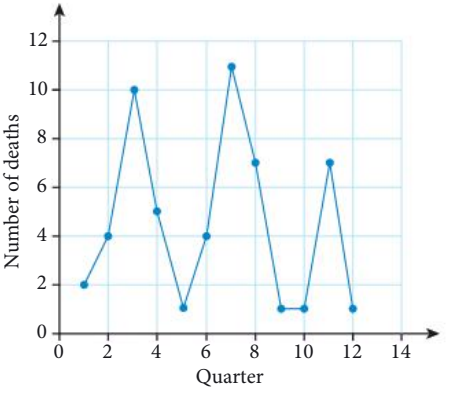
d



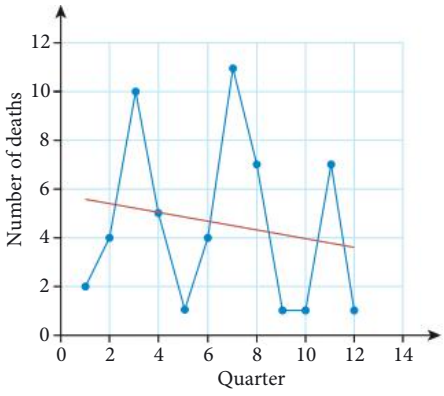
e Screen time appears to increase during the cooler months. Possibly due to it being cold, more time is spent indoors.

6.05

1 a



b



c $y = -0.147\,t + 5.455$

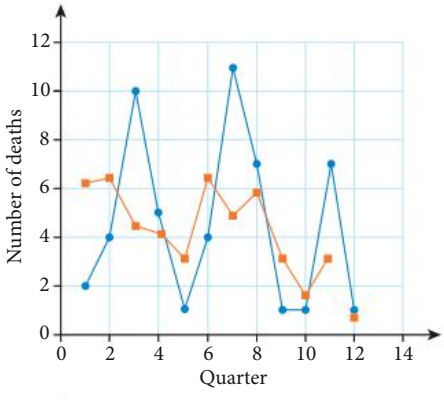
d At $t = 13$, $y = 3.544$. So the number of deaths predicted in the first quarter of 2003 is 4 .

2 a The number of deaths seems to be significantly higher in Q_3, perhaps deaths among the elderly during the winter months.

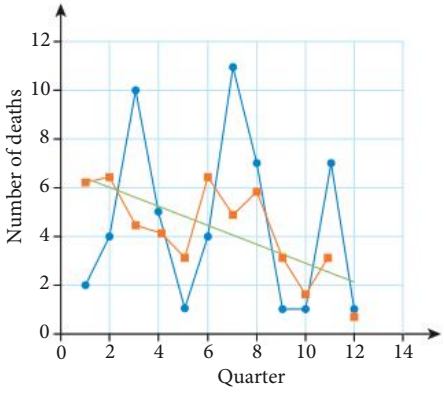
b

Year Deseasonalised	Q1	Q2	Q3	Q4
2000	6.289	6.462	4.533	5.841
2001	3.145	6.462	4.986	8.178
2002	3.145	1.616	3.173	1.168

c



d

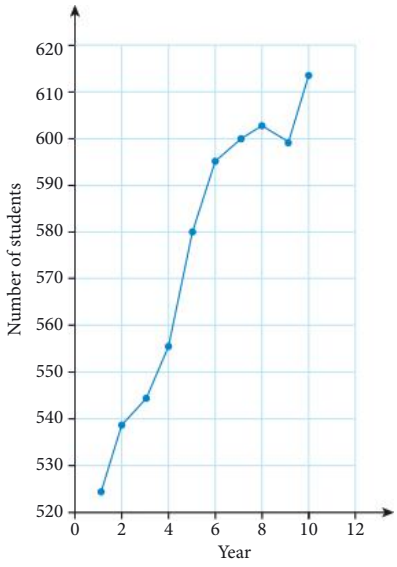


e $y = -0.379\,t + 6.694$

f The answer is 0.56. so between 0 and 1 death.

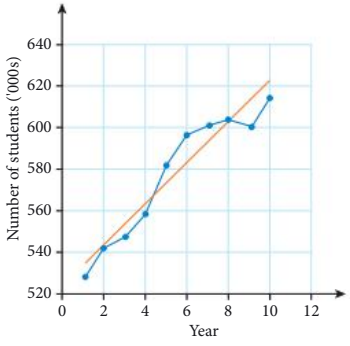
g Discuss with your teacher.

3 a



b positive secular trend

c



$y = 10.091t + 520.4$

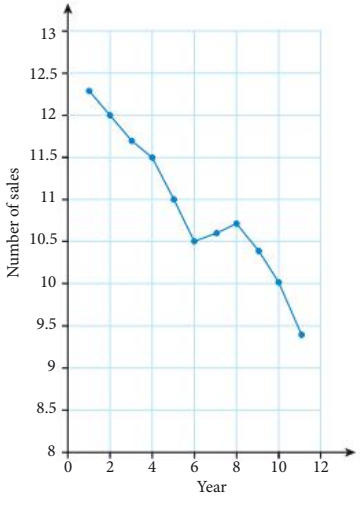
On average the number of students is increasing by 10 000 per year.

d 2008 is year 17. Predicted number of students = 692 000 (rounded).

4 a $y = 1.59\,t + 50.91$

b Forecast sales are approximately 90.

5 a

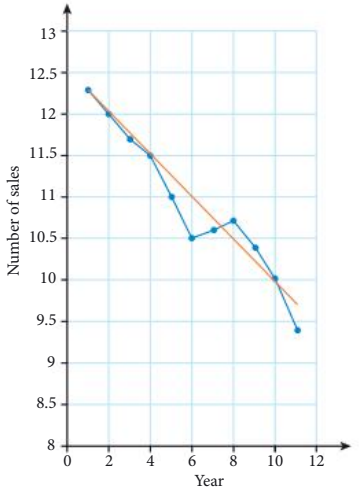


b negative secular trend

9780170250283

c $y = -0.26t + 12.5$
On average the % of sales are decreasing by 0.26% per year.

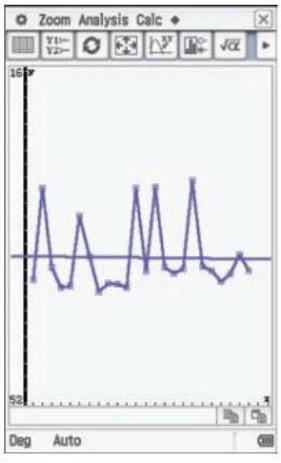
d



e 8.6%

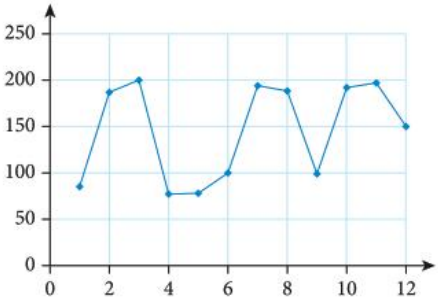
6 a $y = -0.122t + 211.547$

b

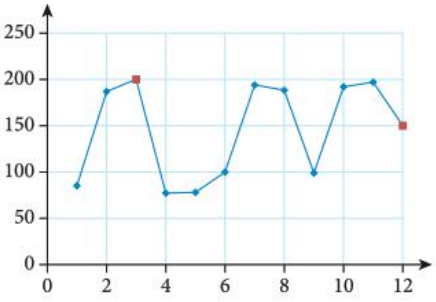


c $180.07

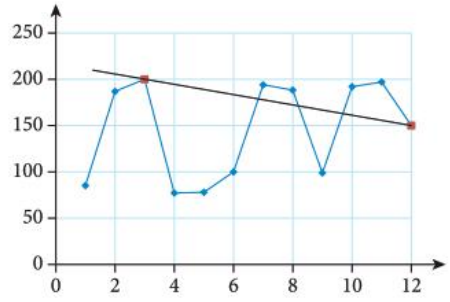
7



8 a



b $y = -5.56t + 216.67$



c 144.4

9 a $y = 4.479t + 116.509$

b Boxing Day in 2014 will be $t = 13$. Sales = $174.74

10 a E b B

CHAPTER 6 REVISION

1 A
2 A
3 E
4 C
5 C
6 E
7 D
8 A
9 E
10 E

11 a

Year	Total deaths	Three point moving average
1910	104	
1911	123	$\frac{104+123+200}{3}=142.33$
1912	200	$\frac{123+200+239}{3}=187.33$
1913	239	$\frac{200+239+204}{3}=214.33$
1914	204	$\frac{239+204+423}{3}=288.67$
1915	423	$\frac{204+423+500}{3}=375.67$
1916	500	$\frac{423+500+469}{3}=464$
1917	469	

b 142.33

12 a

Year, t	Number of tickets sold, y	Two point moving average	2-moving average with centring
Jan	100		
		$\frac{100+120}{2}=110$	
Feb	120		100
		$\frac{120+60}{2}=90$	
Mar	60		72.5
		$\frac{60+50}{2}=55$	
Apr	50		95
		$\frac{50+220}{2}=135$	
May	220		167.5
		$\frac{220+180}{2}=200$	
Jun	180		177.5
		$\frac{180+130}{2}=155$	
Jul	130		127.5
		$\frac{130+70}{2}=100$	
Aug	70		90
		$\frac{70+90}{2}=80$	
Sep	90		87.5
		$\frac{90+100}{2}=95$	
Oct	100		92.5
		$\frac{100+80}{2}=90$	
Nov	80		102.5
		$\frac{80+150}{2}=115$	
Dec	150		

b 90

13

Year, t	Number of tickets sold, y	Four- point moving average	4-moving average with centring
Jan	100		
Feb	120		
		$\frac{100+120+60+50}{4}=82.5$	
Mar	60		97.5
		$\frac{120+60+50+220}{4}=112.5$	
Apr	50		120
		$\frac{60+50+220+180}{4}=127.5$	
May	220		136.25
		$\frac{50+220+180+130}{4}=145$	
Jun	180		147.5
		$\frac{220+180+130+70}{4}=150$	
Jul	130		133.75
		$\frac{180+130+70+90}{4}=117.5$	
Aug	70		107.5
		$\frac{130+70+90+100}{4}=97.5$	
Sep	90		91.25
		$\frac{70+90+100+80}{4}=85$	
Oct	100		95
		$\frac{90+100+80+150}{4}=105$	
Nov	80		
Dec	150		

b 107.5

9780170250283

14

Year, t	Sales, y ($)	Four point moving median	Centred
1	73		
2	138		
		median of (73, 138, 114, 112) = 113	
3	114		119.5
		median of (138, 114, 112, 160) = 126	
4	112		122
		median of (114, 112, 160, 122) = 118	
5	160		
6	122		

The least squares trend line is $y = 2.5t + 112$
At $t = 7$ expect sales to be \$129 500.

15

		Data	3-point moving average	4-point moving average smoothing	4-point moving average centred
1991	J	45			
	F	39	(45 + 39 + 70)/3 = 51.33		
				(45 + 39 + 70 + 30)/4 = 46.00	
	M	70	(39 + 70 + 30)/3 = 46.33		(46.00 + 52.25)/2 = 49.13
				(39 + 70 + 30 + 70)/4 = 52.25	
	A	30	(70 + 30 + 34)/3 = 44.67		(52.25 + 49.75)/2 = 51.00
				(70 + 30 + 34 + 29)/4 = 49.75	
	M	34	(30 + 34 + 29)/3 = 31.00		(49.75 + 43.25)/2 = 46.50
				(30 + 34 + 29 + 44)/4 = 43.25	
	J	29	(34 + 29 + 44)/3 = 35.67		(43.25 + 47.50)/2 = 45.38
				(34 + 29 + 44 + 47)/4 = 47.50	
	J	44	(29 + 44 + 47)/3 = 40.00		(47.50 + 40.25)/2 = 43.88
				(29 + 44 + 47 + 41)/4 = 40.25	
	A	47	(44 + 47 + 41)/3 = 44.00		(40.25 + 41.75)/2 = 41.00
				(44 + 47 + 41 + 35)/4 = 41.75	
1991	S	41	(47 + 41 + 35)/3 = 41.00		(41.75 + 39.50)/2 = 40.63
				(47 + 41 + 35 + 35)/4 = 39.50	
	O	35	(41 + 35 + 35)/3 = 37.00		(39.50 + 41.00)/2 = 40.25
				(41 + 35 + 35 + 53)/4 = 41.00	
	N	35	(35 + 35 + 53)/3 = 41.00		(41.00 + 39.25)/2 = 40.13
				(35 + 35 + 53 + 34)/4 = 39.25	
	D	53	(35 + 53 + 34)/3 = 40.67		(39.25 + 38.75)/2 = 39.00
				(35 + 53 + 34 + 33)/4 = 38.75	

		Data	3-point moving average	4-point moving average smoothing	4-point moving average centred
1992	J	34	(53 + 34 + 33)/3 = 40.00		(38.75 + 38.50)/2 = 38.63
				(53 + 34 + 33 + 34)/4 = 38.50	
	F	33	(34 + 33 + 34)/3 = 33.67		(38.50 + 37.00)/2 = 37.75
				(34 + 33 + 34 + 47)/4 = 37.00	
	M	34	(33 + 44 + 47)/3 = 41.33		(37.00 + 36.00)/2 = 36.5
				(33 + 44 + 47 + 30)/4 = 36.00	
	A	47	(34 + 47 + 30)/3 = 37.00		(36.00 + 35.00)/2 = 35.50
				(34 + 47 + 30 + 29)/4 = 35.00	
	M	30	(47 + 30 + 29)/3 = 35.33		(35.00 + 32.25)/2 = 33.63
				(47 + 30 + 29 + 23)/4 = 32.25	
	J	29	(30 + 29 + 23)/3 = 27.33		(32.25 + 27.00)/2 = 29.63
				(30 + 29 + 23 + 26)/4 = 27.00	
	J	23	(29 + 23 + 26)/3 = 26.00		(27.00 + 26.50)/2 = 26.75
				(29 + 23 + 26 + 28)/4 = 26.50	
	A	26	(23 + 26 + 28)/3 = 25.67		(26.50 + 28.75)/2 = 27.63
				(23 + 26 + 28 + 38)/4 = 28.75	
	S	28	(26 + 28 + 38)/3 = 30.67		(28.75 + 30.50)/2 = 29.63
				(26 + 28 + 38 + 30)/4 = 30.50	
	O	38	(28 + 38 + 30)/3 = 32.00		(30.50 + 34.00)/2 = 32.25
				(28 + 38 + 30 + 40)/4 = 34.00	
	N	30	(38 + 30 + 40)/3 = 36.00		
	D	40			

16 **a** $y = 0.08t + 26$ **b** \$28 000 000
c \$28 960 000

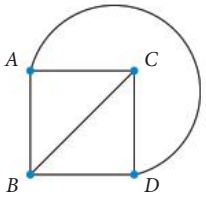
MIXED REVISION 2

Multiple choice

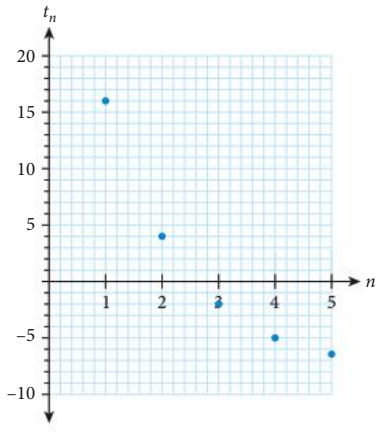
1 E
2 C
3 A
4 B
5 C
6 C
7 D
8 D
9 A

Short answer questions

1 1.26
2 13
3



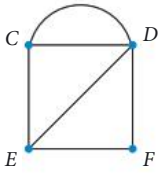
4 **a** $t_{n+1} = t_n - 5$ where $t_1 = -1$
b $t_{n+1} = 1.6t_n$ where $t_1 = 4$

5 a 

b The sequence has a steady state solution of −8.

6 *G–C–B–D–F–A–E*

Analysis answer questions

1 a 

b

	C	*D*	*E*	*F*
C	0	2	1	0
D	2	0	1	1
E	1	1	0	1
F	0	1	1	0

c Vertex *C* or Vertex *E* (Chelsea's house or Evelyn's house)

d *C–D–F–E–D–C–E* or *E–C–D–E–F–D–C*

2 a $B_{n+1} = 1.15B_n$ where $B_0 = 420$

b Region *A*: 360 Region *B*: 483

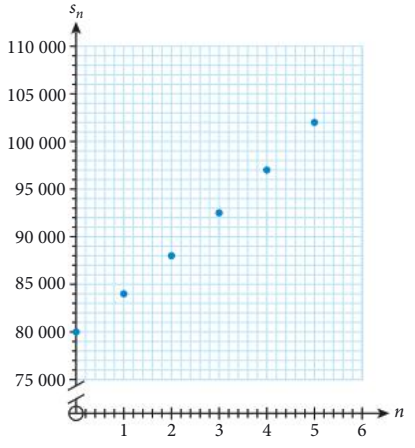
c i 2015 ii 845 iii 236

d $r = 1.15$, $d = -120$, $a = 845$

e increases indefinitely

f 1200

3 a $s_{n+1} = 1.05s_n$ where $s_0 = 80\,000$

b 

c \$97 240.50 d 2019 e \$544 153.03

4 a 1161 b 1.30

c In the summer sales are about half of the average for the year.

d The upward trend is seasonal.

5 a 858 Summer of 2010, 1026 Autumn of 2011.

b deseasonalised sales = 838.80 + 31.94 × quarter number

6 a $N = 839 + 32t$
N for winter of 2013, $t = 15$, gives 1319

b Number of blanket sales for winter of 2013: $N = 1715$

7 a i n
ii $n - 1$ iii 1

b $v - e + f = 2$
LHS = $n - (n - 1) + 1 = 2$ = RHS

7.01

1 a \$387.33 b \$2880

c \$882 d \$278.43

2 \$19 687.50

3 a \$25 875 b \$70 875

4 a \$130 742.50 b \$20 935.20

c \$45 407.93 d \$1290.52

5 C

6 E

7 \$10 000

8 a \$2500 b 9.3%

9 C

10 \$506.25

11 \$114

12 1.7% per quarter

13 0.0156% per day

14 $I = \dfrac{Prn}{100}$
$= \dfrac{12000 \times 7.5 \times n}{100}$
$= \dfrac{90000 \times n}{100}$
$= 900n$

15 The first would give her \$1440 interest whilst the second would give her \$1400. So the first would give her the best return.

16 a 53 days b \$18.47 c \$667.47

17 a \$7800 b \$31 800 c \$244.62

18 \$287.47

7.02

1 a \$7380 b \$ 18.6%

2 E

3 E

4 a i Plan B. They pay \$6942 interest but in plan A only \$5820 interest.

ii Plan B. They have monthly instalments of \$550.88 instead of \$595.

b Plan A requires them to have a deposit of $3900 and whilst the repayments per month are $44.12 more, they are only in debt for 3 years. Plan B allows them to go in with no money but puts them in debt for 4 years and costs $1122 more altogether.

5 a $7836.80 b $6930
c 11.9%
d Popularity of car, excess stock cheaper, temporary price reduction on cars.

6 a A: $6124.44; B: $5227.08
b Option B. She would pay $897.36 less overall.
c The deposit for option A is much smaller. If she didn't have enough money for the deposit for Option B, then she might consider option A to be the better one for her. Also, the repayment is smaller which might be of benefit.

7 $233.03

8 a i $1440 ii $2580.48
b i $1350 ii $1622.70
c Lemon loans: the positive is that you have a whole year to save but the negative is that you need a 10% deposit.
Feelgood finance: the positive is that you don't need a deposit but the negative is that you only have 100 days to pay the loan in full.

9 a $1950 b $1694.80
c $2198.50

10 $388.42

11 a No interest; must have money.
b Don't need cash up front; interest paid.
c Have time to save up money; high interest from day of purchase if not repaid on time.

7.03

1 a $510 b $17 510 c $525.30
d $18 035.30 e $1035.30

2 a $58 879.18 b $8879.18

3 $t_{n+1} = 1.038 \times t_n$ where $t_0 = 12\,000$

4 a $t_{n+1} = 1.08 \times t_n$ where $t_0 = 38\,000$
b $t_{n+1} = 1.046 \times t_n$ where $t_0 = 125\,000$
c $t_{n+1} = 1.065 \times t_n$ where $t_0 = 15\,600$
d $t_{n+1} = 1.075 \times t_n$ where $t_0 = 25\,000$

5 D

6 a $t_{n+1} = 1.071 \times t_n$ where $t_0 = 40\,000$
b $79 424.54

7 a $t_{n+1} = 1.045 \times t_n$ where $t_0 = 45\,000$
b $56 078.19

8 a $9488.39 b $2488.39

9 a 3% b $8741.82
c $741.82

10 a $93 614.69 b $30 614.69

7.04

1 a $25 116.74 b $167 555.66
c $110 739.11 d $25 682.70
e $77 032.91 f $15 315.00
g $186 020.02 h $23 290.57

2 a $8 116.74 b $42 555.66
c $57 239.11 d $15 682.70
e $15 032.91 f $6815.00
g $39 020.02 h $1790.57

3 a 4 years b 5.5% p.a.
c 27 quarters = 6 years 9 months
d $34 314

4 a $102 917.31 b $22 917.31

5 a $56 783.69 b $10 783.69

6 D

7 104 months = 8 years 8 months

8 A

9 E

10 C

11 51 months = 4 years 3 months

12 $6800

13 $5961.13

14 9.95% p.a.

15 116 months = 9 years 8 months

16 27.8% p.a.

17 a $410
b The second investment option is a better choice, earning $4802.44 over the 5 years compared with $4100 over the 5 years if the first option is chosen.

18 a 7.6% b $70 896.75
c $25 896.75

19 C

20 a i $87 570.78 ii $87 409.04
b Account A as he gets $161.74 more at the end of the 2 years.
c Account A: $110 485.73; Account B: $109 976.26.
∴There is no difference as Account A still gives him more.

21 a i $4800 ii $5868.74
b You earn more with compound interest as you earn interest on your principle and your interest whereas with simple interest you only earn interest on your principal.

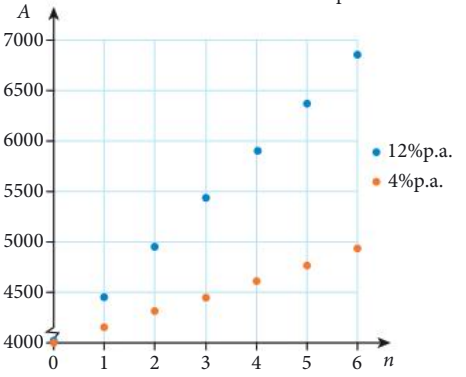
22 a i $58 985.82 ii $59 256.41
iii $59 396.82 iv $59 492.42
v $59 539.24
b The amount of interest increases. The interest is adding on more frequently and hence, you are earning interest on your interest for longer.

7.05

1 a $3080 b $3440 c $3800

2 a 10 years b 13 years c 17 years

3 Value of investment at different simple interest rates



4 a i $5600 ii $7840

b The higher rate of interest has a steeper graph.

c 160

d 480

e The simple interest earned each year.

5 a $2840 b 8% c 9 years

d 12 years e $1200 f 6%

6 a $5440 b $6340

7 a 14 years b 9 years

8 a i

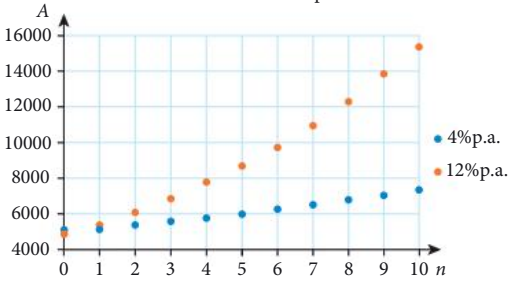
n	0	1	2	3	4	5	6	7	8	9	10
A($)	5000	5200	5408	5624	5849	6083	6327	6580	6843	7117	7401

ii

n	0	1	2	3	4	5	6	7	8	9	10
A($)	5000	5600	6272	7025	7868	8812	9869	11 053	12 380	13 865	15 529

b

Value of investment at different compound interest rates

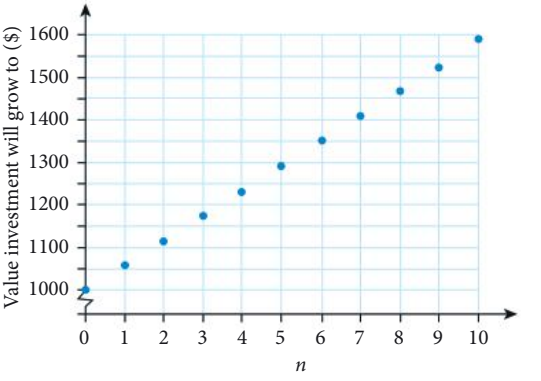


9 a i approximately $9000

ii approximately $19 500

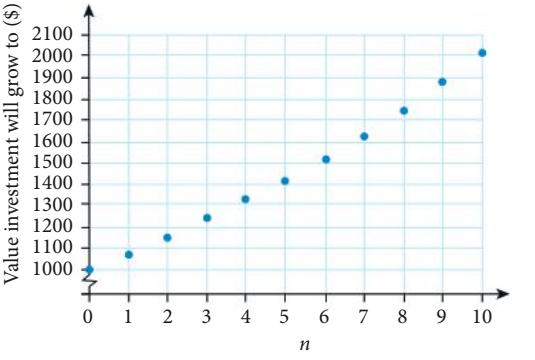
b The higher rate of interest has a steeper graph.

10 a i



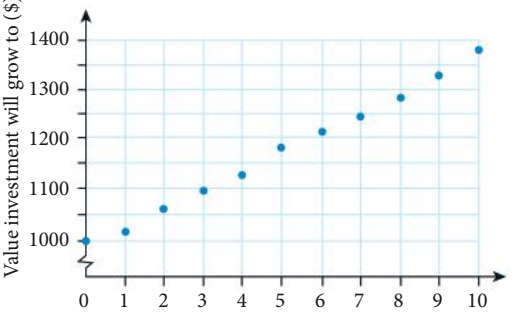
ii $1870

b i



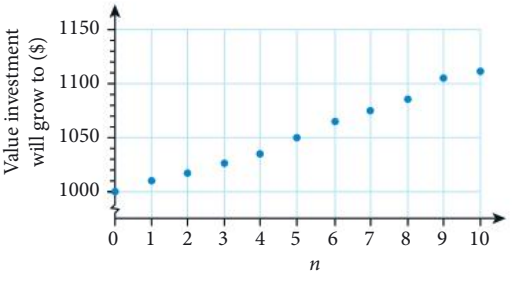
ii $2840

c i



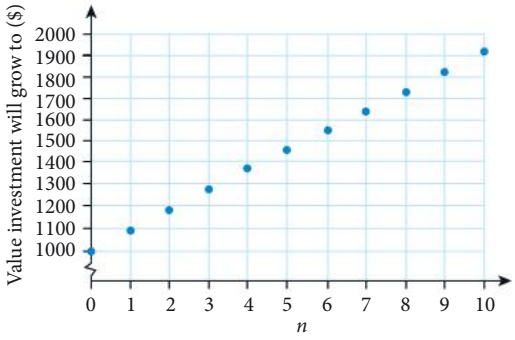
ii $1880

d i



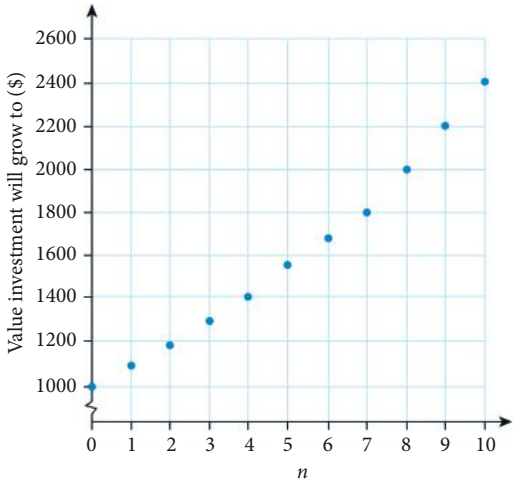
ii $1300

11 a i



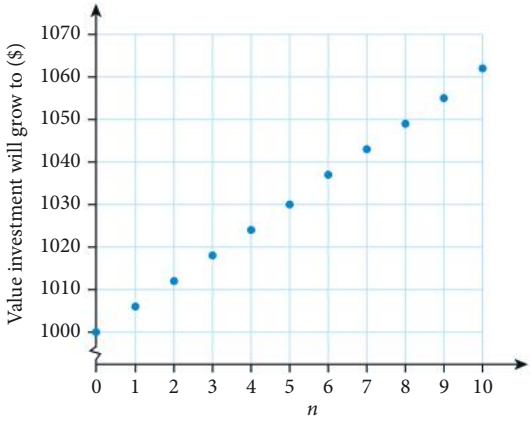
ii 17 years

b i



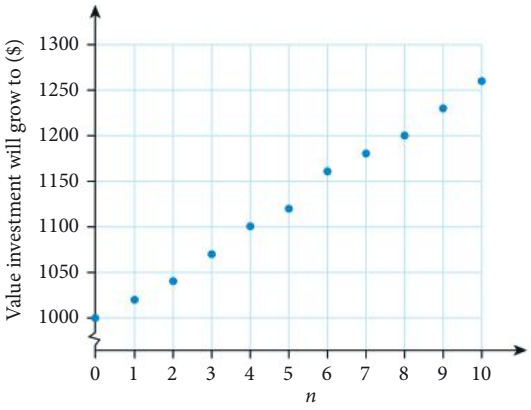
ii 12 years

c i



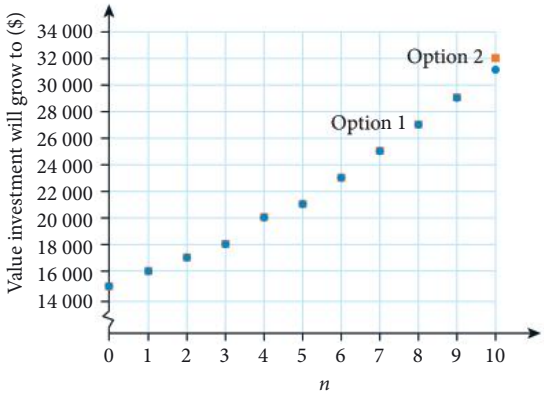
ii 2 years or 24 months

d i



ii 4.5 years or 18 quarters

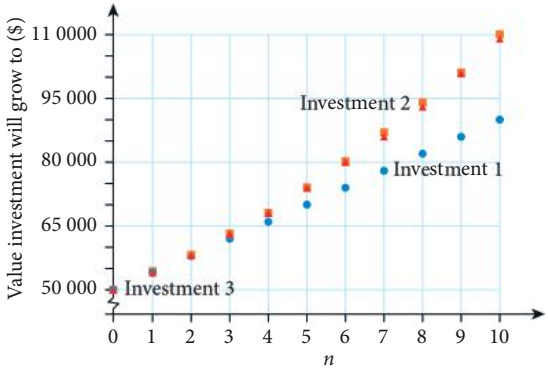
12 a



b Matthew would be best off choosing option 2.

13 a Yes. Investment 2 as they are all at the same interest rate, but it is compounded more frequently and so will grow faster.

b



c Answers will vary.

d Investment 2 is the best option.

9780170250283

14 a i $5000 ii $A = P + \frac{Prn}{100}$

iii $7250 = 5000 + \frac{5000 \times 6 \times r}{100}$, $r = 7.5\%$

iv Annual interest = $375

b i $3000 ii $A = P\left(1 + \frac{r}{100}\right)^n$

iii $7435 = 3000\left(1 + \frac{r}{100}\right)^{10}$, $r = 9.5\%$

iv $18 425

15 a The blue points represent simple interest because they are in a linear pattern – the same interest is earned every year. The green points represent compounding interest because they are in an exponential pattern – the interest earned increases each year as the investment increases.

b The simple interest has a higher interest rate, as the first point (and the first few points) are higher than the compound interest points.

c Simple interest formula: $A = P + \frac{Prn}{100}$;

Compound interest formula: $A = P\left(1 + \frac{r}{100}\right)^n$

d Simple interest: $26\,000 = 10\,000 + \frac{10\,000 \times 20 \times r}{100}$,

Compound interest: $31\,769 = 10\,000\left(1 + \frac{r}{100}\right)^{20}$

e Simple interest rate = 8.0%, Compound interest rate = 5.9%.

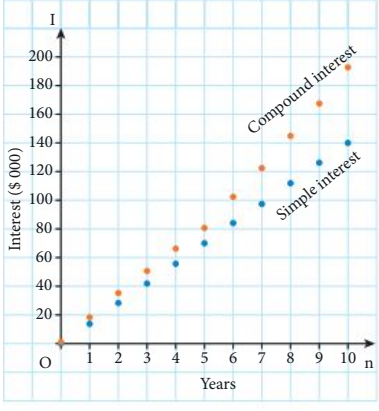
f Simple interest would give an account balance of $34 000 in 30 years, whereas compound interest would give an account balance of $55 831.44 in 30 years.

INVESTIGATION: COMPARING INTEREST

a i $14 000 ii $28 000 iii $70 000
iv $98 000 v $140 000

b i $14 000 ii $28 980 iii $80 510
iv $121 156 v $193 430

c

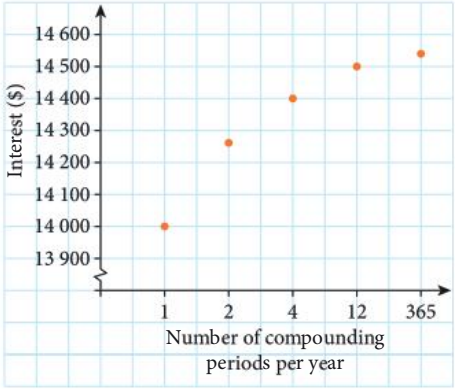


d 7% p.a. compounded annually

e You earn more with compound interest as you earn interest on your principle and your interest whereas with simple interest you only earn interest on your principal.

f i $13 986 ii $14 256 iii $14 397
iv $14 492 v $14 539

g



h 7% p.a. compounded daily

i $14 492 and $14 539

j $14 530. So, in between $14 492 and $14 539.

k The amount of interest increases. The interest is adding on more frequently and hence, you are earning interest on your interest for longer.

l i 7.5% p.a. compounded yearly
ii 6.9% p.a. compounded monthly
iii 5.85% p.a. compounded daily
iv $7\frac{1}{4}$% p.a. compounded half yearly
v 11% p.a. compounded monthly

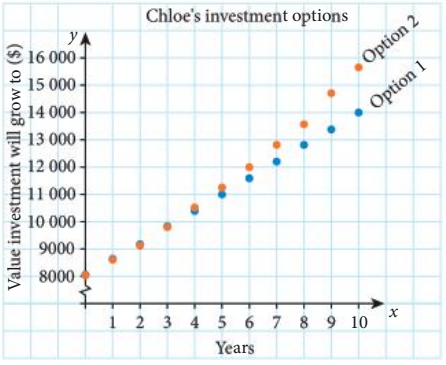
m i $1500 or $1556 ∴ correct
ii $1475 or $1380 ∴ correct
iii $1240 or $1241 ∴ correct
iv $1503 or $1531 ∴ correct
v $2424 or $ 2448 ∴ correct

7.06

1 a 9.38% b 11.61% c 12.36% d 6.18%
2 a 7.76% b 14.34% c 8.67% d 8.32%
3 E
4 7.82%, 7.76%, 7.68%; ∴ 7.6% compounded quarterly is best.
5 8.54%
6 The insurance bonds are best with an effective rate of 7.98%, whereas the others are only 7.79% and 7.9%.
7 a 8.38% b 10.61% c 15.26% d 18.66%
8 The interest rate compounding quarterly is better (6.22% effective).
9 a Interest rates are the same. So one compounded weekly will have the higher effective interest rate as it is compounded more frequently and hence grows faster.
b Weekly is 6.39% whilst quarterly is 6.35%.

CHAPTER 7 REVISION

1 D
2 C
3 D
4 B
5 D
6 A
7 E
8 $142.81
9 $16 200
10 6.07% p.a.
11 a $186.16 b $38.99
12 a $3562.39 b $2949
13 a $t_{n+1} = t_n \times 1.069$ where $t_0 = 15\,000$
b $t_5 = \$20\,940.15$
14 a $2157.25 b $457.25
15 $34 696
16 17 quarters = 4 years 3 months
17 a A b D c C d B
18 a



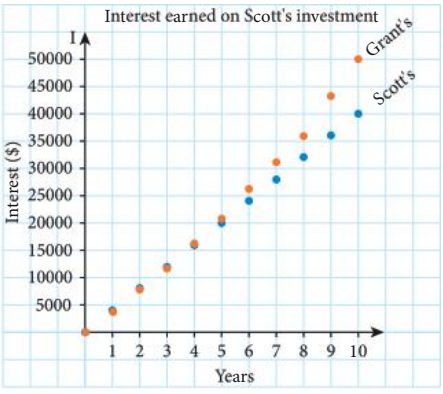
b If Chloe is going to invest for less than 3 years, then she would be better off with option 1, but if she is going to invest for longer than that, then option 2 is better.
19 9.84%
20 a It is an increasing straight line graph.
b i $32 000 ii 2.5 years
c 400 d $400
e $I = 40\,000$ $P = ?$ $r = 8$ $n = 10$

$$I = \frac{Prn}{100}$$
$$40000 = \frac{P \times 8 \times 10}{100}$$
$$P = 50000$$

His initial investment was $50 000.
f $n = 2.62... \approx 3$ years
g i $100 211.57 ii $50 211.57
h

n	0	1	2	3	4	5	6	7	8	9	10
I	0	3600	7500	11 600	16 000	20 800	25 900	31 300	37 200	43 500	50 200

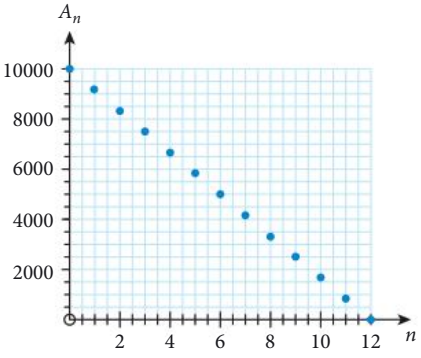
i



j Grant's investment is the best option if invested longer than 5 yrs.
k Effective interest rates: Jodie 7.28%; Grant 7.2%. Jodie made the best investment option.

8.01

1 a $68 454.10 b $1851.30
2 $58 290.84
3 a $A_n = \left(1 + \frac{11}{1200}\right) A_{n-1} - 850$ where $A_0 = 45\,000$
b $18 769.18
4 a $33 276.45 b $10 156.45
5 D
6 a i $48 335.57 ii $47 413.73
iii $46 799.17
b i $2835.57 ii $2813.73
iii $2799.17
c Increasing the size of the repayment reduces the amount of interest paid.
d i $50 179.25
ii The size of the repayment is not sufficient to cover the monthly interest charged on the loan and therefore the amount owing increases instead of decreasing.
7 a Option 1: $34 764.35; Option 2: $34 322.39
b Option 2
8 a



b Over time, the rate at which the loan is paid off increases.

9780170250283

9 a $A_n = \left(1 + \frac{7.5}{1200}\right)A_{n-1} - 180$ where $A_0 = 22\,000$

b 232 months, $121.34

10 a $A_n = \left(1 + \frac{8}{1200}\right)A_{n-1} - 150$ where $A_0 = 16\,500$

b 199 months, $138.56

c $29 838.56

d $A_n = \left(1 + \frac{8}{1200}\right)A_{n-1} - 155$ where $A_0 = 16\,500$

e 187 months, $20.53

f $28 850.53, nearly $1000 less for $5 a month

8.02

1 a $92.63 b $442.63
c $2097.64 d $1514.18

2 a $1948.54 b $607 944.48
c $357 944.48

3 a 96 b 16 c 300 d 48

4 a $38 379.81 b 116

5 a $2771.52
b 5.43 quarters or approximately 16 months

6 a $738.55 b 6 years 11 months

7 a $300.57 b $3034.20 c $12 872.35

8 $965.92

9 $21 808.55

10 $359 367.20

11 C

12 a $185.95 b $4912.80
c Mad Marika's, because the price of $4800 is less.

13 Yes. The quarterly payment required of $1547.42 is less than what she can afford each quarter.

14 a i $1552.96 ii $66 355.20
b i $1253.28 ii $105 590.40
c i $1118.78 ii $148 507.20
d i $1048.68 ii $194 604.00
e The size of the monthly repayment decreases.
f The total amount of interest paid increases.

INVESTIGATION: COMPARING FLAT RATE AND REDUCING BALANCE LOANS

a i $31 000 ii $26 091 iii $4909

b i $395.83 ii $316.69
iii The amount repaid with the reducing balance loan is $9497.20 less.
iv 5.2%

c i Lender 1 ii $1095

8.03

1 a $1829.66 b $2923.20

2 a $1764.29 b $70 571.60
c $20 571.60 d $1744.17
e $482.88

3 a 22 years b $190 100
c $17 400

4 a 12 years (144 months)
b 14 months c $3481.52

5 a $132 998 b 24 years 5 months
c $2096.73 d $169 700

6 a $1276.14 b $306 273.60
c $176 273.60 d $1576.14
e $17 940

7 a i $2418.08 ii $2122.32
iii $1962.78
b i $435 254.40 ii $509 356.80
iii $588 834
c $185 254.40, $259 356.80, $338 834
d i Extending the life of the loan reduces the monthly repayment.
ii Extending the life of the loan increases the amount of interest charged.

8 B

9 a $3955.37 b $1825.56
c One fortnight earlier

10 a $2273.66
b No. The equivalent monthly payment is $2720.40
c Option A is the better option. The total amount repaid is $136 419.60 – $135 955.80 = $463.80 less than with Option B.

8.04

1 a $A_{n+1} = 1.066A_n - 6000$ where $A_1 = 80\,000$
b $74 901.29 c $5280
d 34 years, with a last payment of $1072.02
e 12 years, with a last payment of $7527.99
f 28 years, with a last payment of $3766.27

2 A

3 a $a = 200\,000$, $r = 1.0075$, $d = -2000$
b 187 months
c 93 months, with a last payment of $2299.31
d 166 months, with a last payment of $682.57

4 417 months

5 197 months

6 $426.70

7 $3355.08

8 10 half years or 5 years

9 a $158 120 b $139 858.10 c $13 120

10 $6000

8.05

1 $27 300

2 D

3 $3281.25

4 $66 666.67

5 5.4%

6 $104 166.67

7 $408 000

8 a $525 b $140 000

8.06

1 a

Year	Annual payment ($)	Balance at start of year ($)	Interest ($)	Balance at end of year ($)
1	500	500	40	540
2	500	1040	83.20	1123.20
3	500	1623.20	129.86	1753.06
4	500	2253.06	180.24	2433.30
5	500	2933.30	234.66	3167.96

b $B_{n+1} = 1.08(B_n + 500)$ where $B_1 = 540$.

2 E

3 a $400 b 8.4 c $402.80

4 a $70 139.96 b $17 139.96

5 $7335.93

6 No, only $8962.44

7 C

8 a $39 331.95 b $15 331.95

9 $1456.56

10 $957.07

11 a $45 540.47 b 7 years 8 months

12 $4355.92

CHAPTER 8 REVISION

1 C

2 E

3 B

4 D

5 B

6 D

7 a $19 628.88 b $2201.72

8 a $A_n = \left(1 + \frac{8.25}{1200}\right) A_{n-1} - 300$ where $A_0 = 10\,000$

b $590.73

9 a $3133.74 b $752 097.60

c $302 097.60

10 a 120 months or 10 years

b 9 months c $1073.04

11 a $3454.56 b $1594.41

c 1 fortnight or 2 weeks.

12 a $A_{n+1} = 1.042A_n - 5000$ where $A_1 = 65\,000$

b $22 903.08 c $2730

13 $673.84

14 $42 500

15 a

Year	Annual payment ($)	Balance at start of year ($)	Interest ($)	Balance at end of year ($)
1	5000	5000	240	5240
2	5000	10 240	491.52	10 731.52
3	5000	15 731.52	755.11	16 486.63
4	5000	21 486.63	1 031.36	22 517.99
5	5000	27 517.99	1 320.86	28 838.85
6	5000	33 838.85	1 624.27	35 463.12

b $B_{n+1} = 1.048(B_n + 5000)$ where $B_1 = 5240$.

16 a $64 111.81 b $8111.81

17 a Account B b 3 months

c 41 months d $358.40

e Almost 7 months (6.79 months).

f About 3.3 months longer.

g $1661.21 h $4408.42

18 a $10 625

b Just over 18 years (18.2 years).

c i $134 337.83 ii $276 823.58

d i $67 168.91 ii $129 411.79

e $290 635.62

f i $2 776.39 ii $1740.48

g Just over 8 years (97 months).

9.01

1 D

2 B

3 A

4 a 3 b paths *BD*, *BED*, *BAED*

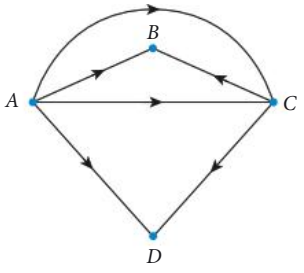
c $R^1 = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$

5 a $R^1 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix} \end{array}$

b $R^2 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix} \end{array}$

9780170250283

6 a



b $R^2 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$

c 3 pathways exist from *A* to *B*.

7

Team	Dominance
W	6
X	1
Y	4
Z	2

b Team *W* was the most dominant.

8 a $D^1 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$

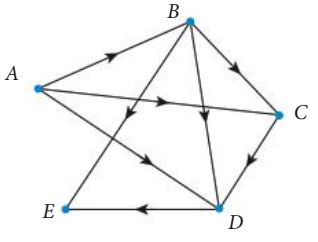
b $D^2 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$

c $D = D^1 + D^2 = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$

d The dominant player is *A*.

9 a $R^1 = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$

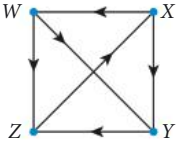
b



c $R^2 = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$

d *ABCDE, ABE, ABDE, ACDE, ADE*

10 a



b $D^1 = \begin{array}{c} \\ W \\ X \\ Y \\ Z \end{array} \begin{array}{c} \begin{array}{cccc} W & X & Y & Z \end{array} \\ \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \end{array}$

c $D^2 = \begin{array}{c} \\ W \\ X \\ Y \\ Z \end{array} \begin{array}{c} \begin{array}{cccc} W & X & Y & Z \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} \end{array}$

d $D = D^1 + D^2 = \begin{array}{c} \\ W \\ X \\ Y \\ Z \end{array} \begin{array}{c} \begin{array}{cccc} W & X & Y & Z \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 2 & 2 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \end{array}$

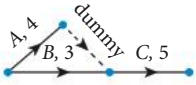
X is the dominant cereal with a dominance score of 5.

9.02

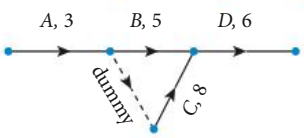
1 B

2 B

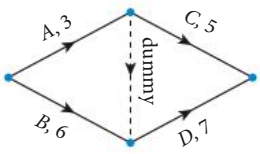
3 a the dummy activity is shown after activity *A*. It can also appear before activity *A*, before activity *B* or after activity *B*.



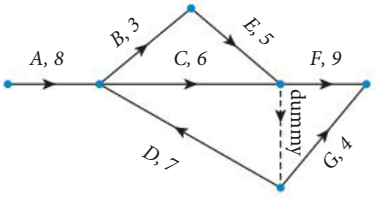
b



c



d

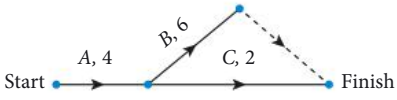


4 **a** A **b** none **c** *C, D* **d** *F, G, H*

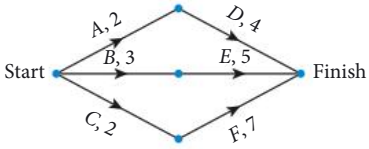
5

Activity	Time (hours)	Predecessor/s
A	3	None
B	6	*A*
C	3	*A*
D	8	*C*
E	7	*B*
F	4	*B*
G	2	*F*
H	2	*E*

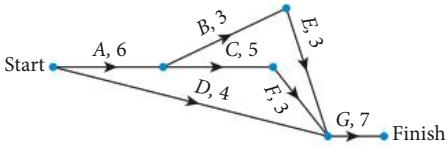
6 **a**



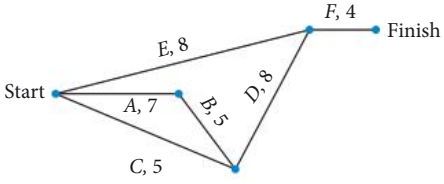
b



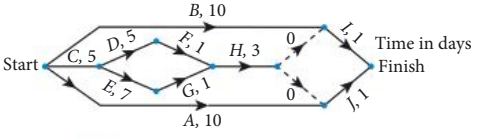
c



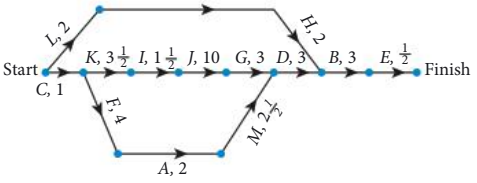
d



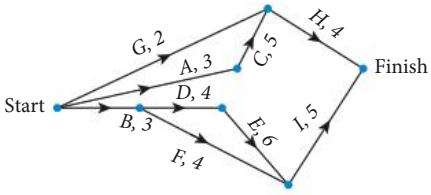
7



8



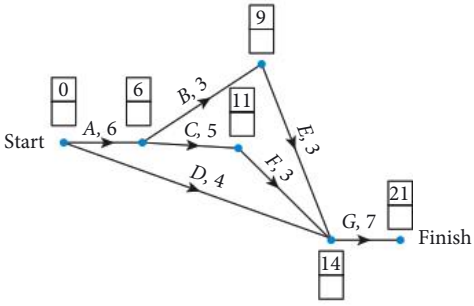
9 **a**



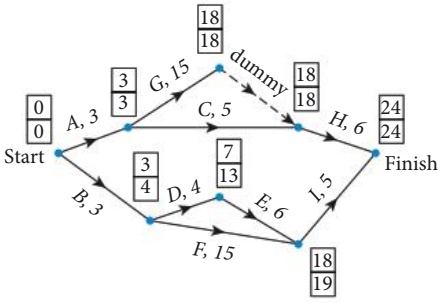
b **i** 18 days **ii** 12 days
iii 12 days **iv** 10 days

9.03

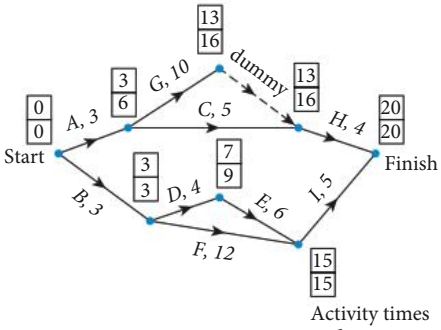
1 D
2 C
3 C
4 **a** 0 **b** 5 **c** 13 **d** 19
5 **a** 9 **b** 3 **c** 16 **d** 18
6



7



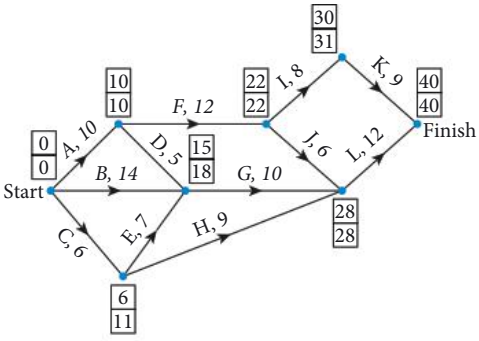
8



Activity times in days

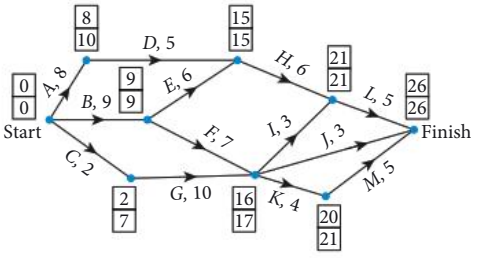
9780170250283

9

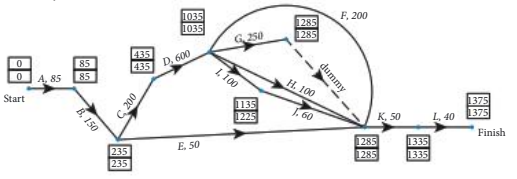


10 This has no impact on the final EST and LST.

11



12 a, b and c

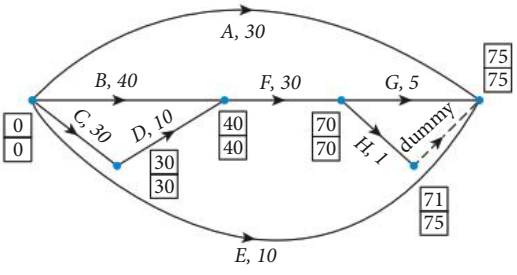


d No change to the final EST.

e Activity *F* changes from 200 to 260 and would add another 10 days to the final EST.

f Activity *G* changes from 250 to 310 and would add another 60 days to the final EST.

13 a, b and c



d No change.

e Add 20 days to the final EST.

f Add 20 days to the final EST.

9.04

1 C

2 C

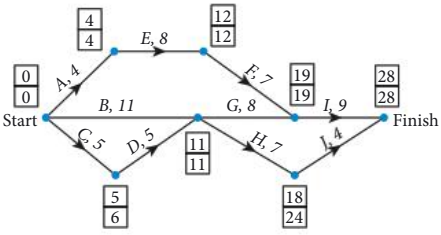
3 B

4 i Path *ABDG* time = 20 h critical path
Path *ACG* time = 19 h
Path *AEFG* time = 14 h

ii Path *AEF* time = 19 h
Path *BG* time = 17 h
Path *CDG* time = 21 h critical path

5 Critical path *ABEG* critical time 34 h

6 a



b Critical paths *AEFI* and *BGI*

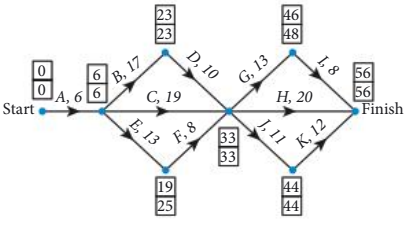
c Minimum project time = 28 h

d Critical path is *CGDI*, critical time = 27 days

7 5 hours

8 10 days

9 a



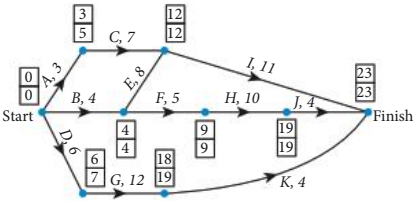
b critical path *ABDJK*

c critical time = 56 hrs

10 Float for activity $C = 6 - 0 - 5 = 1$ h
Float for activity $D = 11 - 5 - 5 = 1$ h
Float for activity $H = 24 - 11 - 7 = 6$ h
Float for activity $J = 28 - 18 - 4 = 6$ h

11 Float for activity $C = 33 - 6 - 19 = 8$ h
Float for activity $E = 25 - 6 - 13 = 6$ h
Float for activity $F = 33 - 19 - 8 = 6$ h
Float for activity $G = 48 - 33 - 13 = 2$ h
Float for activity $H = 56 - 33 - 20 = 3$ h
Float for activity $I = 56 - 46 - 8 = 2$ h

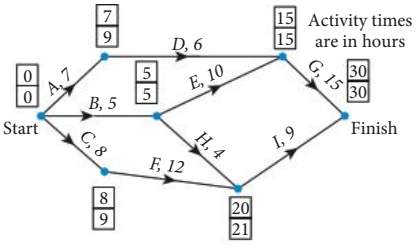
12 a, b and c



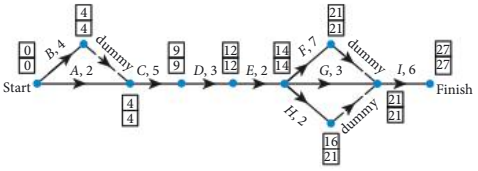
d Two critical paths *BEI* and *BFHJ*.

e The earliest completion time for the project is 23 days.

f Float *A* = 2 days, float *C* = 2 days, float *D* = 1 day, float *G* = 1 day, float *K* = 1 day

13 

c Critical path *BEG*
d Project completion time = 30 h
e Float *A* = 2 h, float *D* = 2 h, float *C* = 1 h, float *F* = 1 h, float *H* = 12 h, float *I* = 1 h
f *ADG* is the critical path , project time = 38 h
g *CFI* is the critical path, project time = 34 h

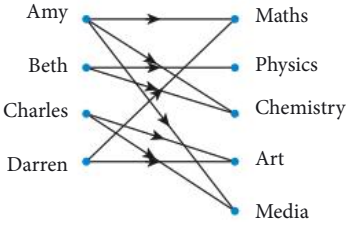
14 

c Critical path *BCDEFI*, project time = 27 days
d Float *A* = 2 days, float *G* = 4 days, float *H* = 5 days
e *A*, *G* and *H*
f Project time = 30 days

9.05

1 B
2 B
3 A
4 a 350 b 450 c 350
5 a 900 b 600 c 600
6 a valid cuts 1 and 2 b cut 1 = 145, cut 2 = 165
7 cut 1 = 13, cut 2 = 26
8 21
9 39
10 70
11 a 8 cuts. Their values are {67, 85, 75, 93, 65, 57, 54, 62}
b Maximum flow = 54
12 a 500 kL/min
b 5 cuts, capacity = {500, 300, 500, 700, 600}
c Maximum = 300
d 300 kL/min
13 a 340 b 195 c 90 d 50

9.06

1 C
2 B
3 

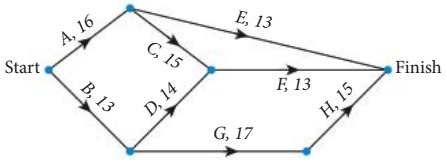
4 a true b false c true
d false e true
5 a April – task 3, Adele – task 2, Alice – task 4, Anne – task 1
b Barry – task 3, Brian – task 1, Ben – task 4, Bill – task 2
c Cathy – task 2, Celeste – task 3, Christen – task 1, Cheryl - task 4
d Dan – task 3, Denise – task 5, Danielle – task 1, Dianne – task 2, Don – task 4
6 a *A*1 = 2, *B*2 = 3, *C*3 = 5, total cost = $10
b *A*2 = 3, *B*3 = 6, *C*1 = 5, total cost = $14
7 a *A*3 = 7, *B*1 = 6, *C*4 = 5, *D*2 = 5, total time = 23 hours
b *A*2 = 5, *B*3 = 5, *C*4 = 2, *D*1 = 4, total time = 16 hrs or *A*4 = 3, *B*3 = 5, *C*2 = 4, *D*1 = 4, total = 16 hrs
8 Rob – task 1, Ben – task 3, Claire – task 2, Dianne – task 4
9 a *A*1, *B*3, *C*2 b Distance = 42 km
10 *A*2 = 19, *B*4 = 21, *C*1 = 16, *D*3 = 16, minimum cost = $72 000
11 a $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 13 & 5 & 5 & 0 \\ 3 & 6 & 4 & 0 \\ 5 & 0 & 7 & 0 \end{bmatrix}$
b $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 13 & 5 & 5 & 0 \\ 3 & 6 & 4 & 0 \\ 5 & 0 & 7 & 0 \end{bmatrix}$ the number of lines does not equal the number of rows
c $\begin{bmatrix} 0 & 0 & 0 & 4 \\ 10 & 2 & 2 & 0 \\ 0 & 3 & 1 & 0 \\ 5 & 0 & 7 & 3 \end{bmatrix}$
d Andrew – task *Y*
Michael – task *Z*
Chris – task *W*
Steph – task *X*
12 a $\begin{bmatrix} 0 & 1 & 5 & 6 \\ 0 & 3 & 11 & 5 \\ 2 & 0 & 0 & 0 \\ 0 & 4 & 1 & 5 \end{bmatrix}$
b Only two lines are required to cover the zeros, however there are 4 rows.
c $\begin{bmatrix} 0 & 0 & 4 & 5 \\ 0 & 2 & 10 & 4 \\ 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 4 \end{bmatrix}$
d *AX*, *BW*, *CZ*, *DY*
e Total distance = 28 + 19 + 23 + 26 = 96 km
13 Darcie 100 m, Shannon 800 m, Chanelle 200 m, Morgan 400 m

9780170250283

14 *A–W, B–Z, C–Y, D–X* for a total of 20 + 18 + 23 + 12 = 73
Or
A–Z, B–Y, C–W, D–X for a total of 21 + 16 + 24 + 12 = 73

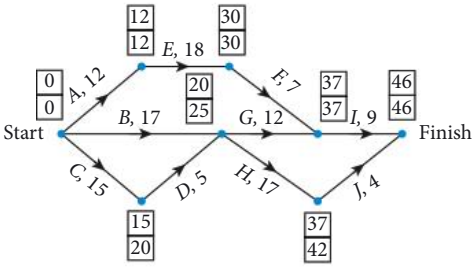
CHAPTER 9 REVISION

1 D
2 C
3 A
4 D
5 E
6 D
7 A
8 E
9 a



b Critical path *BGH*

10 a, b



c Critical path *AEFI*
d Earliest project completion time = 46 hours
e Float $J = 5$, float $H = 5$, float $G = 5$, float $D = 5$, float $C = 5$, float $B = 8$

11 a 0 days b 7 days c 9 days
12 a 380 b 250 c 250
13 a cut 1 = 33, cut 2 = 32, cut 3 = 31
b Maximum flow = 28
14 a 6 cuts
b Flow capacity = {40, 30, 58, 28, 43, 73}
c Maximum flow = 28 litres/sec
15 a John, Elvis
b 2
c guitar, drums, trumpet
d Elvis and Janis
16 *A*1, *B*2, *C*3, *D*4
17 a *A*2, *B*3, *C*4, *D*1 b Total = 140 km
18 a

$$\begin{bmatrix} 15 & 14 & 1 & 0 \\ 14 & 3 & 0 & 3 \\ 0 & 0 & 8 & 0 \\ 15 & 13 & 12 & 0 \end{bmatrix}$$

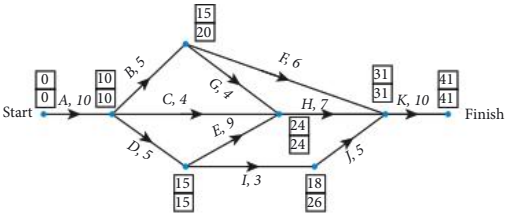
b The zeros can be covered with three lines, which does not equal the number of rows.

c

$$\begin{bmatrix} 11 & 10 & 0 & 0 \\ 11 & 0 & 0 & 4 \\ 0 & 0 & 11 & 4 \\ 11 & 9 & 11 & 0 \end{bmatrix}$$

d *AY*, *BX*, *CW*, *DZ*
e Minimum kilometres = 69

19 a, b, c



d Critical path = *ADEHK*
e Earliest project completion time = 41 days
f Float $B = 5$, float $C = 10$, float $F = 10$, float $G = 5$, float $I = 8$, float $J = 8$
g No impact on the project.
h Delay the project completion time by 2 days.

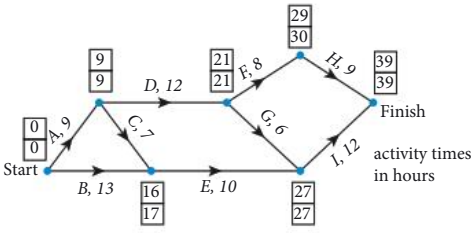
MIXED REVISION 3

Multiple choice

1 D
2 C
3 B
4 B
5 E
6 A
7 D
8 C
9 C

Short answer

1 a $187 524.99 b $57 524.99
2



3 3 years 1 month
4 5.8%
5 8.1% p.a.
6 Ryan – Y, Amber – W, Georgia – Z and Sam – X

Application

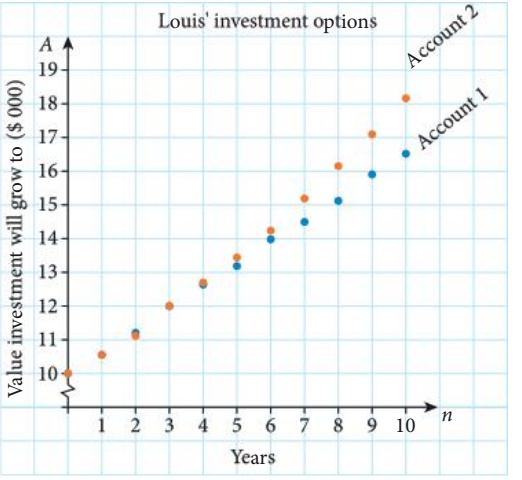
1 a $495.94 per month
b 2 months
c $485.71
d i $991.88
ii $736.58

2 a *AEIJ*
b 44 minutes
c 36 minutes
d activities *D*, *E* and *F*
e 60 minutes

3 a $2325 b $13 175 c $2160.70
d $15 335.70 e $294.92

4 a $A_{n+1} = 1.05A_n - 6000$ where $A_1 = 70\,000$
b $42 433.59 c $3500

5 a, b



c If Louis is investing for less than 3 years then Account 1 is the best option as the value of the investment is greater. However, if he is investing for 3 years or more then Account 2 is the best option.

6 a 330 b 220 c 90 d 50

9780170250283

GLOSSARY

Chapter 1

association/relationship From a scatter plot, a relationship should be described in terms of form, direction and strength.

boxplot Graphical display for numerical data that shows the centre (median) and spread (range and IQR) of a data set. Parallel boxplots can be used when investigating more than one variable.

categorical data Variables represented by qualities or characteristics that can be either nominal or ordinal.

dot plot Graphical display for categorical or numerical discrete data.

explanatory variable A variable which is not affected or changed by another variable.

histogram Graphical display for numerical data (discrete or continuous) with vertical joined columns.

interquartile range (IQR) Measure of spread calculated using IQR = $Q_3 - Q_1$.

mean The average of a data set. It is calculated using $\bar{x} = \frac{\Sigma x}{n}$, and is considered to be a measure of centre.

median The middle value when a data set is ordered from smallest to largest. It is considered to be a measure of centre.

mode The most common data value in a set. It is considered to be a measure of centre.

numerical data Variables represented by quantities or measurements that can be either discrete or continuous

outlier By visual inspection, an extreme (high or low value) in the data, or formally, any value outside the range of $Q_1 - 1.5 \times \text{IQR} \leq x \leq Q_3 + 1.5 \times \text{IQR}$

Pearson's correlation coefficient Numerical value that determines the exact strength of the linear relationship between two variables.

population All items/people of the group being studied.

range Measure of spread calculated using range = highest data value − lowest data value.

response variable A variable which is affected or changed by another variable.

sample Portion of the population; used when the population is large and is randomly selected to eliminate bias.

scatter plot Graphical display used to compare two numerical variables.

segmented column graph Graphical display constructed from a percentage two-way table.

shape of data Described as symmetrical, positively skewed or negatively skewed.

standard deviation Measures the spread of data about the mean; for a sample it is calculated using $S_x = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}}$, but generally a CAS can be used.

stem-and-leaf plot Graphical display for numerical data that is either discrete or continuous. If investigating 2 variables, a back-to-back stem-and-leaf plot is required.

two-way table A table used to display information when comparing two categorical variables.

Chapter 2

association/relationship From a scatter plot, a relationship should be described in terms of form, direction and strength.

coefficient of determination It is a numerical value between 0 and 1 that is used to determine how useful a regression line is as a linear model for a given set of data. More specifically, when it is

converted to a percentage, it tells us the percentage of the variation in the dependent variable that is explained by the independent variable.

explanatory variable A variable that is not affected or changed by another variable.

extrapolation A prediction made outside the original range of the data given.

gradient The amount that the y variable increases/decreases by as the x variable increases by 1 unit.

interpolation A prediction made inside the original range of the data given.

least-squares regression line A least-squares regression line is the linear fit that minimises the square of the vertical distance between the actual values of y and the y value predicted by the least-squares regression line. Its general form is $y = a + bx$, where the a represents the y-intercept and b represents the gradient.

numerical data Variables represented by quantities or measurements that can be either discrete or continuous.

Pearson's correlation coefficient Numerical value, between –1 and 1, that determines the exact strength of the linear relationship between two variables.

predictions Values calculated via substituting into the least-squares regression line modelling the data.

residual A residual value is the difference between the actual y variable value and the predicted y variable value for a given x value. The formula is
Residual = actual y value – predicted y value

residual plot A residual plot is basically a scatter plot with the explanatory variable on the x-axis and the residual values on the y-axis.

response variable A variable which is affected or changed by another variable.

scatter plot Graphical display used to compare two numerical variables.

***y*-intercept** The value of the y variable when the x variable is zero.

Chapter 3

arithmetic progression See **arithmetic sequence**.

arithmetic sequence A sequence in which each term differs from the previous by a constant value d.

common difference The rule for finding d is $d = t_2 - t_1 = t_3 - t_2 = t_4 - t_3 = \cdots$ This gives the general common difference $d = t_n - t_{n-1}$

common ratio The rule for finding r is $r = \frac{t_2}{t_1} = \frac{t_3}{t_2} = \frac{t_4}{t_3} = \cdots$ This gives the general common ratio $r = \frac{t_n}{t_{n-1}}$

depreciation **Straight-line depreciation**: where the book value depreciates by a fixed dollar value each year. **Reducing balance depreciation**: where the book value depreciates by a given percentage each year. **Unit cost method of depreciation**: is dependent on the amount of use the asset has had. An arithmetic sequence can be used to model the straight-line method or the unit cost method of depreciation.

general term, t_n The general term, t_n, represents the nth term of a sequence.

geometric progression See **geometric sequence**.

geometric sequence A sequence in which any term is multiplied by a constant r to get the next term.

infinite sum of a geometric sequence $S_\infty = \frac{a}{1-r}$, where $|r| < 1$ is the necessary condition for the limiting sum to exist.

***n*th partial sum of a geometric sequence** $S_n = \frac{a\left(r^n - 1\right)}{r-1}$ for $|r| > 1$ or $S_n = \frac{a\left(1 - r^n\right)}{1-r}$ for $|r| < 1$.

***n*th term, t_n, of an arithmetic sequence** The rule for the nth term, t_n, of an arithmetic sequence is $t_n = a + (n - 1)d$, where a = the first term and d = common difference.

***n*th term, t_n, of a geometric sequence** The rule for the nth term, t_n, of a geometric sequence is $t_n = ar^{n-1}$, where a = the first term and r = common ratio.

9780170250283

partial sum of an arithmetic sequence The sum of the first n terms of an arithmetic series (nth partial sum) is given by the formula: $S_n = \frac{n}{2}(a + \ell)$, where a = the first term, ℓ = last or nth term.

recursion A recursive formula has two parts:
1 the starting value for a_1
2 the recursion equation for a_n as a function of a_{n-1} (the term before it).

series The sum of terms of a sequence. So for the arithmetic sequence 2, 5, 8, 11, 14, 17, … the equivalent arithmetic series would be $2+5+8+11+14+17+\cdots$ We use the notation S_n to mean the sum of the sequence up to t_n.

series, S_n, of an arithmetic sequence The series, S_n, the sum to n terms, of an arithmetic sequence is $S_n = \frac{n}{2}[2a + (n-1)d]$ where a = the first term, d = common difference, n = number of terms.

series, S_n, of a geometric sequence The series, S_n, of a geometric sequence is $S_n = \frac{a(r^n - 1)}{r - 1}$ where a = the first term, r = common ratio, n = number of terms.

sum of the infinite series, S_∞, of a geometric sequence The sum of the infinite series, S_∞, of a geometric sequence is $S_\infty = \frac{a}{1 - r}$ where a = the first term, r = common ratio, infinite number of terms.

Chapter 4

arithmetic sequence A sequence in which each term differs from the previous by a constant value d.

difference equation A rule that specifies a particular term in a sequence using the previous term.

first order difference equation A rule that specifies how a particular term in a sequence can be found from the previous term in the same sequence.

geometric sequence A sequence in which any term is multiplied by a constant r to get the next term.

recurrence relation A rule that specifies a particular term in a sequence using the previous term.

steady-state solution The value that the terms of a sequence approach in the long run.

Chapter 5

adjacency matrix A representation of how the vertices are connected by edges.

bipartite graph A graph whose set of vertices can be split into two distinct groups, where vertices cannot connect to other vertices from the same group.

bridge An edge, which if removed, leaves a graph disconnected.

complete graph Each vertex in a graph is connected to every other vertex in the graph.

connected graph A graph is connected if there is a path between each pair of vertices.

cycle A cycle is a set of vertices that when connected by edges enclose a space.

degree The number of edges connected to a vertex.

directed graph (digraph) A graph where direction is shown on the edges.

edge The lines connecting the vertices.

Eulerian graph A graph where you can start and finish at the same vertex and every edge can be covered once only.

Euler's formula To test whether a graph is planar:

vertices + faces – edges = 2

face A region bound by edges. The region surrounding a planar graph is also considered to be a face.

graph A graph is a diagram that consists of a set of points, called vertices, that are joined by a set of lines, called edges.

Hamiltonian cycle A cycle that starts and finishes at the same vertex and includes every vertex in a graph once only.

Hamiltonian path A path that starts and finishes at different vertices and includes every vertex once only.

isolated vertex A vertex that has no edges connected to it.

loop A vertex that has an edge connecting it to itself.

multiple edges Two vertices that are connected by more than one edge.

path A set of consecutive vertices that are not repeated.

planar graphs A graph that contains enclosed regions and no edges that cross.

semi-Eulerian graph A graph where you start and finish at different vertices and every edge can be covered once only.

simple graph A graph that contains no loops or multiple edges.

spanning tree A subgraph of a connected graph that connects all vertices by the minimum number of edges.

subgraph A graph that contains the same edges and vertices of another graph.

trail A set of consecutive vertices that are not repeated.

tree A connected graph with no cycles.

vertex (vertices) The points on a graph.

walk A sequence of consecutive vertices that may or may not be repeated.

weighted graph A graph with quantities on the edges representing distance, time, etc.

Chapter 6

centring To smooth data, we can use a two- or four-point moving average with centring. Smoothing with an odd number of points gives a smoothed data point, which is located at the middle time, that is, the data point corresponds to one of the original times. In contrast, when an even number of points is used to smooth the data, there is a problem finding the centred time value. The centre of the set of points is not one of the original points. This is called centring, where we do another step of taking a 2-moving average of the already smoothed values, so they line up with original time values.

cyclic pattern A cyclic pattern exists when significant peaks and troughs occur at irregular intervals, usually with variation in the amplitudes of the peaks and troughs. For example, the earthquakes recorded each year shows significant peaks and troughs, but at unpredictable times. The main difference between cyclic and seasonal patterns is the unpredictability of when peaks and troughs occur.

deseasonalisation Deseasonalisation is a form of smoothing, which takes out the seasonal effects of the data so that a trend line can be fitted and long-term trends can be predicted.

describing a time series plot A time series can have one or more of the components: trend, seasonal pattern, cyclic pattern and random variation.

forecasting The linear model (regression equation) can be used to forecast values in the future (make predictions). Using a trend line fitted to a times series plot to make predictions about future values is known as forecasting.

four-point moving average See **two-point moving average**.

least-squares regression In fitting a straight-line $y = a + bx$ to the relationship between a response variable y and an explanatory variable x, the least-squares line is the line for which the sum of the squared residuals is the smallest.

9780170250283

mean (see moving average) The average of a data set and is calculated using $\bar{x} = \frac{\sum x}{n}$, and is considered a measure of centre.

median Middle value when a data set is in order from smallest to largest, and is considered a measure of centre.

median smoothing Median smoothing uses a similar technique to moving averages; using the median values instead of the average (mean) values. The process of using an odd number of points is very similar. An even number of points, as before, needs the method of centring for median smoothing.

moving average In a time series, a simple moving average is a method used to smooth the time series whereby each observation is replaced by a simple average of the observation and its near neighbours. This process reduces the effect of non-typical data and makes the overall trend easier to see.

negative secular trend We say the time series plot has a negative secular trend when it is moving in an overall negative direction.

outlier By visual inspection, an extreme (high or low value) in the data, or formally, any value outside the range of $Q_1 - 1.5 \times \text{IQR} \leq x \leq Q_3 + 1.5 \times \text{IQR}$

positive secular trend We say that a time series plot has a positive secular trend when it is moving in an overall positive direction.

random variation Random variation or random pattern reflects the unpredictability common in all statistical data. In a time series plot, a random pattern would not show predictable peaks and troughs, nor would there be significant peaks and troughs at unpredictable times. Instead there will be random fluctuations about a relatively stable mean.

regression line The trend line can be used to predict future behaviour over time. A trend exists if there is a long-term increase or decrease in the dependent variable as time passes. If there appears to be a linear trend, then a regression line can be fitted.

residual value The difference between the predicted value and the observed value. The prediction is usually from a statistical model such as least-squares regression.

scatter plot A two-dimensional data plot using Cartesian coordinates to display the values of two variables in a bivariate data set.

seasonal pattern A seasonal pattern arises in the context of economic time series models where the actual seasons of the year affect sales or production of goods. In such a context, there will be a similar pattern of regular peaks and troughs at corresponding months or quarters of successive years. For example, the sales figures of ice-cream in a beach resort is high each summer, and the time series repeats this pattern each year after a regular interval of time.

seasonal adjustment (adjusting for seasonality) A term used to describe a time series from which periodic variations due to seasonal effects have been removed. See also seasonal index.

seasonal index Seasonality can be removed from data by using the seasonal index. An index value is attached to each period of the time series within a year. For months, there are 12 separate seasonal indices, one for each month; For Summer, Autumn, Winter and Spring, the seasons of the year, there are four separate seasonal indices; and so on.

seasonal indices Seasonal indices can be calculated by using a method called the average percentage method. Seasonal indexes are used to deseasonalise a time series.

seasonal variation A regular rise and fall in the time series that recurs each year. Seasonal variation is measured in terms of a seasonal index.

smoothing (time series) Smoothing using moving averages involves taking a fixed number of data points and averaging them to give a smoothed data point.

smoothing See **smoothing (time series)** and **median smoothing**.

statistics menu Found in common graphic and symbolic (CAS) calculators.

three-point moving average To use a three point moving average we use three successive data points to create a smoothed data value.

time series Recorded values of a variable over a period of time, usually at regular intervals. Long-term trends, seasonal variation, and noise or irregular variation are made up from many such series.

time series data If data (observations on y) are collected at a series of times, usually at regular intervals (days, weeks, months, years, …), then the resulting bivariate data set is known as time series data.

time series plot When the data are displayed with the explanatory variable time on the horizontal axis and the response variable y on the vertical axis, the resulting graph is called a time series plot.

trend A trend exists if there is a long-term increase or decrease in the dependent variable as time passes.

trend line A straight line can be fitted to a time series plot either by eye, using the three median method, or using the least squares method.

two- or four-point moving average with centring When an even number of points is used to smooth data, there is a problem finding the centred time value. The centre of the set of points is not one of the original points. This is called centring where we do another step of taking a 2-moving average of the already smoothed values, so they line up with original time values.

Chapter 7

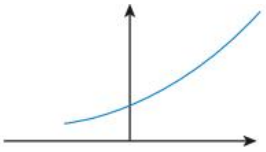
buying on terms (or hire purchase) See **hire purchase**.

compound interest Interest paid on the principal invested as well as any accumulated interest.

compounding periods How often interest is calculated (for example daily, monthly, quarterly, etc.).

effective interest rate Used to compare the interest paid on loans (or investments) with the same nominal annual interest rate i but with different compounding periods (daily, monthly, quarterly, annually, other).

exponential graph A graph of an exponential function. An increasing exponential graph will always have the form:



flat-rate interest See **simple interest**.

future value The value of an investment at the end of the investment period.

hire purchase (or buying on terms) Paying off a loan plus interest using regular repayments.

interest Money earned on an investment or paid on a loan.

nominal interest rate The interest rate quoted for a loan or investment. There is no adjustment for inflation.

per annum (p.a.) Per year.

present value The total amount of the investment now.

principal The original amount of money invested or borrowed.

repayment (instalment) Amount paid regularly to pay off a loan.

simple interest (or flat-rate interest) Interest earned or charged only on the original amount of money (principal) invested or borrowed.

9780170250283

Chapter 8

annuity A compound interest investment from which regular payments are made until the value of the investment is zero.

annuity investment An investment that involves making an initial deposit followed by additional regular payments into an account earning a fixed rate of compound interest.

compound interest Interest paid on the principal invested as well as any accumulated interest.

future value The value of an investment at the end of the investment period.

perpetuity A type of annuity where regular payments are paid from the investment that continue forever.

per annum (p.a.) Per year.

present value The total amount of the investment now.

principal The original amount of money invested or borrowed.

reducing balance loan A loan where interest is calculated on the amount still owing after each repayment is made.

Chapter 9

activity A task required for a project. Activities are the edges in a directed graph.

adjacency matrix The matrix that shows the one stage pathways in a directed graph.

allocation Allocation problems involve the best connection between the items in one group to the items in another group. For example, assigning people to tasks.

assignment The assignment problem is the process of determining the optimum or best allocation.

backward scan The process of finding latest starting times (LST).

bipartite graph A bipartite graph has its vertices in two distinct sets and the edges join elements in the first set to elements in the second set.

capacity The flow capacity is the maximum amount a path can carry.

crashing It is possible to speed up a project by employing more resources and this enables some activities to be completed more quickly. This process is called crashing.

critical activity A critical activity is an activity on the critical path.

critical path The critical path is the longest time path in a network.

cut A cut through a network must stop all flow from the start (*source*) to the end (*sink*).

digraph A graph where direction is indicated along each edge is called a directed graph or digraph.

dominance In a directed graph where an edge moves from vertex A to vertex B, vertex A is dominant over vertex B.

dummy activity A dummy activity has zero time and is added to a network to maintain precedence structure whilst ensuring that no two vertices are connected by multiple edges.

earliest start time The EST is the earliest time an activity can start.

edge Edges are the lines that connect vertices.

event Events are indicated by the vertices or nodes of a directed graph. They show the start or end of an activity.

float The float time for an activity is the maximum time an activity can be delayed without affecting the project completion time.

forward scan The process of determining earliest starting times.

Hungarian algorithm The mathematical process used to solve the assignment problem.

inflow capacity This is the total flow capacity entering a vertex.

latest starting time This is the latest time an activity can start before it affects the project completion time.

matrix A rectangular array of numbers. It is used to show the connections between vertices in a digraph.

network A network is a directed graph and can be used to show the sequence of tasks required to complete a project.

node A node is a vertex.

non-critical activity Non-critical activities are activities that are not on the critical path.

one-stage pathway One-stage pathways connect two vertices along one edge.

optimum The optimum allocation is the best way of assigning items in one group to items in a second group.

outflow capacity This is the total flow capacity leaving the vertex.

reachability Vertex B is reachable from vertex A if a path exists from A to B.

reduction In the assignment problem, row reduction involves subtracting the smallest number in each row from the other numbers in the same row.

resultant matrix To determine reachability, the resultant matrix is found using the formula $R = R^1 + R^2$ if one stage and two stage paths are considered

sink In flow problems, the sink is the final destination of the flows through a digraph.

source In flow problems the source is the start of the flows through a digraph.

two-stage pathway Two-stage pathways connect two vertices along two edges.

valid cut A valid cut through a network must stop all flow from the start (*source*) to the end (*sink*).

vertex A vertex is the point where an edge starts or ends.

9780170250283

INDEX

Aa

adjacency matrix 268–70
allocation problems
    and bipartite graphs 490–6
    Hungarian Algorithm 495–6
    row and column reduction method 493–4
annuities 428–32
    finance solvers for calculating 431
    see also perpetuities
annuities formula 416, 430, 437
annuity investments 436–9
    finance solvers for calculating 437
arithmetic sequences 146–54, 214
    common difference 146–7, 214
    general term (t_n) 148
    graphs of 147
    modelling 161–4, 184–8
    nth term (t_n) 148, 185
    partial sum 156, 157
    recurrence relations 214–15
    series (S_n) of 157, 185
    sum of 156, 159, 185
arithmetic series 155–9, 184–8
assignment problems *see* allocation problems
associations
    between categorical variables 41–6
    looking for 36–9
average percentage method 337

Bb

back-to-back stem-and-leaf plots 50, 51
backward scanning through a network 470–1
bar graphs 23
biased sample 6
bipartite graphs 258, 259, 491
    and allocation problems 490–6
bivariate data 36, 97
boxplots 25–9, 30, 31, 32
    making comparisons 57
    parallel 50, 53–8
bridge 281
buying on terms 366

Cc

categorical data 4, 22, 23
categorical variables
    relationships between 41–6
    relationships with numerical variables 50–8
census 6
centring 326, 327, 328, 331, 333–4
closed trails 280, 281, 283–5
closed walks 280, 281
coefficient of determination (r^2) 110–13
    interpreting 111
collecting data 5–8
column graphs 23, 42, 46
common difference 146–7, 215
common ratio 166–7
complete graphs 258, 259, 260
compound interest formula 376–83
compound interest investments 371–4, 376
    and effective interest rate 399–401
    graphs 388–9
    with regular deposits 436–9
    see also annuities
compound interest loans *see* reducing balance loans
compounding periods 377, 416
connected graphs 258, 259, 283, 292
    see also Eulerian graphs; semi-Eulerian graphs; trees
connections (networks) 459–60
continuous data 5, 29
correlation coefficients, interpreting 73–9
crashing 479–80
critical activities 478

critical paths 475
activity float times 478–9
determining 475–6
finding using EST and LST values 476–7
project crashing 479–80
cycles 280
Hamiltonian 288–90
cyclic pattern (time series) 318

Dd

data
describing 30–2
displaying 22–30
types of 4–5
data collection 6–8
preparation 5–6
deferred payment plans 367–8
degree of a vertex 268
dependent variables *see* response variables
depreciation 162
describing data 30–2
deseasonalising data 337–40
difference equations *see* recurrence relations
directed graphs (digraphs) 258, 450
dominance 454–6
reachability 450–4
discrete data 5, 29
displaying data 22–30
distributions, describing the shape of 30–1
dominance 454–6
dominance graphs 454–5
dominance matrix 455–6
dot plots 24
dummy activities (network diagrams) 461–3

Ee

earliest starting times (EST) 468–9
and activity float times 478–9
for finding critical path 475, 476–7
edges 250–4, 258, 268, 274, 275, 451, 459
effective interest rate 399–401
Eulerian graphs 283–5
Euler's formula, and planar graphs 275–7
explanatory variables 36, 37, 38–9, 62, 63
and coefficient of variation 111
least-squares regression line 96, 97
extrapolation, from least-squares regression lines 118–19

Ff

faces 275
first order recurrence relations 202, 214
graphical representation 219–27
see also recurrence relations
five number summary 25, 26
flat-rate interest 362, 366–7
float times 478–9
flow capacity 484–5
forecasting 343
see also predictions from least-squares regression lines
forward scanning through a network 468–9
four-point moving average 326–9
four-point moving median smoothing 333–4
frequency tables 22–3
future value 377, 416

Gg

geometric sequences (geometric progressions) 166–73, 216
common ratio 166–7
general term (t_n) 167–8
graphs of 168
infinite sum 177–9, 185
modelling 181–3, 184–8
*n*th term (t_n) 167–8
recurrence relations 216–18
series (S_n) of 176, 185
sum of 175–7, 185
sum of the infinite series (S_∞) 177, 185
geometric series 175–9, 184–8
gradient, least-squares regression line 97, 100, 101, 102

9780170250283

graphs
applications 263–6
arithmetic sequences 147
bar 23
basic concepts 250–4
bipartite 258, 259, 490–1
column 23, 42, 46
complete 258, 259, 260
connected 258, 259, 283, 292
directed 258
Eulerian 283–5
geometric sequences 168
interest 386–94
planar 274–5
recurrence relations 219–27
representation through adjacency matrix 268–70
semi-Eulerian 283–5
simple 258, 259
types of 258–61
weighted 296–9
grouped data 23

Hh

Hamiltonian cycles 288–90
Hamiltonian paths 288–90
hire-purchase 366–8
histograms 29–30
Hungarian Algorithm 495–6

Ii

immediate predecessor 459
independent variables *see* explanatory variables
infinite sum of a geometric sequence 177–9
inflow capacity 484
interest
compound 371–4, 376–83
simple 362–4, 366–8
interest graphs 386–94
interest rate 411
effective 399–401
interpolation, from least-squares regression lines 118–19
interquartile range (IQR) 11, 12, 18
investments
annuity 436–9
compound interest 371–4
irregular fluctuations (time series) 314
isolated vertex 251

Ll

latest starting times (LST) 470–1
and activity float times 478–9
for finding critical path 475, 476–7
least-squares regression lines
and coefficient of determination 110–13
fitting 96–107
fitting to time series data 342–4
interpolation and extrapolation from 118–19
making predictions 116–19, 343–4
residual values 122–31
limiting sum 177
linear models, for numerical data 96–131
linear pattern (scatter plots) 63, 64
logic difficulties in the precedence structure 462
long-term decreasing solutions (recurrence relations) 222–6
long-term increasing solutions (recurrence relations) 222
long-term solutions (recurrence relations) 221–7
long-term steady-state solutions (recurrence relations) 222, 232–3
loops 251

Mm

matrix
adjacency 268–70
allocation problems 492, 495–6
dominance 455–6
reachability 452, 453–4
maximum flow (networks) 484–5
capacity of the cut 486
determining 487
minimum cut 485
mean 10–11, 18, 50

measures of centre 10
measures of spread 11–12
median 10–11, 18
median smoothing 323, 331–4
minimum spanning trees 296, 297–9
mode 10–11, 18
modelling
    arithmetic sequences 161–4, 184–8
    geometric sequences 181–3, 184–8
    recurrence relations 235–8
moderate negative relationship (scatter plots) 74
moderate relationship (scatter plots) 65
moving averages, smoothing using 323–9
multiple edges 461–2

Nn

negative correlation (scatter plots) 64
negative secular trend 318
negatively skewed distribution 31
network diagrams
    drawing 459–64
    drawing from an activity table 459, 460–1
    drawing guidelines 460
    dummy activities 461–3
    types of connections 459–60
network flow problems 484
    capacity of a cut 486
    flow capacity and maximum flow 484–5
    maximum flow – minimum cut 485
networks 250–1, 296
    backward scanning 470–1
    forward scanning 468–9
    matrix representation 451–2, 453–4
no relationship (scatter plots) 63, 74
nodes 45
nominal data 4, 5
nominal interest rate 399
non-critical activities 478
non-linear pattern (scatter plots) 63
numerical data 5, 22
numerical variables
    relationships between 62–9
    relationships with categorical variables 50–8

Oo

one-stage pathways (directed graphs) 451
one variable statistics 12–14, 26
open trails 280, 281, 283–5
open walks 280, 281
ordinal data 5
outflow capacity 484
outliers (time series) 214
outstanding balance (reducing balance loans) 416

Pp

parallel boxplots 50, 53–8
partial sum of an arithmetic sequence 156
paths 280
    Hamiltonian 288–90
Pearson's correlation coefficient (r) 73–4, 78, 79
perfect positive relationship (scatter plots) 74
perpetuities 434–5
planar graphs 274–5
    and Euler's formula 275–7
population 6
positive correlation (scatter plots) 64
positive secular trend 318
positively skewed distribution 31
preceding events (network connections) 459–60
predictions, from least-squares regression lines 116–19, 343–4
present value 416
Prim's Algorithm 297–8
principal 362
project networks 459

Qq

questionnaires 5–6

9780170250283

Rr

random pattern (random variation) 318, 319
random sample 7–8
range 11, 12, 18
reachability (directed graphs) 450–4
reachability matrix 452, 453–4
recurrence relations 202
    applications 235–8
    arithmetic sequences 214–15
    compound interest investment 372–4, 376
    finding for a given sequence 211–12
    geometric sequences 216–18
    graphical representations 219–27
    long-term decreasing solution 222–6
    long-term increasing solution 222
    long-term solutions 221–7
    long-term steady-state solutions 222, 232–3
    reducing balance loan 411–13
    to generate sequences 202–9
reducing balance depreciation 162
reducing balance loans 410–14
    changing the terms of the loan 421–5
    finance solvers for calculating 416–18
regression lines *see* least-squares regression lines
relationships
    between categorical variables 41–6
    between numerical and categorical variables 50–8
    between numerical variables 62–9
repayments 367, 410
residual plot 126–31
residual values, from least-squares regression lines 122–31
response variables 36, 37, 38–9, 62, 63
    and coefficient of variation 111
    least-squares regression line 96, 97
row and column reduction method 493–4

Ss

sampling 6–8
scatter plots 62–9
    correlation coefficients 73–9
    direction 64
    form 63–4
    interpreting 73–9
    residual plots 126–31
    strength 65
scholarships 434–5
seasonal index/indices 337, 338, 339–40
seasonal patterns (seasonal variation) 314, 318, 319, 337–440
segmented column graphs 42, 46
semi-Eulerian graphs 283–5
sequences
    finding a recurrence relation for a given sequence 211–12
    generation using recurrence relations 202–9
    see also arithmetic sequences; geometric sequences
series 155, 175
    see also arithmetic series; geometric series
series (S_n)
    of an arithmetic sequence 157, 185
    of a geometric sequence 176, 185
side-by-side column graphs 42
simple graphs 258, 259
simple interest 362–4
    graphs 386–8
    hire-purchase 366–8
smoothing time series data
    using deseasonalisation 337–40
    using medians 323, 331–5
    using moving averages 323–9
spanning trees 292, 294, 296, 297
    minimum 296, 297–9
standard deviation 11, 13, 14, 18, 50
stem-and-leaf plots 24–5

straight-line depreciation 162
strong negative relationship (scatter plots) 74
strong relationship (scatter plots) 65
subgraphs 258, 261, 292
sum
    arithmetic sequence 156, 159, 185
    geometric sequence 175–6
    infinite series (S_∞) of a geometric sequence 177, 185
summary statistics 10–18
symmetrical distribution 30
systematic movements 318

Tt

three-point moving average 324–5
three-point moving median smoothing 331–2
time series data 314
    adjusting for seasonality 314, 318, 319, 337–40
    fitting least-squares regression line to 342–4
    smoothing using medians 323, 331–5
    smoothing using moving averages 323–9
time series plots 314–19, 343
    features and components 314, 318, 319
trails 280, 281, 283–5
trees 292–3
trends (time series) 314, 318, 323, 343
    see also smoothing time series data
two-point moving average 326
two-stage pathways (directed graphs) 451
two-way tables 41–6

Uu

undirected graphs 250–302
ungrouped data 23
unit cost method of depreciation 162
univariate data 36
unsystematic fluctuations 318

Vv

variables 4
    categorical 41–6, 50–8
    explanatory (independent) *see* explanatory variables
    numerical 50–8, 62–9
    response (dependent) *see* response variables
vertex/vertices 250–4, 258, 280, 450, 459
    degree of a vertex 268

Ww

walks 279, 280
weak positive relationship (scatter plots) 74
weak relationship (scatter plots) 65
weighted graphs 296
    and minimum spanning trees 296, 297–9

Xx

x variable 75, 101

Yy

y-intercept, least-squares regression line 97, 100
y variable 75, 101
yearly average 337

9780170250283

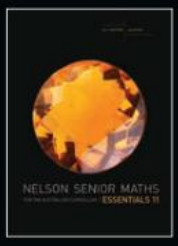
YEAR 12 GENERAL

NELSON SENIOR MATHEMATICS is an innovative series written by an experienced team of teachers. The titles have been written specifically to address the Australian Senior Mathematics Curriculum.

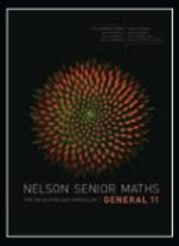
The *Year 12 General Mathematics Student Book* focuses explicitly on the development of content addressing the Australian Curriculum. The chapters are well structured and are broken into lesson-sized sections to best assist the development of student understanding. Simple and clear examples precede each exercise and are matched to relevant questions within the sections. The exercises include a gradation of questions which allow for the range of abilities within the classroom.

An exciting interactive NelsonNetBook, which can be personalised and annotated, accompanies each student text, in addition to a NelsonNet student and a NelsonNet teacher website. The NelsonNet student website provides students with access to all of the resources referenced within the book. The NelsonNet teacher website includes curriculum grids and teaching programs which support the use of the resources and additional resources for assessment purposes.

Titles in the series are shown below.

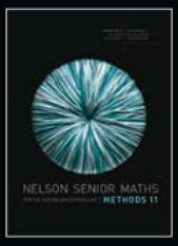


Year 11 Essentials
9780170264013



Year 11 General
9780170250252

WORKED SOLUTIONS
9780170251488



Year 11 Methods
9780170250269

WORKED SOLUTIONS
9780170251495

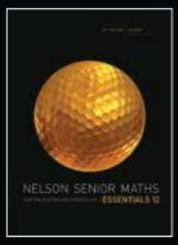


Year 11 Specialist
9780170250276

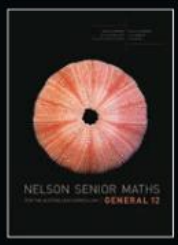
WORKED SOLUTIONS
9780170251501



Year 11
ExamView Package
9780170251549

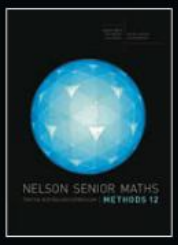


Year 12 Essentials
9780170264112



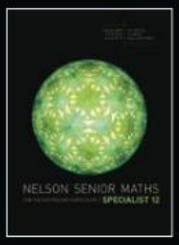
Year 12 General
9780170250283

WORKED SOLUTIONS
9780170254656



Year 12 Methods
9780170250290

WORKED SOLUTIONS
9780170254663



Year 12 Specialist
9780170250306

WORKED SOLUTIONS
9780170254670



Year 12
ExamView Package
9780170254786

ISBN: 978-0170250283
9 780170 250283

NELSON
CENGAGE Learning
For learning solutions, visit **cengage.com.au**