Chapter

8 Statistics and probability

What you will learn

Australian curriculum

STATISTICS AND PROBABILITY

1 Chance

Construct sample spaces for single-step experiments with equally likely outcomes (ACMSPA167)

Assign probabilities to the outcomes of events and determine probabilities for events (ACMSPA168)

2 Data representation and interpretation

Identify and investigate issues involving continuous or large count data collected from primary and secondary sources (ACMSPA169)

Construct and compare a range of data displays including stem-and-leaf plots and dot plots (ACMSPA170)

Calculate mean, median, mode and range for sets of data. Interpret these statistics in the context of data (ACMSPA171)

Describe and interpret data displays and the relationship between the median and mean (ACMSPA172)

NUMBER AND ALGEBRA

3 Linear and non-linear relationships

Investigate, interpret and analyse graphs from authentic data (ACMNA180) AC

Online resources

- Chapter pre-test
- Videos of all worked examples
- Interactive widgets
- Interactive walkthroughs
- Downloadable HOTsheets
- Access to HOTmaths Australian Curriculum courses

Codes and ciphers

Cracking codes and deciphering secret messages involves a lot of maths, from simple addition and subtraction to data handling and logical thinking.

One style of code replaces letters with symbols or pictures. This type of code was used in the sixteenth century by Mary Queen of Scots, when she was plotting against Queen Elizabeth I and sending coded messages about her plans to her co-conspirators. Unfortunately for Mary, there was a simple mathematical way of cracking the code. The letter 'E' is the most commonly used letter in the English language. 'T' is the second most common letter and 'A' is the third most commonly used letter.

All that Elizabeth's spymaster had to do to crack the code was to look through the coded message and count the number of times each symbol came up. The symbol that came up the most would probably stand for the letter 'E'. The next letter that occurred most often was probably 'T', and so forth. This method, called 'frequency analysis', gave the spymaster enough information to decode Mary's messages and foil her plans.

8A Collecting and classifying data

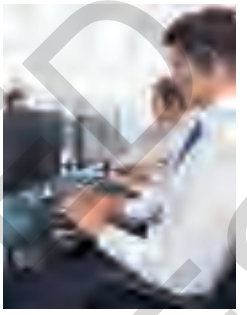
Interactive

Widgets

HOTsheets

Walkthroughs

People collect or use data almost every day. Athletes and sports teams look at performance data, customers compare prices at different stores, investors look at daily interest rates, and students compare marks with other students in their class. Companies often collect and analyse data to help produce and promote their products to customers and to make predictions about the future.



Let's start: Collecting data

Consider, as a class, the following questions and discuss their implications.

- Have you or your family ever been surveyed by a telemarketer at home? What did they want? What time did they call?
- Do you think that telemarketers get accurate data? Why or why not?
- Why do you think companies collect data this way?
- If you wanted information about the most popular colour of car sold in NSW over the course of a year, how could you find out this information?

Key ideas

- In statistics, a **variable** is something measurable or observable that is expected to change over time or between individual observations. It can be numerical or categorical.
  - **Numerical** (**quantitative**) which can be discrete or continuous:
    - **Discrete numerical** – data that can only be particular numerical values, e.g. the number of TV sets in a house (could be 0, 1, 2, 3 but not values in between like 1.3125).
    - **Continuous numerical** – data that can take any value in a range. Variables such as heights, weights and temperatures are all continuous. For instance, someone could have a height of 172 cm, 172.4 cm, or 172.215 cm (if it can be measured accurately).
  - **Categorical** – data that are not numerical such as colours, gender, brands of cars are all examples of categorical data. In a survey, categorical data comes from answers which are given as words (e.g. 'yellow' or 'female') or ratings (e.g. 1 = dislike, 2 = neutral, 3 = like).
- Data can be collected from primary or secondary sources.
  - Data from a **primary source** are firsthand information collected from the original source by the person or organisation needing the data, e.g. a survey an individual student conducts or census data collected and then used by an organisation like the Bureau of Statistics.
  - Data from a **secondary source** have been collected, published and possibly summarised by someone else before we use it. Data collected from newspaper articles, textbooks or internet blogs represent secondary source data.
- Samples and populations
  - When an entire population (e.g. a maths class, all the cars in a parking lot, a company, or a whole country) is surveyed, it is called a **census**.
  - When a subset of the population is surveyed, it is called a **sample**. Samples should be randomly selected and large enough to represent the views of the overall population.
  - When we cannot choose which members of the population to survey, and can record only those visible to us (e.g. people posting their political views on a news website), this is called an **observation**.

Example 1 Classifying variables

Classify the following variables as categorical, discrete numerical or continuous numerical.

a the gender of a newborn baby

b the length of a newborn baby

SOLUTION	EXPLANATION
a categorical	As the answer is 'male' or 'female' (a word, not a number) the data are categorical.
b continuous numerical	Length is a measurement, so all numbers are theoretically possible.

Example 2 Collecting data from primary and secondary sources

Decide whether a primary source or a secondary source is suitable for collection of data on each of the following and suggest a method for its collection.

a the average income of Australian households

b the favourite washing powder or liquid for households in Australia

SOLUTION	EXPLANATION
a primary source by looking at the census data	The population census held every 5 years in Australia is an example of a primary data source collection and will have this information.
b secondary data source using the results from a market research agency	A market research agency might collect these results using a random phone survey. Obtaining a primary source would involve conducting the survey yourself but it is unlikely that the sample will be large enough to be suitable.

Exercise 8A

1–3 | 2 | —

UNDERSTANDING

1 Match each word on the left to its meaning on the right.

a sample	i only takes on particular numbers within a range
b categorical	ii a complete set of data
c discrete numerical	iii a smaller group taken from the population
d primary source	iv data grouped in categories like 'male' and 'female'
e continuous numerical	v data collected firsthand
f population	vi can take on any number in a range

8A

Example 1

UNDERSTANDING

2 Classify the following as categorical or numerical.

- **a** the eye colour of each student in your class
- **b** the date of the month each student was born, e.g. the 9th of a month
- **c** the weight of each student when they were born
- **d** the brands of airplanes landing at Sydney's international airport
- **e** the temperature of each classroom
- **f** the number of students in each classroom period one on Tuesday

3 Give an example of:

a discrete numerical data **b** continuous numerical data **c** categorical data

FLUENCY

4–5(½), 6 4–5(½), 6 4–5(½), 6

4 Classify the following variables as categorical, continuous numerical or discrete numerical data.

- **a** the number of cars per household
- **b** the weights of packages sent by Australia Post of the 20th of December
- **c** the highest temperature of the ocean each day
- **d** the favourite brand of chocolate of the teachers at your school
- **e** the colours of the cars in the school car park
- **f** the brands of cars in the school car park
- **g** the number of letters in different words on a page
- **h** the number of advertisements in a time period over each of the free-to-air channels
- **i** the length of time spent doing this exercise
- **j** the arrival times of planes at JFK airport
- **k** the daily pollutions level in the M5
- **l** the number of text messages sent by an individual yesterday
- **m** the times for the 100 m freestyle event at the world championships over the last 10 years
- **n** the number of Blu-ray discs someone owns
- **o** the brands of cereals available at the supermarket
- **p** marks awarded on a maths test
- **q** the star rating on a hotel or motel
- **r** the censorship rating on a movie showing at the cinema

5 Is observation or a sample or a census the most appropriate way to collect data on each of the following?

- **a** the arrival times of trains at central station during a day
- **b** the arrival times of trains at central station over the year
- **c** the heights of students in your class
- **d** the heights of all Year 7 students in the school
- **e** the heights of all Year 7 students in NSW
- **f** the number of plastic water bottles sold in a year
- **g** the religion of Australian families
- **h** the number of people living in each household in your class
- **i** the number of people living in each household in your school
- **j** the number of people living in each household in Australia
- **k** the number of native Australian birds found in a suburb
- **l** the number of cars travelling past a school between 8 a.m. and 9 a.m. on a school day
- **m** the money spent by students during a week at the canteen
- **n** the ratings of TV shows

8A

Example 2

FLUENCY

6 Identify whether a primary or secondary source is suitable for the collection of data on the following.

- **a** the number of soft drinks bought by the average Australian family in a week
- **b** the age of school leavers in far North Queensland
- **c** the number of cigarettes consumed by school age students in a day
- **d** the highest level of education by the adults in a household
- **e** the reading level of students in Year 7 in Australia

7, 8 | 7, 9–11 | 10–12

PROBLEM-SOLVING

7 Give a reason why someone might have trouble obtaining reliable and representative data using a primary source to find the following.

- **a** the temperature of the Indian Ocean over the course of a year
- **b** the religions of Australian families
- **c** the average income of someone in India
- **d** drug use by teenagers within a school
- **e** the level of education of different cultural communities within NSW

8 Secondary sources are already published data that are then used by another party in their own research. Why is the use of this type of data not always reliable?

9 When obtaining primary source data you can survey the population or a sample.

- **a** Explain the difference between a 'population' and a 'sample' when collecting data.
- **b** Give an example situation where you should survey a population rather than a sample.
- **c** Give an example situation where you should survey a sample rather than a population.

8A

PROBLEM-SOLVING

10 A Likert-type scale is for categorical data where items are assigned a number; for example, the answer to a question could be 1 = dislike, 2 = neutral, 3 = like.

- **a** Explain why the data collected are categorical even though the answers are given as numbers.
- **b** Give examples of a Likert-type scale for the following categorical data. You might need to reorder some of the options.
  - **i** strongly disagree, somewhat disagree, somewhat agree, strongly agree
  - **ii** excellent, satisfactory, poor, strong
  - **iii** never, always, rarely, usually, sometimes
  - **iv** strongly disagree, neutral, strongly agree, disagree, agree

11 A sample should be representative of the population it reports on. For the following surveys, describe who might be left out and how this might introduce a bias.

- **a** a telephone poll with numbers selected from a phone book
- **b** a postal questionnaire
- **c** door-to-door interviews during the weekdays
- **d** a *Dolly* magazine poll
- **e** a Facebook survey

12 Another way to collect primary source data is by direct observation. For example, the colour of cars travelling through an intersection (categorical data) is best obtained in this way rather than through a questionnaire.

- **a** Give another example of a variable for which data could be collected by observation.
- **b** Explain how you could estimate the proportion of black cars parked at a large shopping centre car park without counting every single one.

13 | 14, 15 | 14, 16

REASONING

13 When conducting research on Indigenous Australians, the elders of the community are often involved. Explain why the community is needed to be involved in the research process.

14 Television ratings are determined by surveying a sample of the population

- **a** Explain why a sample is taken rather than conducting a census.
- **b** What would be a limitation of the survey results if the sample included 50 people nationwide?
- **c** If a class census was taken on which (if any) television program students watched from 7.30–8.30 last night, why might the results be different to the official ratings?
- **d** Research how many people are sampled by Nielsen Television Audience Measurement in order to get an accurate idea of viewing habits and stick within practical limitations.

8A

REASONING

15 Australia's census surveys the entire population every five years.

- **a** Why might Australia not conduct a census every year?
- **b** Over 40% of all Australians were born overseas or had at least one of their parents born overseas. How does this impact the need to be culturally sensitive when designing and undertaking a census?
- **c** The census can be filled out on a paper form or using the internet. Given that the data must be collated in a computer eventually, why does the government still allow paper forms to be used?
- **d** Why might a country like India or China conduct their national census every 10 years?

16 Write a sentence explaining why two different samples taken from the same population can produce different results. How can this problem be minimised?

Collecting a sample — — 17

ENRICHMENT

17 a Use a random number generator on your calculator or computer to record the number of times the number 1 to 5 appears (you could even use a die by re-rolling whenever you get a 6) out of 50 trials. Record these data.

- **i** Tabulate your results.
- **ii** Compare the results of the individuals in the class.
- **iii** Explain why differences between different students might occur.

b Choose a page at random from a novel or an internet page and count how many times each vowel (A, E, I, O, U) occurs. Assign each vowel the following value
A = 1, E = 2, I = 3, O = 4, U = 5 and tabulate your results.

- **i** Why are the results different from those in part **a**?
- **ii** How might the results for the vowels vary depending on the webpage or novel chosen?

8B Summarising data numerically

Although sometimes it is important to see a complete set of data, either as a list of numbers or as a graph, it is often useful to summarise the data with a few numbers.

For example, instead of listing the height of every Year 7 student in a school, you could summarise this by stating the median height and the difference (in cm) between the tallest and shortest people.

Let's start: Class summary

For each student in the class, find their height (in cm), their age (in years), and how many siblings they have.

- Which of these three sets would you expect to have the largest range?
- Which of these three sets would you expect to have the smallest range?
- What do you think is the mean height of students in the class? Can you calculate it?

Key ideas

- The **range** of a set of data is given by:
  Range = highest number – lowest number.

$$\underset{\text{lowest}}{1}\quad 6\quad \underset{\text{highest}}{7}\quad 1\quad 5 \longrightarrow \text{range} = 7 - 1 = 6$$

- The **mean** of a set of data is given by:
  Mean = (sum of all the values) ÷ (total number of values)

$$1 + 6 + 7 + 1 + 5 = 20 \longrightarrow \text{mean} = 4$$

- The **median** is the middle value if the values are sorted from lowest to highest. If there are two middle values, then add them together and divide by 2.

$$1\quad 1\quad \underset{\text{middle}}{5}\quad 6\quad 7 \longrightarrow \text{median} = 5$$

- The **mode** is the most common value. It is the value that occurs most frequently. We also say that it is the value with the highest frequency. There can be more than one mode.

$$1\quad 1\quad 5\quad 6\quad 7 \longrightarrow \text{mode} = 1$$

Example 3 Finding the range, mean, median and mode

Consider the ages (in years) of seven people who are surveyed in a shop:
15, 31, 12, 47, 21, 65, 12

a Find the range of values.
b Find the mean of this set of data.
c Find the median of this set of data.
d Find the mode of this set of data.
e If another person is surveyed who is 29 years old, what will be the new median?

SOLUTION	EXPLANATION
a range = 65 – 12 = 53	Highest number = 65, lowest number = 12 The range is the difference.
b mean = 203 ÷ 7 = 29	Sum of values = 15 + 31 + 12 + 47 + 21 + 65 + 12 = 203 Number of values = 7
c values: 12, 12, 15, **21**, 31, 47, 65 median = 21	Place the numbers in ascending order and see that the middle value is 21.
d mode = 12	The most common value is 12.
e values: 12, 12, 15, **21**, **29**, 31, 47, 65 new median = 25	Place the numbers in ascending order. Now there are two middle values (21 and 29), so the median is $\frac{21+29}{2} = \frac{50}{2} = 25$.

Exercise 8B

1–3 | 1, 3 | —

UNDERSTANDING

Example 3a

1 Consider the set of numbers 1, 5, 2, 10, 3.
a Write the numbers from smallest to largest.
b State the largest number.
c State the smallest number.
d What is the range?

2 State the range of the following sets of numbers.
a 2, 10, 1, 3, 9
b 6, 8, 13, 7, 1
c 0, 6, 3, 9, 1
d 3, 10, 7, 5, 10

Example 3b–d

3 For the set of numbers 1, 5, 7, 7, 10, find the:
a total of the numbers when added
b mean
c median
d mode

4(½), 5 | 4(½), 5, 6 | 4(½), 5, 6

8B

FLUENCY

Example 3e

4 For each of the following sets of data, calculate the:

i range **ii** mean **iii** median **iv** mode

a 1, 7, 1, 2, 4
b 2, 2, 10, 8, 13
c 3, 11, 11, 14, 21
d 25, 25, 20, 37, 25, 24
e 1, 22, 10, 20, 33, 10
f 55, 24, 55, 19, 15, 36
g 114, 84, 83, 81, 39, 12, 84
h 97, 31, 18, 54, 18, 63, 6

5 The number of aces that a tennis player serves per match is recorded over eight matches.

Match	1	2	3	4	5	6	7	8
Number of aces	11	18	11	17	19	22	23	12

a What is the mean number of aces the player serves per match? Round your answer to 1 decimal place.
b What is the median number of aces the player serves per match?
c What is the range of this set of data?

6 Brent and Ali organise their test marks for a number of topics in Maths, in a table.

	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Test 10
Brent	58	91	91	75	96	60	94	100	96	89
Ali	90	84	82	50	76	67	68	71	85	57

a Which boy has the higher mean?
b Which boy has the higher median?
c Which boy has the smaller range?
d Which boy do you think is better at tests? Explain why.

7, 8 | 8–10 | 9–11

PROBLEM-SOLVING

7 The set 3, 7, 9, 10 has one extra number added to it, and this causes the mean to be doubled. What is the number?

8 Alysha's tennis coach records how many double faults Alysha has served per match over a number of matches. Her coach presents the results in a table.

Number of double faults	0	1	2	3	4
Number of matches with this many double faults	2	3	1	4	2

a In how many matches does Alysha have no double faults?
b In how many matches does Alysha have 3 double faults?
c How many matches are included in the coach's study?
d What is the total number of double faults scored over the study period?
e Calculate the mean of this set of data, correct to 1 decimal place.
f What is the range of the data?

8B

PROBLEM-SOLVING

9 A soccer goalkeeper recorded the number of saves he makes per game during a season. He presents his records in a table.

Number of saves	0	1	2	3	4	5
Number of games	4	3	0	1	2	2

a How many games did he play that season?
b What is the mean number of saves this goalkeeper made per game?
c What is the most common number of saves that the keeper had to make during a game?

10 Give an example of a set of data with:
a a mean of 10 and a range of 2
b a median of 10 and a range of 5
c a range of 100 and a mean of 50
d a mean of 6, a median of 7 and a mode of 5

11 The set 1, 2, 5, 5, 5, 8, 10, 12 has a mode of 5 and a mean of 6.
a If a set of data has a mode of 5 (and no other modes) and a mean of 6, what is the smallest number of values the set could have? Give an example.
b Is it possible to make a dataset for which the mode is 5, the mean is 6 and the range is 20? Explain your answer.

12 | 12, 13 | 14–16

REASONING

12 Evie surveys all the students in her class to find the distance from their homes to school. One of the students is on exchange from Canada and reports a distance of 16 658 km. Would this very large value have a greater effect on the mean or median distance? Explain your answer.

8B

REASONING

13 Consider the set of values 1, 3, 5, 10, 10, 13.

a Find the mean, median, mode and range.

b If each number is increased by 5, state the effect this has on the:

i mean **ii** median **iii** mode **iv** range

c If each of the original numbers is doubled, state the effect this has on the:

i mean **ii** median **iii** mode **iv** range

d Is it possible to include extra numbers and keep the same mean, median, mode and range? Try to expand this set to at least 10 numbers, but keep the same values for the mean, median, mode and range.

14 **a** Two whole numbers are chosen with a mean of 10 and a range of 6. What are the numbers?

b Three whole numbers are chosen with a mean of 10 and a range of 2. What are the numbers?

c Three whole numbers are chosen with a mean of 10 and a range of 4. Can you determine the numbers? Try to find more than one possibility.

15 Prove that for three consecutive numbers, the mean will equal the median.

16 A **frequency table** shows how common an event is. For instance, Magda has presented her weekly spelling test scores as a frequency table.

Score	4	5	6	7	8	9	10
Frequency	1	5	5	11	12	5	1

This is easier than writing out the results as $4, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, \ldots, 9, 10$.

a State the range of scores and the mean score, correct to 2 decimal places.

b If the frequency of each score is reduced by 1, describe the effect this will have on the range and the mean of the scores.

Mean challenges — — 17, 18

ENRICHMENT

17 For the set of numbers $1, 2, 3, \ldots, 100$, find the mean.

18 **a** Give an example of a set of numbers with the following properties.

i mean = median = mode

ii mean > median > mode

iii mode > median > mean

iv median < mode < mean

b If the range of a set of data is 1, is it still possible to find datasets for each of parts **i** to **iv** above?

An important aspect of scientific investigation is collecting data and summarising it numerically.

8C Dot plots and column graphs

Interactive

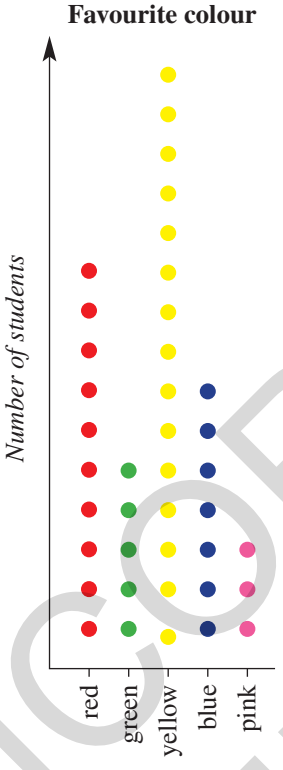
Widgets

HOTsheets

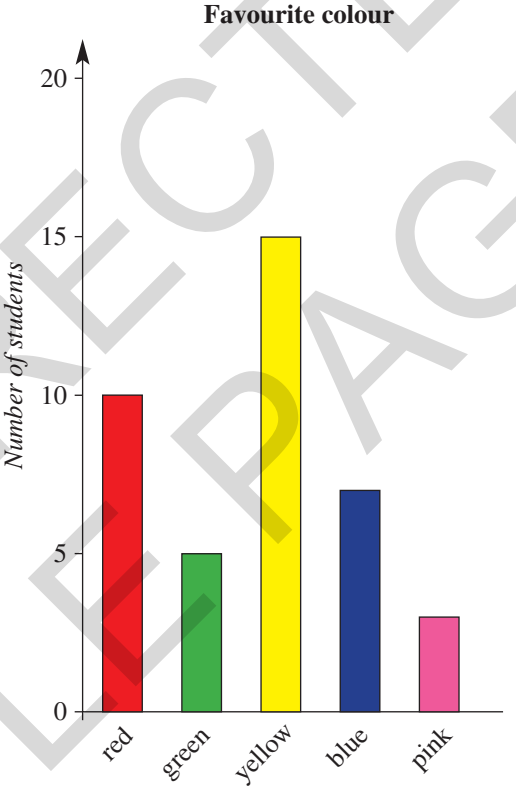
Walkthroughs

Numerical and categorical data can be shown graphically using **dot plots**, where each value is represented as a filled circle. More commonly, it is represented using **column graphs**, where the height of each column represents a number. Column graphs can be drawn vertically or horizontally.

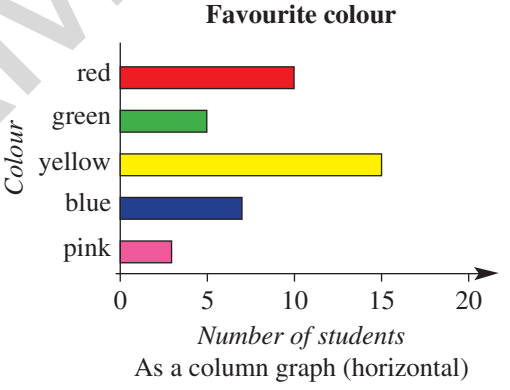
Consider a survey of students who are asked to choose their favourite colour from five possibilities. The results could be represented as a dot plot or as a column graph.



Represented as a dot plot



As a column graph (vertical)



As a column graph (horizontal)

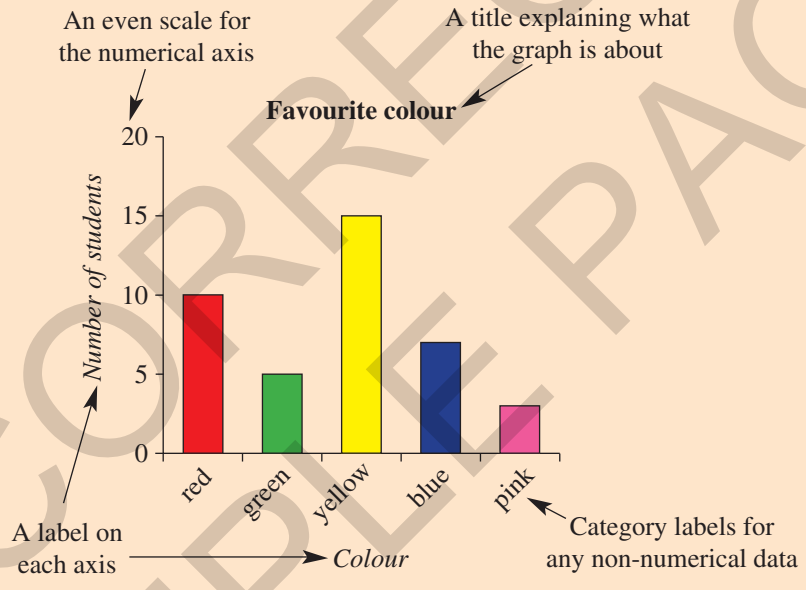
Let's start: Favourite colours

Survey the class to determine each student's favourite colour from the possibilities red, green, yellow, blue and pink.

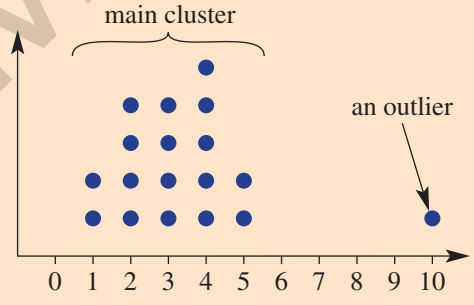
- Each student should draw a column graph or a dot plot to represent the results.
- What are some different ways that the results could be presented into a column graph? (There are more than 200 ways.)

Key ideas

- A **dot plot** can be used to display data, where each dot represents one **datum**.
- A **column graph** is an alternative way to show data in different categories, and is useful when more than a few items of data are present.
- Column graphs can be drawn vertically or horizontally.
- Graphs should have the following features:

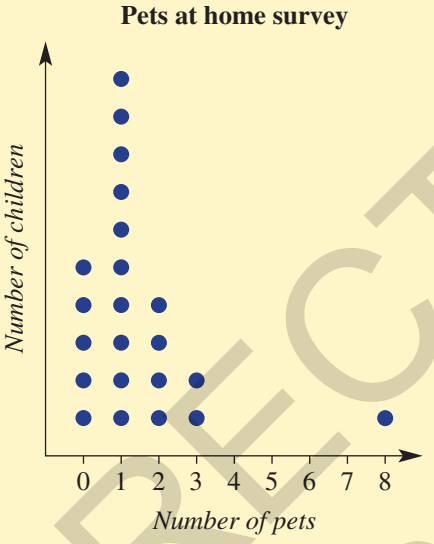


- Any numerical axis must be drawn to scale.
- An **outlier** is a value that is noticeably distinct from the main cluster of points.



Example 4 Interpreting a dot plot

The dot plot below represents the results of a survey that asked some children how many pets they have at home.



a Use the graph to state how many children have 2 pets.
b How many children participated in the survey?
c What is the range of values?
d What is the median number of pets?
e What is the outlier?
f What is the mode?

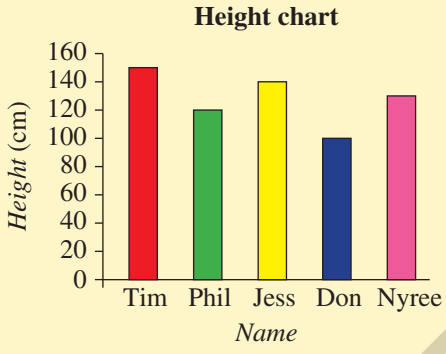
SOLUTION	EXPLANATION
a 4 children	There are 4 dots in the '2 pets' category, so 4 children have 2 pets.
b 22 children	The total number of dots is 22.
c 8 – 0 = 8	Range = highest – lowest In this case, highest = 8, lowest = 0.
d 1 pet	As there are 22 children, the median is the average of the 11th and 12th value. In this case, the 11th and 12th values are both 1.
e the child with 8 pets	The main cluster of children has between 0 and 3 pets, but the person with 8 pets is significantly outside this cluster.
f 1 pet	The most common number of pets is 1.

Example 5 Constructing a column graph

Draw a column graph to represent the following people's heights.

Name	Tim	Phil	Jess	Don	Nyree
Height (cm)	150	120	140	100	130

SOLUTION



EXPLANATION

First decide which scale goes on the vertical axis.

Maximum height = 150 cm, so axis goes from 0 cm to 160 cm (to allow a bit above the highest value).

Remember to include all the features required, including axes labels and a graph title.

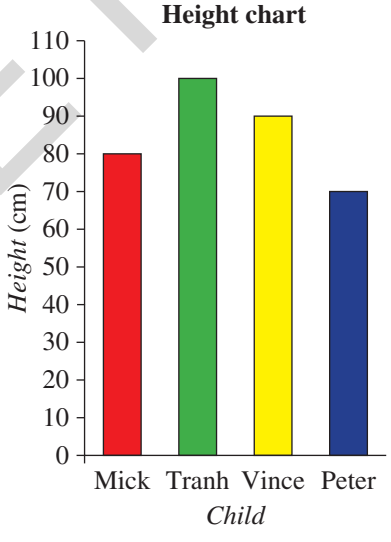
Exercise 8C

1, 2 2 —

UNDERSTANDING

1 The graph opposite shows the height of four boys.
Answer true or false to each of the following statements.

a Mick is 80 cm tall.
b Vince is taller than Tranh.
c Peter is the shortest of the four boys.
d Tranh is 100 cm tall.
e Mick is the tallest of the four boys.

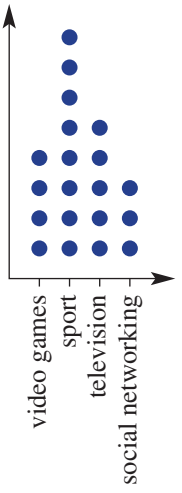


8C

Example 4a,b

UNDERSTANDING

2 The favourite after-school activity of a number of Year 7 students is recorded in the dot plot below.

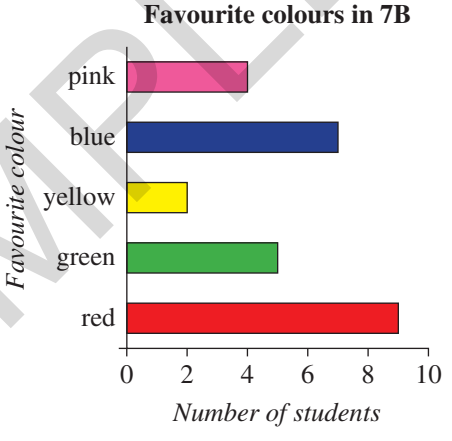


a How many students have chosen television as their favourite activity?
b How many students have chosen social networking as their favourite activity?
c What is the most popular after-school activity for this group of students?
d How many students participated in the survey?

3–7 | 3–8 | 4, 6–8

FLUENCY

3 From a choice of pink, blue, yellow, green or red, each student of Year 7B chose their favourite colour. The results are graphed below.



a How many students chose yellow?
b How many students chose blue?
c What is the most popular colour?
d How many students participated in the class survey?
e Represent these results as a dot plot.

Example 4c–f

4 In a Year 4 class, the results of a spelling quiz are presented as a dot plot.

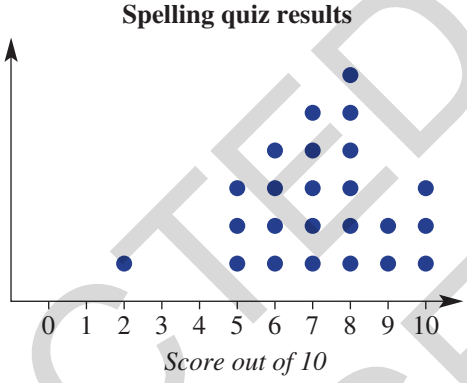
a What is the most common score in the class?

b How many students participated in the quiz?

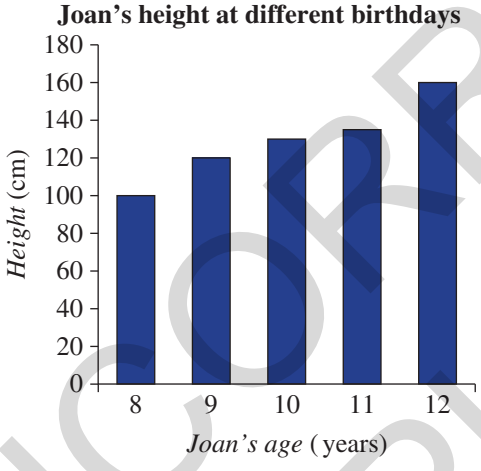
c What is the range of scores achieved?

d What is the median score?

e Identify the outlier.



5 Joan has graphed her height at each of her past five birthdays.



a How tall was Joan on her 9th birthday?

b How much did she grow between her 8th birthday and 9th birthday?

c How much did Joan grow between her 8th and 12th birthdays?

d How old was Joan when she had her biggest growth spurt?

Example 5

6 Draw a column graph to represent each of these boys' heights at their birthdays.

a Mitchell

Age (years)	Height (cm)
8	120
9	125
10	135
11	140
12	145

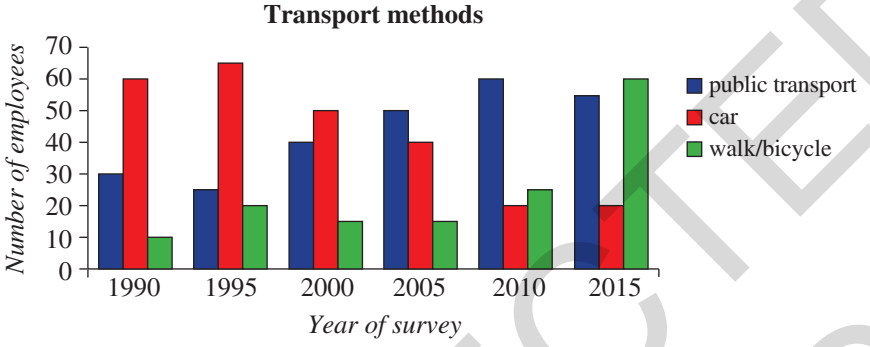
b Fatu

Age (years)	Height (cm)
8	125
9	132
10	140
11	147
12	150

8C

FLUENCY

7 Every five years, a company in the city conducts a transport survey of people's preferred method of getting to work in the mornings. The results are graphed below.



a Copy the following table into your workbook and complete it, using the graph.

	1990	1995	2000	2005	2010	2015
Use public transport	30					
Drive a car	60					
Walk or cycle	10					

b In which year(s) is public transport the most popular option?

c In which year(s) are more people walking or cycling to work than driving?

d Give a reason why the number of people driving to work has decreased.

e What is one other trend that you can see from looking at this graph?

8 **a** Draw a column graph to show the results of the following survey of the number of boys and girls born at a certain hospital. Put time (years) on the horizontal axis.

	2000	2001	2002	2003	2004	2005
Number of boys born	40	42	58	45	30	42
Number of girls born	50	40	53	41	26	35

b During which year(s) were there more girls born than boys?

c Which year had the fewest number of births?

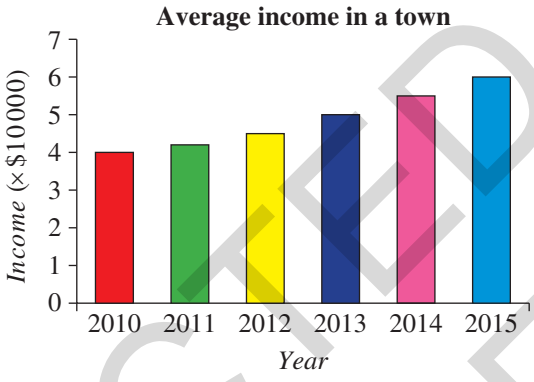
d Which year had the greatest number of births?

e During the entire period of the survey, were there more boys or girls born?

9 | 9, 10 | 10, 11 | 8C

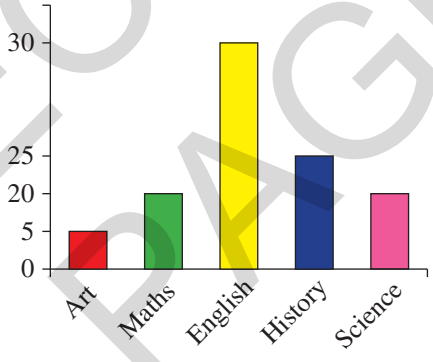
PROBLEM-SOLVING

9 The average (mean) income of adults in a particular town is graphed over a 6-year period.



- **a** Describe in one sentence what has happened to the income over this period of time.
- **b** Estimate what the income in this town might have been in 2009.
- **c** Estimate what the average income might be in 2025 if this trend continues.

10 A survey is conducted of students' favourite subjects from a choice of Art, Maths, English, History and Science. Someone has attempted to depict the results in a column graph.

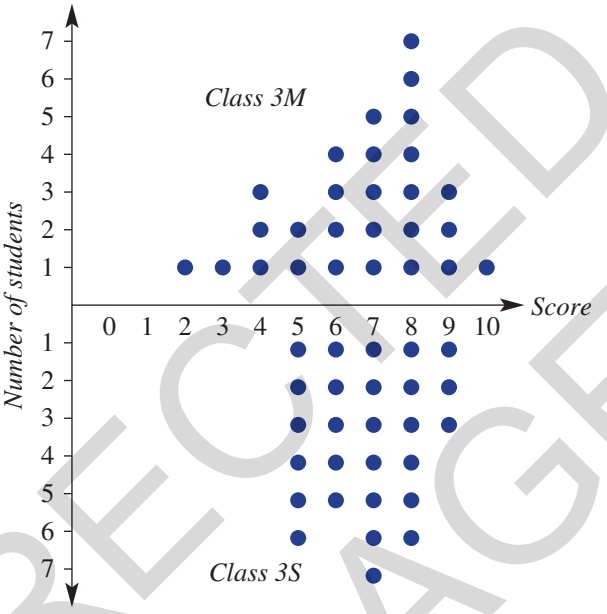


- **a** What is wrong with the scale on the vertical axis?
- **b** Give at least two other problems with this graph.
- **c** Redraw the graph with an even scale and appropriate labels.
- **d** The original graph makes Maths look twice as popular as Art, based on the column size. According to the survey, how many times more popular is Maths?
- **e** The original graph makes English look three times more popular than Maths. According to the survey, how many times more popular is English?
- **f** Assume that Music is now added to the survey's choice of subjects. Five students who had previously chosen History now choose Music, and 16 students who had previously chosen English now choose Music. What is the most popular subject now?

8C

PROBLEM-SOLVING

11 Mr Martin and Mrs Stevensson are the two Year 3 teachers at a school. For the latest arithmetic quiz, they have plotted their students' scores on a special dot plot called a parallel dot plot, shown below.

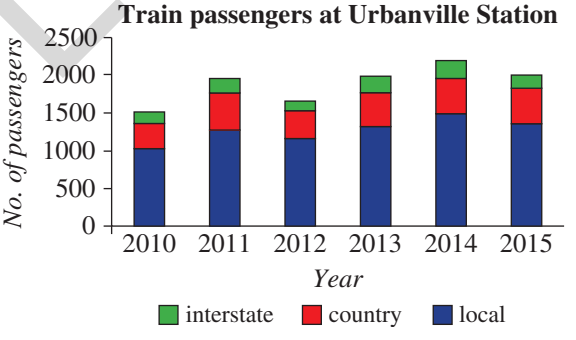


- **a** What is the median score for class 3M?
- **b** What is the median score for class 3S?
- **c** State the range of scores for each class.
- **d** Based on this test, which class has a greater spread of arithmetic abilities?
- **e** If the two classes competed in an arithmetic competition, where each class is allowed only one representative, which class is more likely to win? Justify your answer.

12 | 12, 13 | 12, 13

REASONING

12 At a central city train station, three types of services run – local, country and interstate. The average number of passenger departures during each week is shown in the stacked column graph.



- **a** Approximately how many passenger departures per week were there in 2010?
- **b** Approximately how many passenger departures were there *in total* during 2015?
- **c** Does this graph suggest that the total number of passenger departures has increased or decreased during the period 2010–2015?
- **d** Approximately how many passengers departed from this station in the period 2010–2015? Explain your method clearly and try to get your answer within 10 000 of the actual number.

8C

REASONING

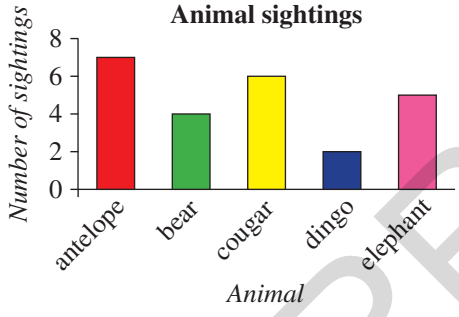
13 Explain why it is important to align dot points in a dot plot. Illustrate your explanation with two dot plots of the set of data below.

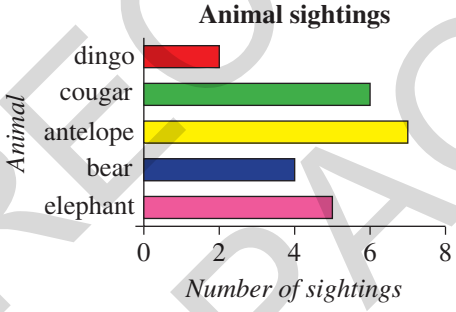
Activity	Netball	Dancing	Tennis	Chess
Number of students	5	3	2	4

How many ways? — — 14

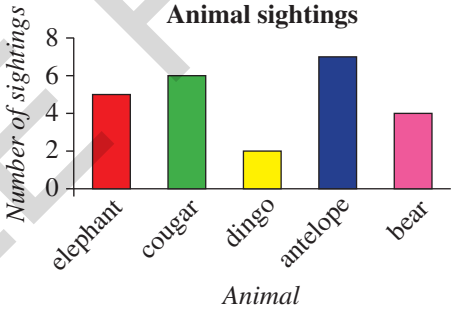
ENRICHMENT

14 As well as being able to draw a graph horizontally or vertically, the order of the categories can be changed. For instance, the following three graphs all represent the same data.





How many different column graphs could be used to represent the results of this survey? (Assume that you can only change the order of the columns, and the horizontal or vertical layout.) Try to list the options systematically to help with your count.

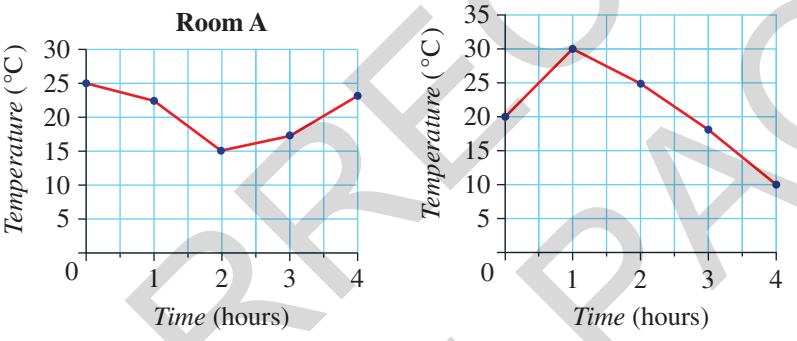


8D Line graphs

A **line graph** is a connected set of points joined with straight line segments. The variables on both axes should be continuous numerical data. It is often used when a measurement varies over time, in which case time is conventionally listed on the horizontal axis. One advantage of a line graph over a series of disconnected points is that it can often be used to estimate values that are unknown.

Let's start: Room temperature

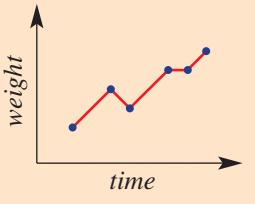
As an experiment, the temperature in two rooms is measured hourly over a period of time. The results are graphed below.



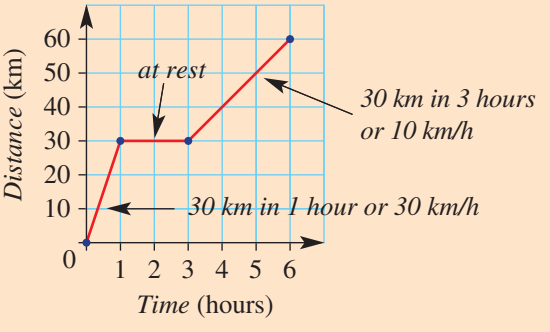
- Each room has a heater and an air conditioner to control the temperature. At what point do you think these were switched on and off in each room.
- For each room, what is the approximate temperature 90 minutes after the start of the experiment?
- What is the proportion of time that room A is hotter than room B?

Key ideas

- The variables on both axes should be **continuous** numerical data.
- A **line graph** consists of a series of points joined by straight line segments.
- Time is often shown on the horizontal axis.



- A common type of line graph is a travel graph.
  - Time is shown on the horizontal axis.
  - Distance is shown on the vertical axis.
  - The slope of the line indicates the rate at which the distance is changing over time. This is called **speed**.



Example 6 Drawing a line graph

The temperature in a room is noted at hourly intervals.

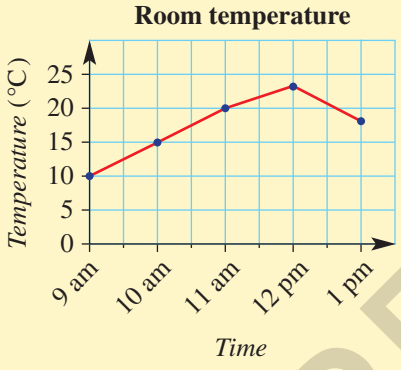
Time	9:00 am	10:00 am	11:00 am	12:00 pm	1:00 pm
Temperature (°C)	10	15	20	23	18

a Present the results as a line graph.

b Use your graph to estimate the room temperature at 12:30 pm.

SOLUTION

a



b About 20°C

EXPLANATION

- The vertical axis is from 0 to 25. The scale is even (i.e. increasing by 5 each time).
- Dots are placed for each measurement and joined with straight line segments.

By looking at the graph halfway between 12 pm and 1 pm an estimate is formed.

Example 7 Interpreting a travel graph

This travel graph shows the distance travelled by a cyclist over 5 hours.

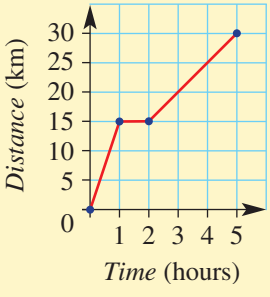
a How far did the cyclist travel in total?

b How far did the cyclist travel in the first hour?

c What is happening in the second hour?

d When is the cyclist travelling the fastest?

e In the fifth hour, how far does the cyclist travel?



SOLUTION

a 30 km

b 15 km

c at rest

d in the first hour

e 5 km

EXPLANATION

The right end point of the graph is at (5, 30).

At time equals 1 hour, the distance covered is 15 km.

The distance travelled does not increase in the second hour.

This is the steepest part of the graph.

In the last 3 hours, the distance travelled is 15 km, so in 1 hour, 5 km is travelled.

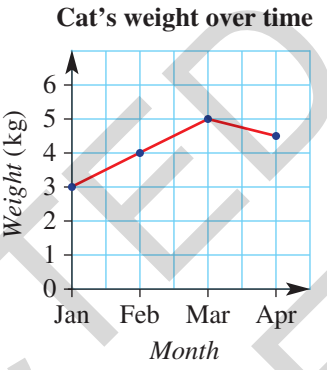
Exercise 8D

1, 2 | 1, 2 | —

UNDERSTANDING

1 The line graph shows the weight of a cat over a 3-month period. It is weighed at the start of each month. State the cat's weight at the start of:

- **a** January
- **b** February
- **c** March
- **d** April



Example 6a

2 A dog is weighed over a period of 3 months.
Draw a line graph of its weight.
January: 5 kg, February: 6 kg, March: 8 kg, April: 7 kg.

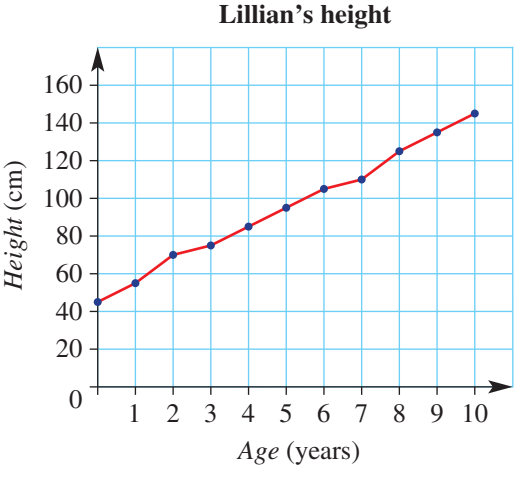
3–6 | 3–7 | 3, 5–7

FLUENCY

Example 6b

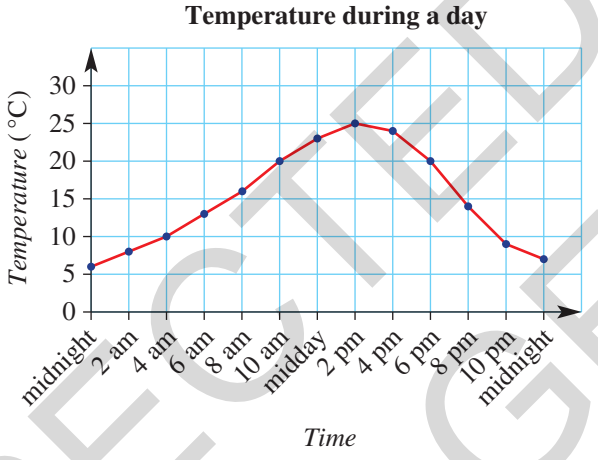
3 The graph shows Lillian's height over a 10-year period from when she was born.

- **a** What was Lillian's height when she was born?
- **b** What was Lillian's height at the age of 7 years?
- **c** At what age did she first reach 130 cm tall?
- **d** How much did Lillian grow in the year when she was 7 years old?
- **e** Use the graph to estimate her height at the age of $9\frac{1}{2}$ years.



4 Consider the following graph, which shows the outside temperature over a 24-hour period that starts at midnight.

- **a** What was the temperature at midday?
- **b** When was the hottest time of the day?
- **c** When was the coolest time of the day?
- **d** Use the graph to estimate the temperature at these times of the day:
  - **i** 4:00 am
  - **ii** 9:00 am
  - **iii** 1:00 pm
  - **iv** 3:15 pm



5 Oliver measures his pet dog's weight over the course of a year, by weighing it at the start of each month. He obtains the following results.

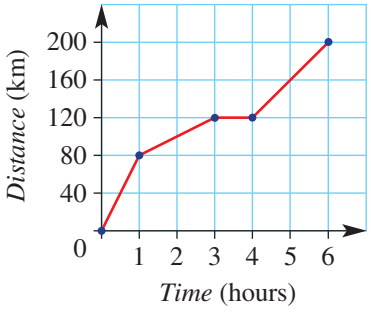
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Weight (kg)	7	7.5	8.5	9	9.5	9	9.2	7.8	7.8	7.5	8.3	8.5

- **a** Draw a line graph showing this information, making sure the vertical axis has an equal scale from 0 kg to 10 kg.
- **b** Describe any trends or patterns that you see.
- **c** Oliver put his dog on a weight loss diet for a period of 3 months. When do you think the dog started the diet? Justify your answer.

Example 7

6 This travel graph shows the distance travelled by a van over 6 hours.

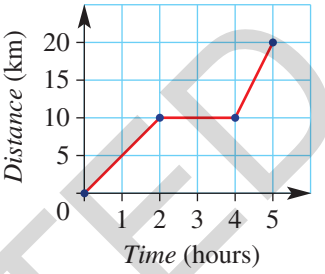
- **a** How far did the van travel in total?
- **b** How far did the van travel in the first hour?
- **c** What is happening in the fourth hour?
- **d** When is the van travelling the fastest?
- **e** In the sixth hour, how far does the van travel?



8D

FLUENCY

7 This travel graph shows the distance travelled by a bushwalker over 5 hours.

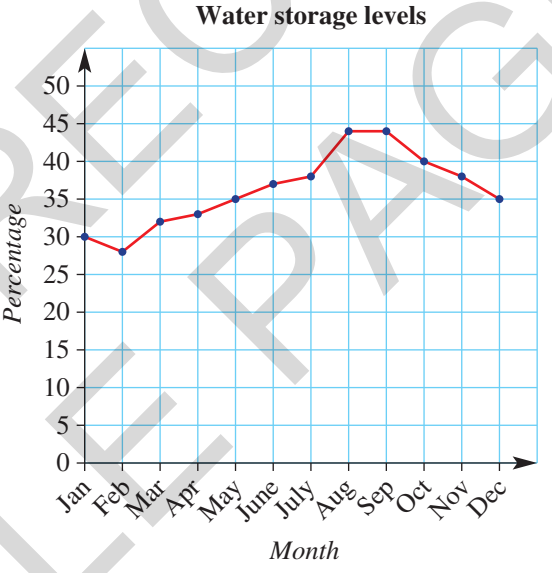


- **a** For how long was the bushwalker at rest?
- **b** How far did the bushwalker walk in the second hour?
- **c** During which hour did the bushwalker walk the fastest?

8 8 8, 9

PROBLEM-SOLVING

8 The water storage levels for a given city are graphed based on the percentage of water available on the first day of each month. For this question, assume that the amount of water used does not change from month to month.

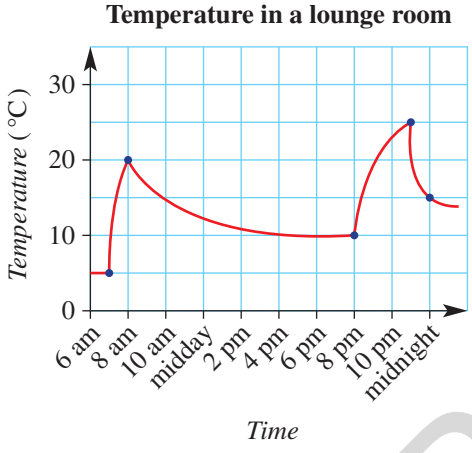


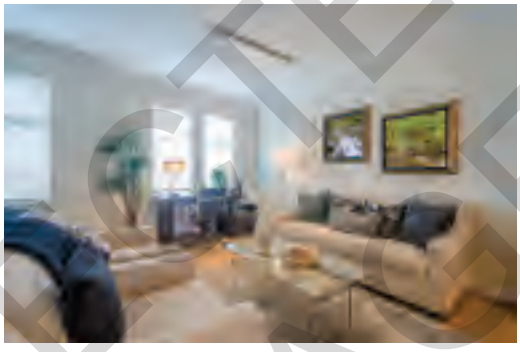
- **a** During which month did it rain the most in this city?
- **b** At what time(s) in the year is the water storage below 40%?
- **c** From August to September, it rained a total of 20 megalitres of water. How much water did the people in the city use during this period?
- **d** Is it more likely that this city is located in the Northern Hemisphere or the Southern Hemisphere? Justify your answer.

8D

PROBLEM-SOLVING

9 The temperature in a lounge room is measured frequently throughout a particular day. The results are presented in a line graph, as shown below. The individual points are not indicated on this graph to reduce clutter.





a Twice during the day the heating was switched on. At what times do you think this happened? Explain your reasoning.

b When was the heating switched off? Explain your reasoning.

c The house has a single occupant, who works during the day. Describe when you think that person is:

i waking up **ii** going to work **iii** coming home **iv** going to bed

d These temperatures were recorded during a cold winter month. Draw a graph that shows what the lounge room temperature might look like during a hot summer month. Assume that the room has an air conditioner, which the person is happy to use when at home.

10a 10 10

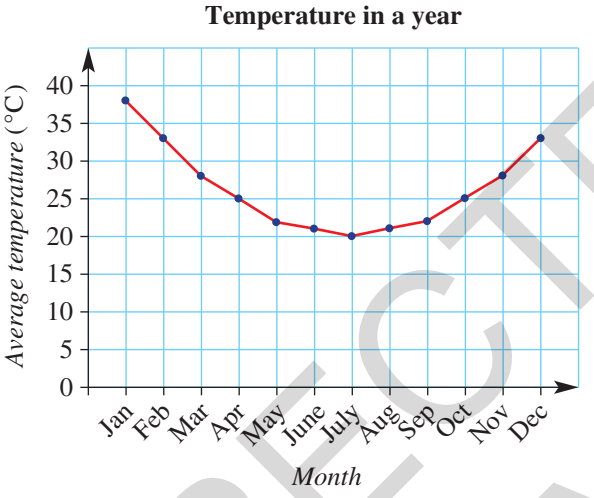
REASONING

10 Draw travel graphs to illustrate the following journeys.

a A car travels:
- 120 km in the first 2 hours
- 0 km in the third hour
- 60 km in the fourth hour
- 120 km in the fifth hour

b A jogger runs:
- 12 km in the first hour
- 6 km in the second hour
- 0 km in the third hour
- at a rate of 6 km per hour for 2 hours

11 The following line graph shows the average monthly temperature in a city.



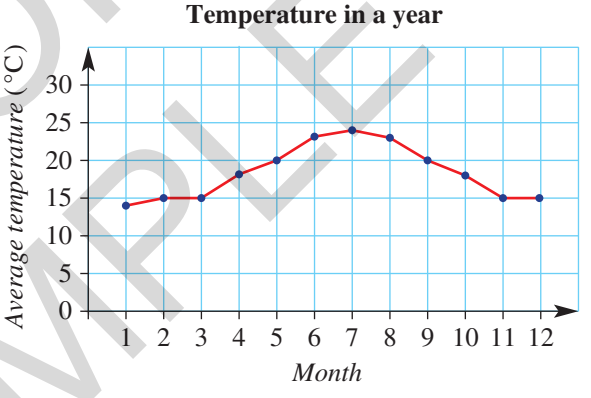
a Is this city in the Northern or Southern Hemisphere? Explain why.

b Is this city close to the equator or far from the equator? Explain why.

c Redraw the graph to start the 12-month period at July and finish in June.

d Describe how the new graph's appearance is different from the one shown above.

e In another city, somebody graphs the temperature over a 12-month period, as shown below. In which hemisphere is this city likely to be? Explain your answer.



8E Stem-and-leaf plots

Interactive

Widgets

HOTsheets

Walkthroughs

A stem-and-leaf plot is a useful way of presenting numerical data in a way that allows trends to be spotted easily. Each number is split into a stem (the first digit or digits) and a leaf (the last digit).

	Stem	Leaf
53 is	5	3
78 is	7	8
125 is	12	5

By convention, leaves are shown in increasing order as you work away from the stem, and stems are shown in increasing order going down the page.

The advantage of presenting data like this comes when multiple numbers have the same stem. For example, the list 122, 123, 124, 124, 127, 129 can be represented as shown opposite.

Stem	Leaf						
12	2	3	4	4	7	7	9

Let's start: Test score analysis

In a class, students' most recent test results out of 50 are recorded.

Test 1 results
43, 47, 50, 26, 38, 20, 25, 20, 50, 44,
33, 47, 47, 50, 37, 28, 28, 22, 21, 29

Test 2 results

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

- For each test, try to find how many students:
  - achieved a perfect score (i.e. 50)
  - failed the test (i.e. less than 25)
  - achieved a mark in the 40s.
- If there are 100 test results that you wish to analyse, would you prefer a list or a stem-and-leaf plot?
- What is it that makes a stem-and-leaf plot easier to work with? Discuss.

Key ideas

- A stem-and-leaf plot is a way to display numerical data.
- Each number is usually split into a **stem** (the first digit or digits) and a **leaf** (the last digit). For example:

	Stem	Leaf
The number 7 is	0	7
The number 31 is	3	1
The number 152 is	15	2

- Leaves should be aligned vertically, listed in ascending order as you move away from the stem.
- Any outliers can be identified by looking at the lowest value or highest value to see if they are far away from all the other numbers.

Example 8 Interpreting a stem-and-leaf plot

Average daily temperatures are shown for some different countries.

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

a Write out the temperatures as a list.
b How many countries' temperatures are represented?
c What are the minimum and maximum temperatures?
d What is the range of temperatures recorded?
e What is the median temperature recorded?

SOLUTION	EXPLANATION
a 13, 16, 16, 20, 20, 21, 22, 25, 25, 26, 28, 29, 30, 32	Each number is converted from a stem and a leaf to a single number. For example, $1\vert 3$ is converted to 13.
b 14	The easiest way is to count the number of leaves – each leaf corresponds to one country.
c minimum = 13 maximum = 32	The first stem and leaf is $1\vert 3$ and the last stem and leaf is $3\vert 2$.
d range = 19	Range = maximum – minimum = 32 – 13 = 19.
e median = 23.5	The middle value is halfway between the numbers $2\vert 2$ and $2\vert 5$, so median = $\frac{1}{2}(22 + 25) = 23.5$.

Example 9 Creating a stem-and-leaf plot

Represent this set of data as a stem-and-leaf plot:
23, 10, 36, 25, 31, 34, 34, 27, 36, 37, 16, 33

SOLUTION

Sorted: 10, 16, 23, 25, 27, 31, 33, 34, 34, 36, 36, 37

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

EXPLANATION

Sort the list in increasing order so that it can be put directly into a stem-and-leaf plot.

Split each number into a stem and a leaf. Stems are listed in increasing order and leaves are aligned vertically, listed in increasing order down the page.

Exercise 8E

1–3 | 1, 3 | —

UNDERSTANDING

1 The number 52 is entered into a stem-and-leaf plot.
- **a** What digit is the stem?
- **b** What digit is the leaf?

2 What number is represented by the following combinations?
- **i** 3|9
- **ii** 2|7
- **iii** 13|4

3 In this stem-and-leaf plot, the smallest number is 35. What is the largest number?

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

4–6 | 4, 5, 6–7(½) | 4, 5, 6–7(½)

FLUENCY

Example 8a–c

4 This stem-and-leaf plot shows the ages of people in a group.

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

- **a** Write out the ages as a list.
- **b** How many ages are shown?
- **c** Answer true or false to each of the following.
  - **i** The youngest person is aged 10.
  - **ii** Someone in the group is 17 years old.
  - **iii** Nobody listed is aged 20.
  - **iv** The oldest person is aged 4.

8E

FLUENCY

Example 8d,e

5 For each of the stem-and-leaf plots below, state the range and the median.

a

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

b

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

c

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

Example 9

6 Represent each of the following sets of data as a stem-and-leaf plot.

a 11, 12, 13, 14, 14, 15, 17, 20, 24, 28, 29, 31, 32, 33, 35

b 20, 22, 39, 45, 47, 49, 49, 51, 52, 52, 53, 55, 56, 58, 58

c 21, 35, 24, 31, 16, 28, 48, 18, 49, 41, 50, 33, 29, 16, 32

d 32, 27, 38, 60, 29, 78, 87, 60, 37, 81, 38, 11, 73, 12, 14

7 Represent each of the following datasets as a stem-and-leaf plot. (Remember: 101 is represented as 10|1.)

a 80, 84, 85, 86, 90, 96, 101, 104, 105, 110, 113, 114, 114, 115, 119

b 120, 81, 106, 115, 96, 98, 94, 115, 113, 86, 102, 117, 108, 91, 95

c 192, 174, 155, 196, 185, 178, 162, 157, 173, 181, 158, 193, 167, 192, 184, 187, 193, 165, 199, 184

d 401, 420, 406, 415, 416, 406, 412, 402, 409, 418, 404, 405, 391, 411, 413, 413, 408, 395, 396, 417

8, 9 | 9, 10 | 10–12 | 8E

PROBLEM-SOLVING

8 This back-to-back stem-and-leaf plot shows the ages of all the people in two shops. The youngest person in shop 1 is 15 (not 51).

Shop 1	Stem	Shop 2
5	1	6 7
7 7 5 3	2	4 5
	3	1
2	4	5

For each statement below, state whether it is true in shop 1 only (1), shop 2 only (2), both shops (B) or neither shop (N).

- **a** This shop has a 31-year-old person in it.
- **b** This shop has six people in it.
- **c** This shop has a 42-year-old person in it.
- **d** This shop has a 25-year-old person in it.
- **e** This shop has two people with the same age.
- **f** This shop has a 52-year-old person in it.
- **g** This shop has a 24-year-old person in it.
- **h** This shop's oldest customer is an outlier.
- **i** This shop's youngest customer is an outlier.

9 A company recorded the duration (in seconds) that visitors spent on its website's home page.

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

- **a** How many visitors spent less than 20 seconds on the home page?
- **b** How many visitors spent more than half a minute?
- **c** How many visitors spent between 10 and 30 seconds?
- **d** What is the outlier for this stem-and-leaf plot?
- **e** The company wishes to summarise its results with a single number. 'Visitors spend approximately ____ on our home page.' What number could it use?

8E

PROBLEM-SOLVING

10 Two radio stations poll their audience to determine their ages.

a Find the age difference between the oldest and youngest listener polled for:

i station 1

ii station 2

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

b One of the radio stations plays contemporary music that is designed to appeal to teenagers and the other plays classical music and broadcasts the news. Which radio station is most likely to be the one that plays classical music and news?

c Advertisers wish to know the age of the stations' audiences so that they can target their advertisements more effectively (e.g. to 38 to 58 year olds). Give a 20-year age range for the audience majority who listen to:

i station 1 **ii** station 2

11 A group of boys and girls have their heights recorded in a back-to-back stem-and-leaf plot, shown here.

a State the range of heights for:

i boys **ii** girls

b Which gender has a bigger range?

c State the median height for:

i boys **ii** girls

d Which gender has the larger median height?

e Which year level do you think these boys and girls are in? Justify your answer.

f Describe how you might expect this back-to-back stem-and-leaf plot to change if it recorded the heights of male and female Year 12 students.

Girls	Stem	Boys
	10	
	11	
6 3 1	12	
8 4 3 2	13	8
7 6 5 4 0	14	3 4 7 9
6 4 1	15	0 1 2 4 6
2 0	16	2 3 6 8
3	17	

8E

PROBLEM-SOLVING

12 A teacher has compiled her students' recent test scores out of 50 as a stem-and-leaf plot. However, some values are missing, as represented by the letters a, b, c and d.

Stem	Leaf
1	5
2	4 5 a 6 7 9
3	b 0 1 5
4	2 8 c
5	d

a How many students took the test?

b How many students passed the test (i.e. achieved a mark of 25 or higher)?

c State the possible values for each of the missing digits a to d.

13 13, 14 13, 14

REASONING

13 **a** Explain why it is important that leaves are aligned vertically. (Hint: Consider how the overall appearance could be helpful with a large dataset.)

b Why might it be important that data values are sorted in stem-and-leaf plots?

14 A stem-and-leaf plot is constructed showing the ages of all the people who attended a local farmer's market at a certain time of the day. However, the plot's leaves cannot be read.

Stem	Leaf
1	?
2	?
3	? ? ? ? ? ? ? ? ? ?
4	? ? ? ? ? ? ? ? ? ? ? ? ? ? ?
5	? ? ?

a For each of the following, either determine the exact answer or give a range of values the answer could take.

i How many people were at the market?

ii How many people aged in their 30s were at the market?

iii How old is the youngest person?

iv What is the age difference between the youngest and oldest person?

v How many people aged 40 or over were at the market?

vi How many people aged 35 or over were at the market?

8E

REASONING

b Classify each of the following as true or false.
- **i** The majority of people at the market were aged in their 30s or 40s.
- **ii** There were five teenagers present.
- **iii** Exactly two people were aged 29 years or under.
- **iv** Two people in their 40s must have had the same age.
- **v** Two people in their 30s must have had the same age.
- **vi** Two people in their 20s could have had the same age.

c Explain why it is possible to determine how many people were aged 40 or over, but not the number of people who are aged 40 or under.

d It is discovered that the person under 20 years of age is an outlier for this market. What does that tell you about how old the next oldest person is?

Negative stem-and-leaf plots — — 15

ENRICHMENT

15 Negative numbers can also be displayed in stem-and-leaf plots. This stem-and-leaf plot gives the average winter temperatures in 15 different cities.

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

a What are the minimum and maximum temperatures listed?

b Find how many cities had average temperatures:
- **i** between –10°C and 10°C
- **ii** between –25°C and 5°C
- **iii** below 5°C

c Why is there a 0 row and a –0 row, even though 0 and –0 are the same number?

d What is the average (or mean) of all the listed temperatures in the 15 cities? Give your answer correct to 1 decimal place.

e What is the median of all the listed temperatures? Compare this to the mean found in part **d**.

8F Pie charts and divided bar graphs EXTENDING

Interactive

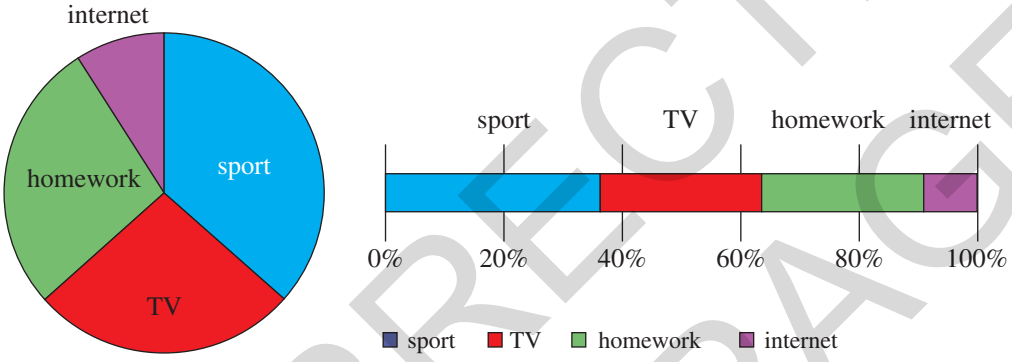
Widgets

HOTsheets

Walkthroughs

A pie chart (also called sector graph) consists of a circle divided into different sectors or 'slices of pie', where the size of each sector indicates the proportion occupied by any given item. A divided bar graph is a rectangle divided into different rectangles or 'bars', where the size of each rectangle indicates the proportion of each item.

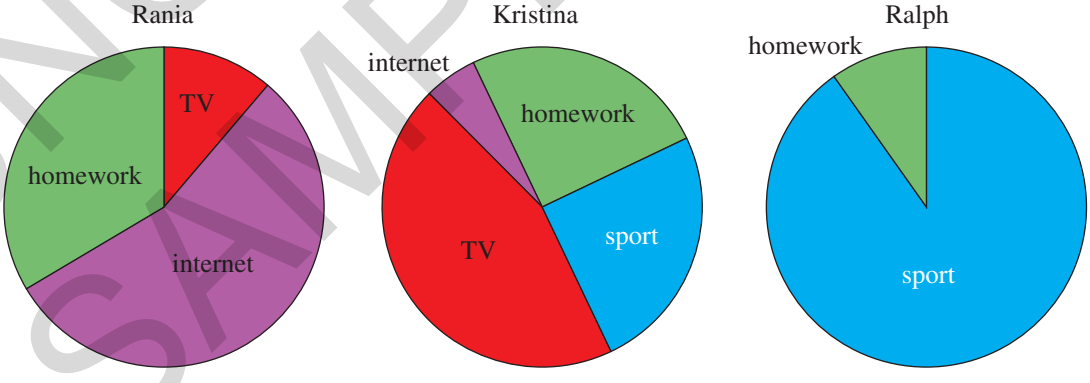
If a student is asked to describe how much time they spend each evening doing different activities, they could present their results as either type of graph:



From both graphs, it is easy to see that most of the student's time is spent playing sport and the least amount of time is spent using the internet.

Let's start: Student hobbies

Rania, Kristina and Ralph are asked to record how they spend their time after school. They draw the following graphs.



- Based on these graphs alone, describe each student in a few sentences.
- Justify your descriptions based on the graphs.

Key ideas

- To calculate the size of each section of the graph, divide the value in a given category by the sum of all category values. This gives the category's proportion or fraction.
- To draw a **pie chart** (also called a **sector graph**), multiply each category's proportion or fraction by 360° and draw a sector of that size.
- To draw a **divided bar graph**, multiply each category's proportion or fraction by the total width of the rectangle and draw a rectangle of that size.

Example 10 Drawing a pie chart and a divided bar graph

On a particular Saturday, Sanjay measures the number of hours he spends on different activities.

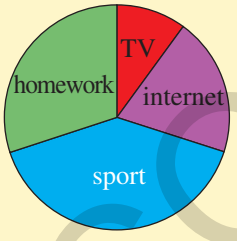
Television	Internet	Sport	Homework
1 hour	2 hours	4 hours	3 hours

Represent the table as:

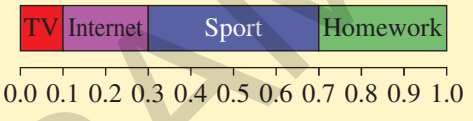
a a pie chart **b** a divided bar graph

SOLUTION

a



b



EXPLANATION

The total amount of time is 1 + 2 + 4 + 3 = 10 hours. Then we can calculate the proportions and sector sizes:

Category	Proportion	Sector size (°)
Television	$\frac{1}{10}$	$\frac{1}{10} \times 360 = 36$
Internet	$\frac{2}{10} = \frac{1}{5}$	$\frac{1}{5} \times 360 = 72$
Sport	$\frac{4}{10} = \frac{2}{5}$	$\frac{2}{5} \times 360 = 144$
Homework	$\frac{3}{10}$	$\frac{3}{10} \times 360 = 108$

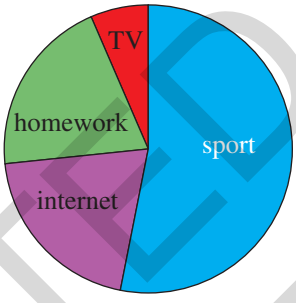
Using the same proportions calculated above, make sure that each rectangle takes up the correct amount of space. For example, if the total width is 15 cm, then sport occupies $\frac{2}{5} \times 15 = 6$ cm.

Exercise 8F

1, 2 | 2 | —

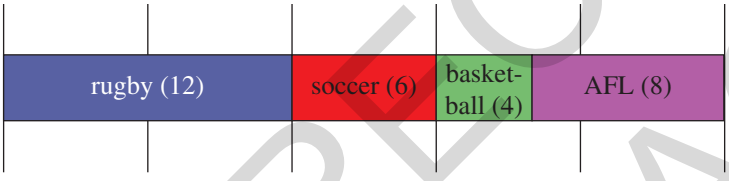
UNDERSTANDING

1 Jasna graphs a pie chart of how she spends her leisure time.



- **a** What does Jasna spend the most time doing?
- **b** What does Jasna spend the least time doing?
- **c** Does she spend more or less than half of her time playing sport?

2 Thirty students are surveyed to find out their favourite sport and their results are graphed below.



- **a** What is the most popular sport for this group of students?
- **b** What is the least popular sport for this group of students?
- **c** What fraction of the students has chosen soccer as their favourite sport?
- **d** What fraction of the students has chosen either rugby or AFL?

3, 4 | 3–5 | 3–5

FLUENCY

Example 10a

3 A group of passengers arriving at an airport is surveyed to establish which countries they have come from. The results are presented below.

Country	China	United Kingdom	USA	France
No. of passengers	6	5	7	2

- **a** What is the total number of passengers who participated in the survey?
- **b** What proportion of the passengers surveyed have come from the following countries? Express your answer as a fraction.
  - **i** China **ii** United Kingdom **iii** USA **iv** France
- **c** On a pie chart, determine the angle size of the sector representing:
  - **i** China **ii** United Kingdom **iii** USA **iv** France
- **d** Draw a pie chart showing the information calculated in part **c**.

8F

Example 10b

4 A group of students in Years 7 and 8 is polled on their favourite colour, and the results are shown at right.

Colour	Year 7 votes	Year 8 votes
Red	20	10
Green	10	4
Yellow	5	12
Blue	10	6
Pink	15	8

- **a** Draw a pie chart to represent the Year 7 colour preferences.
- **b** Draw a different pie chart to represent the Year 8 colour preferences.
- **c** Describe two differences between the charts.
- **d** Construct a divided bar graph that shows the popularity of each colour across the total number of Years 7 and 8 students combined.

5 Consider the following results of a study on supermarket shopping habits.

Items	Food	Drinks	Household items	Other
Proportion of money spent	50%	25%	20%	5%

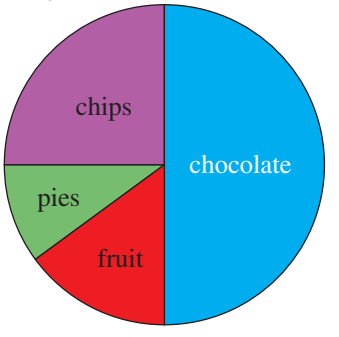
- **a** Represent this information in a divided bar graph.
- **b** Graph this information as a sector graph.

FLUENCY

6 6, 7 7, 8

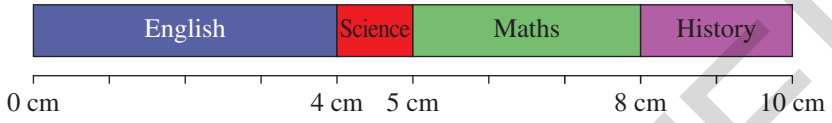
PROBLEM-SOLVING

6 A group of Year 7 students was polled on their favourite foods, and the results are shown in this pie chart.



- **a** If 40 students participated in the survey, find how many of them chose:
  - **i** chocolate
  - **ii** chips
  - **iii** fruit
  - **iv** pies
- **b** Health experts are worried about what these results mean. They would like fruit to appear more prominently in the pie graph, and to not have the chocolate sector next to the chips. Redraw the pie chart so this is the case.
- **c** Another 20 students were surveyed. Ten of these students chose chocolate and the other 10 chose chips. Their results are to be included in the pie graph. Of the four sectors in the graph, state which sector will:
  - **i** increase in size
  - **ii** decrease in size
  - **iii** stay the same size

7 Yakob has asked his friends what is their favourite school subject, and he has created the following divided bar graph from the information.



a If Yakob surveyed 30 friends, state how many of them like:

i Maths best **ii** History best **iii** either English or Science best

b Redraw these results as a pie chart.

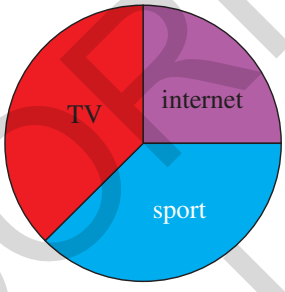
8 Friends Krishna and Nikolas have each graphed their leisure habits, as shown below.

a Which of the two friends spends more of their time playing sport?

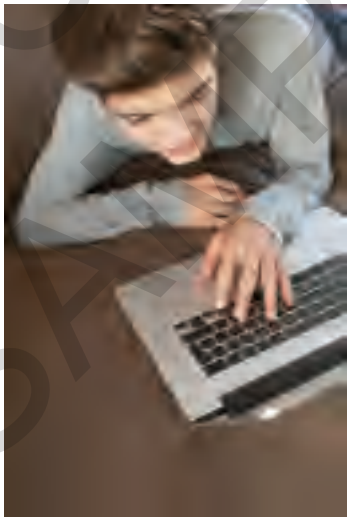
b Which of the two friends does more intellectual activities in their leisure time?

c Krishna has only 2 hours of leisure time each day because he spends the rest of his time doing homework. Nikolas has 8 hours of leisure time each day. How does this affect your answers to parts **a** and **b** above?

Krishna's leisure time



Nikolas' leisure time

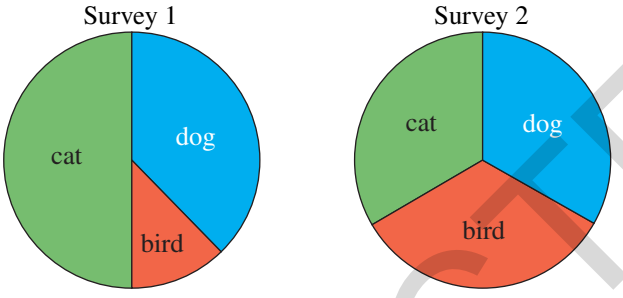


8F

9a–c | 9a–d | 9

REASONING

9 In two surveys, people were asked what is their favourite pet animal.



a If 16 people participated in survey 1, how many chose dog?

b If 30 people participated in survey 2, how many chose bird?

c Jason claims that 20 people participated in survey 1. Explain clearly why this cannot be true.

d Jaimee claims that 40 people participated in survey 2. Explain clearly why this cannot be true.

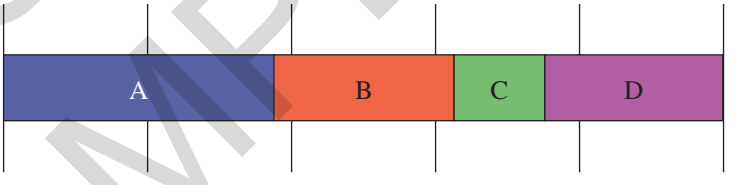
e In actual fact, the same number of people participated for each survey. Given that fewer than 100 people participated, how many participants were there? Give all the possible answers.

Rearranging graphs

— | — | 10

ENRICHMENT

10 Consider the divided bar graph shown below.



a Show how this graph will look if the segments are placed in the order C, D, A, B (from left to right).

b In how many different ways could this divided bar graph be drawn (counting ABCD and CDAB as two different ways)?

c If this bar graph is redrawn as a pie chart, how many ways could the segments be arranged? Try to list them systematically. Do not consider two pie charts to be different if one is just a rotation of another.

Progress quiz

8A **1** Classify the following variables as categorical, discrete numerical or continuous numerical.

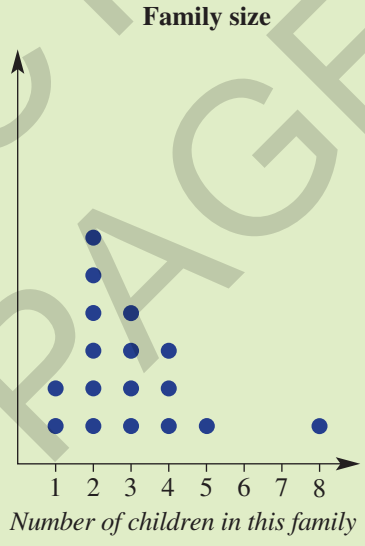
a eye colour
b animal weight (kg)
c number of siblings
d time to run 100 metres (sec)

8B **2** For each of the following sets of data, calculate the:

i range **ii** mean **iii** median **iv** mode

a 5, 12, 3, 8, 2, 9, 3
b 15, 24, 22, 28, 16, 15

8C **3** This dot plot represents the number of children in each family of some Year 7 students.



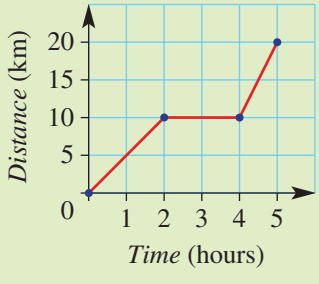
a What is the most common family size in this class? (family size means the number of children)
b How many families are shown by this graph, assuming there are no siblings within the class?
c What is the range of family sizes?
d What is the median family size?
e Identify the outlier.

8D **4** The temperature outside a classroom was recorded four times during one school day. The following results were obtained.

Time	9 am	11 am	1 pm	3 pm
Temperature	15°C	20°C	28°C	25°C

a Draw a line graph showing this information.
b Use your graph to estimate the temperature at noon.

8D **5** This travel graph shows the distance travelled by a cyclist over 5 hours.



a How far did the cyclist travel in total?
b How far had the cyclist travelled after 3 hours?
c What is happening in the fourth hour?
d In the fifth hour, how far did the cyclist travel?
e During what hour was the cyclist travelling the fastest?

8E **6** This stem-and-leaf table shows the noon temperatures (in °C) of different towns around Australia on one particular day.

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

a How many towns have their temperatures listed in this stem-and-leaf table?
b What is the maximum and minimum noon temperature recorded?
c What is the range of temperatures recorded?
d What is the median temperature recorded?

8E **7** Represent this set of data as a stem-and-leaf plot: 10, 21, 16, 18, 7, 19, 18, 9, 20, 12

8F **8** Some Year 7 students were asked how they travelled to school. The results are shown in this table.

Ext

Public transport	Bicycle	Car
14	2	4

a Represent the data as a pie chart.
b Represent the data as a divided bar graph of total length 15 cm.

8G Describing chance

CONSOLIDATING

Often, there are times when you may wish to describe how likely it is that an event will occur. For example, you may want to know how likely it is that it will rain tomorrow, or how likely your sporting team will win this year's premiership, or how likely it is that you will win a lottery. Probability is the study of chance.

The probability of winning first price in a lottery is very, very low.

Let's start: Likely or unlikely?

Try to rank these events from least likely to most likely. Compare your answers with other students in the class and discuss any differences.

- It will rain tomorrow.
- Australia will win the soccer World Cup.
- Tails landing uppermost when a 20-cent coin is tossed.
- The Sun will rise tomorrow.
- The king of spades is at the top of a shuffled deck of 52 playing cards.
- A diamond card is at the bottom of a shuffled deck of 52 playing cards.

Key ideas

- When using the English language to describe chance, there are a number of phrases that can be used.

100%	certain	
	likely	more likely
50%	even chance	
	unlikely	less likely
0%	impossible	

- If two events have the same chance of occurring, then you say that it is **equally likely** they will occur.

Example 11 Describing chance

Classify each of the following statements as either true or false.

a It is likely that children will go to school next year.

b It is an even chance for a fair coin to display tails.

c Rolling a 3 on a 6-sided die and getting heads on a coin are equally likely.

d It is certain that two randomly chosen odd numbers will add to an even number.

SOLUTION	EXPLANATION
a true	Although there is perhaps a small chance that the laws might change, it is (very) likely that children will go to school next year.
b true	There is a 50-50, or an even chance, of a fair coin displaying tails. It will happen, on average, half of the time.
c false	These events are not equally likely. It is more likely to flip heads on a coin than to roll a 3 on a 6-sided die.
d true	No matter what odd numbers are chosen, they will always add to an even number.

Exercise 8G

1, 2 | 1 | —

UNDERSTANDING

1 Match each of the events a to d with a description of how likely they are to occur (A to D).

a A tossed coin landing heads up.
b Selecting an ace first try from a fair deck of 52 playing cards.
c Obtaining a number other than 6 if a fair 6-sided die is rolled.
d Obtaining a number greater than 8 if a fair 6-sided die is rolled.

A unlikely
B likely
C impossible
D even chance

2 Fill in the blanks, using the appropriate terminology.

a If an event is guaranteed to occur, we say it is _____.
b An event that is equally likely to occur or not occur has an _____ _____.
c A rare event is considered _____.
d An event that will never occur is called _____.

Example 11

3 Consider a fair 6-sided die with the numbers 1 to 6 on it. Answer true or false to each of the following.

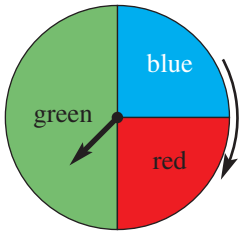
a Rolling a 3 is unlikely.
b Rolling a 5 is likely.
c Rolling a 4 and rolling a 5 are equally likely events.
d Rolling an even number is likely.
e There is an even chance of rolling an odd number.
f There is an even chance of rolling a multiple of 3.

4 Match up each of the events **a** to **d** with an equally likely event **A** to **D**.

a Rolling a 2 on a 6-sided die
b Selecting a heart card from a fair deck of 52 playing cards
c Flipping a coin and tails landing face up
d Rolling a 1 or a 5 on a 6-sided die

A Selecting a black card from a fair deck of 52 playing cards
B Rolling a number bigger than 4 on a 6-sided die
C Selecting a diamond card from a fair deck of 52 playing cards
D Rolling a 6 on a 6-sided die

5 Consider the spinner shown, which is spun and could land with the arrow pointing to any of the three colours. (If it lands on a boundary, it is re-spun until it lands on a colour.)



a State whether each of the following is true or false.
i There is an even chance that the spinner will point to green.
ii It is likely that the spinner will point to red.
iii It is certain that the spinner will point to purple.
iv It is equally likely that the spinner will point to red or blue.
v Green is twice as likely to occur as blue.

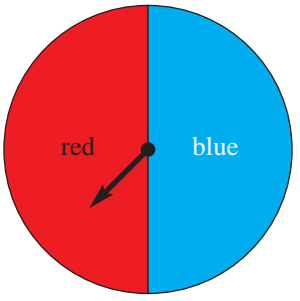
b Use the spinner to give an example of:
i an impossible event
ii a likely event
iii a certain event
iv two events that are equally likely

8G

6 6, 7 7, 8

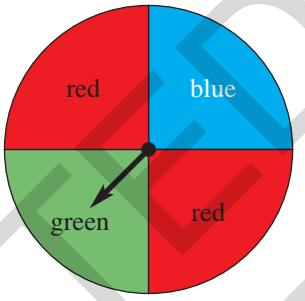
PROBLEM-SOLVING

6 Three spinners are shown below. Match each spinner with the description.



spinner 1

spinner 2



spinner 3

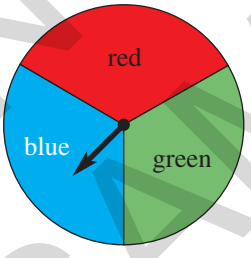
- **a** Has an even chance of red, but blue is unlikely.
- **b** Blue and green are equally likely, but red is unlikely.
- **c** Has an even chance of blue, and green is impossible.

7 Draw spinners to match each of the following descriptions, using blue, red and green as the possible colours.

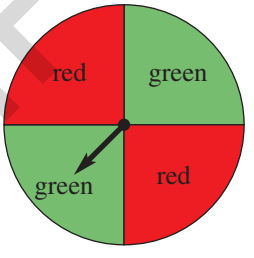
- **a** Blue is likely, red is unlikely and green is impossible.
- **b** Red is certain.
- **c** Blue has an even chance, red and green are equally likely.
- **d** Blue, red and green are all equally likely.
- **e** Blue is twice as likely as red, but red and green are equally likely.
- **f** Red and green are equally likely and blue is impossible.
- **g** Blue, red and green are all unlikely, but no two colours are equally likely.
- **h** Blue is three times as likely as green, but red is impossible.

8 For each of the following spinners, give a description of the chances involved so that someone could determine which spinner is being described. Use the colour names and the language of chance (i.e. 'likely', 'impossible' etc.) in your descriptions.

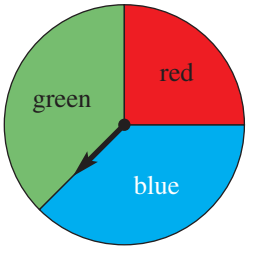
a



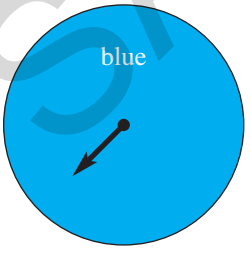
b



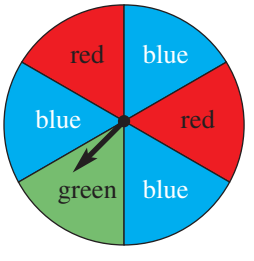
c



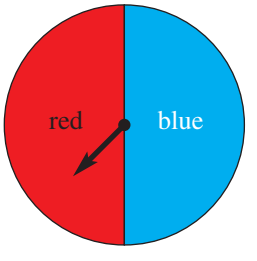
d



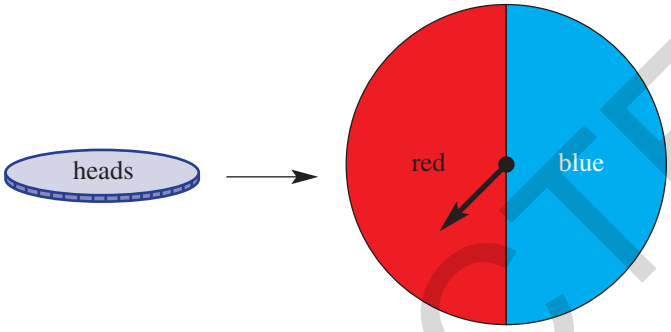
e



f



9 A coin consists of two sides that are equally likely to occur when tossed. It is matched up with a spinner that has exactly the same chances, as shown below.



Tossing the coin with heads landing uppermost is equally likely to spinning red on the spinner. Tossing the coin with tails landing uppermost is equally likely to spinning blue on the spinner. Hence, we say that the coin and the spinner are **equivalent**.

a Draw a spinner that is equivalent to a fair 6-sided die. (Hint: The spinner should have six sections of different colours.)

b How can you tell from the spinner you drew that it is equivalent to a fair die?

c A die is 'weighted' so that there is an even chance of rolling a 6, but rolling the numbers 1 to 5 are still equally likely. Draw a spinner that is equivalent to such a die.

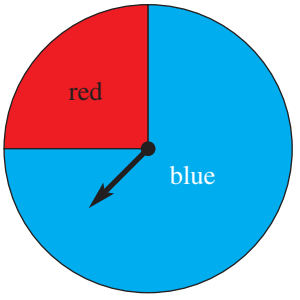
d How could you make a die equivalent to the spinner shown in the diagram?

e Describe a spinner that is equivalent to selecting a card from a fair deck of 52 playing cards.

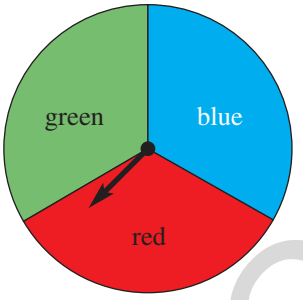
8G Spinner proportions — — 10

ENRICHMENT

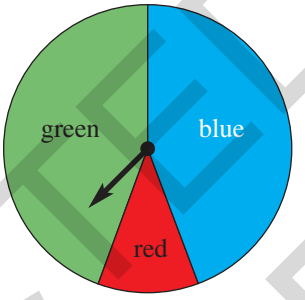
10 The language of chance is a bit vague. For example, for each of the following spinners it is 'unlikely' that you will spin red, but in each case the chance of spinning red is different.



spinner 1



spinner 2



spinner 3

Rather than describing this in words we could give the fraction (or decimal or percentage) of the spinner occupied by a colour.

a For each of the spinners above, give the fraction of the spinner occupiedby red.

b What fraction of the spinner would be red if it had an even chance?

c Draw spinners for which the red portion occupies:

i 100% of the spinner **ii** 0% of the spinner

d For the sentences below, fill in the gaps with appropriate fraction or percentage values.

i An event has an even chance of occurring if that portion of the spinner occupies _____ of the total area.

ii An event that is impossible occupies _____ of the total area.

iii An event is unlikely to occur if it occupies more than _____ but less than _____ of the total area.

iv An event is likely if it occupies more than _____ of the total area.

e How can the fractions help determine if two events are equally likely?

f Explain why all the fractions occupied by a colour must be between 0 and 1.

The game of 'Twister' uses a spinner (being held by the girl at the back) to determine the positions to be attempted.

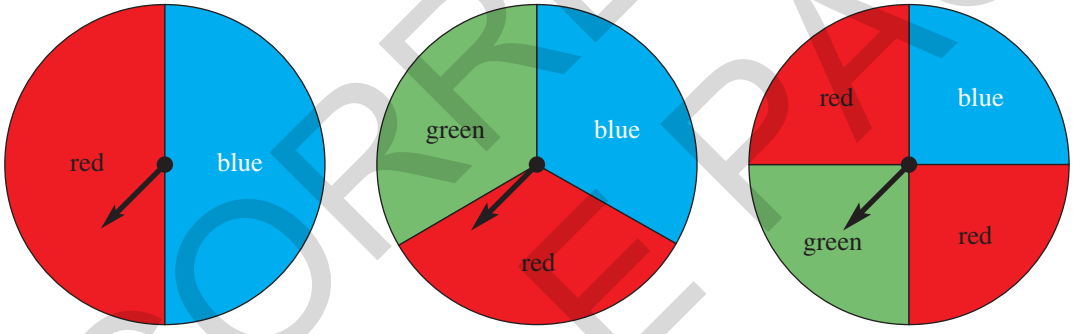
8H Theoretical probability

The **probability** of an event occurring is a number between 0 and 1. This number states precisely how likely it is for an event to occur. It is often written as a fraction and can indicate how frequently the event would occur over a large number of trials. For example, if you toss a fair coin many times, you would expect heads to come up half the time, so the probability is $\frac{1}{2}$. If you roll a fair 6-sided die many times, you should roll a 4 about one-sixth of the time, so the probability is $\frac{1}{6}$.

To be more precise, we should list the possible outcomes of rolling the die: 1, 2, 3, 4, 5, 6. Doing this shows us that there is a 1 out of 6 chance that you will roll a 4 and there is a 0 out of 6 (= 0) chance of rolling a 9.

Let's start: Spinner probabilities

Consider the three spinners shown below.



- What is the probability of spinning blue for each of these spinners?
- What is the probability of spinning red for each of these spinners?
- Try to design a spinner for which the probability of spinning green is $\frac{4}{7}$ and the probability of spinning blue is 0.

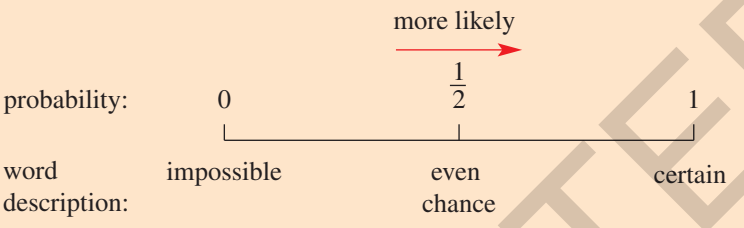
Key ideas

- An **experiment** or **trial** could be tossing a coin, rolling a die or spinning a spinner.
- An **outcome** is a possible result of the experiment, like rolling a 5 or a coin showing tails.
- An **event** is either a single outcome (e.g. rolling a 3) or a collection of outcomes (e.g. rolling a 3, 4 or 5).
- The **probability** of an event is a number between 0 and 1 inclusive, to represent the chance of the event occurring. The probability of an event occurring, if all the outcomes are equally likely, is:

$$\frac{\text{number of outcomes where the event occurs}}{\text{total number of outcomes}}$$

Key ideas

- Probability is often written as a fraction, but it can be written as a decimal or as a percentage.
- The **sample space** is the set of all possible outcomes of an experiment.
- We write Pr(green) to mean 'the probability that a spinner shows green'.



Example 12 Calculating probability

A fair 6-sided die is rolled.

a List the sample space.

b Find the probability of rolling a 3, giving your answer as a fraction.

c Find the probability of rolling an even number, giving your answer as a decimal.

d Find the probability of rolling a number less than 3, giving your answer as a percentage.

SOLUTION	EXPLANATION
a sample space = $\{1, 2, 3, 4, 5, 6\}$	For the sample space, we list all the possible outcomes. Technically, the sample space is {roll a 1, roll a 2, roll a 3, roll a 4, roll a 5, roll a 6}, but we do not usually include the additional words.
b $\Pr(3) = \frac{1}{6}$	The event can occur in one way (rolling a 3) out of six possible outcomes.
c $\Pr(\text{even}) = \frac{1}{2} = 0.5$	The event can occur in three ways (i.e. 2, 4 or 6). So the probability is $\frac{3}{6} = \frac{1}{2}$. As a decimal this is 0.5.
d $\Pr(\text{less than } 3) = \frac{1}{3} = 33.3\%$	The event can occur in two ways (1 or 2). So the probability is $\frac{2}{6} = \frac{1}{3}$. As a percentage this is 33.3%, rounded to 1 decimal place.

Exercise 8H

1–3 | 3 | —

UNDERSTANDING

1 Match up each experiment **a** to **d** with the list of possible outcomes **A** to **D**.

a tossing a coin	**A** 1, 2, 3, 4, 5, 6
b rolling a die	**B** red, white, blue
c selecting a suit from a fair deck of 52 playing cards	**C** heads, tails
d choosing a colour on the French flag	**D** hearts, diamonds, clubs, spades

The French flag is divided into three sections of different colours and equal size.

2 Complete the following sentences.

a The _____ _____ is the set of possible outcomes.

b An impossible event has a probability of _____.

c If an event has a probability of 1, then it is _____.

d The higher its probability, the _____ likely the event will occur.

e An event with a probability of $\frac{1}{2}$ has an ____ ____ of occurring.

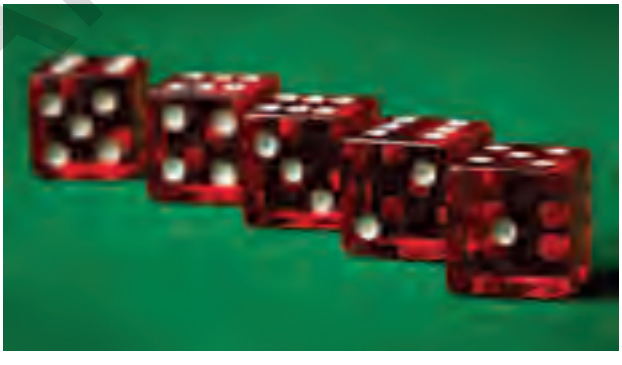
Example 12a

3 Consider a fair 6-sided die.

a List the sample space.

b List the odd numbers on the die.

c State the probability of throwing an even number.



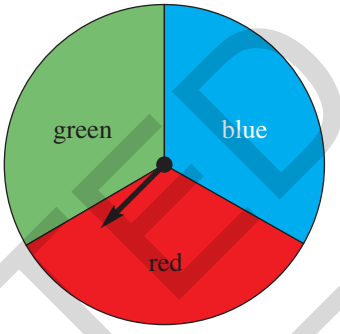
8H

4–6 | 4–7 | 4, 5, 7

FLUENCY

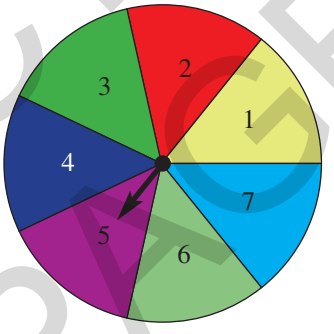
Example 12b–d

4 Consider the spinner shown.



- **a** How many outcomes are there? List them.
- **b** Find Pr(red); i.e. find the probability of the spinner pointing to red.
- **c** Find Pr(red or green).
- **d** Find Pr(not red).
- **e** Find Pr(yellow).

5 A spinner with the numbers 1 to 7 is spun. The numbers are evenly spaced.



- **a** List the sample space.
- **b** Find Pr(6).
- **c** Find Pr(8).
- **d** Find Pr(2 or 4).
- **e** Find Pr(even).
- **f** Find Pr(odd).
- **g** Give an example of an event having the probability of 1.

6 The letters in the word MATHS are written on 5 cards and then one is drawn from a hat.

- **a** List the sample space.
- **b** Find Pr(T), giving your answer as a decimal.
- **c** Find Pr(consonant is chosen), giving your answer as a decimal.
- **d** Find the probability that the letter drawn is also in the word TAME, giving your answer as a percentage.

7 The letters in the word PROBABILITY are written on 11 cards and then one is drawn from a hat.

- **a** Find Pr(P).
- **b** Find Pr(P or L).
- **c** Find Pr(letter chosen is in the word BIT).
- **d** Find Pr(not a B).
- **e** Find Pr(a vowel is chosen).
- **f** Give an example of an event with the probability of $\frac{3}{11}$.

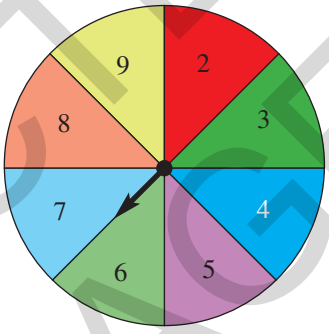
8 8, 9 9, 10 8H

PROBLEM-SOLVING

8 A bag of marbles contains 3 red marbles, 2 green marbles and 5 blue marbles. They are all equal in size and weight. A marble is chosen at random.

- **a** What is the probability that a red marble is chosen? (Hint: It is not $\frac{1}{3}$ because the colours are not all equally likely.) Give your answer as a percentage.
- **b** What is the probability that a blue marble is chosen? Give your answer as a percentage.
- **c** What is the probability that a green marble is *not* chosen? Give your answer as a percentage.

9 Consider the spinner opposite, numbered 2 to 9.



- **a** List the sample space.
- **b** Find the probability that a prime number will be spun, giving your answer as a decimal. (Remember that 2 is a prime number.)
- **c** Giving your answers as decimals, state the probability of getting a prime number if each number in the spinner opposite is:
  - **i** increased by 1
  - **ii** increased by 2
  - **iii** doubled

  (Hint: It will help if you draw the new spinner.)
- **d** Design a new spinner for which the Pr(prime) = 1.

10 A bag contains various coloured marbles – some are red, some are blue, some are yellow and some are green. You are told that Pr(red) = $\frac{1}{2}$, Pr(blue) = $\frac{1}{4}$ and Pr(yellow) = $\frac{1}{6}$. You are not told the probability of selecting a green marble.

- **a** If there are 24 marbles:
  - **i** Find how many there are of each colour.
  - **ii** What is the probability of getting a green marble?
- **b** If there are 36 marbles:
  - **i** Find how many there are of each colour.
  - **ii** What is the probability of getting a green marble?
- **c** What is the minimum number of marbles in the bag?
- **d** Does the probability of getting a green marble depend on the actual number of marbles in the bag? Justify your answer.

8H

11 | 11 | 11, 12

REASONING

11 a State the values of the pronumerals in the following table.

Event	Pr(event occurs)	Pr(event does not occur)	Sum of two numbers
Rolling a die, get a 3	$\frac{1}{6}$	$\frac{5}{6}$	a
Tossing a coin, get H	$\frac{1}{2}$	b	c
Rolling a die, get 2 or 5	d	$\frac{2}{3}$	e
Selecting a letter from 'HEART', getting a vowel	f	g	h

b If the probability of selecting a vowel in a particular word is $\frac{3}{13}$, what is the probability of selecting a consonant?

c If the probability of spinning blue with a particular spinner is $\frac{4}{7}$, what is the probability of spinning a colour other than blue?

12 A box contains different coloured counters, with Pr(purple) = 10%, Pr(yellow) = $\frac{2}{3}$ and Pr(orange) = $\frac{1}{7}$.

a Is it possible to obtain a colour other than purple, yellow or orange? If so, state the probability.

b What is the minimum number of counters in the box?

c If the box cannot fit more than 1000 counters, what is the maximum number of counters in the box?

Designing spinners — | — | 13

ENRICHMENT

13 a For each of the following, design a spinner using only red, green and blue sectors to obtain the desired probabilities. If it cannot be done, then explain why.

i Pr(red) = $\frac{1}{2}$, Pr(green) = $\frac{1}{4}$, Pr(blue) = $\frac{1}{4}$

ii Pr(red) = $\frac{1}{2}$, Pr(green) = $\frac{1}{2}$, Pr(blue) = $\frac{1}{2}$

iii Pr(red) = $\frac{1}{4}$, Pr(green) = $\frac{1}{4}$, Pr(blue) = $\frac{1}{4}$

iv Pr(red) = 0.1, Pr(green) = 0.6, Pr(blue) = 0.3

b If Pr(red) = x and Pr(green) = y, write a formula using x and y to determine what Pr(blue) must equal.

8I Experimental probability EXTENDING

Interactive

Widgets

HOTsheets

Walkthroughs

Although the probability of an event tells us how often an event should happen in theory, we will rarely find this being exactly right in practice. For instance, if you toss a coin 100 times, it might come up heads 53 times out of 100, which is not exactly $\frac{1}{2}$ of the times you tossed it. Sometimes we will not be able to find the exact probability of an event, but we can carry out an experiment to estimate it.

Tossing a coin 100 times does not mean it will come up heads 50 times.

Let's start: Tossing coins

For this experiment, each class member needs a fair coin that they can toss.

- Each student should toss the coin 20 times and count how many times heads occurs.
- Tally the total number of heads obtained by the class.
- How close is this total number to the number you would expect that is based on the probability of $\frac{1}{2}$? Discuss what this means.

Key ideas

- The **experimental probability** of an event occurring based on a particular experiment is defined as:

$$\frac{\text{number of times the event occurs}}{\text{total number of trials in the experiment}}$$

- The **expected number** of occurrences = probability $\times$ number of trials.
- If the number of trials is large, then the experimental probability is likely to be close to the actual probability of an event.

Example 13 Working with experimental probability

When playing with a spinner with the numbers 1 to 4 on it, the following numbers come up: 1, 4, 1, 3, 3, 1, 4, 3, 2, 3.

a What is the experimental probability of getting a 3?

b What is the experimental probability of getting an even number?

c Based on this experiment, how many times would you expect to get a 3 if you spin 1000 times?

SOLUTION	EXPLANATION
a $\frac{2}{5}$ or 0.4 or 40%	$\frac{\text{number of 3s}}{\text{number of trials}} = \frac{4}{10} = \frac{2}{5}$
b $\frac{3}{10}$	$\frac{\text{number of times with even result}}{\text{number of trials}} = \frac{3}{10}$
c 400 times	probability $\times$ number trials $= \frac{2}{5} \times 1000 = 400$

Exercise 8I

1, 2 | 1 | —

UNDERSTANDING

Example 13a,b

1 A 6-sided die is rolled 10 times and the following numbers come up: 2, 4, 6, 4, 5, 1, 6, 4, 4, 3.

a What is the experimental probability of getting a 3?

b What is the experimental probability of getting a 4?

c What is the experimental probability of getting an odd number?

2 When a coin is tossed 100 times, the results are 53 heads and 47 tails.

a What is the experimental probability of getting a head?

b What is the experimental probability of getting a tail?

c What is the theoretical probability of getting a tail if the coin is fair?

3–7 | 3–8 | 4–8

FLUENCY

3 A survey is conducted on people's television viewing habits.

Number of hours per week	0–	5–	10–	20–	30+
Number of people	20	10	15	5	0

a How many people participated in the survey?

b What is the probability that a randomly selected participant watches less than 5 hours of television?

c What is the probability that a randomly selected participant watches 20–30 hours of television?

d What is the probability that a randomly selected participant watches between 5 and 20 hours of television?

e Based on this survey, the experimental probability of watching 30+ hours of television is 0. Does this mean that watching 30+ hours is impossible?

Example 13c

4 A fair coin is tossed.

- **a** How many times would you expect it to show tails in 1000 trials?
- **b** How many times would you expect it to show heads in 3500 trials?
- **c** Initially, you toss the coin 10 times to find the probability of the coin showing tails.
  - **i** Explain how you could get an experimental probability of 0.7.
  - **ii** If you toss the coin 100 times, are you more or less likely to get an experimental probability close to 0.5?

5 A fair 6-sided die is rolled.

- **a** How many times would you expect to get a 3 in 600 trials?
- **b** How many times would you expect to get an even number in 600 trials?
- **c** If you roll the die 600 times, is it possible that you will get an even number 400 times?
- **d** Are you more likely to obtain an experimental probability of 100% from two throws or to obtain an experimental probability of 100% from 10 throws?

6 Each time a basketball player takes a free throw there is a 4 in 6 chance that the shot will go in. This can be simulated by rolling a 6-sided die and using numbers 1 to 4 to represent 'shot goes in' and numbers 5 and 6 to represent 'shot misses'.

- **a** Use a 6-sided die over 10 trials to find the experimental probability that the shot goes in.
- **b** Use a 6-sided die over 50 trials to find the experimental probability that the shot goes in.
- **c** Working with a group, use a 6-sided die over 100 trials to find the experimental probability that the shot goes in.
- **d** Use a 6-sided die over just one trial to find the experimental probability that the shot goes in. (Your answer should be either 0 or 1.)
- **e** Which of the answers to parts **a** to **d** above is closest to the theoretical probability of 66.67%? Justify your answer.

7 The colour of the cars in a school car park is recorded.

Colour	Red	Black	White	Blue	Purple	Green
Number of cars	21	24	25	20	3	7

Based on this sample:

- **a** What is the probability that a randomly chosen car is white?
- **b** What is the probability that a randomly chosen car is purple?
- **c** What is the probability that a randomly chosen car is green or black?
- **d** How many purple cars would you expect to see in a shopping centre car park with 2000 cars?

8I

FLUENCY

8 The number of children in some families is recorded in the table shown.

Number of children	0	1	2	3	4
Number of families	5	20	32	10	3

a How many families have no children?
b How many families have an even number of children?
c How many families participated in the survey?
d Based on this experiment, what is the probability that a randomly selected family has 1 or 2 children?
e Based on this experiment, what is the probability that a randomly selected family has an even number of children?
f What is the total number of *children* considered in this survey?

9 | 9, 10 | 9, 10

PROBLEM-SOLVING

9 A handful of 10 marbles of different colours is placed into a bag. A marble is selected at random, its colour recorded and then returned to the bag. The results are:

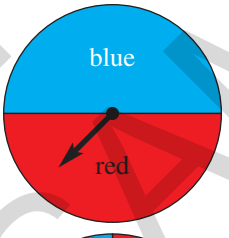
Red marble chosen	Green marble chosen	Blue marble chosen
21	32	47

a Based on this experiment, how many marbles of each colour do you think there are? Justify your answer in a sentence.
b For each of the following, state whether or not they are possible outcomes for the 10 marbles.
i 3 red, 3 green, 4 blue
ii 2 red, 4 green, 4 blue
iii 1 red, 3 green, 6 blue
iv 2 red, 3 green, 4 blue, 1 purple
v 2 red, 0 green, 8 blue

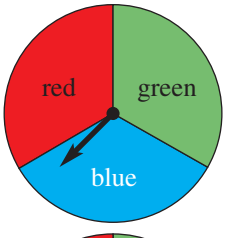
10 Match each of the experiment results **a** to **d** with the most likely spinner that was used (**A** to **D**).

	Red	Green	Blue
a	18	52	30
b	27	23	0
c	20	23	27
d	47	0	53

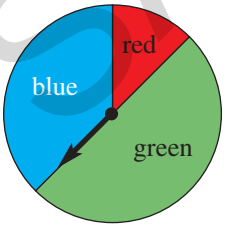
A



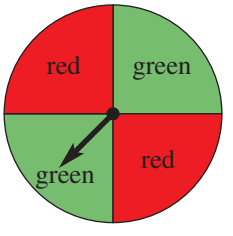
B



C



D



11 11 11, 12

81

REASONING

11 Assume that any baby has a 50% chance of being a boy or a girl, and use a coin to simulate a family with four children. Toss the coin four times, using heads to represent boys and tails to represent girls. Count the number of girls in the family. Repeat this experiment 20 times and present your results in a table like the one below.

Number of girls	0	1	2	3	4	Total
Number of families						20

- **a** Based on your simulation, what is the experimental probability that a family will have just one girl?
- **b** Based on your simulation, what is the experimental probability that a family will have four girls?
- **c** Explain why you might need to use simulations and experimental probabilities to find the answer to parts **a** and **b** above.
- **d** If you had repeated the experiment only 5 times instead of 20 times, how might the accuracy of your probabilities be affected?
- **e** If you had repeated the experiment 500 times instead of 20 times, how might the accuracy of your probabilities be affected?

12 Classify the following statements as true or false. Justify each answer in a sentence.

- **a** If the probability of an event is $\frac{1}{2}$, then it must have an experimental probability of $\frac{1}{2}$.
- **b** If the experimental probability of an event is $\frac{1}{2}$, then its theoretical probability must be $\frac{1}{2}$.
- **c** If the experimental probability of an event is 0, then the theoretical probability is 0.
- **d** If the probability of an event is 0, then the experimental probability is also 0.
- **e** If the experimental probability is 1, then the theoretical probability is 1.
- **f** If the probability of an event is 1, then the experimental probability is 1.

Improving estimates

— — 13

ENRICHMENT

13 A spinner is spun 500 times. The table below shows the tally for every 100 trials.

	Red	Green	Blue
First set of 100 trials	22	41	37
Second set of 100 trials	21	41	38
Third set of 100 trials	27	39	34
Fourth set of 100 trials	25	46	29
Fifth set of 100 trials	30	44	26

- **a** Give the best possible estimate for Pr(red), Pr(green) and Pr(blue) based on these trials.
- **b** If your estimate is based on just one set of trials, which one would cause you to have the most inaccurate results?
- **c** Design a spinner that could give results similar to those in the table. Assume you can use up to 10 sectors of equal size.
- **d** Design a spinner that could give results similar to those in the table if you are allowed to use sectors of different sizes.

Investigation

Monopoly risk

In the game of Monopoly, two 6-sided dice are rolled to work out how far a player should go forward. For this investigation, you will need two 6-sided dice or a random number simulator that simulates numbers between 1 and 6.

a Roll the two dice and note what they add up to. Repeat this 100 times and complete this table.

Dice sum	2	3	4	5	6	7	8	9	10	11	12	Total
Tally												100

b Represent the results in a column graph. Describe the shape of the graph. Do you notice any patterns?

c Use the results of your experiment to give the experimental probability of two dice adding to:
i 3 **ii** 6 **iii** 8 **iv** 12 **v** 15

d What is the most likely sum for the dice to add to, based on your experiment? Is this the mean, median or mode that you are describing?

e If the average Monopoly game involves 180 rolls, find the expected number of times, based on your experiment, that the dice will add to:
i 3 **ii** 6 **iii** 8 **iv** 12 **v** 15

f Why do you think that certain sums happen more often than others? Explain why this might happen by comparing the number of times the dice add to 2 and the number of times they add to 8.

g What is the mean dice sum of the 100 trials you conducted above?

To conduct many experiments, a spreadsheet can be used. For example, the spreadsheet below can be used to simulate rolling three 6-sided dice. Drag down the cells from the second row to row 1000 to run the experiment 1000 times.

	A	B	C	D
1	Die 1	Die 2	Die 3	=MODE(D2:D1001)
2	=INT(6*RAND())+1	=INT(6*RAND())+1	=INT(6*RAND())+1	=A2+B2+C2

h Investigate what the most likely dice sums are when you roll more than two dice. You should use a spreadsheet like the one above to find the most likely values. (Note: Instead of using the MODE function to help you, you can also use the AVERAGE function.)

Problems and challenges

Up for a challenge? If you get stuck on a question, check out the 'Working with unfamiliar problems' poster at the end of the book to help you.

1 Six numbers are listed in ascending order and some are removed. The mean and median are both 6, the mode is 2 and the range is 10. Fill in the missing numbers.

?, ?, 5, ?, ?, ?

2 A survey is conducted at a school and the results are presented as a pie chart. Find the minimum number of people who participated in the survey if the smallest sector has an angle of:

a 90° **b** 36° **c** 92° **d** 35°

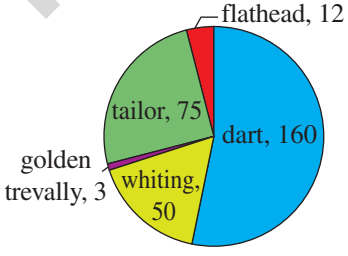
3 In a class of 20 students, a poll was taken of the number of cars owned by each family. The median number of cars owned is 1.5 and the mean number is 1.4 cars. Complete the following table of the results.

Number of cars	0	1	2	3
Number of students	4			

4 Each of the 8 letters of a word is written on a separate card. Given the following probabilities, what is the word?

Pr(letter P) = Pr(letter R) = 12.5%, Pr(letter B)= $\frac{1}{4}$, Pr(vowel) = 0.375

5 Frank the fisherman enjoys beach fishing on Fraser Island, Qld. One year he kept a count of the fish types that he caught and displayed these numbers as a pie chart in his fish shop. Calculate the answers to these questions, showing all steps.



a What angle did Frank use to represent his whiting catch?

b Find the difference between the smallest and largest angles in this pie chart.

c What is the probability of Frank catching a tailor?

d Find the probability that Frank does *not* catch a golden trevally.

e What is the probability of Frank catching a flathead or golden trevally?

6 A circular spinner is made using the colours red, green, purple and yellow in four sectors with two sectors being equal in size. The spinner is spun 120 times and the results obtained are shown in the table below.

Sector colour	Red	Green	Purple	Yellow
Frequency	40	32	19	29

Design a spinner that is likely to give these results, labelling the sector colours and angles. Explain the mathematical reason for your answer and show relevant calculations.

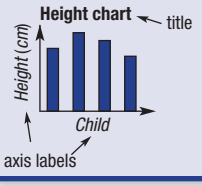
Chapter summary

Data types
- Numerical
  e.g. number of people (discrete), height of trees (continuous)
- Categorical
  e.g. colours, gender

Data collection
- Primary source (firsthand) or secondary source (collected by someone else)
- Census (whole population) or sample (selection)

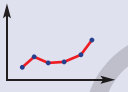
Graphical representations

Column graphs



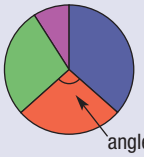
Dot plots

Line graphs



Time is often displayed on the horizontal axis.

Pie charts

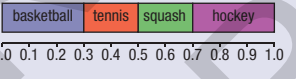


$\text{proportion} = \frac{\text{number}}{\text{total}}$

angle = 360° × proportion

Divided bar graphs

width of bar = proportion × total width



Stem-and-leaf plots

Stem	Leaf
2	3 6 7

⇒ 23, 26, 27

Statistics and probability

Summarising data numerically

range = 10 − 1
= 9

1, 2, 2, 3, 4, 4, 4, 5, 5, 6, 7, 8, 8, 9, 10

mode = 4 (most common value)

median = 5 (middle value)

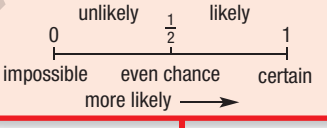
$\text{mean} = \frac{\text{sum of values}}{\text{number of values}}$

$= \frac{78}{15} = 5.2$

Experimental probability (Ext)

Use an experiment or survey or simulation to estimate probability.
e.g. Spinner lands on blue 47 times out of 120 →
Experimental probability = $\frac{47}{120}$

Probability: how likely an event is

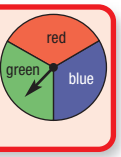


Outcome: possible result of an experiment
Event: either a single outcome or a collection of outcomes

Experiment/trial: e.g. roll a fair die
Sample space: {1, 2, 3, 4, 5, 6}
Pr(roll a 5) = $\frac{1}{6}$
Pr(roll odd number) = $\frac{3}{6} = \frac{1}{2}$

Expected number is
Pr(event) × number of trials
e.g. Flip coin 100 times, expected number of heads
$= \frac{1}{2} \times 100 = 50$
e.g. Roll die 36 times, expected number of 5s
$= \frac{1}{6} \times 36 = 6$

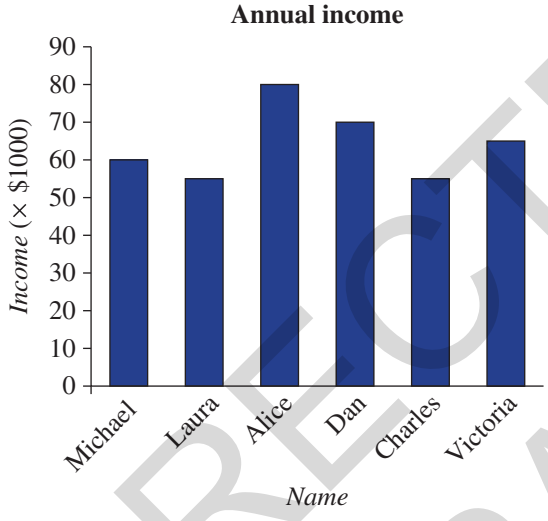
Sample space: e.g. {red, green, blue}
Pr(spin red) = $\frac{1}{3}$
Pr(don't spin blue) = $\frac{2}{3}$



Multiple-choice questions

8C **1** In the column graph shown, the highest income is earned by:

A Michael **B** Alice **C** Dan **D** Laura **E** Victoria



*Questions **2** and **3** relate to the following information.*

The results of a survey are shown below.

Instrument learned	piano	violin	drums	guitar
Number of students	10	2	5	3

8F **2** If the results above are presented as a pie chart, then the angle occupied by the drums sector is:

A 360° **B** 180° **C** 120° **D** 90° **E** 45°

8I Ext **3** Based on the survey, the experimental probability that a randomly selected person learns the guitar is:

A $\frac{1}{4}$ **B** $\frac{1}{2}$ **C** 3 **D** $\frac{3}{5}$ **E** $\frac{3}{20}$

8A **4** Which one of the following variables is continuous numerical?

A the gender of newborn babies
B the number of babies born in a given years
C the number of hairs on a baby's head
D the weight (in kg) of newborn babies
E the length (in letters) of a baby's first name

Chapter review

*Questions **5** and **6** relate to the following information.*

In a class of 20 students, the number of days each student was absent over a 10-week period is recorded.

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

8B **5** The mode is:

A 0 **B** 1 **C** 2 **D** 3 **E** 4

8B **6** The mean number of days a student was absent is:

A 1 **B** 2 **C** 1.95 **D** 3 **E** 39

8B **7** The range of the numbers 1, 5, 3, 9, 12, 41, 12 is:

A 40 **B** 41 **C** 12 **D** 3 **E** 1

8H **8** Which of the following events has the same probability as rolling an odd number on a fair 6-sided die?

A rolling a number greater than 4 on a fair 6-sided die
B choosing a vowel from the word CAT
C tossing a fair coin and getting heads
D choosing the letter T from the word TOE
E spinning an odd number on a spinner numbered 1 to 7

8H **9** Each letter of the word APPLE is written separately on five cards. One card is then chosen at random. Pr(letter P) is:

A 0 **B** 0.2 **C** 0.4 **D** 0.5 **E** 1

8I Ext **10** A fair 6-sided die is rolled 600 times. The expected number of times that the number rolled is either a 1 or a 2 is:

A 100 **B** 200 **C** 300 **D** 400 **E** 600

Short-answer questions

8C **1** Draw a column graph to represent the following people's ages.

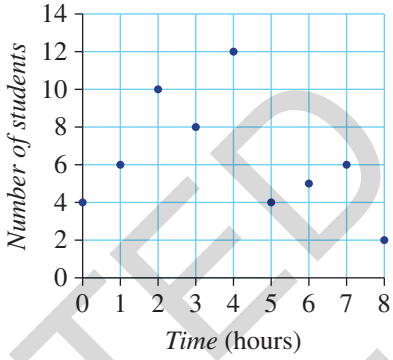
Name	Sven	Dane	Kelly	Hugo	Frankie
Age (years)	20	12	15	22	25

8C **2** A Year 7 group was asked how many hours of television they watch in a week. The results are given in the table.

TV Watched (hours)	No. of students
8	5
9	8
10	14
11	8
12	5

a How many students participated in the survey?
b What is the total number of hours of television watched?
c Find the mean number of hours of television watched.
d Show this information in a column graph.

8D **3** The number of students in the library is recorded hourly, as displayed in the graph.



- **a** How many students entered the library when it first opened?
- **b** How many students were in the library at 8 hours after opening?
- **c** If the library opens at 9:00 am, at what time are there the most number of students in the library?
- **d** How many students were in the library at 4:00 pm?

8F Ext **4** 120 people were asked to nominate their favourite take-away food from the list: chicken, pizza, hamburgers, Chinese. The results are given in the table.

Food	Frequency
Chicken	15
Pizza	40
Hamburgers	30
Chinese	35

- **a** If you want to show the data in a pie chart, state the angle needed to represent Chinese food.
- **b** What percentage of people prefer hamburgers?
- **c** Represent the results in a pie chart.

8B **5** Consider the data 1, 4, 2, 7, 3, 2, 9, 12. State the:

a range **b** mean **c** median **d** mode

8B **6** Consider the data 0, 4, 2, 9, 3, 7, 3, 12. State the:

a range **b** mean **c** median **d** mode

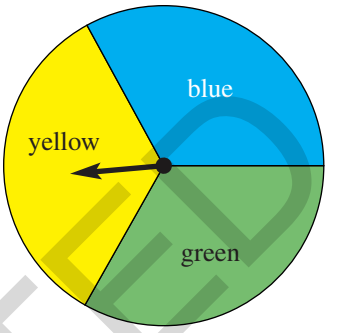
8G **7** For each of the following descriptions, choose the probability from the set $0, \frac{1}{8}, \frac{3}{4}, 1, \frac{19}{20}$ that matches best.

a certain **b** highly unlikely **c** highly likely
d likely **e** impossible

Chapter review

8H **8** List the sample space for each of the following experiments.

- **a** A fair 6-sided die is rolled.
- **b** A fair coin is tossed.
- **c** A letter is chosen from the word DESIGN.
- **d** Spinning the spinner shown opposite.



8H **9** Vin spins a spinner with nine equal sectors, which are numbered 1 to 9.

- **a** How many outcomes are there?
- **b** Find the probability of spinning:
  - **i** an odd number
  - **ii** a multiple of 3
  - **iii** a number greater than 10
  - **iv** a prime number less than 6
  - **v** a factor of 8
  - **vi** a factor of 100

8H **10** One card is chosen at random from a standard deck of 52 playing cards. Find the probability of drawing:

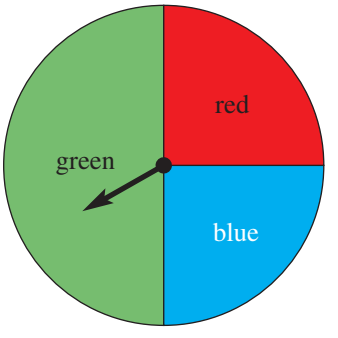
- **a** a red king
- **b** a king or queen
- **c** a jack of diamonds
- **d** a picture card (i.e. king, queen or jack)

8I Ext **11** A coin is tossed 100 times, resulting in 42 heads and 58 tails.

- **a** What is the experimental probability of getting heads? Give your answer as a percentage.
- **b** What is the actual probability of getting heads if the coin is fair? Give your answer as a percentage.

8H **12** Consider the spinner shown.

- **a** State the probability that the spinner lands in the green section.
- **b** State the probability that the spinner lands in the blue section.
- **c** Tanya spins the spinner 100 times. What is the expected number of times it would land in the red section?
- **d** She spins the spinner 500 times. What is the expected number of times it would land in the green section?



Extended-response questions

1 The number of rainy days experienced throughout a year in a certain town is displayed below.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
No. of rainy days	10	11	3	7	2	0	1	5	6	9	7	5

a Show this information in a column graph.

b For how many days of the year did it rain in this town?

c What is the probability that it will rain in any day during winter (i.e. during June, July and August)?

d What type of variable (e.g. continuous numerical) is the number of rainy days?

e What type of variable is the month?

2 At a school camp, a survey was conducted to establish each student's favourite dessert.

Ice-cream	Yoghurt	Danish pastry	Jelly	Pudding	Cheesecake
10	5	2	7	4	12

a How many students participated in the survey?

b What is the most popular dessert selected?

c What is the probability that a randomly selected student chooses jelly as their favourite dessert?

d For each of the following methods listed below, state whether it would be a reasonable way of presenting the survey's results.

i column graph **ii** line graph

iii pie chart **iv** divided bar graph

e If the campers attend a school with 800 students, how many students from the entire school would you expect to choose pudding as their preferred dessert?