

9



Statistics and probability

9

Caution required when exercising. We all know that exercise is good for us but we also need to be careful, especially when using gym equipment.

Hospital records indicate that treadmill-related injuries are on the rise. Out of the total number of treadmill-related injuries that occurred in a five-year period, it was found that 38% of the injuries were reported in the fifth year. The most worrying statistic to emerge was that children and teenagers are those most at risk. The age range 0–14 years accounted for 60% of all injuries, with the 0–4 age group accounting for 38% of the injuries alone. Of the 12% of injuries that required hospital admission most were in this youngest age group. The high proportion of

young children with these injuries is related to the fact that 70% of the injuries were inflicted in the home.

Forum

What sport do you think would result in the most injuries? How would you research this issue? Is the sport with the highest number of reported injuries necessarily the most dangerous?

Why learn this?

In the media, we are often presented with some sort of statistics or probability, such as the average temperature for the month or claims by advertisers about the worth of their products. Understanding probability and statistics helps us to make sense of such figures and to be aware when someone might be trying to mislead us.

After completing this chapter you will be able to:

- organise, present and interpret categorical and numerical data
- determine the mean, median, mode and range of a data set using various techniques
- present data appropriately using dot plots, stem-and-leaf plots, bar graphs, histograms, line graphs and sector graphs
- analyse univariate and simple bivariate data using dot plots and stem-and-leaf plots
- construct sample spaces for single-step experiments and use them to assign probabilities.

Recall 9

Prepare for this chapter by attempting the following questions. If you have difficulty with a question, go to Pearson Places and download the Recall Worksheet from Pearson Reader.

- 1** (a) List the numbers you get if you count by fives from 0 to 30.
 (b) List the numbers you get if you count by twos from 16 to 28.
 (c) List the numbers you get if you count by hundreds from 1300 to 1700.
 (d) List the numbers you get if you count by twenties from 60 to 200.
- 2** (a) List the multiples of 50 from 50 to 250.
 (b) List the multiples of 25 from 100 to 200.
 (c) List the multiples of 200 from 1400 to 2000.
- 3** Write the following lists of numbers in order from smallest to largest.
 (a) 12, 22, 14, 20, 10, 16, 24, 6
 (b) 15, 0, 50, 20, 55, 100, 25, 5
 (c) $5, 1\frac{1}{2}, 6\frac{1}{2}, 1, \frac{1}{2}, 0, 3\frac{1}{2}, 5\frac{1}{2}, 2, 3, 2\frac{1}{2}, 4$
- 4** Copy and complete the pattern by adding the same number each time.
 (a) 12, ___, ___, 21, ___, 27
 (b) 120, ___, ___, ___, ___, 220
 (c) 225, ___, ___, 300, ___, 350
- 5** Write the following fractions in simplest form.
 (a) $\frac{2}{6}$ (b) $\frac{6}{8}$

Key Words

average	continuous data	measures of centre	sample space
bar graph	data	median	stem-and-leaf plot
bimodal	discrete data	mode	tally
bivariate data	dot plot	multimodal	theoretical probability
categorical data	frequency	outcomes	trend
chance	frequency table	outlier	univariate data
class interval	histogram	probability	
column graph	mean	range	

Collecting data

9.1

The study of statistics is the collection, interpretation and presentation of **data**. Data is information such as the colour of the family car or the amount of time Year 7 students spend on the internet every day. In this chapter, you will develop skills so that you can understand how to collect data, present it in useful forms, and interpret it.

Types of data

We can collect three different types of data.

The first type is **categorical data** because this information is sorted into groups of the same type. An example of categorical data would be the eye colour of a group of students because the data collected would be the number of students with blue eyes, brown eyes, green eyes etc. Another example could be the favourite fruit of a group of students because the data collected would be answers such as apples, bananas, pears, grapes, oranges etc.

The second and third types of data are numerical data, but they are quite different.

The second type of data is **discrete data** and this is data that is counted. Examples of this are the number of runs scored by a cricketer over a season or the number of emails you receive each day.

The third type of data is **continuous data** and this is data that is measured. Examples of continuous data are the weights of newborn babies at a hospital or the volume of water in a dam each day. The accuracy of the measurements depends on the measuring equipment used.

Suppose you asked pet owners the following questions about their pets.

- 1 What type of pet do you own?
- 2 How many pets do you own?
- 3 What does each of your pets weigh?

The first question will collect categorical data because the answers will be cat, dog, rat, tortoise, rabbit etc.

The second question will collect discrete data because the answers will be 1, 2, 3, 4 etc.

This third question will collect continuous data such as 5.3 kg, 34.7 kg, 24.3 kg, 2.6 kg etc.

Worked Example 1

WE1

For each of the following, state whether the data from a survey asking these questions is categorical, discrete or continuous.

- (a) How many television sets are in your house?
- (b) What is the height of the tallest person in your family?
- (c) How far do you have to travel to get to school?
- (d) What is your favourite dessert?

9.1

Thinking

(a) Are you counting, measuring or putting things in categories? (Here, we are counting the number of televisions.)

(b) Are you counting, measuring or putting things in categories? (Here, we are measuring height.)

(c) Are you counting, measuring or putting things in categories? (Here, we are measuring distance.)

(d) Are you counting, measuring or putting things in categories? (Here, we are putting things in categories.)

Working

(a) Discrete data

(b) Continuous data

(c) Continuous data

(d) Categorical data

Frequency tables

Any data that has been collected can be recorded in a **frequency table**. Here, each time a piece of data occurs, it is recorded in the **tally** column. We count the data in the tally column and write this number in the **frequency** column. Frequency means the number of times an event occurs.

Worked Example 2

WE2

A Year 7 class of 25 students was surveyed about the number of brothers and sisters they have. The results were: 2, 3, 5, 0, 1, 1, 2, 0, 3, 1, 4, 0, 1, 1, 3, 2, 2, 2, 1, 1, 0, 1, 1, 2, 4

- (a) Construct a frequency table showing this information.
 (b) Is this data categorical, discrete or continuous?

Thinking

- (a) 1 Construct a table with three columns that shows the item being surveyed, the tally and the frequency.
 2 Enter the data into the tally column. Start at the left of the list of numbers and work to the right. (Don't count all the 1s, then all the 2s etc.) Notice **||||** is used to represent 5.
 3 Enter these tallies as numbers in the frequency column. Add the numbers in this column to check that you have tallied correctly.

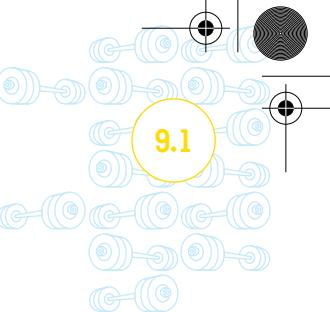
- (b) Is the data categorical, discrete or continuous?

Working

(a) Frequency table

Number of brothers and sisters	Tally	Frequency
0		4
1		9
2		6
3		3
4		2
5		1
	Total	25

- (b) It is discrete data.



9.1

Grouped data

When there is a large amount of numerical data spread out over many values, data is often grouped so that there are between five and 10 different groups. The spread of values is called the **range**.

Range = largest data value – smallest data value.

Data is usually grouped in intervals of 2, 5, 10 etc., so that we get between five and 10 groups when we divide our chosen interval into the range. The spread of the values in each group gives us the **class interval**. The class intervals are the groups in the first column of the frequency table. The class intervals for discrete data are presented in discrete groups; e.g. 0–9, 10–19, 20–29. The class intervals for continuous data are presented in continuous groups; e.g. 0–<10, 10–<20, 20–<30.

Worked Example 3

WE3

The Exceptional Felines basketball team keeps track of the points they score on the way to the State championship. The point totals are:

73, 84, 68, 45, 52, 44, 45, 52, 66, 42, 43, 40, 53, 47, 82, 76, 42, 57, 65, 81, 80, 40, 56, 72, 74, 83, 41, 66, 76, 75, 68, 81, 82, 79, 58, 81, 78, 80, 78, 76

Group the results and then construct a frequency table showing this information.

Thinking

- Decide on how to group the data by subtracting the smallest score from the highest to find the range. Then divide this number by 2, 5, 10 etc. until you get an interval size that gives between five and 10 groups. (If you get a decimal number, round to the nearest whole number.)
- Construct a table with three columns. Write the grouped values in the first column. Be careful with your groupings so that you don't miss values or include them twice. (In this example, we will use 40–44, 45–49, 50–54 and so on.)
- Use the tally column to help with the count. Remember to go through the data list systematically from the first piece of data to the last, tallying as you go.
- Enter the tallies as numbers in the frequency column. Add the numbers in this column to check that you have tallied correctly.

Working

$$\text{Range} = 84 - 40 \\ = 44$$

$$44 \div 5 = 8.8$$

$$\text{Number of groups} = \frac{44}{5} \\ = 8.8$$

So, we will group with an interval size of 5 and have 9 groups.

Number of points scored	Tally	Frequency
40–44		7
45–49		3
50–54		3
55–59		3
60–64		0
65–69		5
70–74		3
75–79		7
80–84		9
	Total	40

9.1

9.1 Collecting data

Navigator

**Answers
page 674**

Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8,
Q12, Q14, Q16

Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8,
Q9, Q10, Q12, Q14, Q16

Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8,
Q9, Q10, Q11, Q12, Q13, Q15

Fluency

WE1

- For each of the following, state whether the data from a survey asking these questions is categorical, discrete or continuous.
 - How do you travel to school?
 - What colour hair do you have?
 - How much time did you spend at the beach last summer?
 - What is your height?
 - How many cars passed by the school in a day?
 - What is your favourite book?
 - How many trees are in your front yard?

WE2

- A Year 7 class of 25 students was surveyed about the number of pets they own. The results were:

0, 0, 2, 1, 1, 0, 6, 9, 4, 3, 0, 1,
3, 2, 1, 1, 4, 5, 4, 3, 2, 1, 4, 7

- Construct a frequency table showing this information.
- Is this data categorical, discrete or continuous?

WE3

- The number of calls per day to the emergency help department of an automobile association was recorded for the month of June. The results were:

120	100	75	60	66	100
110	197	55	80	155	179
154	43	75	98	172	130
25	55	110	99	72	90
113	65	87	94	112	18

Group the results and then construct a frequency table showing this information.



- Students at a school noted the colour of teachers' cars in the car park. They found there were 13 black cars, 21 white cars, 25 silver cars, 7 red cars, 5 blue cars and 11 other colours.
 - How many cars were in the car park?
 - Construct a frequency table showing this information.
 - Is this data categorical, discrete or continuous?
- A Year 7 class was surveyed about how many hours of television they watched on a normal week night. The results were:

3, 2, $1\frac{1}{2}$, $\frac{1}{2}$, 1, $2\frac{1}{2}$, 1, 0, $4\frac{1}{2}$, $2\frac{1}{2}$, 2, $2\frac{1}{2}$, 2, 0, 4, 2, $1\frac{1}{2}$, $2\frac{1}{2}$, 3, $\frac{1}{2}$, 0, $\frac{1}{2}$, 1, $2\frac{1}{2}$, 2, 2, $\frac{1}{2}$

Construct a frequency table showing this information.

- 6 Class 7C was surveyed about their favourite takeaway food. The results were:

Sarah—pizza
Whitney—hamburgers
Morgan—Chinese
Kimiko—pizza
Liam—hamburgers
Anne—chicken
David—never eats takeawa
Malcolm—pizza
Gary—hamburgers
Dieter—chicken
Yang—pizza

Brent—fish & chips
Ruby—pizza
Romano—fish & chips
Gareth—chicken
Fleur—never eats takeaway
Jonty—pizza
Carol—Chinese
Tracy—hamburgers
Marcia—pizza
Douglas—pizza
Yusef—fish & chips

Copy and complete the frequency table below to show this information.

Type of takeaway	Tally	Frequency
Pizza		
Fish & chips		
?		
?		
?		
None		



- 7 A Year 7 class was surveyed about how much money they had in their pockets at that time. The results were:

The results were:

\$1.20	\$0.75	\$1.00	\$2.55	\$1.50	\$4.00	\$0.00	\$0.00
\$0.80	\$2.40	\$1.25	\$0.00	\$3.60	\$4.40	\$4.30	\$2.90
\$1.80	\$2.35	\$2.20	\$0.50	\$0.40	\$2.50	\$1.50	\$1.75

Group the results and then construct a frequency table showing this information.

Understanding

- ### **8** Look at the following data list.

2	8	9	15	22	24	17	16	28	31	40
15	52	16	51	42	39	27	26	42	55	46
17	22	18	5	47	46	33	21	19	21	17

Group the data using 0–9, 10–19, 20–29 etc. as the groups.

- (a) The frequency for the 10–19 group is:

A 6 **B** 8 **C** 9 **D** 10

- (b) The frequency for the last group is:

A 2 B 3 C 4 D 5



9.1

- 9** A class was surveyed about the number of hours spent exercising over a weekend. The results were as follows:
 1, 3, 2, 1, 1, 2, 0, 2, 0, 1, 1, 2, 6, 3, 4, 2, 1, 0, 2, 8, 3, 5, 1, 2, 0, 2
 (a) How many students were surveyed?
 (b) Construct a frequency table of this data.
 (c) How many students did not exercise at all?
 (d) How many students spent more than 2 hours exercising?
 (e) What number of hours was most commonly spent exercising by the members of the class?
 (f) What is the greatest number of hours spent exercising by a student?
 (g) One student hoped to qualify for the next Olympics, one student wanted to be drafted into the AFL and another student's netball team was in the finals. What data values are mostly likely to belong to these three students? Explain your answer.
- 10** The cast and crew of a school musical were to be given T-shirts advertising the production. There were five sizes available, XS (extra small), S (small), M (medium), L (large) and XL (extra large). The orders are recorded below:
 S, M, S, XL, L, M, L, S, XS, M, L, S, M, L, M, XL, S, M, M, XL, L, M, S, XS, L, S, XL, M, L
 Five students, all in Year 7, forgot to order a T-shirt.
 (a) How many students were in the cast and crew of the musical?
 (b) Construct a frequency table for this data.
 (c) What type of data is this?
 (d) What was the most popular T-shirt size ordered?
 (e) How many T-shirts are smaller than the M size?
 (f) The music teacher, Ms Campbell, could not decide whether to order one extra T-shirt in each size or five more M size T-shirts for the students who had not submitted an order. What would you advise her to do, explaining your advice?



Reasoning

- 11** Look at the following data set.

20	32	42	53	63	21	33	42	53	65	22	34	43	57	61
24	34	48	50	61	24	39	40	50	62	29	31	41	52	62

- (a) Construct a frequency table using class intervals of 20–29, 30–39 etc.
 (b) Describe what you find in your frequency table.
 (c) Construct a frequency table using class intervals of 20–24, 25–29, 30–34 etc.
 (d) Describe what you find in your frequency table.
 (e) What does this show?

- 12** Sixty 13-year-old students were tested to find their pulse rate (heart beat) when resting. The following beats per minute were measured.

70	68	68	76	79	68	76	55	55	60	60	94	72	65	64
93	71	62	67	82	76	65	77	82	81	59	74	74	67	68
78	76	63	82	81	82	74	70	66	63	84	81	69	84	79
71	70	54	68	64	78	58	84	61	75	72	73	71	91	66

- (a) Construct a frequency table using class intervals of 51–60, 61–70 etc.
 (b) Construct a frequency table using class intervals of 51–55, 56–60, 61–65 etc.
 (c) Which table shows the spread of results more clearly? Give a brief reason to support your answer.



- 13 The following is a list of birth weights (g) of 30 babies.

2900	2805	2925	2010	2720	3125	2670
2555	2963	2972	3151	2515	3529	3098
2126	2417	3000	3254	2997	2986	2719
2842	3519	3509	3218	3002	2437	2222
2019	2113					

Construct a frequency table for this data using a suitable class interval. Briefly explain why you chose this class interval.



Open-ended

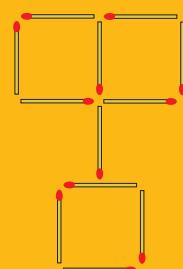
- 14 (a) A die is rolled 20 times. Construct a possible frequency table for the outcomes 1, 2, 3, 4, 5, 6 with no outcome having a frequency greater than 4.
- (b) Fifty dogs were weighed. Their weights ranged from 1 kg to 25 kg. Construct a possible frequency table if the data is grouped using class intervals of 1–5, 6–10, 11–15, 16–20 and 21–25 with no data group having a frequency greater than 15.
- 15 Find out how many hours of television each of the students in your class watches on a normal week night by surveying your class. Construct a frequency table of your results. Round your answers to the nearest half hour.
- (a) How many of your class members watch no television on a normal week night?
- (b) How many of your class members watch less than 1 hour of television on a normal week night?
- (c) How many of your class members watch 3 or more hours of television on a normal week night?
- (d) How much television do the majority of the members of your class watch on a normal week night?
- (e) Do you watch more or less television than the majority of the members of your class?
- 16 Find out the favourite school subject of each of the students in your class by surveying your class. Construct a frequency table of your results.
- (a) What is the favourite subject of the members of your class?
- (b) What is the second favourite subject of the members of your class?
- (c) Did any student have no favourite subjects?

Outside the Square Puzzle

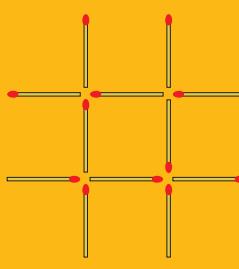
Matchstick puzzles



Remove six of the fifteen matches to leave ten.



Reposition four matches to make exactly five squares.



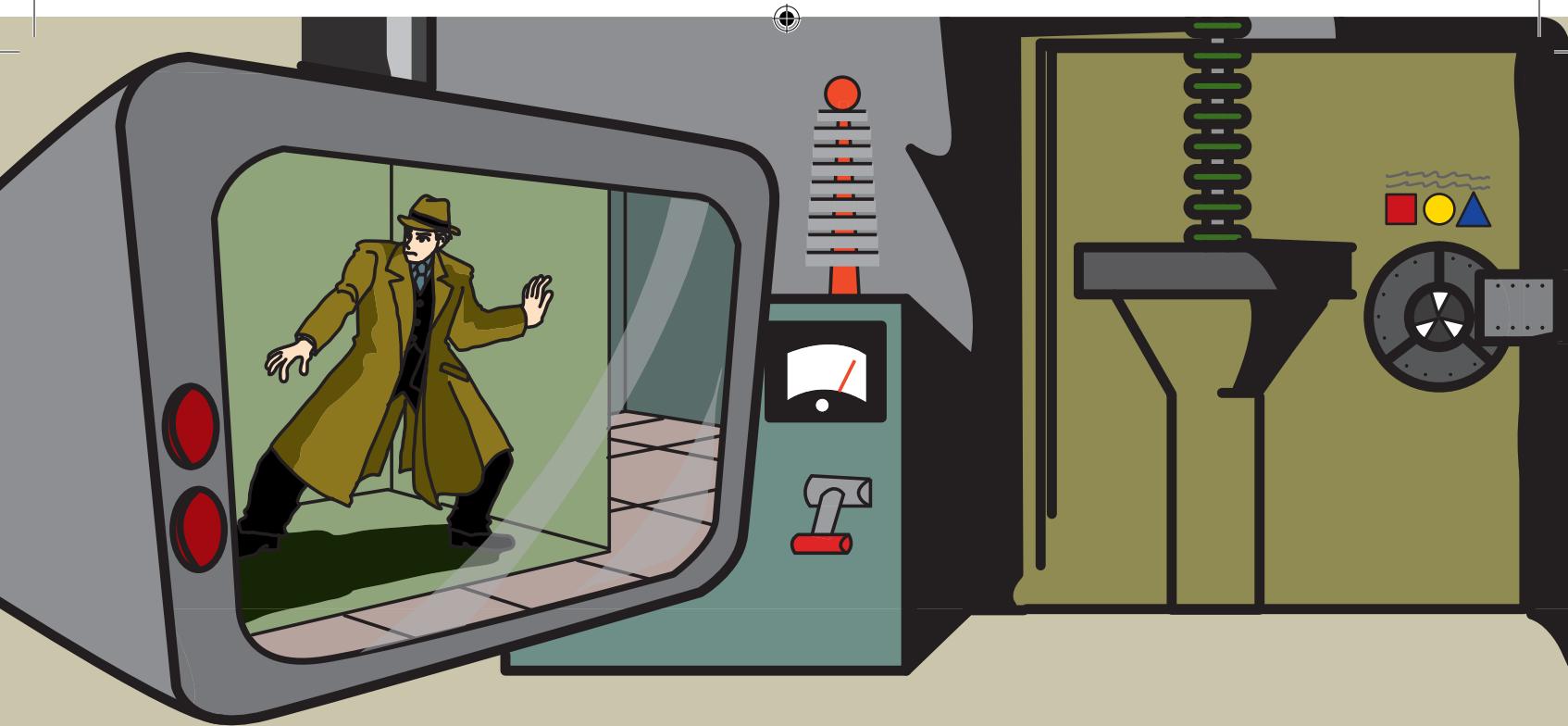
Reposition three matches to make exactly three squares.

Mathspace

Crack the Code

You have been captured by your archenemy Dr Code and find yourself in a locked room with the following message written on the door above three symbols—a red square, a yellow circle and a blue triangle. You know the coded message tells you the order in which to press the symbols to open the door, but how do you crack the code?





Code Box #1

UWJXX YMJ DJQQTB HNWHQJ YMJS YMJ WJI
XVZFWJ FSI KNSFQQD YMJ GQZJ YWNFSLQJ



You decide to try a simple letter shift code, knowing that E is the most frequently used letter in the English language. (In a letter shift code, each letter is moved forwards or backwards by the same amount. For example, if the shift is 3, then A would be written as D, B as E, C as F ... X as A, Y as B and Z as C.)

- Which letter in the code appears the most?
- What are the two possible shifts that would produce this match?
- Try the simpler of the two shifts. What does the message say?

Now you can escape the locked room!

So far, so good, but you now find yourself facing a much more difficult code to escape the next room. You know this because written above the new code Dr Code has written, 'Aha—not just a letter shift this time!'

You rightly suspect that this time the letter substitution is just random, so you will need to know the frequency with which letters are used in English. Happily, you know this and you write out the following table from memory. It is based on the number of times each letter would appear for every one hundred letters counted.

Frequency of letters used in English per 100 letters

A	B	C	D	E	F	G	H
8.17	1.49	2.78	4.25	12.70	2.23	2.02	6.09
I	J	K	L	M	N	O	P
6.97	0.15	0.77	4.03	2.41	6.75	7.51	1.93
Q	R	S	T	U	V	W	X
0.10	5.99	6.33	9.06	2.76	0.98	2.36	0.15
Y	Z						
1.97	0.07						

Again, the code requires you to press some buttons that look the same as the last ones.

Code Box #2

Aha—not just a letter shift this time!

NRCD NCJG EYGD NRG QOSG NYCFXUOG NRGX
NRG VGOOWT ACYAOG NTCAG FXP BCXFOOV
NRG YGP DUSFYG BWSY NCJGD



- Looking at the table above, what are the top five most frequently used letters?

- Draw up a frequency table for Code #2. Which letter is used most frequently? Put this against E in an alphabet chart. You may like to label the second-most frequently used letter too.
- Using your knowledge from Code #1 about the words Dr Code often uses in his messages, see if you can guess a few of the words. Use these to solve more of the letters in the code.
- What is the button sequence this time?

If you have broken the code you can escape.

9.2

Measures of centre

In everyday life, people talk about ‘the average temperature’ and ‘the average number of runs scored in an over in a cricket match’. The **mean**, commonly known as the **average** of a set of numbers, is a value we use to represent the centre of the data. However, this is not the only measure of the centre that we can use. The **median** and **mode** are two other ways of finding a central value of a whole group of data values. The mean, median and mode are all **measures of centre**, but each is calculated in a different way and gives us different information about the data we are interpreting.

Mean

The mean is the sum of all the values in the data set divided by the number of values in that data set.

$$\text{Mean} = \frac{\text{sum of data values}}{\text{number of data values}}$$

Median

The median is the *middle* value of an ordered data set.

To calculate the median, we arrange the data set in ascending order, then count the number of data values.

If the number of values is odd, select the middle value.

If the number of values is even, select the two middle values. Add them together and divide by 2.



Mode

The mode is the value that occurs *most often* in a data set.

Sometimes, a set of results can have more than one result that occurs most frequently. As an example, consider the following data: 1, 1, 3, 3, 3, 4, 5, 5, 6, 7, 7, 7, 8, 8, 9.

Here, we can see that both 3 and 7 occur three times. When there are two values that occur most frequently, we say that the results are **bimodal**. If it happens that there are three or more values that each have the maximum frequency, we say that the data is **multimodal**. The mode is not very useful in a multimodal set of data. We sometimes say that such data sets have no mode.

Worked Example 4

WE4

Find the (i) mean, (ii) median and (iii) mode of each of the following sets of test results.

(a) 9, 4, 5, 7, 8, 7, 2

(b) 10, 7, 3, 4, 8, 2, 6, 5

Thinking

- (a) (i) 1 Write the rule for finding the mean.
- 2 Substitute data into the rule.
- 3 Simplify and evaluate.
- 4 Write down the mean.
- (ii) 5 Order the values from smallest to largest. Count the number of values. If the number is odd, select the middle value.
- 6 Write down the median.
- (iii) 7 Does any value appear more often than any other? This is the mode. (Yes)
- 8 Write down the mode.

Working

$$\begin{aligned} \text{(a) (i) Mean} &= \frac{\text{sum of data values}}{\text{number of data values}} \\ &= \frac{9+4+5+7+8+7+2}{7} \\ &= \frac{42}{7} \end{aligned}$$

Mean = 6

$$\begin{array}{c} \text{(ii) } 2, 4, 5, 7, 8, 9 \\ \uparrow \\ 7 \end{array}$$

Median = 7

(iii)

Mode = 7

- (b) (i) 1 Write the rule for finding the mean.
- 2 Substitute data into the rule.
- 3 Simplify and evaluate.
- 4 Write down the mean.
- (ii) 5 Order the values from smallest to largest. Count the number of values. If the number is even, select the two middle values.
- 6 Find the mean of these two numbers to find the median.
- 7 Write down the median.
- (iii) 8 Does any value appear more often than any other? This is the mode. (No)
- 9 Write that there is no mode.

$$\text{(b) (i) Mean} = \frac{\text{sum of data values}}{\text{number of data values}}$$

$$\begin{aligned} &= \frac{10+7+3+4+8+2+6+5}{8} \\ &= \frac{45}{8} \end{aligned}$$

= 5.625

Mean = 5.625

(ii) 2, 3, 4, 5, 6, 7, 8, 10

2, 3, 4, 5, 6, 7, 8, 10

Median = $\frac{5+6}{2}$

Median = 5.5

(iii)

There is no mode.

9.2

9.2 Measures of centre

Navigator

**Answers
page 676**Q1 Column 1, Q2 (a)–(d), Q3, Q4,
Q5, Q6, Q9Q1 Column 2, Q2, Q3, Q4, Q5,
Q6, Q8, Q9Q1 Column 2, Q2, Q3, Q4, Q5,
Q6, Q7, Q8, Q10**Equipment required:** Calculator for Question 2

Fluency

WE4

- 1 Find the (i) mean, (ii) median and (iii) mode of each of the following sets of test results.

- | | |
|----------------------------------|----------------------------------|
| (a) 7, 1, 7 | (b) 5, 8, 5 |
| (c) 6, 6, 6, 6, 6 | (d) 4, 4, 4, 4, 4, 4, 4 |
| (e) 7, 6, 7, 5, 7, 9, 7, 8 | (f) 6, 5, 5, 5, 9, 9, 10 |
| (g) 8, 7, 9, 5, 8, 2, 3 | (h) 2, 8, 2, 6, 8, 6, 1, 2, 1, 4 |
| (i) 4, 3, 6, 4, 8, 2, 5, 7, 6, 4 | (j) 5, 3, 8, 7, 2, 5, 9, 5 |
| (k) 1, 6, 2, 7, 9, 8, 9 | (l) 6, 3, 2, 8, 9, 2, 5 |

- 2 Find the (i) mean, (ii) median and (iii) mode of each of the following sets of data. Round your answers to two decimal places if necessary.

- | |
|--|
| (a) 29, 36, 29, 23, 32, 25 |
| (b) 67, 73, 72, 82, 50, 45, 76 |
| (c) 56, 72, 39, 47, 89, 81, 63, 81 |
| (d) 12.8, 6.5, 19.3, 14.2, 4.7, 6.8, 6.5, 11.3 |
| (e) 16.3, 15.8, 11.4, 15.2, 12.3, 11.4, 13.5 |
| (f) 11.01, 11.001, 10.1, 10.11, 10.01, 1.11, 11.11, 1.01, 11 |

Understanding

- 3 *Without doing any calculations*, in each case, choose which one of the alternatives could be the mean.

- | | |
|------------------------------------|---|
| (a) 9, 7, 8, 7, 7, 8, 6, 6, 9 | A 3 B 5 C 7 D 10 |
| (b) 6, 6, 6, 7, 6, 7, 4, 7, 2, 7 | A 2 B 4 C 6 D 8 |
| (c) 12, 19, 14, 18, 11, 17 | A 10 B 12 C 16 D 19 |
| (d) 86, 74, 85, 68, 91, 76, 80, 90 | A 74 B 81 C 89 D 91 |



- 4 The Korea–Japan 2002 World Cup in soccer was played in more stadiums than ever before. This record was not beaten in Germany in 2006 (12 stadiums) or in South Africa in 2010 (10 stadiums). The following table shows the city and capacity of the different stadiums.

Korea	Capacity
Seoul	63 961
Incheon	52 179
Suwon	43 188
Daejeon	40 407
Daegu	68 014
Jeonju	42 391
Gwangju	42 880
Ulsan	43 550
Busan	55 982
Seogwipo	42 256



Japan	Capacity
Sapporo	42 000
Miyagi	49 000
Niigata	42 300
Ibaraki	42 000
Saitama	63 000
Shizuoka	50 600
Kobe	42 000
Osaka	50 000
Oita	43 000
Yokohama	70 000

- (a) Find the (i) mean and (ii) median capacity for the stadiums in Korea.
- (b) Which stadium has closest to the mean capacity in Korea?
- (c) Find the (i) mean and (ii) median capacity for the stadiums in Japan.
- (d) Which stadium has closest to the mean capacity in Japan?
- (e) Use the mean and median to compare the sizes of the stadiums in the two countries.
- (f) Find the (i) mean and (ii) median capacity for all 20 stadiums.
- (g) Which stadium is closest to the mean capacity for all 20 stadiums?

Reasoning

- 5 Consider the following test results: 7, 5, 6, 8, 3
- (a) Do you expect the mean to be an actual data value? Explain your answer.
 - (b) Find the mean of these results. Is the mean an actual data value?
 - (c) Do you expect the median to be an actual data value? Explain your answer.
 - (d) Find the median for the test results above.
- 6 (a) Explain what is meant by the statement ‘The average family has 2.3 children’.
- (b) Explain why the median of a group of whole-number values can only ever either end in .5 or be a whole number.
 - (c) If the data values are all whole numbers, can the mode ever be a decimal? Explain why or why not.

9.2

- 7 A small company has a manager, an assistant manager, two office workers and 10 factory workers. The manager is paid \$125 000 per year, the assistant manager \$85 000, the office workers \$55 000 each, and the factory workers \$50 000 each.
- Find the (i) mean, (ii) median and (iii) mode for the annual income of all the people in the company, rounding your answer to the nearest cent.
 - Which measure of centre gives you the best idea about the salaries paid by this company?
 - Omit the manager's salary and recalculate the (i) mean, (ii) median and (iii) mode.
 - Comment on what you found.

Open-ended

- 8 As well as giving a set of results, make sure you include a short explanation of how you found your answers for each of the following.
- Find five test results for which the mean, median and mode are all equal to 6 if the test results are not all the same.
 - Find five test results for which the mean and the median are 7 and the mode is 9.
 - Find five test results for which the mean is 7 and the median and mode are 6.
- 9 Three students are finding the median of the following data values.
20, 17, 28, 23, 19, 21, 22, 20
Sam says 'The median is the fourth value, so is 23.'
Veran says 'No, you need to put the data values in order first, so the median is 20.'
Lakme says 'No, you are both wrong.'
Help Lakme explain to Sam and Veran why they are incorrect.
- 10 Nine friends went tree planting at their local creek. The median number of trees planted was 5, the mean was 6 and the mode was 4. Find a data set for the number of trees each person planted to match the given statistics.

Outside the Square Problem solving

Running the race

- In a sprint, the first three runners had times of 12.1, 12.1 and 10.9 seconds. Find the mean, median and mode for the three times.
- In a different race, the first three runners had a mean time of 11.7 seconds, a median time of 10.8 seconds and a modal time of 10.8 seconds. Find the times for the three runners.



- In a third race, the first three runners had a mean time of 11.8 seconds and a median time of 11.9 seconds. If all three times are different what can you say about the first and third placegetters?

Give at least three pairs of possible times for these two runners, using one decimal place accuracy in your solutions. Explain how you found your answers.



Strategy options

- Work backwards.
- Test all possible combinations.

Graphing univariate data

9.3

Information presented in tables is useful, but we often find it much easier to interpret this information if it is displayed in a graph. There are many types of graphs we can use for displaying data. In this section we will explore graphs of **univariate data**, data that consists of only one variable. A variable in statistics describes data that can change or vary.

Dot plots

A **dot plot** is a very simple statistical graph where, for each data value, a dot is placed on a scale. When drawing a dot plot it is important to space the dots evenly to give an accurate picture of the data set. A dot plot can be used for categorical or discrete data such as car colour or the number of people who live in each of the houses on a street.

Worked Example 5

WE 5

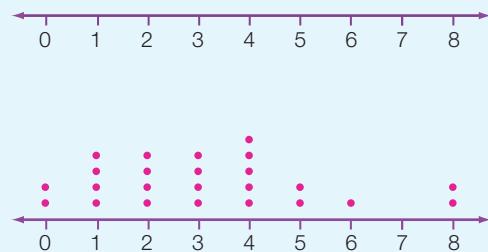
Draw a dot plot to illustrate the following data that represents the number of pets owned by each student in a class.

2, 5, 8, 6, 2, 4, 3, 4, 2, 1, 8, 3, 4, 0, 3, 1, 0, 3, 5, 1, 4, 2, 1, 4

Thinking

Working

- Identify the lowest and the highest values and make these the end points of a scale.
- Complete the plot by marking a dot for each of the data values, being careful to space them evenly.



Stem-and-leaf plots

When there is a large amount of numerical data it is often convenient to combine several pieces of data into a group. **Stem-and-leaf plots** are useful because they can group data without losing individual data values.

A stem-and-leaf plot consists of two parts, a stem and a leaf. Each piece of data must be divided into two parts so it can be plotted. For example, a score of 23 would be written as:

STEM	LEAF
2	3

The smallest place value is used as the leaf with the remaining part of the number as the stem. The leaf can only contain the digits 0 to 9. For example, 34.7 would be written as:

STEM	LEAF
34	7

A stem-and-leaf plot should always include a key; e.g. $34 | 7 = 34.7$

Note that a decimal point is not used in the plot.

9.3

As another example, 536 would be written as:

STEM	LEAF
53	6

$$53 | 6 = 536$$

Before stem-and-leaf plots can be used to find statistics, we need to arrange the numbers in the leaf section in ascending order to construct an ordered stem-and-leaf plot.

Worked Example 6

WE6

A class obtained the following percentage results (rounded to the nearest whole number) in a science test.

Express the following data as an ordered stem-and-leaf plot with a class interval of 10.

69, 54, 91, 42, 73, 75, 81, 96, 88, 76, 65, 83, 74, 76, 82, 61, 57, 64, 58, 79, 83, 90, 73, 60, 51

Thinking

- Decide upon the stem you will use for your stem-and-leaf plot. This will be the key for your stem-and-leaf plot.
- List the stems down the page in ascending order. In this case, the tens digit is the stem (4, 5, 6, 7, 8, 9). Write the key you are using beside the stem-and-leaf plot.
- Work through the data in the order given. List the leaf of each individual result in the appropriate row. The leaf is the units digit.
- Order the data in the leaf part of the stem-and-leaf plot, putting the lowest value closest to the stem.

Working

STEM	LEAF	Key
4	2	5 4 = 54
5	4 7 8 1	
6	9 5 1 4 0	
7	3 5 6 4 6 9 3	
8	1 8 3 2 3	
9	1 6 0	

STEM	LEAF	Key
4	2	5 4 = 54
5	1 4 7 8	
6	0 1 4 5 9	
7	3 3 4 5 6 6 9	
8	1 2 3 3 8	
9	0 1 6	

It is important that there are enough stems to spread the data. When data values are very close together, each division of the stem is divided into two parts, a lower and an upper section. The lower part contains the digits 0 to 4 in the leaf and the upper part contains the digits 5 to 9 in the leaf. As an example of this, consider the following data set.

21, 32, 41, 22, 25, 29, 33, 42, 33, 44, 21, 28, 37, 40, 39, 36, 24, 32, 44, 30, 34

If we used the tens value as the stem we would have only three stems (2, 3 and 4). To spread out the data we can use 2_L for the lower values 20, 21, 22, 23, 24 and 2_U for the upper values 25, 26, 27, 28, 29. The stem-and-leaf plot would look like this:

STEM	LEAF	Key
2_L	1 1 2 4	3 2 = 32
2_U	5 8 9	
3_L	0 2 2 3 3 4	
3_U	6 7 9	
4_L	0 1 2 4 4	

Outliers

Sometimes, in a data set, there are one or two values that are quite different from the other data values. These may be called **outliers**. Outliers may simply be errors made in recording or in measuring incorrectly or they may be the result of extremely unusual circumstances. If outlier values can be explained reasonably, they are often excluded from data analysis, as the statistics obtained if they are included can be misleading. Values can often be identified as possible outliers in both dot plots and stem-and-leaf plots.

Using dot plots and stem-and-leaf plots

Both dot and ordered stem-and-leaf plots can be used to find the median and the range of the data set because the data is organised in both types of plots. The range can also be found, as the maximum (largest) and the minimum (smallest) values are retained in both types of plots.

Worked Example 7

WE7

Calculate the median and range for each of the following data sets. Comment on whether there are any possible outliers in the data.

- (a) The students in a class were asked to indicate the number of brothers and sisters they have. These are the results obtained:

1 1 2 5 2 3 2 0 0 4 2 3 9 1 0 2 1 4 1 5 3 2 0

Draw a dot plot and use it to help you answer the question.

- (b) A new display at a museum attracts varying numbers of people. A survey was conducted to count the number of people visiting the attraction each day. These are the results obtained for the first 20 days:

235 195 156 188 177 165 189 190 177 166
199 186 165 170 169 173 184 177 192 200

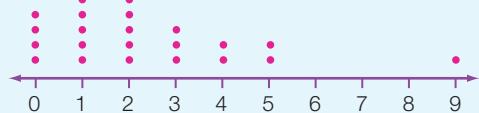
Draw an ordered stem-and-leaf plot and use it to help you answer the question.

Thinking

Working

- (a) 1 Construct the dot plot in the normal way.

(a)



Median = 2

- 2 Identify the middle value to find the median. Count from the left. (As there are 23 values here, the middle value is the 12th value.)

- 3 Find the range by subtracting the lowest value from the highest value.

- 4 Comment on any possible outliers.

Range: $9 - 0 = 9$

It would seem that 9 is an outlier for this data set.

9.3

- (b) 1 Construct the (ordered) stem-and-leaf plot.

(b) STEM	LEAF
15	6
16	5 5 6 9
17	0 3 7 7 7
18	4 6 8 9
19	0 2 5 9
20	0
21	
22	
23	5

Key 15 | 5 = 155

- 2 Identify the middle values and use them to find the median. Count from the left.
 3 Find the range by subtracting the lowest value from the highest value.
 4 Comment on any possible outliers.

$$\text{Median} = \frac{177 + 184}{2} \\ = 180.5$$

$$\text{Range: } 235 - 156 = 79$$

It would seem that 235 is an outlier for this data set.

Column and bar graphs

Column graphs present data in vertical columns, with a space between each column. **Bar graphs** are the same as column graphs except the bars are horizontal. Column or bar graphs should be used for categorical data and for ungrouped discrete data.

Drawing a column graph

All column graphs should include:

- 1 a title
- 2 categories or a numerical scale with equally spaced intervals (gaps) between the numbers on the horizontal axis
- 3 a numerical scale with equally spaced intervals (gaps) between the numbers on the vertical axis
- 4 a label for the horizontal axis (x -axis) that explains the variable being represented
- 5 a label for the vertical axis (y -axis) that shows the frequency
- 6 a small gap between each column (usually half a column width) to show that this is categorical or discrete (counted) data.

Worked Example 8

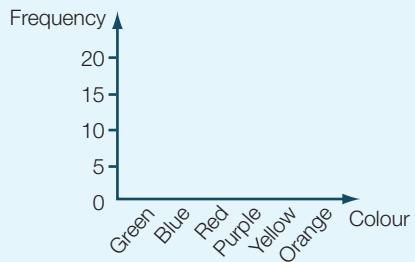
WE8

The frequency table below shows the responses from a group of children who were asked 'What is your favourite colour?' Only one response per student was recorded. Draw a column graph to represent this information.

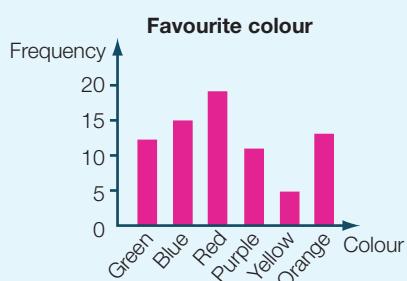
Colour	Green	Blue	Red	Purple	Yellow	Orange
Frequency	12	15	19	11	5	13

Thinking

- 1 Draw axes. Scale and label the x - and y -axes appropriately.

Working

- 2 Use the frequency values to draw a rectangular column for each category. (Here, each rectangle represents a different colour.) The length of the rectangle indicates the frequency. Ensure that each rectangle has the same width on the x -axis. Insert a small space between each category. Add a title.

**Histograms**

Histograms present data in vertical columns, but they are different from column graphs. The columns have no space between them and each column deals with a range of values rather than a single value. Histograms should be used for continuous data.

Drawing a histogram

All histograms should include:

- 1 a title
- 2 a continuous scale on the x -axis for the variable being represented
- 3 a numerical scale with equally spaced intervals between the numbers on the y -axis
- 4 a label for the horizontal axis that explains the variable that is being represented
- 5 a label for the vertical axis that shows the frequency
- 6 no gap between each column
- 7 a half-column width before the first column even if the first set of data starts at zero.

Worked Example 9

WE9

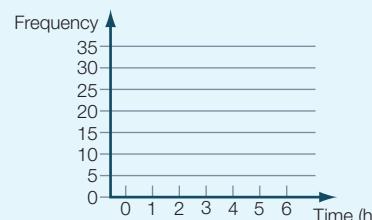
100 Year 7 students were surveyed about the number of hours they spent watching TV over the weekend. The data is presented below.

Hours spent watching TV over the weekend	0-<1	1-<2	2-<3	3-<4	4-<5	5 or more hours
No. of students	8	17	27	32	10	6

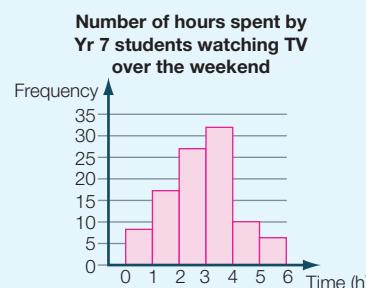
Draw a histogram to represent this data.

9.3**Thinking**

- 1** Draw axes; scale and label the x - and y -axes appropriately.

Working

- 2** Ensure rectangles each have the same width on the x -axis but with no gaps between the x -axis categories. Add a title.



Section 9.1 described three different types of data. The table below shows the appropriate graph for each data type.

Question	Data type	Examples of answers	Appropriate graph
What kind of animal do you own?	Categorical	Cat, dog, rat, tortoise, rabbit	Column/bar
How many pets do you own?	Discrete	1, 2, 3, 4	Column
What does each of your pets weigh?	Continuous	5.3 kg, 34.7 kg, 24.3 kg, 2.6 kg	Histogram

9.3 Graphing univariate data

Navigator

**Answers
page 677**

Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8,
Q9, Q10, Q11, Q14

Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8,
Q9, Q11, Q12, Q13, Q14

Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8,
Q9, Q10, Q11, Q12, Q13, Q14

Fluency

WE5

- 1** Draw a dot plot to illustrate the following data that represents the number of birds counted by a bird watcher in the bush in 15-minute intervals over the course of 6 hours.

12, 18, 12, 16, 19, 13, 19, 20, 17, 15, 20, 21, 18, 15, 20, 18, 21, 21, 17, 12, 17, 19, 12, 21

- 2 A class obtained the following percentage results (rounded to the nearest whole number) in a science test.

Express each of the following data as an ordered stem-and-leaf plot with a class interval of 10.

WE6

- (a) 57, 48, 42, 41, 31, 22, 37, 30, 23, 28, 34, 52
 (b) 25, 44, 15, 29, 37, 26, 33, 35, 14, 8, 2, 27, 30, 42, 34, 21, 18, 3

- 3 Calculate the median and range for each of the following data sets. Comment on whether there are any possible outliers in the data.

WE7

- (a) The students in a class were asked to indicate their shoe size. These are the results obtained:

5, 5.5, 7, 4.5, 4.5, 6, 10, 5.5, 4, 4, 5.5, 4.5, 6.5, 7.5, 6, 5.5, 4.5, 4, 6.5

Draw a dot plot and use it to help you answer the question.

- (b) The number of points scored by each of the teams in the opening round of the local Australian Rules Football competition was recorded. These are the results obtained:

110, 75, 88, 56, 59, 89, 112, 115, 121, 99, 67, 88, 77, 88, 100, 55

Draw an ordered stem-and-leaf plot and use it to help you answer the question.

- 4 The frequency table below shows the colour of cars passing the school in a particular timed period. Draw a column graph to represent this information.

WE8

Colour	White	Red	Green	Blue	Silver	Other
Frequency	42	6	5	12	18	17

- 5 The owner of a car dealership wants a graph of the number of cars sold plotted against the value of the cars sold in the last year. The data is presented below.

WE9

Value of car	Under \$10000	\$10000–< \$20000	\$20000–< \$30000	\$30000–< \$40000	\$40000–< \$50000	\$50000 or more
Number of cars sold	25	73	132	43	18	9

Draw a histogram to represent this data.

- 6 Express the following data as ordered stem-and-leaf plots with class intervals of 5.

- (a) 23, 28, 34, 27, 38, 42, 31, 21, 22, 27, 20
 (b) 48, 57, 56, 53, 68, 64, 47, 58, 52, 51, 61, 42, 47

- 7 The number of chocolate bars produced at a factory every hour over a 20-hour period is 146, 128, 133, 138, 140, 142, 125, 129, 136, 130, 133, 142, 126, 129, 150, 134, 126, 121, 128, 143.

- (a) Using class intervals of 5, produce an ordered stem-and-leaf plot of the data.
 (b) Draw an ordered stem-and-leaf plot of the data if it is expressed in intervals of 10.

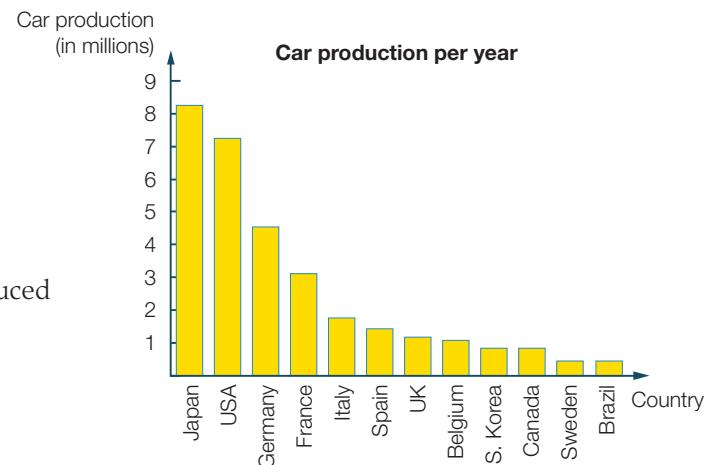
9.3

Understanding

- 8 This column graph shows the numbers of cars produced in a year by the world's major car-producing countries.

- (a) Which country produces the most cars?

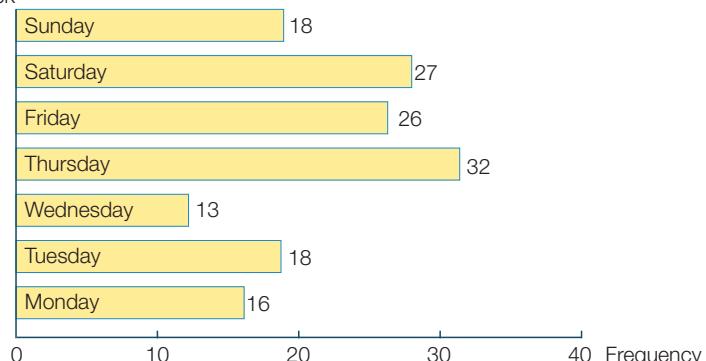
- A USA
 - B Germany
 - C Japan
 - D Brazil
- (b) The number of cars produced by Germany is:
- A 4.5
 - B 450 000
 - C 4 500 000
 - D 45 000 000



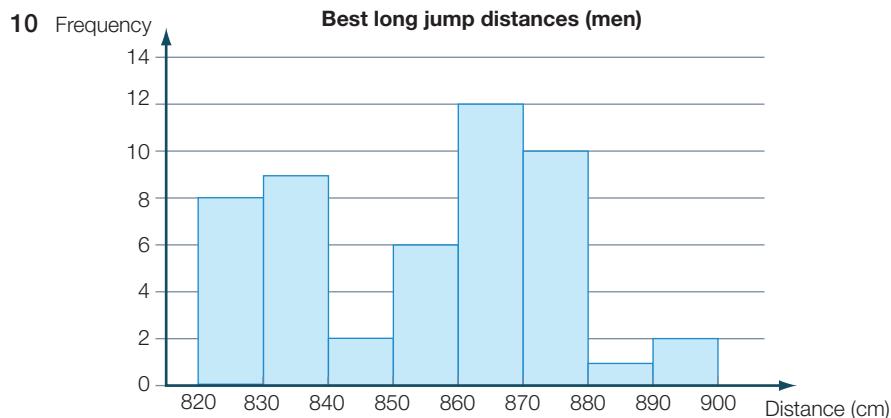
- 9 The bar graph at right shows the results for a survey of 150 people as to which day they were most likely to do their supermarket shopping. The actual numbers are included on the graph to make it easier for you to calculate your answers.

Days of the week

Most likely shopping day



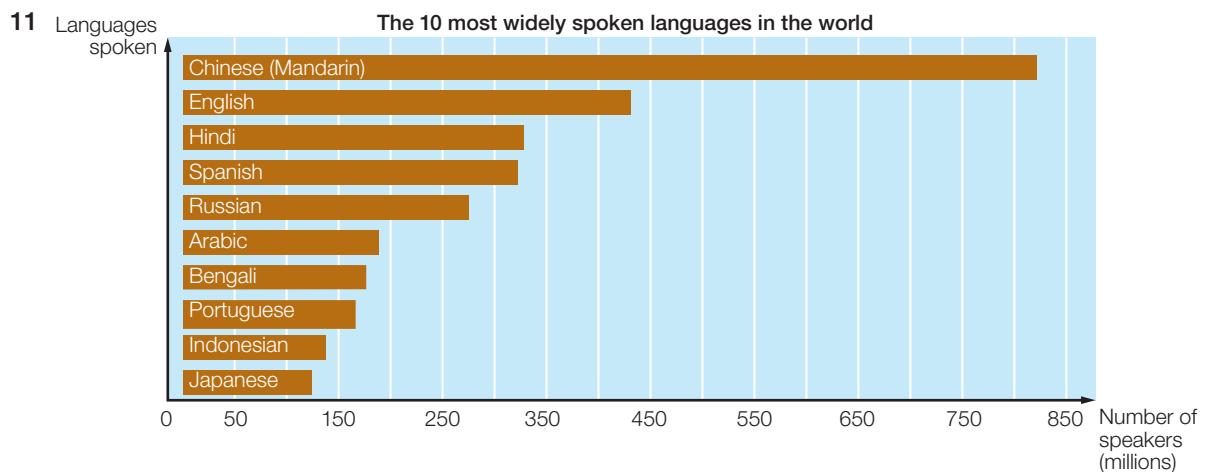
- (a) Which day was most popular?
 (b) How many people chose the least popular day?
 (c) How many people chose the weekend?



The histogram above shows the best long jump distances for men recorded from 1960 to 2009.

- What is the class interval?
- What is the range of the recorded distances?
- Which class interval had six recorded jumps?
- How many jumps over 870 cm were recorded in this time period?
- How many jumps have been recorded? (This will be the sum of all the frequencies.)

Reasoning



The horizontal bar graph above shows the 10 most widely spoken languages in the world.

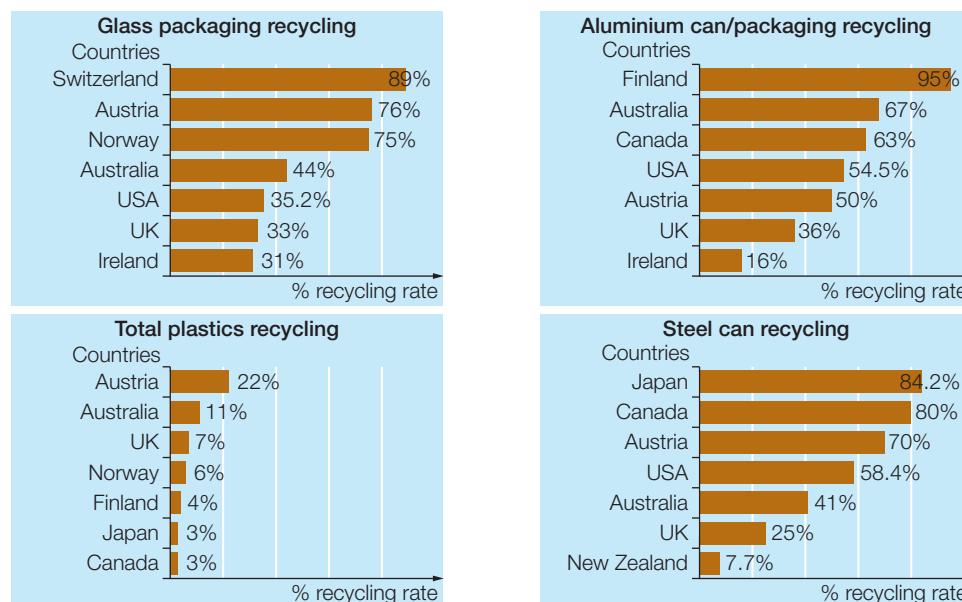
- What are the differences between this and a column graph?
 - Approximately how many people in the world speak Mandarin?
 - Approximately how many people in the world speak English?
 - What is the fourth most-spoken language in the world?
 - What is unusual about the scale at the bottom of the graph?
 - Could this have been drawn as a column graph?
- 12** A class of Year 7 students were asked how long it took them to get to school each day. The times, to the nearest minute, are listed below:

34, 13, 26, 38, 23, 15, 19, 30, 28, 40, 36, 25, 14, 7, 10, 11, 8, 24, 32, 69, 29, 32, 44, 21, 18, 22

- Is this discrete or continuous data?
- Group the data with appropriate class intervals.
- Draw an appropriate graph to represent the data.
- Is there a possible outlier? Why or why not?

9.3

- 13 Compare the four recycling graphs to answer the following questions.



- Does Australia or Canada have a better recycling rate for aluminium?
- Does Australia or Canada have a better recycling rate for steel cans?
- Australia, the UK and Austria appear in all four graphs. Which country seems to be the worst recycler of the three?
- Which kind of recycling does Australia seem to be best at?
- Which kind of recycling does Australia seem to be worst at?
- Compare Australia's recycling rate in all four areas with the best recycling countries on the graphs given.
- Compare Australia's recycling rate in all four areas with the worst recycling countries on the graphs given.

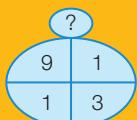
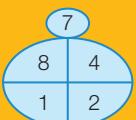
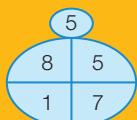
Open-ended

- 14 Collect some **(a)** categorial, **(b)** discrete and **(c)** continuous data from the students in your class and draw an appropriate graph to represent each.

Outside the Square Puzzle

Number puzzles A

- 1 Find the number that completes the sequence of diagrams:



- 2 Find as many solutions as you can for the following, where A , B and C are all different and are all positive whole numbers less than ten. Each question can have new values for A , B and C .

$$A \times A = B \times C$$

$$A \times A = B + C$$

$$A \times A = B \div C$$

$$A + A = B \times C$$

Comparing parts of a whole with graphs

9.4

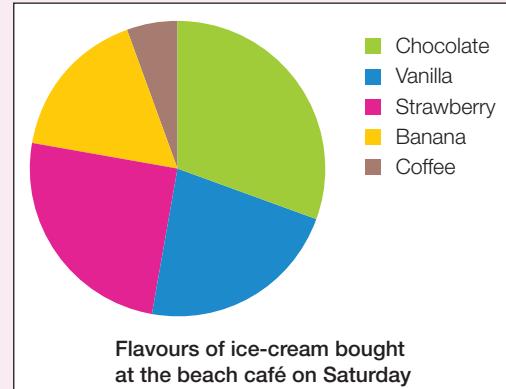
Often, data given to us are the various parts of a whole. For example, a budget shows all the separate items on which a total amount of available money will be spent. The two types of graphs we use to display this type of data are sector graphs and divided bar graphs. Both of these break a whole into parts.

Sector graphs

A sector graph is often called a pie graph because it resembles a pie divided into pieces from the centre of the circle.

Drawing sector graphs

- 1 Add the various parts to find a total.
- 2 Divide each part by the total to obtain fractional parts of a whole.
- 3 Multiply each fractional part by 360° to find the sector angle.
- 4 Use a protractor to divide a circle into the various parts and draw your graph.
- 5 Sector graphs should have titles and a key that identifies each of the sectors.



Divided bar graphs

Another way of showing parts of a whole is to divide a rectangle according to the size of the pieces.

Drawing divided bar or column graphs

- 1 Draw a rectangle, making the length a multiple or a factor of the total.
- 2 Multiply each data value by the same factor or multiple to find the length of each section.

The sector graph above can be represented by a divided bar graph as shown below.



9.4

Worked Example 10

We 10

For the data in the table opposite:

- (a) draw a sector graph
- (b) draw a divided bar graph.

All Year 7 students at a certain school were surveyed about their favourite type of pet and the responses were recorded in the table opposite.

Type of pet	Number recorded
Dog	78
Cat	65
Bird	39
Rabbit	26
Other	52

Thinking

- (a) 1 Add the number of responses to find a total.
- 2 Find the fraction of the whole by dividing the number recorded for each type by the total and record your results in a table.

Working

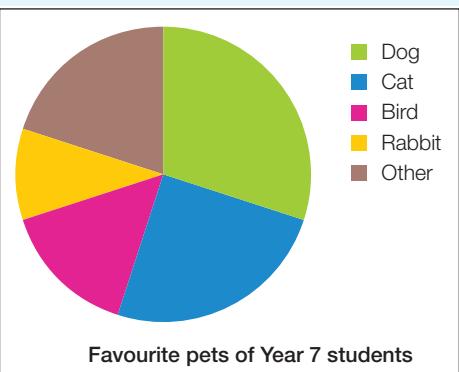
(a) 260

Type of pet	Number recorded	Fraction of a whole
Dog	78	$\frac{78}{260}$
Cat	65	$\frac{65}{260}$
Bird	39	$\frac{39}{260}$
Rabbit	26	$\frac{26}{260}$
Other	52	$\frac{52}{260}$
Total	260	1

- 3 Multiply the fraction for each type by 360° and round your answer to the nearest degree. Check that the sum of your angles is 360° . (It may be very slightly more or less after rounding.)

Type of pet	Number recorded	Fraction of a whole	Sector angle ($^\circ$)
Dog	78	$\frac{78}{260}$	108
Cat	65	$\frac{65}{260}$	90
Bird	39	$\frac{39}{260}$	54
Rabbit	26	$\frac{26}{260}$	36
Other	52	$\frac{52}{260}$	72
Total	260	1	360

- 4 Rule a line from the centre to the edge of the circle. Use this line and a protractor to measure the sector angles and rule each sector. Give your graph an informative title and provide a key for each of the sectors.



(b) 1 Add the number of responses to find a total.

- 2 Choose a length that is a multiple or a factor of the total. (As 130 mm is half of 260, draw a line 13 cm long.) Multiply by the same factor or multiple to find the length of each section of the graph. (Multiply each recorded response by $\frac{1}{2}$.)

- 3 Draw a divided bar graph using a ruler to measure the lengths of each section.

(b) 260

Type of pet	Number recorded	Length of line (mm)
Dog	78	39
Cat	65	32.5
Bird	39	19.5
Rabbit	26	13
Other	52	26
Total	260	130

Favourite pets of Year 7 students



9.4

9.4 Comparing parts of a whole with graphs

Navigator

**Answers
page 679**

Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8,
Q12

Q1, Q2, Q3, Q4, Q6, Q7, Q8, Q9,
Q11, Q12

Q1, Q2, Q3, Q4, Q6, Q8, Q9,
Q10, Q11, Q12, Q13

Equipment required: A protractor for Questions 8, 10 and 11, a computer containing Excel software or similar may be used for Question 13

Fluency

W.E.10

- 1 For the data in each of the following tables:

- (i) draw a sector graph
- (ii) draw a divided bar graph.

- (a) Belinda picked some vegetables from her garden and recorded the number of each type of vegetable.

Vegetable	Number recorded
Tomatoes	20
Carrots	12
Cucumbers	8
Lettuce	10

- (b) A group of tourists was asked to state their nationality and their responses were recorded.

Nationality	Number recorded
Italian	22
French	18
Japanese	12
Indonesian	8

- (c) The number of pizzas purchased one evening at Abracadabra Pizza Parlour was recorded.

Pizzas	Number recorded
Super Supreme	6
Capricciosa	8
Hawaiian	4
Italiana	3
Australiana	2



- (d) The number of each make of car in the staff car park at a small secondary school was recorded.

Car make	Number recorded
Ford	4
Holden	5
Mazda	4
Toyota	3
Mitsubishi	2
Other	5



- 2 For each of the following data tables:

- (a) draw a sector graph
 (b) draw a divided bar graph.

(i) Fifty people were asked their favourite music style.

Music style	Rock & Roll	Hip Hop	Techno	Rap	Punk	Pop
Number	10	9	6	3	5	17

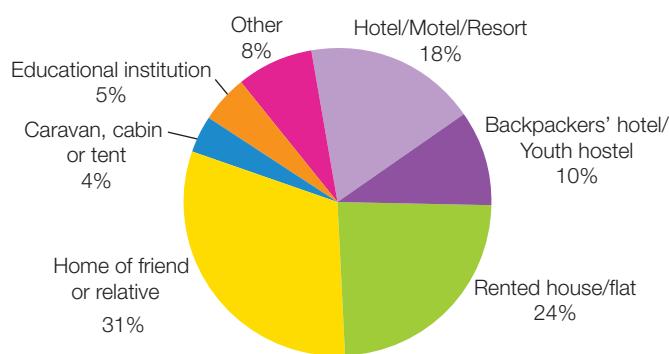
(ii) Seventy-five people were asked their favourite vegetable.

Vegetable	Potato	Carrot	Pumpkin	Cabbage	Capsicum	Cauliflower
Number	33	22	9	4	4	3

(iii) Justin has been working in a full-time job for 6 months. He was asked how he spent his money. He gave the following estimates.

Item	Rent	Food	Entertainment	Clothes	Other
Percentage	47	30	10	10	3

- 3 The sector graph below represents information about the type of accommodation used by overseas visitors to Australia.



- (a) One of the sectors uses an angle of 64.8° . Which one is it?

- A Hotel/Motel/Resort B Backpackers' hotel
 C Rented house/flat D Other

- (b) The Educational Institution sector uses an angle closest to:

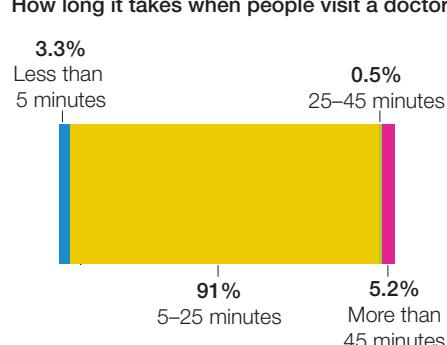
- A 5° B 10° C 13° D 18°

9.4

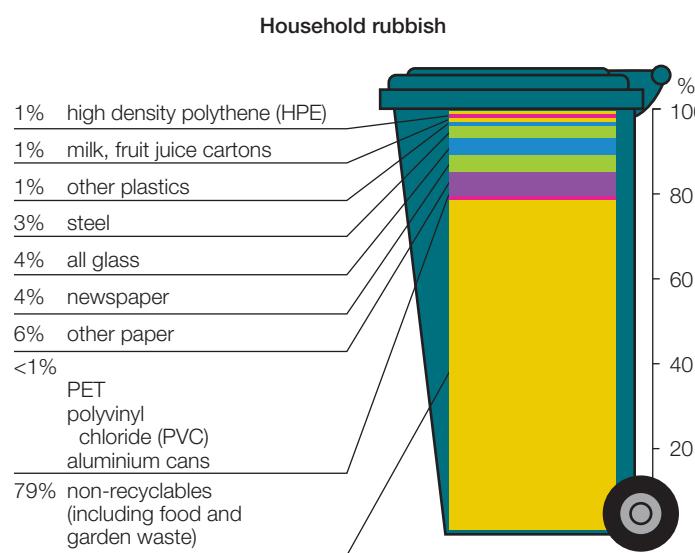
- 4 As part of a project to tackle global warming at a local level, a Victorian local council conducted a survey to see what percentage of greenhouse gas emissions was produced by each of its facilities and services. The following data was recorded: Buildings 28%, Streetlights 47%, Waste 1%, Vehicle fleet 24%.
- Create a sector graph to display the data.
 - Create a divided bar graph to display the data.
 - Which graph do you find easier to understand?

Understanding

- 5 The following graph shows how long people take when they visit a doctor.

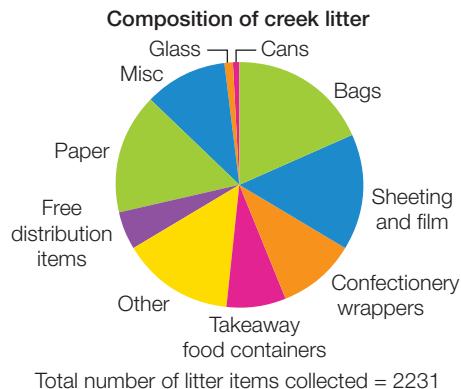


- What percentage of people spend less than 5 minutes when they visit a doctor?
 - What percentage of people spend more than 45 minutes when they visit a doctor?
 - How long do the majority of people take when they visit the doctor?
 - What percentage of people spend less than 25 minutes when they visit the doctor?
- 6 Look at the following divided bar graph, which shows how much of different types of rubbish Australians throw out.

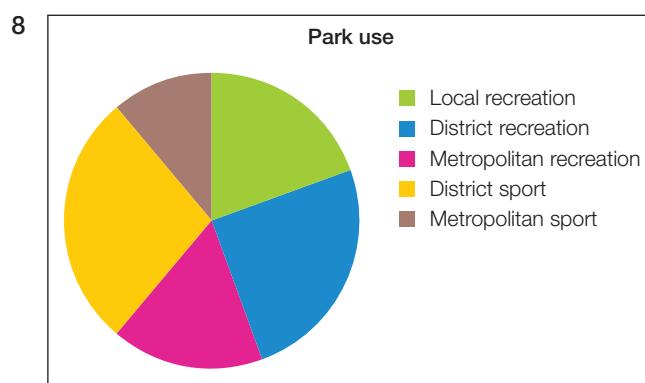


- What percentage of our rubbish is all glass?
- What percentage of our rubbish is milk and fruit juice cartons?
- What percentage of our rubbish is newspaper?
- (i) What is the largest category of rubbish?
(ii) What percentage of rubbish is in this category?

- (e) Do we throw out more high density polythene or more paper?
 (f) Do we throw out more steel or more glass?
 (g) Is the picture of the rubbish bin actually part of the graph?
- 7 Look at the following sector graph showing the type of litter in a particular creek.



- (a) What was the most common type of litter?
 (b) Was there more cans or more sheeting and film?
 (c) Was there more takeaway food containers or more confectionery wrappers?
 (d) When you compare this sector graph with the one in Question 3, what is missing?
 (e) Misc stands for miscellaneous. What kinds of rubbish would be classified as miscellaneous?
 (f) We are told how many litter items were collected in the survey. Is this important? Why or why not?
 (g) Why are sector graphs like divided bar graphs?



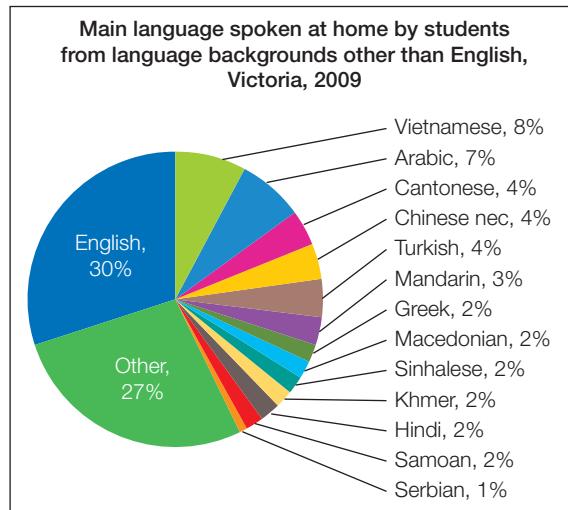
Above is a sector graph showing park use in a local council area of an Australian city.

- (a) Use a protractor to find the angle size of each sector.
 (b) Use your answer from part (a) to find the percentage of the park that is used for metropolitan recreation.
 (c) Use your answer from part (a) to find the percentage of the park that is used for district sport, to the nearest whole number.
 (d) Use your answer from part (a) to find the percentage of the park that is used for sport, to the nearest whole number.
 (e) Is more parkland used for sport or for recreation? State which is larger and by how much.
 (f) What type of park use is represented by the smallest sector?
 (g) Draw a divided bar graph of the data.

9.4

Reasoning

- 9 Using the graph below, answer the following questions.



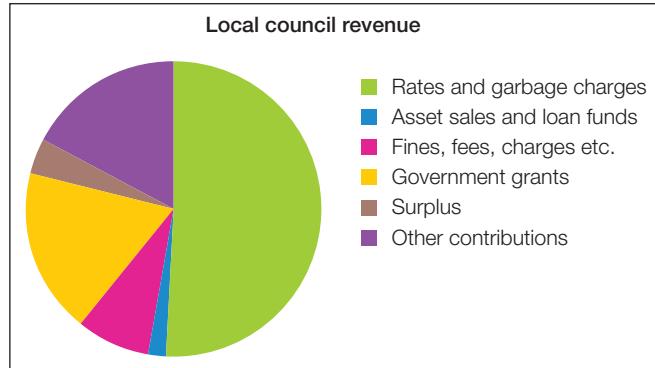
- (a) What is the main language spoken at home by students from non-English-speaking backgrounds?
- (b) What is the total percentage of students whose main language spoken at home is some form of Chinese (Cantonese, Mandarin or Chinese nec (not elsewhere classified))?
- (c) List all the languages that are spoken in 2% of students' homes.
- (d) Do more students speak Turkish or Arabic at home?
- (e) Why is the 'Other' section so large?
- (f) Suggest some of the languages in the 'Other' sector.

- 10 The following table shows vehicle sales for Australia in 2008, divided by state or territory.

State/Territory	Number	State/Territory	Number
ACT	5 462	South Australia	21 492
New South Wales	106 250	Tasmania	5 940
Northern Territory	3 756	Victoria	89 826
Queensland	75 228	Western Australia	39 560

- (a) How many vehicles were sold in total in 2008?
- (b) Find the sector angles that represent each state or territory (correct to the nearest degree).
- (c) Using a protractor, draw a sector graph for this information, clearly labelling each sector.
- (d) If you were to draw a divided bar chart for this information, how long would you make your graph?
- (e) Use this length to draw the divided bar chart and clearly show the length of each sector, correct to the nearest mm.
- (f) Which graph do you think is easier to draw? Give a brief explanation to support your point of view.

11



The sector graph above shows the revenue (money) a local council expected to receive in the 2010/2011 financial year. Use this graph to answer the following questions.

- What is the major source of income for the council?
- Use a protractor to find the sector angle for the fines, fees, charges etc. sector.
- Use the angle you found in part (b) to find the percentage of revenue (to the nearest %) raised through fines, fees, charges etc.
- Does the council expect to receive more money from government grants or from other contributions? Explain your answer.

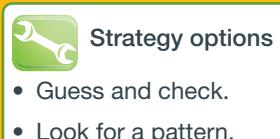
Open-ended

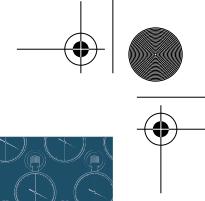
- (a) Draw a sector graph to represent one characteristic of the students in your class.
(b) Draw a divided bar graph to represent a different characteristic of the students in your class.
- Generate sector graphs for Question 10 using Chart Wizard in Excel (or similar) if you have access to a computer. Compare your graphs with the computer-generated ones (you could use 3D graphs also) and comment on any differences.

Outside the Square Problem solving

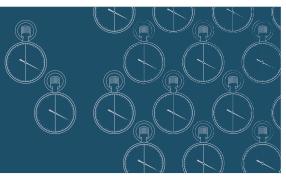
Missing pieces of the pie

- A sector graph has been drawn and labelled, but there is a 60° gap in it. This is to be divided between two pieces, A and B , so that A is twice the size of B . How big will the two pieces be? Give your answers correct to one decimal place if necessary.
- This time, the 60° piece is to be divided between three pieces, A , B and C , where A is twice as big as B and B is twice as big as C . How big will the three pieces be? Give your answers correct to one decimal place if necessary.
- This time, the 60° piece is to be divided between four pieces, A , B , C , and D , where A is twice as big as B , B is twice as big as C and C is twice as big as D . How big will the four pieces be? Give your answers correct to one decimal place if necessary.





Half-time 9



Fx 92

- 1** Find the **(i)** mean (correct to one decimal place) **(ii)** median and **(iii)** mode for each of the following sets of test results.

(a) 3, 2, 4, 8, 6, 2, 4, 6, 3, 2, 1 **(b)** 4, 6, 3, 4, 3, 6, 5, 9, 8, 7, 6, 3, 2, 4, 5, 6, 4

Ex. 9.4

Victoria's greenhouse gas emissions (end use allocation of emissions)	
Residential	17%
Commercial	16%
Manufacturing	32%
Passenger transport	14%
Primary production net	12%
Freight transport	5%
Waste	3%
Other	1%

Using the table given above:

Ex. 9.1

- 3** The height of each of the 25 students in a class was recorded.

142	147	151	142	146	151	157	138	141
140	129	135	142	151	150	160	141	152
148	157	130	143	147	149	154		

- (a)** Why are class intervals of 125–129, 130–134 etc. appropriate?
(b) Using these class intervals, draw a frequency table showing this information.

Ex. 9.1

- 4 The number of students absent from class was recorded for each of the 21 school days in a month.

$$\begin{array}{cccccccccc} 3 & 0 & 1 & 2 & 4 & 1 & 0 & 1 & 0 & 2 & 0 \\ 1 & 0 & 0 & 2 & 1 & 3 & 0 & 1 & 2 & 0 \end{array}$$

Ex 0 3

- 5 Draw a dot plot to illustrate the following data that represents the number of times the students in a class have travelled interstate.

$$\begin{array}{cccccccccccc} 6 & 4 & 0 & 1 & 2 & 3 & 1 & 2 & 2 & 0 & 0 & 0 \\ 1 & 2 & 4 & 5 & 3 & 2 & 5 & 3 & 0 & 1 & 1 & 2 \end{array}$$

Ex. 9.3

- 6 The number of marks taken by the top 12 players in the AFL after round 15, 2010, are given below.

86 74 76 91 132 84 67 80 82 62 85 90

- (a) Draw a stem-and-leaf plot to represent this data.
 - (b) Find the median for this data set.
 - (c) Can you identify any possible outliers? Justify your answer.



Technology Exploration Excel



Equipment required: 1 brain, 1 computer with an Excel spreadsheet

Versions of this Exploration are available for other technologies in Pearson Reader.

Finding statistics using technology

When we have a very large data set it is best to use some form of technology to calculate statistics such as mean, mode, median and range. Here, we will use Excel to find some of these statistics.

The following table shows the height, in cm, of the 47 players on the list for the Geelong Cats in the AFL for the 2010 playing season. We will use this data to find the mean, median, mode and range using Excel.

Note that Excel uses the word 'Average' to find the mean.

182	187	187	200	195	175	179	176	191	184
176	191	187	178	187	184	192	197	182	185
183	195	189	183	189	189	197	192	188	189
195	182	202	193	189	192	182	206	177	182
193	179	180	198	198	189	180			

- 1 Open an Excel spreadsheet and enter the data in column A.
 - 2 Enter the names of the statistics to be found in column B, Mean in B1, Median in B2 and Mode in B3.
 - 3 As the range is the difference between the highest (maximum) value and the lowest (minimum) value, Excel is used to find the maximum and the minimum values. Enter Max in B4, Min in B5 and Range in B6.
 - 4 In cell C1, enter the formula =Average (A1:A47) to find the mean.
 - 5 In cell C2, enter the formula to find the median, in cell C3, enter the formula to find the mode, in cell C4, enter the maximum formula and, in cell C5, enter the minimum formula.
 - 6 In cell C6, subtract the minimum value from the maximum value to find the range.

Any of these formulas can be found by selecting Formulas and Insert Function from the toolbar. The formulas are entered as shown in the screenshot at the bottom right of the page.

Taking it further

- 7 The weights, in kg, of the 47 players on the playing list for the Geelong Cats in the AFL for the 2010 playing season are listed below.

Use Excel (or another form of technology) to find the mean, median, mode and range of the weights.

88	84	88	102	85	77	88	77
89	72	84	87	82	81	91	85
88	101	83	100	84	95	87	88
85	92	99	87	79	92	99	70
108	102	93	94	86	105	80	73
92	68	80	94	104	88	80	

	A	B	C
1	182	mean	187.787234
2	187	median	188
3	187	mode	189
4	200	max	206
5	195	min	175
6	175	range	31
7	179		
8	176		
9	191		
10	184		
11	176		
12	191		
13	187		
14	178		
15	187		
16	184		
17	192		
18	197		
19	182		
20	185		
21	183		
22	195		
23	189		
24	183		
25	189		
26	189		
27	197		
28	192		
29	188		
30	189		
31	195		
32	182		
33	202		
34	193		
35	189		
36	192		
37	182		
38	206		
39	177		
40	182		
41	193		
42	179		
43	180		
44	198		
45	198		
46	189		
47	180		
48			

B	C
mean	=AVERAGE(A1:A47)
median	=MEDIAN(A1:A47)
mode	=MODE(A1:A47)
max	=MAX(A1:A47)
min	=MIN(A1:A47)
range	=C4-C5

9.5

Graphing bivariate data

Bivariate data consists of two variables, an independent variable and a dependent variable. The dependent variable changes as the independent variable changes. An example is the distance travelled by a car in a given time period. The time of the journey is the independent variable and the distance travelled is the dependent variable.

Line graphs

A line graph is used when we have data that shows changes over time. Line graphs are useful if we are trying to identify a **trend** in the data. That means we are looking for a general pattern being followed that we could use to predict future values in the data set. Line graphs are used only for continuous data, such as mass, height, profits and temperatures. These are all data that can be measured over time.

Drawing a line graph

All line graphs should include:

- 1 a title
- 2 the independent variable always on the horizontal axis
- 3 a scale across the horizontal axis that uses equally spaced intervals and is clearly labelled, including any relevant unit of measurement
- 4 a scale on the vertical axis that also uses equally spaced intervals and is clearly labelled, including any relevant unit of measurement.

When drawing line graphs you need to choose a suitable scale. The following data set will be used to illustrate the process.

Worked Example 11

WE 11

The mean daily maximum temperature ($^{\circ}\text{C}$) for each month in Hobart (Botanical Gardens) is shown in the table below:

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mean max. temp. ($^{\circ}\text{C}$)	23.1	22.2	20.7	17.0	13.9	11.4	10.5	12.2	14.9	17.4	19.4	21.9

Bureau of Meteorology

- Draw a line graph to represent this data.
- For which months was the mean maximum temperature below 17°C ?

Thinking

- (a) 1 Look for the highest value to be graphed.
- 2 Decide how tall you want your graph to be (a vertical height of 10 cm is good) and divide the maximum value by the chosen height to get an idea of the scale required.
- 3 Use this number to choose an appropriate scale.
- 4 Draw the axes, label them and give the graph a title.

- 5 Scale both axes and plot the pairs of values. (The first one to plot would be January, 23.1. Choose Jan on the horizontal axis and 23.1 on the vertical axis and plot the point.) Join each point to the next with a straight line.

- (b) 1 Draw a horizontal line across from the vertical axis for the given value (17°).

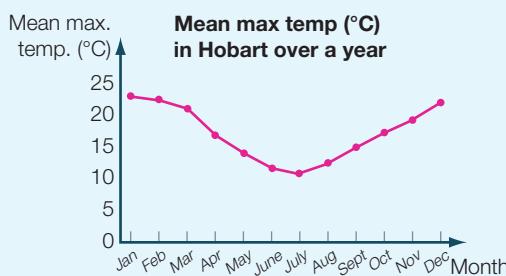
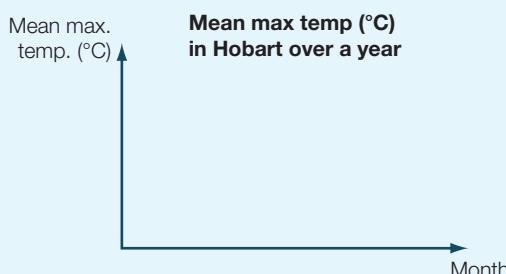
- 2 Find the points on the line graph under this line and read off the values on the horizontal axis that correspond to these points.

Working

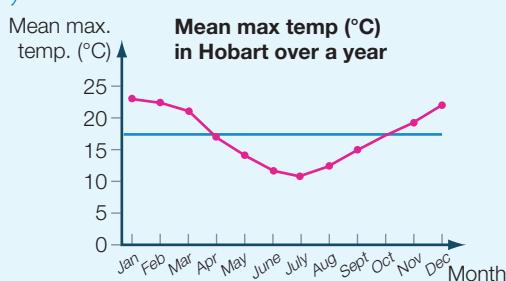
(a) Highest value = 23.1°

$$\text{Approximate scale} = \frac{23.1}{10} = 2.31$$

Scale 1 cm = 2°C



(b)



May, June, July, August, September

9.5 Graphing bivariate data

Navigator

**Answers
page 680**

Q1, Q2, Q3, Q4, Q5, Q8, Q9,
Q10

Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q9,
Q10

Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8,
Q9, Q10

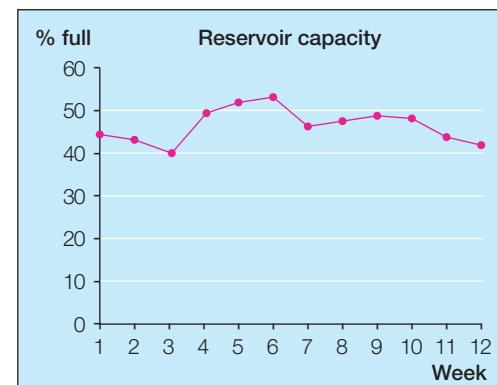
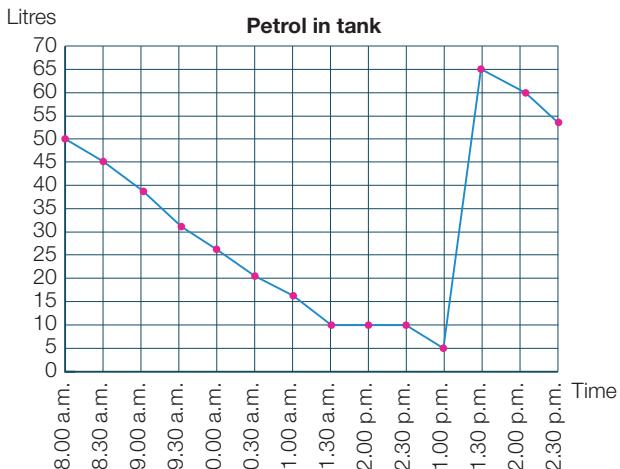
Fluency

W.E11

- 1 At an athletics competition, the wind speed is measured so that new records can be accepted. The following table shows the wind speed (metres per second) at 10-minute intervals for the hour over which the 100 m sprints were run.

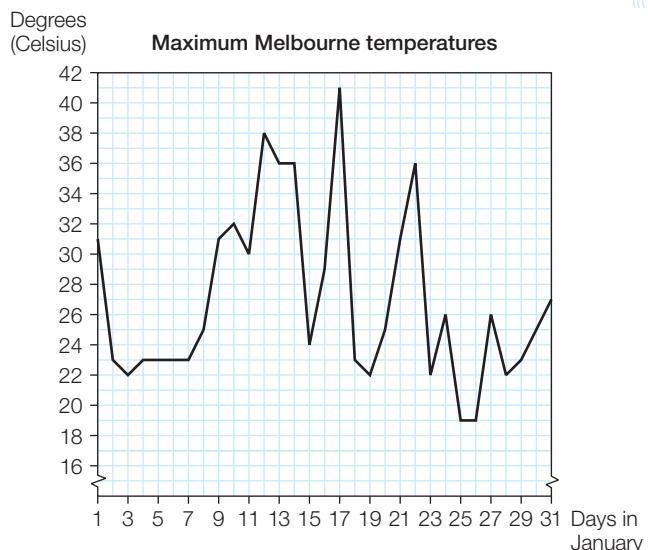
Time	1.00 p.m.	1.10 p.m.	1.20 p.m.	1.30 p.m.	1.40 p.m.	1.50 p.m.	2.00 p.m.
Wind speed (metres per second)	0.9	0.95	1.3	1.2	1.15	1.1	0.95

- (a) Draw a line graph to represent this data.
- (b) When was the wind speed greater than 0.95 metres per second?
- 2 The graph shows the number of litres of petrol in a car at various times during the day.
- (a) How many litres were in the tank at the start of the day?
- (b) When did the driver visit the petrol station?
- (c) What was the minimum number of litres of petrol purchased?
- (d) Between which times is it likely the driver had lunch?
- 3 Each week, the amount of water in a reservoir is recorded as a percentage of its maximum capacity. The graph on the right shows the results for a number of weeks.
- (a) How full was the reservoir when measured in Week 4?
- A 40% B 45%
C 50% D 55%
- (b) There was heavy rain during a particular week. The measurement for which week shows the result of this rain?
- A 2 B 4 C 5 D 6
- (c) A large quantity of water was released from the reservoir to flush out the downstream areas. The measurement for which week most likely shows the result of this action?
- A 2 B 3 C 7 D 10

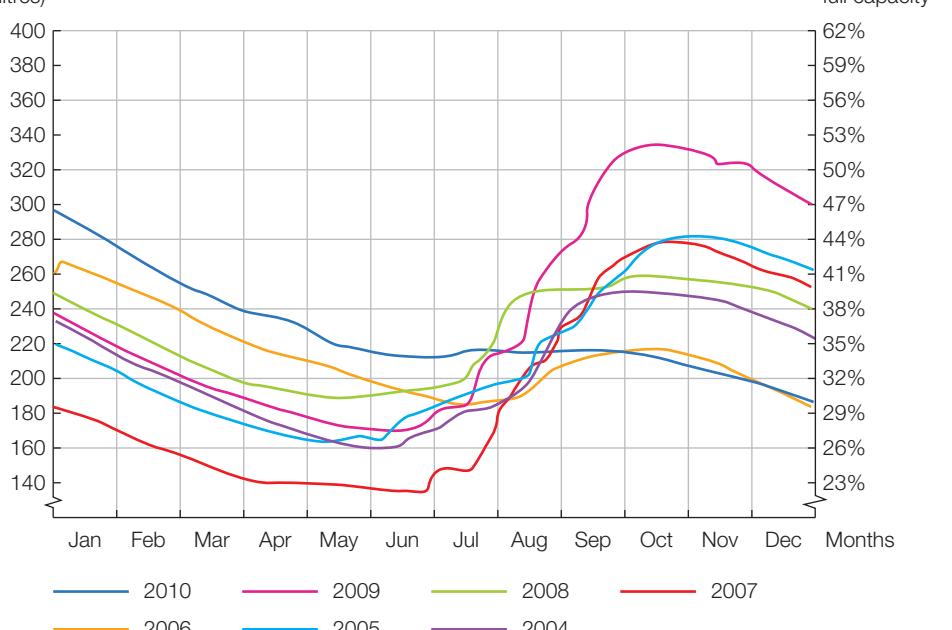


Understanding

- 4 The line graph opposite shows the maximum temperatures in Melbourne for January of a particular year.
- (i) What was the highest maximum temperature reached?
(ii) On which date was this temperature reached?
 - (i) What was the lowest maximum temperature reached?
(ii) On which date(s) was this temperature reached?
 - Up until 15 January, the maximum temperature was always between which two values?
 - After 23 January, the maximum temperature was always between which two values?



- 5 Storage in gigalitres (millions of kilolitres)



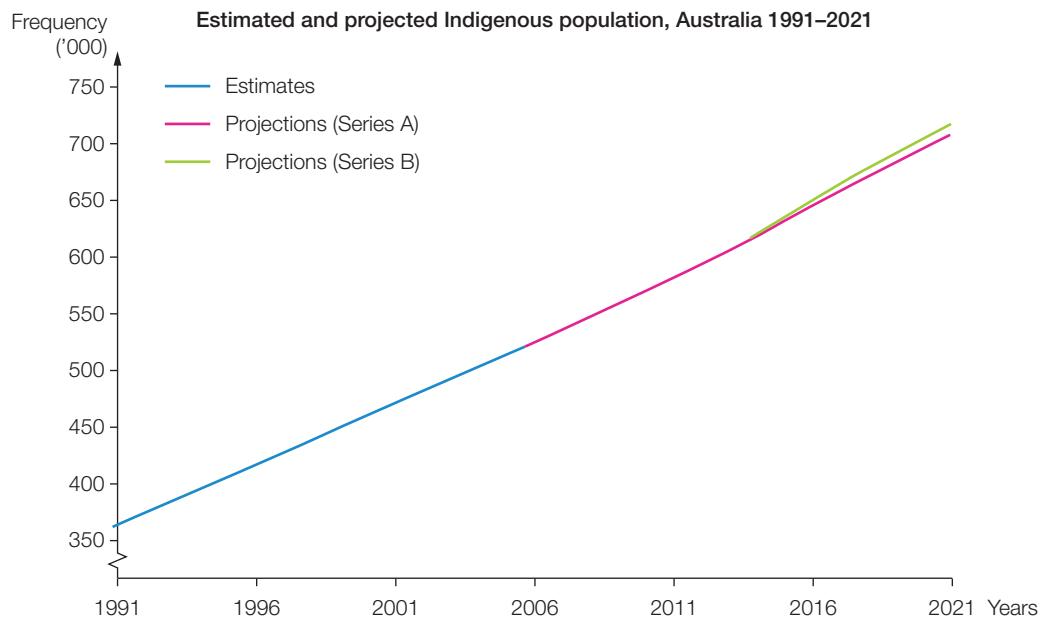
Above is a graph of Perth's water storage from 2004 to 2010. Use this graph to answer the following questions.

- What was the storage percentage at the start and end of 2008?
- In which year was there almost no change in storage levels from the beginning of June to the end of September?
- Which 6-month period over the years 2004–2010 had the lowest water storage? What does this tell you about the rainfall in that period?
- By what percentage did the water storage levels decrease in December 2009? Check the percentage drop in water storage levels for other years in December. What does this tell you about the rainfall in December in Perth?
- Did the water storage levels increase or decrease over the year 2009? Find the increase or decrease percentage.
- From the data shown here, in which month is Perth likely to be the wettest?
- From the data shown here, in which month is Perth likely to be the driest?

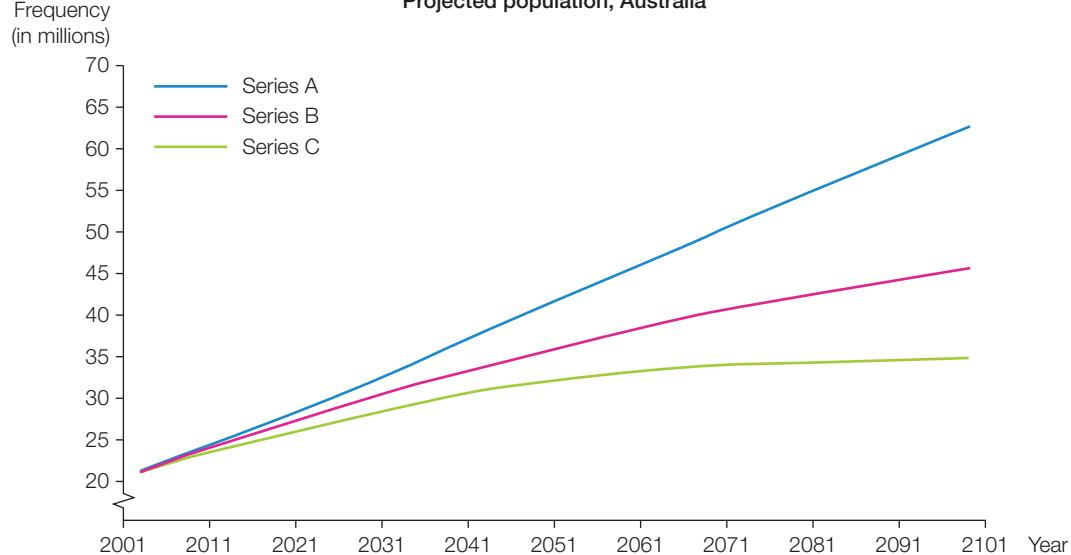
9.5

Reasoning

- 6 This graph shows the Indigenous population of Australia from 1991 to 2006 and the projected (estimated) population for 2006 to 2021. Different projections are made using different estimation methods for future populations.

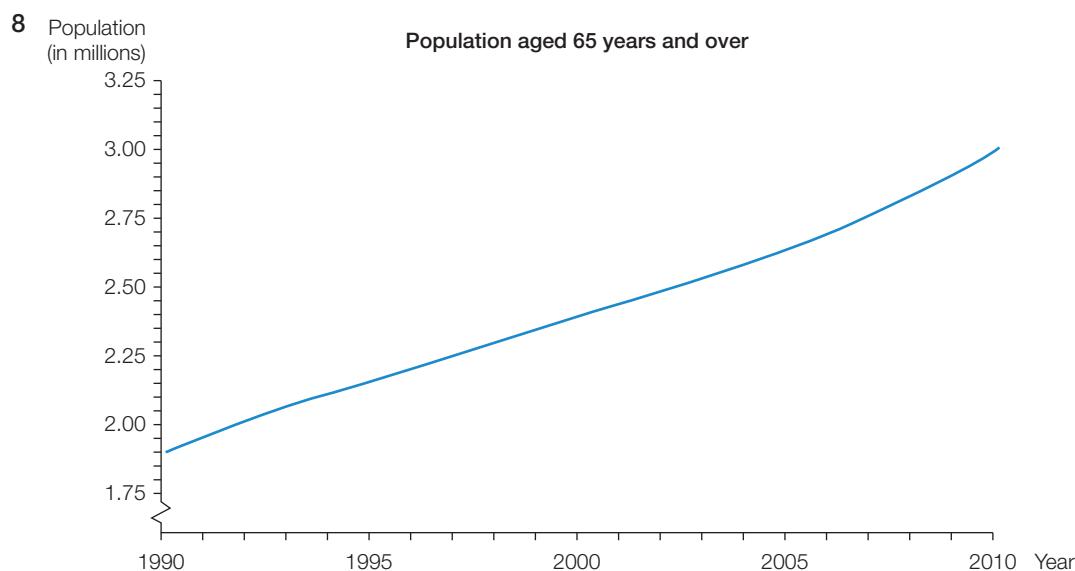


- (a) What was the estimated Indigenous population of Australia in 1991?
 - (b) What was the estimated Indigenous population of Australia 15 years later?
 - (c) Why do you think the graph from 1991 to 2006 is an estimated graph?
 - (d) What are the two projected values for the Indigenous population in 2021?
 - (e) Give two reasons why you think the Indigenous population has increased so much in the 30 years from 1991 to 2021.
- 7 Statisticians use data from the past to predict trends for the future. Below are three graphs that show different ways of predicting (or projecting) the future population of Australia.



- (a) Use the graph to find the population of Australia in 2007.
- (b) What is the projected population of Australia in 2041 on the Series A graph?

- (c) Find the increase in population predicted over the time period shown on the Series A projection.
- (d) What are the two main factors that would cause the population to increase according to the Series A projection?
- (e) List at least two problems Australia would face if the Series A projection was correct.
- (f) What would be the projected population of Australia in 2061 on the Series B projection?
- (g) If the number of births equals the number of deaths in the Series B projection, why would the population continue to grow?
- (h) The Series C projection flattens out by 2081. What would be the projected population over the period 2081–2101?
- (i) What needs to happen if the population of Australia is to stay stable (the same year after year)?



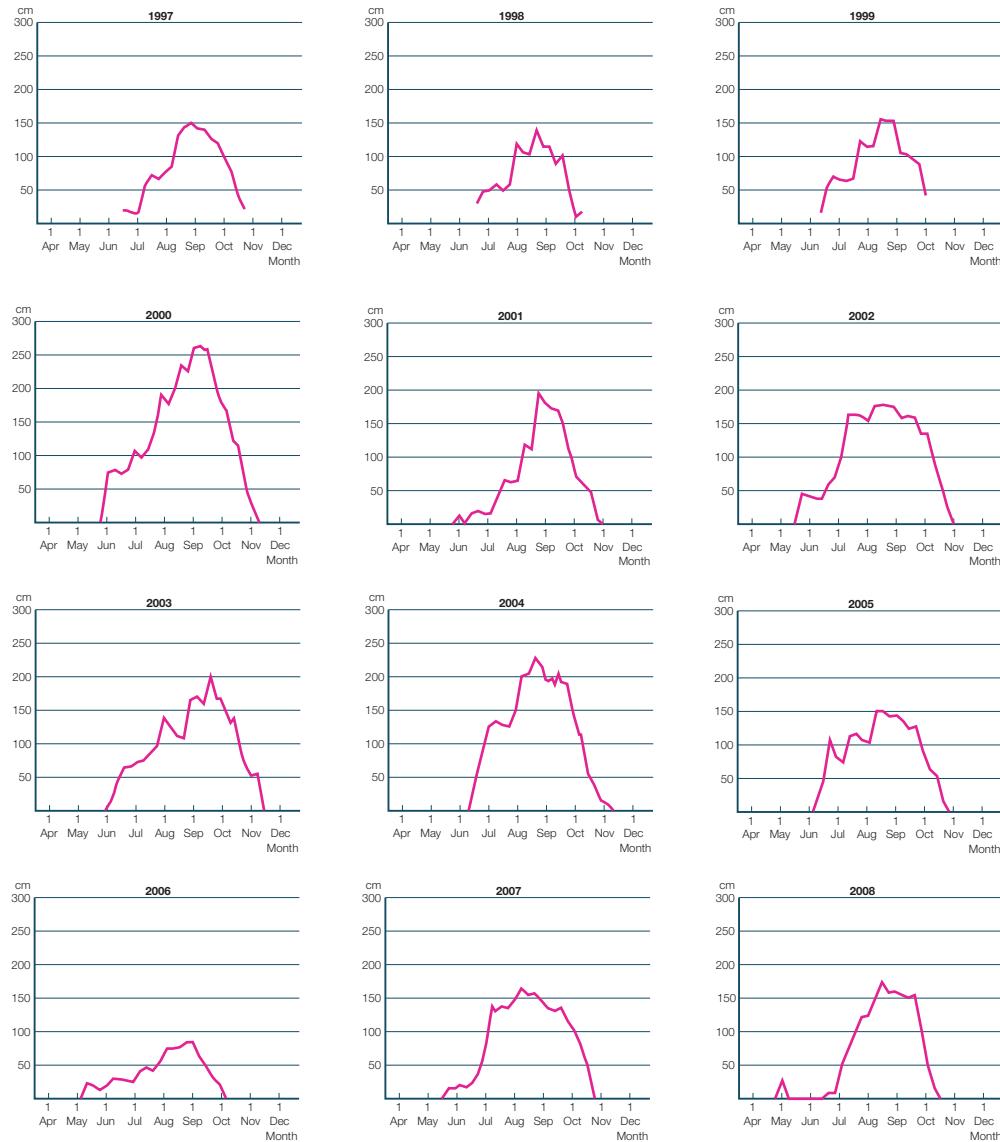
This line graph shows the number of Australians 65 years and older in the time period 1990–2010.

- (a) In 2000, approximately how many Australians were 65 years or older?
- (b) In which year did the number of Australians 65 years and older reach 2.75 million?
- (c) Between which years does the graph follow an almost perfectly straight line? (Use a ruler to help you.)
- (d) Find the population aged 65 years and over in 2002, 2006 and 2010. Use these findings to describe the shape of the graph in the period 2002–2010.
- (e) Use your answers to parts (c) and (d) to describe the shape of the graph overall.
- (f) After World War II ended in 1945, Australia's birth rate increased dramatically. In addition, Australia's immigration policies have changed greatly since then, accepting immigrants from all around the world. Use these facts and your answer to part (e) to give two reasons why the shape of the graph has changed.
- (g) What do you expect the population aged 65 years and over will be in 2015 if the present trend continues?

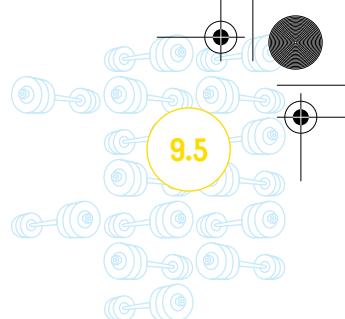
9.5

- 9 Look at the following line graphs that show the average snow depths (in centimetres) at Spencers Creek, Mt Kosciuszko, from 1997 to 2008. (The deeper the snow, the better it is for skiing.)

Spencers Creek Snow Depths



- Did the snow depth reach 2.5 metres at any time during the 12 years? If so, which year(s) did this happen?
- What was the greatest snow depth reached in 1997?
- What was the greatest snow depth reached in 2008?
- Which year had the most snow in June?
- Which year had the most snow through the whole of July?
- Which year had the worst October skiing?
- Which year had the best October skiing?
- Which year would have probably had the best November skiing?
- Which year had the worst skiing overall?
- Which month usually has the best skiing?



- (k) Which month is usually the second best skiing month?
- (l) It has been claimed that New South Wales' overall temperature has been getting higher due to climate change. Is there any evidence in these graphs to back this up?
- (m) Why do we use line graphs rather than bar graphs to show this information?



Open-ended

- 10 When do you think you feel hungriest during the day? Draw a line graph that shows your level of hunger during the day from the time you wake up until the time you go to sleep. Do you think the shape of the graph would be similar for most people? Explain any differences you can think of.

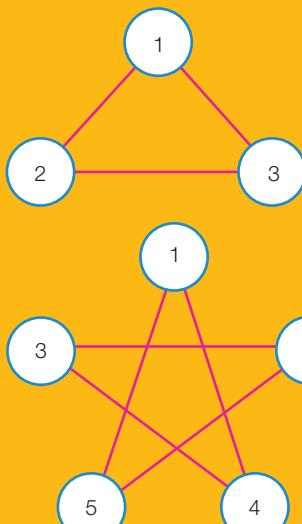
Outside the Square Puzzle

Add it up

A game board consists of the numbers 1, 2 and 3 arranged in a triangle as shown. The first player puts a counter on any of the three numbers and takes that number as their score. The other player then moves the counter to one of the other numbers and takes that as their score. The first player moves the counter and adds the value to their total. Play continues in this way until a player gets a total of 12. If you go over 12 you lose.

Explain the strategy you would follow if you were playing this game. You might like to play the game with a friend to help formulate your strategy.

Now, make up a strategy for the second game board shown. What would be a suitable target number for this board?



**Maths
4 Real**

The 1-watt initiative

To reduce our carbon footprint (and to save money) we should try to reduce the amount of electricity we use in our households. The unit that measures how much electrical power we use is called the watt and has the symbol W.

When an electronic device such as a television is not on it is often not completely off either. It is usually in what we call standby mode, and can be turned on quickly, often by using a remote control. It has been estimated that standby power is responsible for 5–10 per cent of total electricity use in most homes.

- 1** How many devices can you think of in your home that would be in standby mode when not being used?
Make a list of them now.

How big an impact does this have?

As a simple example, you probably have a microwave oven in your house that has a digital clock built into it. If you used the microwave for an average of 10 minutes each day to cook food, then about 21 per cent of its annual energy consumption would be just to display the clock!

The following table shows how much power some electronic devices use per hour.

Appliance	Off (W)	Passive standby (W)	Active standby (W)	Delay start (W)
air conditioner	0.2	1.7		3.0
clothes dryer	0.2		2.4	3.8
dishwasher	0.6		2.7	3.0
front-loading washing machine	0.9		3.5	4.0
top-loading washing machine	1.2		3.0	3.5
microwave oven		3.0		
laptop computer	1.4			
computer monitor	0.9	1.9		
inkjet printer	0.8	3.8		
LCD TV	0.5	1.6		
DVD recorder		7.3	21.5	
set-top box		8.4	13.5	

'Off' means the device is switched off at the device but not at the wall. Some appliances continue to use electricity even when you think they are off!

'Passive standby' means the device is not performing its main function but is ready to be switched on (in most cases with a remote control), or is performing a secondary function, like the clock on the microwave oven.

'Active standby' means the device is on but not performing its main function. For example, the DVD recorder may be on but is not playing or recording.

'Delay start' means the device has been programmed to start at a later time.

- 2** As you go across the table, energy use increases. Suggest a reason for this.

How much does all of this cost us?

As an example, we will look at an air conditioner in passive standby mode. According to the table this uses 1.7 W per hour, or 1.7 W h (watt hours). We can multiply this by 24 to get a daily use and then by 365 to get a yearly use.

$$1.7 \times 24 \times 365 = 14\,892 \text{ W h per year}$$

To avoid using such large numbers we usually deal with kilowatt hours (kW h). $1000 \text{ W h} = 1 \text{ kW h}$.

Now, let's assume electricity costs 12.2 c/kW h. The air conditioner uses 14.892 kW h, which costs us $14.892 \times 12.2 = 181.6824$ cents per year.

This is an annual cost of approximately \$1.82. This may not seem like a lot but it is only one device and one of the more efficient ones. Imagine how much this would be for every air conditioner in Australia.

- 3** Calculate the following annual costs.

- (a)** a clothes dryer on active standby
- (b)** an inkjet printer on passive standby
- (c)** a computer monitor switched off
- (d)** the difference in cost between a front-loading washing machine and a top-loading washing machine both switched off
- (e)** the total cost of an LCD TV connected to a DVD recorder and a set-top box all in passive standby mode

- 4** In 2006 the Australian government introduced a policy stating that by 2012 all new appliances with a standby mode should use no more than 1 watt of electricity for the standby component. What would be the maximum annual standby cost per machine if the 1-watt initiative was implemented?

Research

Find out more about James Watt, the person after whom the unit of power is named.

See how many electronic devices you have in your house that are on standby at any given time. Come up with some energy-saving (and therefore money-saving) suggestions for your home.

9.6

Comparing data sets

Comparing column graphs

We can use the shape of column graphs to compare different variables. If the graph shapes are similar, there may be a connection between the variables. If the graph shapes are very different, we would consider that there was no connection.

Worked Example 12

We 12

Rajit, Steve, Kenny and Dave are keen joggers. The tables below show some data on the four boys. Draw column graphs for each set of data and comment on the likelihood that there is a connection between any of the variables.

	Distance jogged in a week (km)
Rajit	20
Steve	28
Kenny	35
Dave	24

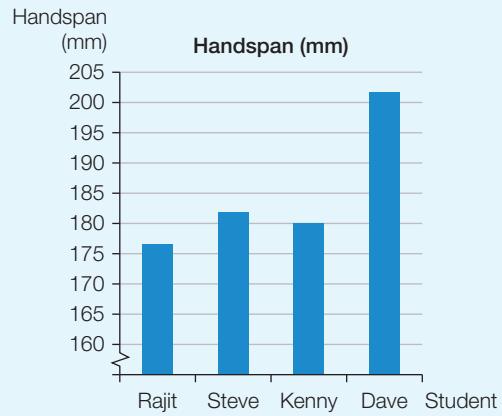
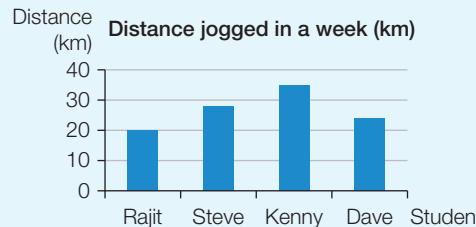
	Height (cm)
Rajit	152
Steve	160
Kenny	157
Dave	168

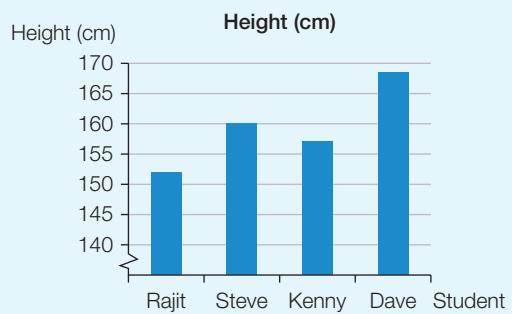
	Handspan (mm)
Rajit	177
Steve	182
Kenny	180
Dave	202

Thinking

- Draw three separate column graphs.

Working





- 2 Look at the shape of the graphs and make a comment about similarities and differences.

The shape of the graph of distance jogged each week is quite different from the shape of the other two graphs. There seems to be no connection between the distance they jog and either their height or their handspan.

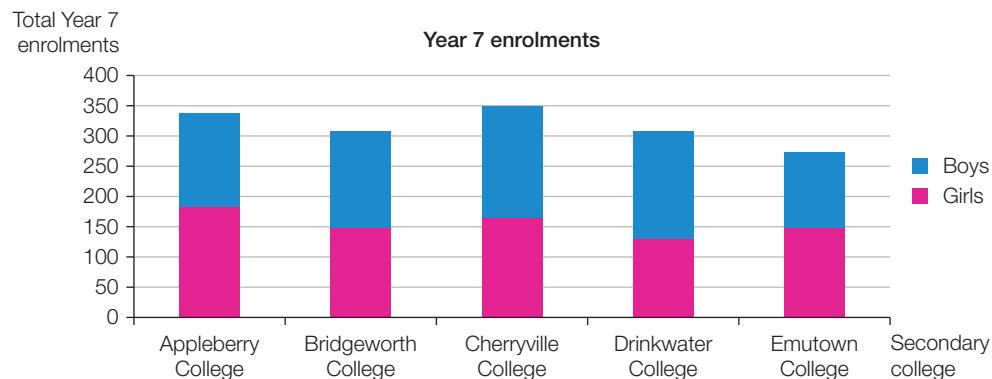
The shapes of the graphs showing the height of the students and their handspan are very similar, so there could be a connection between these variables.

Comparing divided column graphs

We often use several divided column graphs to compare data.

Worked Example 13

WE 13



The graph above shows the Year 7 enrolment in five secondary colleges. Use the graph to answer the following questions.

- Which school has the highest Year 7 enrolment and what is its total Year 7 enrolment?
- Which part of the graph displays the number of girls and which school has the highest number of girls enrolled?
- Which school has the lowest number of boys enrolled?
- Which school has almost the same number of boys as girls?
- Which schools have about the same number of students enrolled in Year 7?

Thinking

- (a) Identify the variable that has the longest column and find the number that corresponds to the length of the column.

Working

- (a) Cherryville College
350

9.6

- (b) Identify which part of the column corresponds to the data required and find the column in which this part of the column is longest.
- (c) Identify which part of the column corresponds to the data required and find the column in which this part of the column is shortest.
- (d) Identify a column where the lengths of each part of the column are about the same.
- (e) Identify all the columns that have the same length.
- (b) The lower part of each column corresponds to girls.
Appleberry College has the highest number of girls enrolled.
- (c) Emutown College has the lowest number of boys enrolled.
- (d) Bridgeworth College has about the same number of boys and girls enrolled.
- (e) Bridgeworth College and Drinkwater College have the same number of students.

In Section 9.2, we found measures of centre to give us information about data sets. However, they do not give us the complete picture. Consider the following data sets.

Set A: 11, 12, 14, 14, 14, 16, 17

Set B: 1, 2, 14, 14, 14, 26, 27

In both of these sets, the mean = median = mode, but they are clearly different data sets. To help us get a better picture of the data, we need to use the range. In Section 9.1, the range was defined as the difference between the highest and the lowest values in the data set.

For Set A: range = $17 - 11 = 6$

For Set B: range = $27 - 1 = 26$

As the range is larger in Set B, we can see that the values are spread out much more than in Set A.

Using tables of information

Worked Example 14

WE14

The following table gives the monthly mean maximum temperatures, measured in °C, for Adelaide (Kent Town) and Brisbane (Brisbane Aero).

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Adelaide	29.1	29.3	26.3	22.6	19.0	16.1	15.3	16.7	19.0	21.8	25.1	27.0
Brisbane	29.0	29.1	28.0	26.0	23.6	21.3	20.9	21.8	24.2	25.2	26.8	28.3

Compare the monthly mean maximum temperatures for the two cities using measures of centre and the range.

Thinking

- 1 Calculate the mean, median and range (correct to one decimal place) in the normal way.

Working

	Mean	Median	Range
Adelaide	22.3	22.2	14.0
Brisbane	25.4	25.6	8.2

- 2 Compare the means and medians for the two cities and use them to describe any differences or similarities. Also, compare the ranges to explain the pattern of temperatures in each city over the year.

Brisbane is hotter according to both the mean and the median temperatures, although both cities have a very similar highest mean maximum temperature. The range of temperatures is much greater in Adelaide, which means there is a much greater change in the mean maximum temperature from summer to winter than in Brisbane.

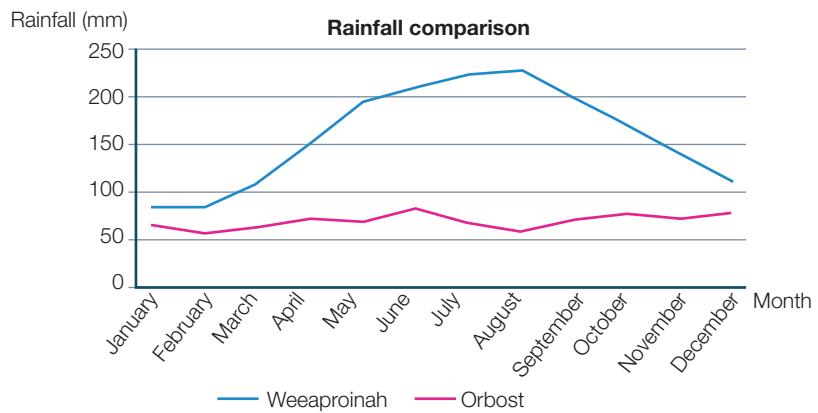
Using graphical data

If we are presented with a graph we can make some reasonable estimates for the key statistics in which we are interested. The range is the easiest to estimate as you just need to find the difference between the highest and lowest values of each data set. For the median, estimate the value of each data point from the graph, order these values and find the middle value. For the mean you need to draw a horizontal line that would flatten out the graph without changing the mean. Look at the distance between each data value and the corresponding point on the horizontal line. The sum of these distances should be approximately zero.

Worked Example 15

WE15

Use the following graph to compare the monthly mean rainfall (mm) for the two Victorian towns of Weeaproinah and Orbost. Estimate the mean, median and range to assist your comparison.



Thinking

- 1 To find the range, look at the highest and lowest value for each data point in the set. To find the median, use the graph to find the middle value. (Here, we find the 6th and 7th lowest rainfall for April and October and average them.) To find the mean, draw a horizontal line that would flatten out the line graph. (Here, if a horizontal line was drawn for rainfall = 160 mm and the difference each month was calculated between each value and 160, the sum of these differences would be approximately zero.)

Working

	Mean	Median	Range
Weeaproinah	160	$\frac{150 + 170}{2} = 160$	$230 - 90 = 140$
Orbost	70	$\frac{75 + 75}{2} = 75$	$85 - 55 = 30$

9.6

- 2** Compare the means and medians for both sets of data and use them to describe any differences or similarities. Compare the ranges to explain what the shapes of the graphs are telling you about the data.

Weeaproinah is clearly much wetter than Orbost with both the mean and median for Weeaproinah being more than twice the mean and median for Orbost. The range for Weeaproinah is also much larger as it is much wetter in winter than in summer, whereas there is not much difference from month to month in Orbost's rainfall data.

9.6 Comparing data sets

Navigator

**Answers
page 681**

Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8,
Q9, Q10, Q11, Q13

Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8,
Q9, Q10, Q11, Q12, Q13

Q1, Q2, Q3, Q4, Q5, Q7, Q8, Q9,
Q10, Q11, Q12, Q13, Q14

Fluency

W.E12

- 1** Veena, Fay, Kylie and Deanne are keen swimmers. The tables below show some data on the four girls. Draw column graphs for each set of data and comment on the likelihood that there is a connection between any of the variables.

	Hours spent training each week
Veena	15
Fay	20
Kylie	24
Deanne	18

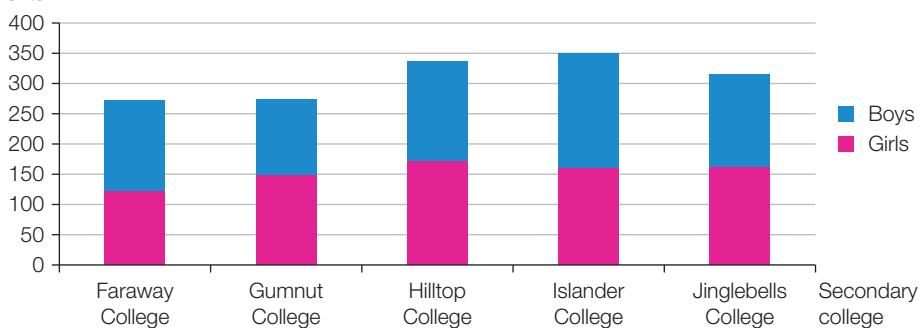
	Height (cm)
Veena	147
Fay	153
Kylie	150
Deanne	162

	Foot length (mm)
Veena	217
Fay	236
Kylie	230
Deanne	242

W.E13

- 2** Total Year 7 enrolments

Year 7 enrolments



The graph above shows the Year 7 enrolment in five secondary colleges. Use the graph to answer the following questions.

- Which school has the highest Year 7 enrolment and what is its total Year 7 enrolment?
- Which part of the graph displays the number of girls and which school has the highest number of girls enrolled?
- Which school has the lowest number of boys enrolled?
- Which school has almost the same number of boys as girls?
- Which schools have about the same number of students enrolled in Year 7?

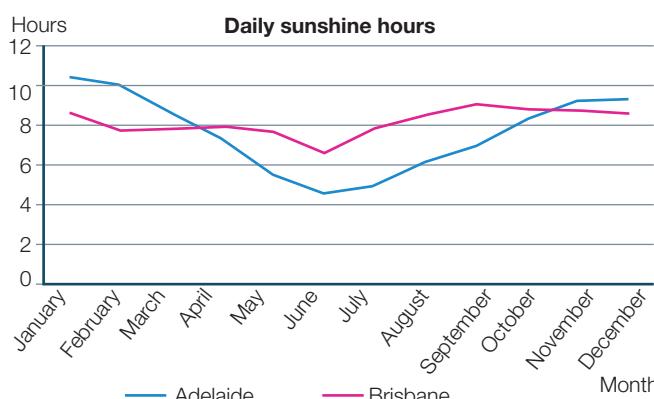
WE14

- 3 The following table gives the monthly mean minimum temperatures, measured in °C, for Adelaide (Kent Town) and Brisbane (Brisbane Aero).

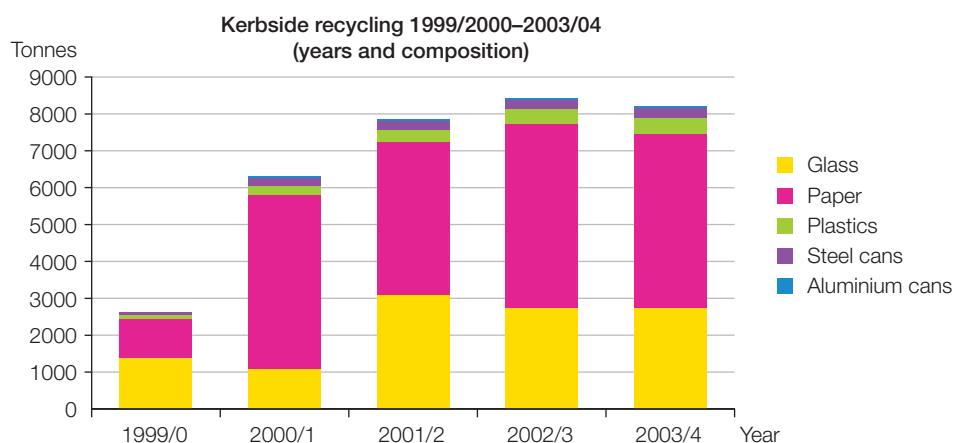
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Adelaide	17.0	17.1	15.2	12.4	10.2	8.1	7.5	8.2	9.7	11.4	14.0	15.5
Brisbane	21.1	21.1	19.3	16.2	12.8	10.5	8.8	9.6	12.7	15.7	18.0	20.1

Compare the monthly mean minimum temperatures for the two cities using measures of centre and the range.

- 4 Use the following graph to compare the monthly mean sunshine (hours) for Adelaide (Airport) and Brisbane (Airport). Estimate the mean, median and range to assist your comparison.



5



The graph above shows the types of material found in kerbside recycling by Shoalhaven Council over the time period 1999/2000–2003/04. Use the graph to answer the following questions.

- What material (by weight) was recycled the most in 1999/2000?
- What material (by weight) was recycled the most over the whole time period shown?
- Which time period was the best for recycling?
- Which was the best time period for glass recycling?
- Comment on the recycling of steel cans in the time period 2001–2004.
- Comment on the recycling of plastics over the time period shown.

Understanding

- 6 Compare the following data sets using as many statistics as possible.

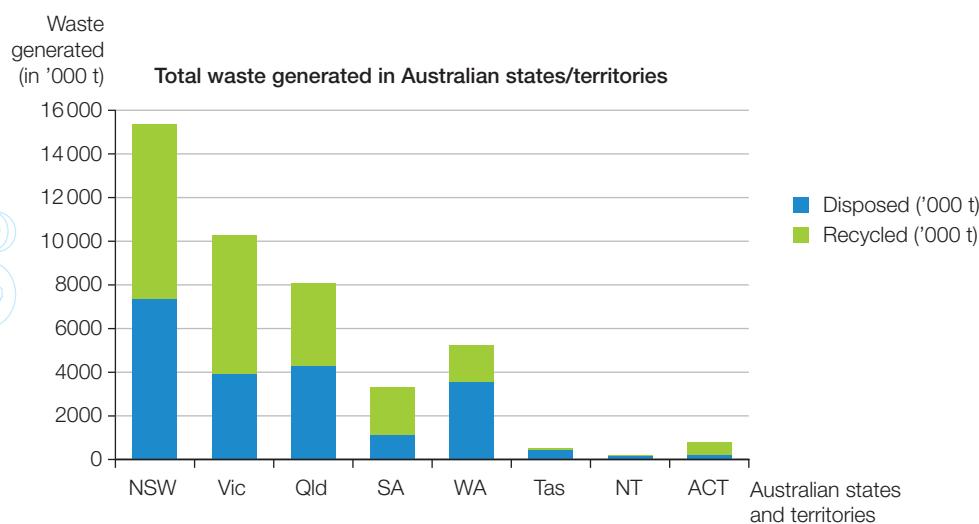
- | | |
|---|---|
| (a) Set A: 1, 2, 4, 5, 6, 7, 7, 7, 8, 10 | Set B: 2, 2, 2, 2, 4, 5, 8, 9, 9, 9, 9 |
| (b) Set A: 2, 4, 1, 5, 6, 2, 3, 7, 8, 2, 4, 1 | Set B: 5, 2, 7, 4, 7, 1, 8, 3, 9, 8, 9, 6 |

9.6

- 7 April went to the market and purchased 5 boxes of bananas at \$40 a box, 3 boxes of apples at \$55 a box and a box of mangoes at \$65 a box. Her friend Izzy bought 3 boxes of bananas at \$45 a box, 3 boxes of apples at \$50 a box and a box of mangoes for \$70. Compare the purchases of the two friends using as many statistics as possible.

Reasoning

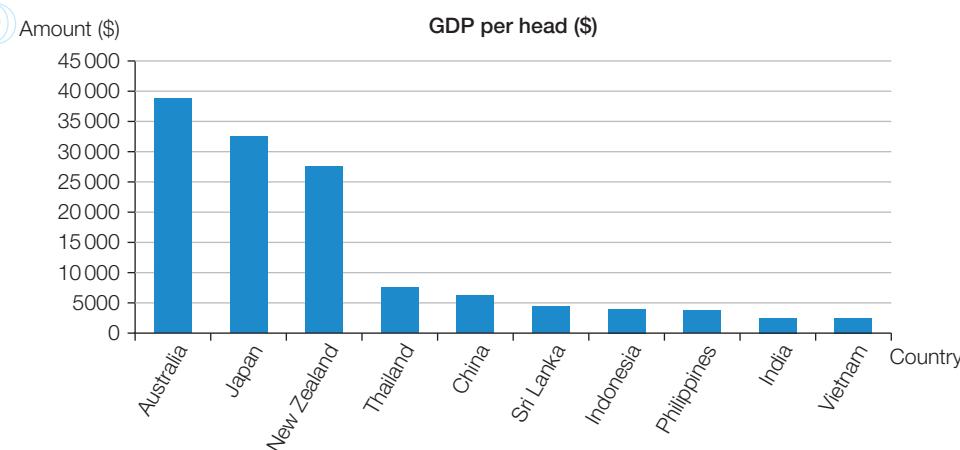
8

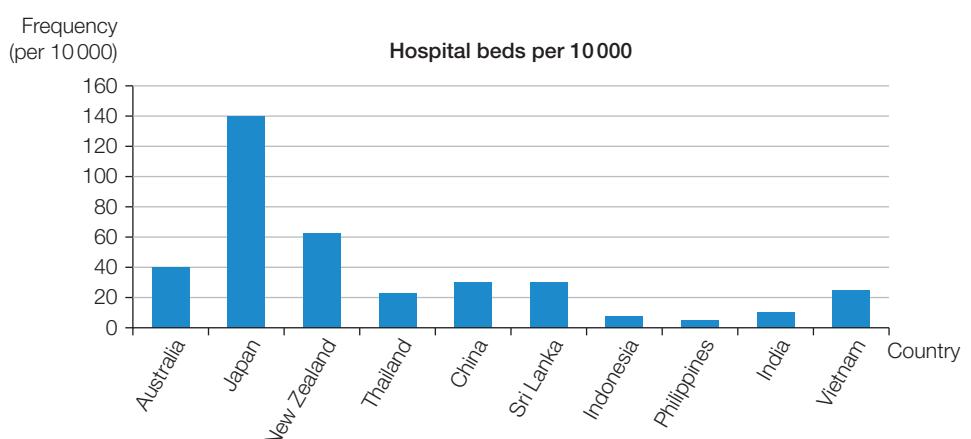
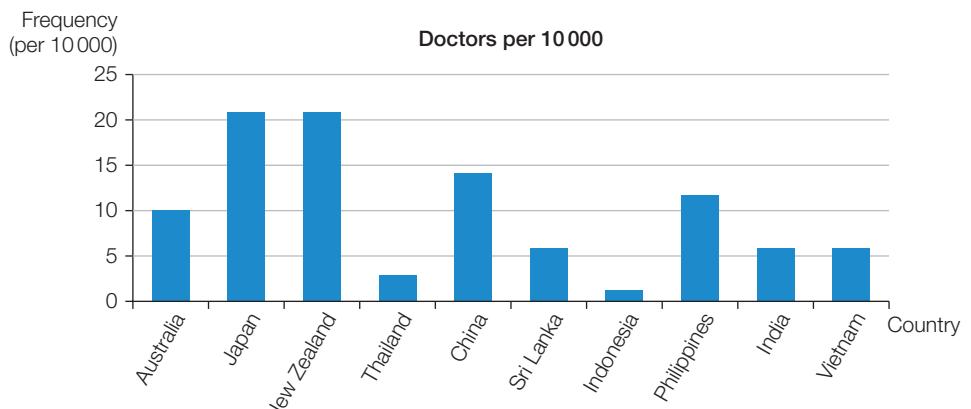


Above is a graph of the estimated total waste generated in Australia 2006/07. The divisions show the amount of waste disposed of (sent to landfill) and the amount of waste recycled. Use the graph to answer the following questions.

- Which state/territory generated the most waste?
- How much waste did Queensland generate?
- Which states/territories recycled more than they disposed of?
- Which state/territory recycled the largest proportion of its waste?
- Which state/territory recycled the smallest proportion of its waste?
- Did South Australia or Victoria recycle the larger proportion of its waste?

9



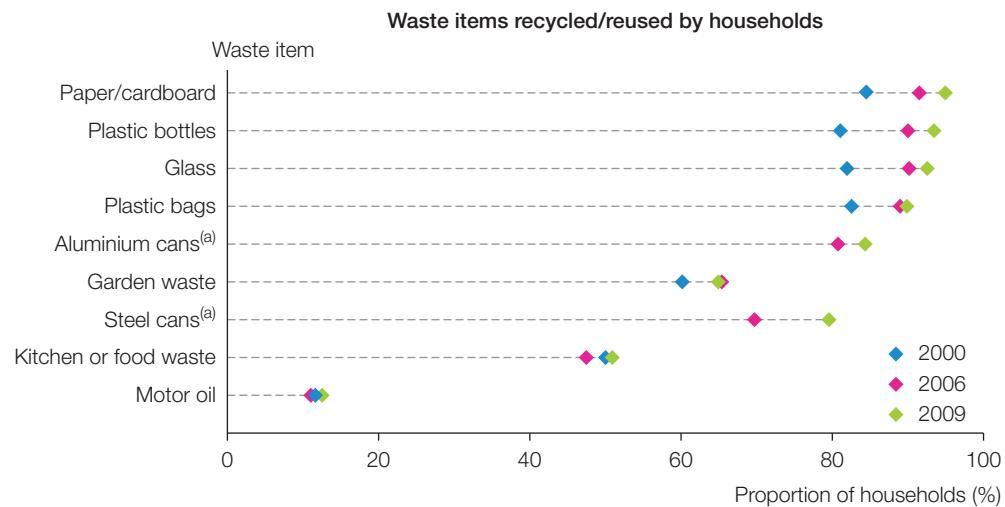


The Gross Domestic Product (GDP) of a country is one measure of a country's wealth. The number of doctors and hospital beds are measures of the health-care facilities of a country. Use the graphs above to answer the following questions.

- Do any of the graphs have a similar shape?
- Which country appears to have the best health-care facilities? Give a reason for your answer.
- Compare the health-care facilities of this country with its GDP per person.
- Which country appears to have the worst health-care facilities?
- Compare the health-care facilities of this country to its GDP per person.
- Comment on Australia's health-care facilities compared with its GDP per person.
- Comment on Vietnam's health-care facilities compared with its GDP per person.

9.6

- 10 This graph shows different types of waste items and the percentage of Australian households that recycle them.

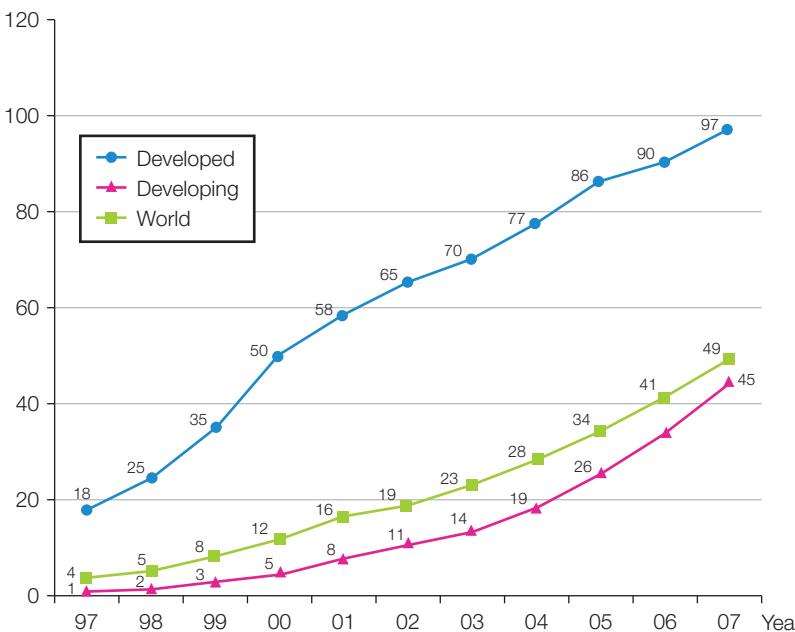


(a) Data not available for 2000.

- (a) (i) In 2009, what were the top three types of waste recycled by Australian households?
(ii) Suggest a reason for the answer to (i).
- (b) Which item had a large increase in the level of recycling between 2006 and 2009?
- (c) Which three items had the most significant increase in the level of recycling between 2000 and 2009?

- 11 Frequency

Mobile phone subscribers per 100 inhabitants 1997–2007



This graph shows the number of people out of every hundred who have a mobile phone, worldwide. The 'developed' countries include wealthier countries such as the United States, Australia, Japan and much of Europe. The 'developing' countries include poorer countries, including many in Africa and Asia.

- (a) Compare the number of people who owned a mobile phone in developed and developing countries in the year 2000.
- (b) What happened in the developed world over the years 2000–2007 to the number of people who own a mobile phone?
- (c) What happened in the developing world in the years 2000–2007 to the number of people who own a mobile phone?
- (d) Why is the world graph closer in shape to the developing world graph than the developed world graph?
- (e) In which year did the biggest jump in mobile phone ownership in the developed world occur?
- (f) If the trend of mobile phone ownership in the developing world shown by the graph has continued over the years since 2007, what do you think the number of people per 100 who own a mobile phone would have been in 2010?

- 12 The following table shows the number of members for 16 AFL teams for 3 different years. (The V in brackets shows Victorian-based teams.)

	2009	2005	2000
Adelaide	46 472	43 256	42 896
Brisbane Lions	24 873	28 913	20 295
Carlton (V)	42 408	33 534	27 571
Collingwood (V)	45 972	38 612	28 932
Essendon (V)	40 412	32 734	34 278
Fremantle	39 206	34 178	24 925
Geelong (V)	37 160	30 821	25 595
Hawthorn (V)	52 496	29 261	26 879
Melbourne (V)	31 506	24 805	18 227
North Melbourne (V)	28 340	24 154	22 156
Port Adelaide	30 605	36 834	34 925
Richmond (V)	36 981	28 029	26 869
St Kilda (V)	31 906	32 043	17 855
Sydney Swans	26 269	24 955	30 177
West Coast Eagles	43 927	42 406	38 868
Western Bulldogs (V)	28 125	21 974	18 056



- (a) Compare the club membership statistics across the 3 years. What comment could you make about the number of people who are members of AFL clubs?
- (b) Why is the mode not useful at all in these comparisons?
- (c) For 2009 only, compare the membership numbers for the Victorian-based teams against the non-Victorian-based teams. On the basis of this comparison, what comment could you make about the popularity of AFL clubs.
- (d) Comparing 2000 to 2009, which club has been most successful at increasing its membership base?
- (e) Is there any other information you would like access to before stating which club has been the most successful at turning its fans into members?

Open-ended

13 Sahrita measured the heights, in metres, of the 10 girls in her mathematics class. They were:

1.68 1.55 1.67 1.64 1.43 1.77 1.67 1.45 1.59 1.66

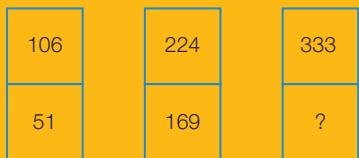
- (a) Work out the mean, median and modal heights.
- (b) Two new girls joined the class. The modal height did not change, but the mean changed by 0.15 m. Find two possible pairs of values for the heights of the new girls.
- (c) Using the two pairs of values you used in (b) calculate the new median height.

14 Eric is comparing two sets of data. He has written that the mean of Set A is 1.5 greater than that for Set B, but the median for Set A is 2 less than the median for Set B. The mode and the range are the same for both sets. Construct two data sets that meet these descriptions, given that there are at least 15 results in each data set.

Outside the Square Puzzle

Number puzzles B

1 In each of the following sets of three diagrams one of the numbers is missing. To find it you will need to work out the connection between the numbers in the other two parts of the diagram. As an example:



the ? here would be replaced by 278, as $106 - 51 = 55$ and $224 - 169 = 55$, so we need the number $333 - 55$, which is 278.

- (a)

405	549	693
45	61	?
- (b)

275	638	1001
25	58	?
- (c)

301	413	?
43	59	76

2 On each of the following sets of three triangles there is a number missing. To find it you will need to work out the connection between the outside numbers and the inside number in each of the first two triangles.

- (a)

--	--	--
- (b)

--	--	--
- (c)

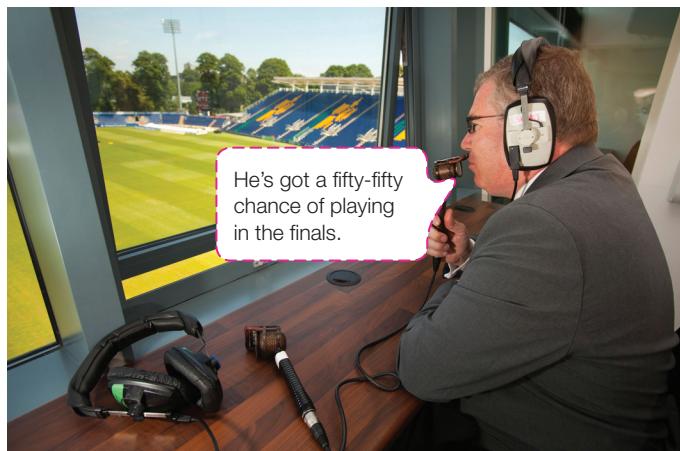
--	--	--

Probability and sample space

9.7

We use the words **probability** and **chance** to mean the same thing.

We hear and make statements that involve probability all the time.



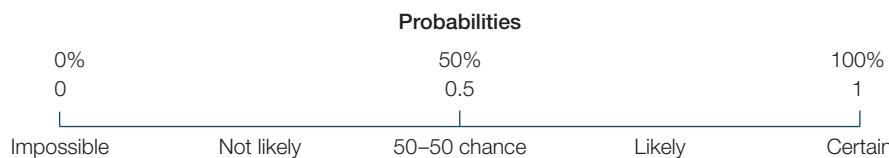
Theoretical probability

Probabilities can be written as fractions or as decimals from 0 to 1, written in simplest form, or as a percentage from 0% to 100%.

- Events that have no chance of happening (impossible) have a probability of 0 or 0%.
- Events that are certain to happen have a probability of 1 or 100%.
- Events that have a 50–50 chance of happening have a probability of $\frac{1}{2}$ or 0.5 or 50%.

9.7

There is a whole range of probabilities between 0 and 1. The easiest way to think of it is as a number line where the closer you get to 0 the less likely an event is to occur and the closer you get to 1 the more likely an event is to occur.



Whenever we determine a probability without actually doing an experiment and collecting data, we are working with **theoretical probability**.

We often use Pr as an abbreviation for probability. The event is then written in brackets. So, if we were spinning a spinner that had different coloured sections, $\text{Pr}(\text{white})$ would be written to stand for the probability of the pointer landing in a white section.

To find the probability of an event occurring we do the following calculation.

$$\text{Pr}(\text{event}) = \frac{\text{number of favourable outcomes}}{\text{total number of outcomes}}$$

Sample space

To determine the probability of a particular event occurring, it is useful to list the **sample space**. The sample space is the complete list of all the possible **outcomes** of a probability experiment, such as rolling a die, tossing a coin, or selecting a card from a pack.

When the outcomes of an experiment are equally likely, we can use the sample space to determine the probability of a particular event occurring. A favourable outcome is one where the event for which we are calculating the probability occurs.

Worked Example 16

WE16

Find the probability of:

- rolling an odd number with a normal die
- selecting a vowel (a, e, i, o, u) from a bag containing 26 identical pieces of paper, each containing a different letter of the alphabet.

Thinking

- List the sample space and count the number in the sample space.
- Identify the favourable outcomes and count the number of favourable outcomes.
- Find the probability using the rule

$$\text{Pr}(\text{event}) = \frac{\text{number of favourable outcomes}}{\text{total number of outcomes}}$$
and simplify the fraction if possible.

Working

- Sample space: 1, 2, 3, 4, 5, 6*
There are 6 outcomes in the sample space.

*Favourable outcomes are: 1, 3 and 5.
 There are 3 favourable outcomes.*

$$\begin{aligned}\text{Pr}(\text{odd number}) &= \frac{3}{6} \\ &= \frac{1}{2}\end{aligned}$$

- (b) 1 List the sample space and count the number in the sample space.

- 2 Identify the favourable outcomes and count the number of favourable outcomes.

- 3 Find the probability using the rule

$$\text{Pr(event)} =$$

$$\frac{\text{number of favourable outcomes}}{\text{total number of outcomes}}$$

and simplify the fraction if possible.

(b) Sample space: a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z

There are 26 outcomes in the sample space.

Favourable outcomes are: a, e, i, o, u

There are 5 favourable outcomes.

$$\text{Pr(vowel)} = \frac{5}{26}$$

Fair games

There are many sorts of games. Some rely mainly on skill to win, such as chess. Some require a mixture of skill and luck, such as poker. Some rely purely on luck, such as lotteries.

When we play a game we expect it to be fair. By this, we mean that each person playing has the same chance of winning. Rules are written to make games fair.

We can use the sample space to help us decide whether a game is fair. For instance, if the game consists of rolling a single die, then the sample space consists of the numbers 1, 2, 3, 4, 5, 6, as these are the events that can occur when the die is rolled. If one player wins if an even number is rolled and the other player wins if a multiple of 3 is rolled, then the game is not fair as the first player has three chances of winning (2, 4 and 6) out of 6, whereas the second player has only two chances (3 and 6) out of 6.

9.7 Probability and sample space

Navigator

Q1, Q2, Q3, Q4, Q5 (a), Q7, Q9,
Q10, Q11

Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q9,
Q10, Q11

Q1, Q2, Q3, Q4, Q5, Q6, Q8, Q9,
Q10, Q12

**Answers
page 683**

Equipment required: A die for Question 6

Fluency

- 1 Find the probability of:

- (a) rolling an even number with a normal die
- (b) choosing an 'a' from the letters of the word *mathematics*
- (c) getting a 4 on a roll of a normal die
- (d) drawing a card from a normal pack and getting a diamond
- (e) drawing a card from a normal pack and getting a king
- (f) drawing a card from a normal pack and getting a black 4
- (g) drawing a card from a normal pack and getting a red 7
- (h) drawing a card from a normal pack and getting the 9 of spades.

WE16

9.7

- 2** Decide whether you would describe the following events as 'impossible', 'not likely', 'likely' or 'certain'.

- (a) Tomorrow will have a maximum temperature of -20°C .
- (b) Christmas Day will be on 25 December.
- (c) Sri Lanka will win the next cricket World Cup.
- (d) The Sun will set tonight.
- (e) You toss a coin and either a head or a tail turns up.
- (f) You roll a die and a 1 turns up.
- (g) You roll a die and a 7 turns up.
- (h) You draw a card from a normal pack of 52 playing cards and it's the ace of clubs.
- (i) You draw a card from a normal pack of playing cards and it's either red or black.



- 3** Choose the value that you think is closest to the correct answer.

- (a) The probability of winning first prize in Tattslotto is:

A 0	B 0.5
C 0.6	D 0.75
- (b) The probability of Australia winning a gold medal in swimming at the next Olympics is:

A 0.3	B 0.4
C 0.9	D 1
- (c) The probability of rain at some time during the next two months is:

A 0	B 0.2
C 0.5	D 1
- (d) The probability of tossing a coin and getting a tail is:

A 0.2	B 0.3	C 0.5	D 0.7
--------------	--------------	--------------	--------------
- (e) The probability of drawing a heart from a pack of cards is:

A 0.0001	B 0.25	C 0.5	D 1
-----------------	---------------	--------------	------------



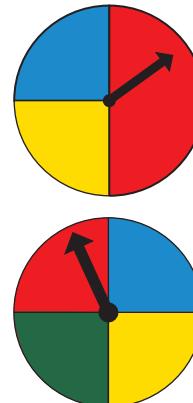
- 4** Write out the sample space for each of the following games.

- (a) A ten-sided die, labelled 1 to 10, is rolled once.
 - (b) A letter is drawn at random from the letters of the word *probability*.
 - (c) A coin is tossed twice.
- 5 A bag of jelly beans contains 4 blue, 5 red, 3 green, 2 purple and 1 black jelly beans.
- (a) Rachel puts her hand in the bag and selects one jelly bean without looking. What is the probability of the jelly bean being:

(i) red	(ii) green	(iii) blue or purple?
----------------	-------------------	------------------------------
 - (b) Rachel pulls out one red jelly bean and eats it. She selects another one from the bag without looking. Recalculate the three probabilities you calculated in (a).

Understanding

- 6 Determine whether the following games are fair by listing the sample space and the successful outcomes for each player. If you are unsure of the answer, or if you want to check if you are right, use a die and play the game with a partner.
- (i) Joe and Henry roll a die. Joe wins if a 1, 2 or 3 turns up, and Henry wins if a 4, 5 or 6 turns up. Is this a fair game?
 (ii) If Joe and Henry play their game 100 times, how many wins would they probably have each?
 - (i) Edward and Irene roll a die. Edward wins if a 3 turns up, and Irina wins if a 3 doesn't turn up. Is this a fair game?
 (ii) If Edward and Irene play their game a large number of times, and Edward wins 40 times, how many times would Irene probably have won?
 - (i) Ravi and Simon roll a die. Branka wins if a multiple of 3 turns up, and Simon wins if a multiple of 3 doesn't turn up. Is this a fair game?
 (ii) Ravi and Simon score a point each time one of them wins. Suppose the game is the first one to 60 points. Simon wins, but a week later they argue about what the final score was. Simon says he won 60 : 31, while Branka says Simon only won 60 : 55. Who do you think was probably right?
- 7 Steve and Nicky play a game based on the spinner shown. They each spin it once. Steve wins if he spins blue, and Nicky wins if she spins red.
- Is the game fair?
 - Why is there a problem with the 'fairness' of this game?
- 8 Maria and Walter are playing a game that is based on the spinner shown. Are these games fair? Think about the sample space and the number of chances of winning.
- Maria wins if the pointer finishes on blue and Walter wins if the pointer finishes on red. If the pointer finishes on anything else, they spin again until one of them wins.
 - Maria wins if the pointer finishes on blue or yellow and Walter wins if the pointer finishes on red. If the pointer finishes on green, they spin again until one of them wins.
 - Maria wins if the pointer finishes on blue or green and Walter wins if the pointer finishes on red or yellow.
 - Maria wins if the pointer finishes on blue, yellow or green, and Walter wins if the pointer finishes on red.
 - Maria wins if the pointer finishes on yellow and Walter wins if the pointer finishes on red. If the pointer finishes on green or blue, they spin again until one of them wins. However, if green comes up three times in a row, Maria wins, and if blue comes up four times in a row, Walter wins.



Reasoning

- 9 The following is a list of words you may have heard associated with probability.

even chance	very likely	virtually impossible	often
sure thing	most unlikely	almost always	quite often
likely	impossible	sometimes	usually
quite frequently	almost never	maybe	more often than not

Put these words in order from impossible to certain. Be prepared to defend your order in a class discussion.

9.7

- 10 A drawer contains six identical blue socks and two identical brown socks. How many socks do you have to take from the drawer to be certain of having:

- (a) a pair of blue socks
- (b) a pair of brown socks
- (c) a pair of socks
- (d) two odd socks?

Explain your answers in each case.

Open-ended

- 11 (a) Write down two events of your own that you think are impossible.
 (b) Write down two events of your own that you think are not likely.
 (c) Write down two events of your own that you think are likely.
 (d) Write down two events of your own that you think are certain.
- 12 Make up a game between two players involving two dice that has a number of conditions for winning, but that you think is fair.

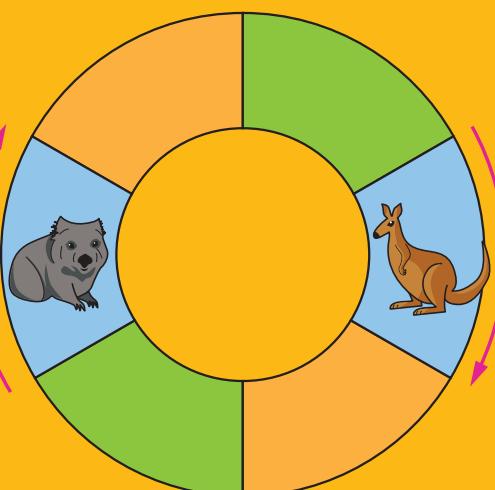
Outside the Square Game

Catch me if you can

Equipment required: 1–2 brains, 2 dice, 2 unique counters

How to win:

Land on the segment the other player's counter is on.



How to play:

Decide which player is the 'kangaroo' and which is the 'wombat' and place your counter on the appropriate start position. Roll the two dice and move the appropriate counter one space, given the following rules.

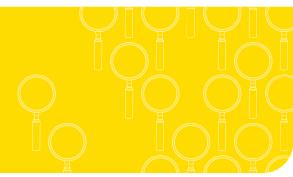
The wombat moves if the sum of the dice is equal to: 2, 3, 4, 5, 10, 11, 12.

The kangaroo moves if the sum of the dice is equal to: 6, 7, 8 or 9.

Continue rolling the two dice until one of the animals catches the other.

After you have played a few times, decide whether you think the game is fair. Explain your reasons. If you don't think it is fair, see if you can change the rules to make it fair. Play the new game a few times to test it.

Investigation



Experimental probability

Equipment required: 2 brains, 1 calculator, 1 drawing pin, 1 coin



The Big Question

Can we use probability to decide whether a game of chance is fair?

Engage

- 1 A drawing pin can land two possible ways. Try dropping the drawing pin a few times and see if it lands point up or point down. Do you think you need to be careful in the way you drop the pin? Should you give it a bit of a 'flip' to help it on its way? Why do you think we might be using the pin instead of a coin or a die?

point up point down



Explore

Don and Jacob play a game that involves dropping a drawing pin.

If the drawing pin falls point up, Don scores a point. If the drawing pin falls point down, Jacob scores a point. The first one to get to 50 wins.

- 2 (a) Do you think this is a fair game? Who do you think will win?
 (b) Predict the final score. Is it harder to make this sort of prediction than the predictions in the previous exercise?

- (c) Play the game with a partner. Decide who gets a score for point up and who gets a score for point down. Tally your results in a table like this:

Point up			
Point down			

How close were your predictions?

- (d) Compare your results with other pairs of students by constructing a table like this:

	Score	
	Point up	Point down
Pair 1	32	50
Pair 2	50	44
Pair 3		

- (e) Do you notice any pattern in the results of the games?
 (f) (i) Find the point up total by adding all the numbers in the point up column.
 (ii) Find the point down total by adding up all the numbers in the point down column.
 (iii) Then, find the total number of throws by adding the point up and point down totals.
 (g) Find the probability that a drawing pin will fall point up by calculating
 $\frac{\text{point up total}}{\text{total number of throws}}$:
 (i) as a fraction
 (ii) as a decimal
 (iii) as a percentage.
 (iv) Could the game still be fair if the percentage is not exactly 50%?
 (h) Find the probability that a drawing pin will fall point down by calculating
 $\frac{\text{point down total}}{\text{total number of throws}}$:
 (i) as a fraction
 (ii) as a decimal
 (iii) as a percentage.

- (iv) Could the game still be fair if the percentage is not exactly 50%?
- (v) Look at the two probabilities you have found in parts (g) (ii) and (h) (ii). Add them together. What do you notice?
- (vi) If the game is fair, what do you think each of the two probabilities should be equal to? Is the game fair? Is it close to being a fair game?



Strategy options

- Make a table.
- Act it out.
- Look for a pattern.

Explain

- 3 How could you change the rules for the drawing pin game to make it fair?

Elaborate

- 4 Omar and Michelle take it in turns to toss a coin four times. If four heads come up, or if two heads and two tails come up, Omar scores a point. If four tails come up, or if three tails and one head come up, Michelle scores a point. If three heads and one tail come up then they both score a point. The first one to 50 points is the winner.
- (a) Do you think this is a fair game? Who do you think will win?
 - (b) Make a prediction about what you think the final score will be.
 - (c) Play the game with a partner. Decide who is Omar and who is Michelle.
Tally your results in a table. How close were your predictions?
 - (d) Compare your results with other pairs of students by constructing a table.
 - (e) Do you notice any pattern in the results of the games?
 - (f)
 - (i) Find Omar's total by adding all the numbers in the Omar column.
 - (ii) Find Michelle's total by adding all the numbers in the Michelle column.

- (iii) Then, find the total number of throws by adding Omar's and Michelle's totals.

- (g) (i) Find the fraction of points that Omar gets by calculating:

$$\frac{\text{Omar's total}}{\text{total number of throws}}$$

- (ii) Change this to a decimal. Write the probability that Omar will win as $\Pr(\text{Omar wins}) =$

- (iii) Write the percentage of times Omar will win.

- (h) (i) Find the fraction of points that Michelle gets by calculating:

$$\frac{\text{Michelle's total}}{\text{total number of throws}}$$

- (ii) Change this to a decimal. Write the probability that Michelle will win as $\Pr(\text{Michelle wins}) =$

- (iii) Write the percentage of times Michelle will win.

- (iv) Find $\Pr(\text{Omar wins}) + \Pr(\text{Michelle wins})$.

- (v) By looking at your values for $\Pr(\text{Omar wins})$ and $\Pr(\text{Michelle wins})$, would you say that the game is fair? Is it close to being fair?

Evaluate

- 5 How could you change the rules for the coin game to make it fair?

Extend

- 6 What other sorts of probability games do you think you could play? Try to come up with some that do not involve coins or dice.
- 7 (a) Did you affect the result by the way you dropped the pin or tossed the coin?
(b) Do you think the tasks in this investigation have helped you understand probability? Explain why or why not?
(c) Do you think it is important for you to know whether a game is fair? Why?

Challenge 9



- 1 The dots on opposite faces of a die total 7.

Two normal dice are placed next to each other on a table as shown.

Kenny walks around the table.



The greatest number of dots that Kenny can see is:

- A 30 B 34 C 36 D 42

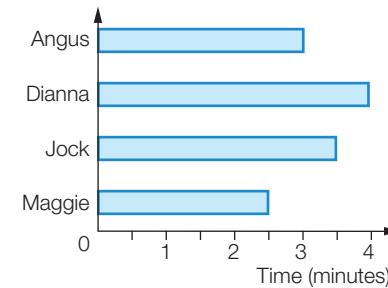
- 2 You meet three men at a meeting: Mr Black, Mr Green and Mr White. They are wearing black, green and white bow ties. Mr Black says, "Did you notice that the colour of our ties are different from our names?" The person in the green bow tie says, "Yes, you are correct!" What colour tie is each man wearing?

- 3 A die is constructed by labelling the faces of a wooden cube with the numbers 1, 1, 2, 3, 3 and 4. When it is tossed, each face has an equal chance of occurring. If this die is rolled once, the probability of rolling an odd number is:

- A $\frac{1}{6}$ B $\frac{1}{3}$ C $\frac{1}{2}$ D $\frac{2}{3}$

- 4 Four Jack Russell terriers had a race swimming across a pond. The graph shows the time that each dog took to complete the race. Which dog won the race?

- A Angus B Dianna
C Jock D Maggie

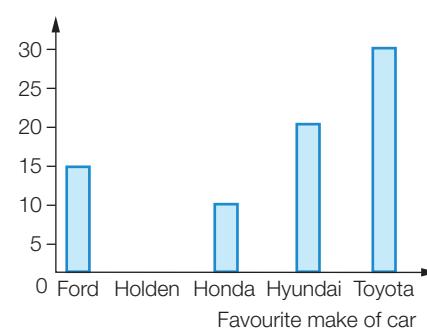


- 5 Vijay has a bag that contains 3 blue marbles, 6 green marbles, 2 red marbles and 6 yellow marbles. Vijay adds a number of white marbles to the bag and tells Samantha that if he now draws a marble from the bag, the probability of it being blue or green is $\frac{3}{7}$. The number of white marbles that Vijay added to the bag is:

- A 3 B 4 C 5 D 6

- 6 In a survey, 1000 people were asked 'What is your favourite make of car?' Their responses were recorded and then graphed as percentages in a column graph. The column representing Holden has been omitted. What percentage of the people selected a Holden as their favourite make of car?

- A 25 B 20
C 15 D 5



- 7 Jade has played 20 games of tennis and has won 95% of the games. If she does not lose any more games, how many more games in a row must she win in order to have then won 96% of the games?

- A 1 B 4 C 5 D 10

Chapter review

9

D.I.Y. Summary

Key Words

average	continuous data	measures of centre	sample space
bar graph	data	median	stem-and-leaf plot
bimodal	discrete data	mode	tally
bivariate data	dot plot	multimodal	theoretical probability
categorical data	frequency	outcomes	trend
chance	frequency table	outlier	univariate data
class interval	histogram	probability	
column graph	mean	range	

Copy and complete the following using the words and phrases from this list, where appropriate, to write a summary for this chapter. A word or phrase may be used more than once.

- The graphs in this chapter are all ways of presenting _____.
- The listing of all the outcomes possible in a probability experiment is called the _____.
- Of the three measures of central tendency considered, the one that represents the value occurring most frequently is called the _____.
- The probability of an event occurring is another way of saying what _____ there is of the event happening.
- In a frequency table for grouped data, each group is called a _____.
- The _____ is the sum of all the values in the data set divided by the number of values in that data set.

Equipment required: Calculator for Question 2

Fluency

- A Year 7 class was surveyed about how many television sets their families had at home. The results were:

1, 1, 1, 1, 2, 1, 1, 3, 2, 5, 2, 1, 1, 0, 1, 2, 0, 4, 1, 2, 0, 3, 1

Construct a frequency table for this information.

Ex. 9.1

- Find the mean, median and mode for the following sets of data. Use a calculator when necessary, and give your answers correct to two decimal places.

(a) 2, 4, 2, 7, 3, 5, 4, 2, 3, 6, 1

(b) 40, 20, 60, 30, 50, 10

(c) 3.3, 4.1, 3.9, 2.9, 4.4, 4.1, 3.8, 3.3, 4.4, 3.2, 3.9, 4.1, 2.9

Ex. 9.2

- 3 Draw a bar graph to show the following data that represents the favourite style of TV program for a sample of 100 people.

Ex. 9.3

Style	Number of people
Comedy	35
Drama	21
Reality	17
Documentary	8
Game	19

- 4 (a) Draw a dot plot to represent the number of customers in a cafe recorded in 15-minute intervals over the course of 8 hours.

4	4	7	3	6	4	5	9	4	8	8	5	7	9	10	12
11	8	9	8	6	2	4	4	5	5	2	5	3	0	1	1

- (b) Calculate the median of the data set.

- (c) Do there appear to be any outliers?

- 5 (a) Draw an ordered stem-and-leaf plot to represent the number of runs scored by each of the players at a local cricket club in one round of the season.

46	64	31	0	50	42	11	7	4	0	5
21	17	115	0	42	61	15	18	5	3	1
40	0	52	56	77	50	41	20	7	2	2
88	17	58	63	14	0	18	19	17	6	2

- (b) Calculate the median of the data set.

- (c) Do there appear to be any outliers? If so, which values are they?

- 6 The sector graph at right represents information about the money spent by overseas visitors to Australia.

Ex. 9.3

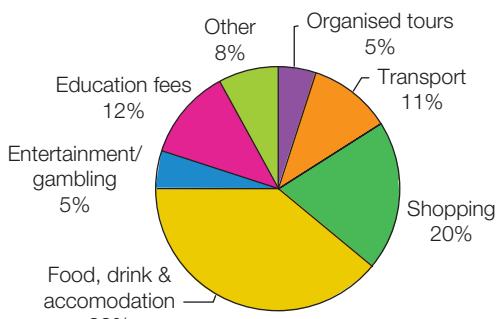
- (a) One of the sectors uses an angle of 28.8° . Which one is it?

- A Organised tours
- B Transport
- C Entertainment/gambling
- D Other

- (b) The Food, drink & accommodation sector uses an angle closest to:

- A 95°
- B 110°
- C 130°
- D 140°

- 7 The following are the sales figures for Beeper computer games from 2006 to 2011.

Ex. 9.4

Year	2006	2007	2008	2009	2010	2011
Units sold	10 000	25 000	35 000	30 000	40 000	45 000

Draw a line graph for this information.

Ex. 9.5

- 8 The following table gives the mean rainfall, measured in mm, for Cairns (Aero) and Brisbane (Aero).

Ex. 9.6

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Cairns	396.8	452.5	426.1	198.2	91.4	46.1	29.3	27.0	33.4	39.7	93.4	178.5
Brisbane	109.3	121.6	76.1	63.5	122.8	70.6	27.1	35.2	32.7	65.2	102.0	120.9

Compare the mean rainfall for the two cities using as many statistics as you can.

- 9 The following table shows the handspan, measured in cm, of twenty 16-year-old students, divided into boys and girls.

Ex. 9.6

Boys	19.5	21.5	21	22.5	19	19.5	21	21.5	20.5	20
Girls	18	18.5	17.5	17	17	17.5	18	18.5	17	19.5

Draw a parallel dot plot and use it to help you describe the data set.

- 10 (a) When a normal die is rolled, the probability of getting a number less than 5 is equal to:

A $\frac{1}{5}$

B $\frac{2}{3}$

C $\frac{5}{6}$

D 1

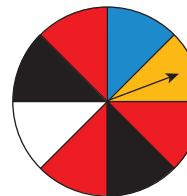
- (b) If the spinner shown on the right is spun once, the probability of the pointer landing on red is:

A $\frac{1}{4}$

B $\frac{3}{8}$

C $\frac{1}{2}$

D $\frac{3}{5}$

**Ex. 9.7**

Understanding

- 11 A number of Year 7 students were surveyed about their shoe size. The results were:

3, 5, $5\frac{1}{2}$, 5, $4\frac{1}{2}$, 4, $2\frac{1}{2}$, $2\frac{1}{2}$, $3\frac{1}{2}$, 6, 2, $5\frac{1}{2}$, $1\frac{1}{2}$, 1, $3\frac{1}{2}$, $4\frac{1}{2}$, 3, 4, 5, $5\frac{1}{2}$, $2\frac{1}{2}$, 2, $3\frac{1}{2}$, $4\frac{1}{2}$, 3, 1, $3\frac{1}{2}$, $1\frac{1}{2}$, 4, $3\frac{1}{2}$

- (a) Construct a frequency table to show this information.

- (b) Which shoe size occurred most frequently?

- (c) Which shoe size occurred least frequently?

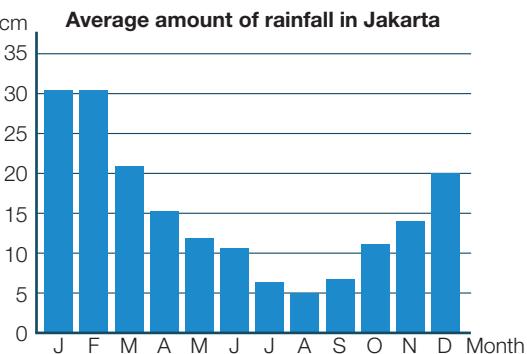
- 12 Look at the following information.

How much Australians spend on pets each year (\$ million)				
Expenditure type	Dogs	Cats	Other	Total
Food	560	431	90	1081
Vet charges and prescriptions	230	148	26	404
Pet care products/equipment	153	85	12	250
Pet services	116	35	—	151
Other expenses	109	20	12	141
Total	1168	719	140	2027

- (a) Draw a bar chart showing the various expenditure types for dogs.
 (b) Draw a divided bar chart to show the total expenditure for the three categories of pets.
 (c) Draw a sector graph to show the various expenditure types for cats.
 (d) Explain why a line graph could not be used for any of these sets of data.

- 13 Look at the following graph, which shows the average amount of rainfall (cm) in the city of Jakarta over 1 year.

- (a) What type of graph is this?
 (b) What is Jakarta's second driest month?
 (c) What is the highest average monthly rainfall?
 (d) Which 4 months form the rainy season?



- 14 Shelley and Douglas play a game in which they toss two coins. Shelley wins if two heads turn up, and Douglas wins if two tails turn up. If both a head and a tail turn up, they throw one of the coins again. If this comes up heads, Shelley wins. If it comes up tails, Douglas wins. Do you think this is a fair game?

- 15 (a) What value would the probability that you will be struck by lightning be close to?
 (b) What value would the probability that it will rain at some time in the next six months be close to?
 (c) If something has a probability of $\frac{1}{2}$ and something else has a probability of $\frac{3}{4}$, which one of the two is more likely to occur?
 (d) If someone tells you something has a probability of 1.2 of happening, how do you know that they have made a mistake?

Reasoning

- 16 Twenty-five Year 7 boys and 25 Year 7 girls had their height measured to the nearest cm. The results are as follows.

Boys:	140	143	144	142	148	148	152	140	146	144
	154	151	145	140	141	147	153	150	149	140
	147	152	148	151	149					
Girls:	142	151	157	138	145	148	143	141	146	154
	156	154	150	143	150	144	148	144	157	156
	146	157	154	145	152					

- (a) Construct an ungrouped frequency table for the boys.
 (b) Find the (i) mode (ii) median and (iii) mean height for the boys.
 (c) Construct an ungrouped frequency table for the girls.
 (d) Find the (i) mode (ii) median and (iii) mean height for the girls.
 (e) Describe the similarities and differences between the heights of the boys and the girls.
 (f) Draw a combined ungrouped frequency table for boys and girls.
 (g) Find the (i) mode (ii) median and (iii) mean height for the students.
 (h) Describe the heights of the students.

- 17 In this game you have 12 counters and three dice. You place as many counters as you like on any number of squares, although you must place all 12. You throw the dice and add the numbers shown. If you have any counters on that number you remove them. You continue until the board is clear.

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

- (a) On which numbers should you never place any counters? Explain why.
- (b) On which numbers should you place most counters? Explain why.
- (c) Once you have decided on a spread of counters, will every game take the same number of turns to clear the board? Explain your answer.

NAPLAN practice 9

Numeracy: Non-calculator

- 1 Some students were surveyed about the number of hours spent using a computer over the last weekend. The following results were recorded: 4, 7, 3, 2, 4, 6, 8, 5, 9.

The median result is:

- A 4 B 5 C 6 D 9

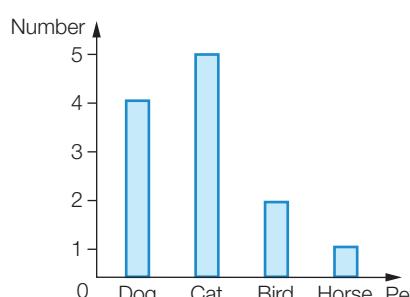
- 2 A jar contains 16 red, 12 black, 9 yellow and 8 green lollies. Without looking, Jen takes one lolly from the jar. What is the chance that the lolly is yellow?

- A $\frac{1}{2}$ B $\frac{1}{3}$ C $\frac{1}{4}$ D $\frac{1}{5}$

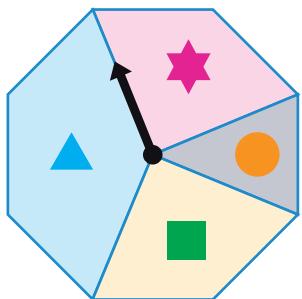
- 3 Ricky surveyed the members of his cricket team to find out what pets they had. The results are shown in the graph on the right.

How many pets does the team have altogether?

- A 4 B 5 C 12 D 15



- 4 Kenny spins the arrow 80 times.



Which table is *most* likely to show his results?

Shape section	Number of spins
★	10
●	30
■	20
▲	20

Shape section	Number of spins
★	20
●	30
■	10
▲	20

Shape section	Number of spins
★	20
●	10
■	20
▲	30

Shape section	Number of spins
★	20
●	20
■	10
▲	30

Numeracy: Calculator allowed

- 5 A business recorded the number of customers each day for a working week.

- (a) Which day was the busiest?
- (b) What was the mean number of customers per day?
- (c) On which day was the number of customers closest to the mean number of customers per day?

- 6 The goal attack for the under 13 Panthers netball team scored the following number of goals in the first six games of the season: 14, 19, 23, 20, 28, 19.

The mean number of goals scored is closest to:

A 6

B 19.5

C 20.5

D 126

