



Department of
Education

AUSTRALIAN CURRICULUM

MATHEMATICS YEAR 7

Analysing data

MATHEMATICS YEAR 7

Analysing data

Student's name: _____

Teacher's name: _____

First published 2012

ISBN 9780730744504

SCIS 1564106

© Department of Education WA 2012 (Revised 2020)



<https://creativecommons.org/licenses/by-nc/4.0/>

Requests and enquiries concerning copyright should be addressed to:

Manager Intellectual Property and Copyright
Department of Education
151 Royal Street
EAST PERTH WA 6004

Email: copyright@education.wa.edu.au



Department of
Education

This resource contains extracts from The Australian Curriculum Version 3.0 © Australian Curriculum, Assessment and Reporting Authority 2012.

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

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

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



creativecommons.org/licenses/by-nc-sa/3.0/au/

Graphics used in this resource are sourced from <http://openclipart.org> under the creative commons license <http://creativecommons.org/publicdomain/zero/1.0>



This product will be registered through the National Copyright Unit for use in all Australian schools without remuneration.

Contents

Signposts	2
Introduction.....	3
Curriculum details	4
1. Finding the range and mode	7
2. Finding the mean and median	15
3. The effect of outliers	25
4. Measures of central tendency	33
5. Calculations from graphs	43
6. Use in everyday life	55
7. Calculations with provided data.....	67
8. Summary.....	75
9. Review tasks	77
Solutions.....	87

Signposts

Each symbol is a sign to help you.

Here is what each one means:



The recommended time you should take to complete this section.



An explanation of key terms, concepts or processes.



A written response.
Write your answer or response in your journal.



Correct this task using the answers at the end of the resource.



Calculators may not be used here.



Make notes describing how you attempted to solve the problem. Keep these notes to refer to when completing the Self-evaluation task. Your teacher may wish you to forward these notes.

Introduction

This resource should take you approximately two weeks to complete. It comprises seven learning sections, a summary section and a review task section.

The learning sections have the following headings:

- **Key words**
These are the main words that you need to understand and use fluently to explain your thinking.
- **Warm-up**
Warm-up tasks should take you no longer than 10 minutes to complete. These are skills from previous work you are expected to recall from memory, or mental calculations that you are expected to perform quickly and accurately. If you have any difficulties in answering these questions, please discuss them with your teacher.
- **Review**
Some sections have reviews immediately after the warm-up. The skills in these reviews are from previous work and are essential for that section. You will use these to develop new skills in mathematics. Please speak to your teacher immediately if you are having any trouble in completing these activities.
- **Focus problem**
Focus problems are designed to introduce new concepts. They provide examples of the types of problems you will be able to solve by learning the new concepts in this resource. Do not spend too long on these but do check and read the solutions thoroughly.
- **Skills development**
These help you consolidate new work and concepts. Most sections include skills development activities which provide opportunities for you to become skilled at using new procedures, apply your learning to solve problems and justify your ideas. Please mark your work after completing each part.

Correcting your work

Please mark and correct your work as you go. Worked solutions are provided to show how you should set out your work. If you are having any difficulty in understanding them, or are getting the majority of the questions wrong, please speak to your teacher immediately.

Journal

Please keep an exercise book to record your notes and to summarise your learning. At the end of each section, write definitions for the key words that were introduced for that section.

Curriculum details

Content Descriptions

This resource provides learning and teaching to deliver the Australian Curriculum: Mathematics for the following Year 7 Content Descriptions.

Calculate mean, median, mode and range for sets of data. Interpret these statistics in the context of data (ACMSP171)

Describe and interpret data displays using median, mean and range (ACMSP172)

Content Descriptions	1	2	3	4	5	6	7	R
ACMSP171								
ACMSP172								



Indicates the content description is explicitly covered in that section of the resource.

Previous relevant Content Descriptions

The following Content Descriptions should be considered as prior learning for students using this resource.

At Year 5 level

Describe and interpret different data sets in context (ACMSP120)

At Year 6 level

Interpret secondary data presented in digital media and elsewhere (ACMSP148)

Proficiency strand statements at Year 7 level

At this year level:

Understanding includes describing patterns in uses of indices with whole numbers, recognising equivalences between fractions, decimals, percentages and ratios, plotting points on the Cartesian plane, identifying angles formed by a transversal crossing a pair of lines, and connecting the laws and properties of numbers to algebraic terms and expressions

Fluency includes calculating accurately with integers, representing fractions and decimals in various ways, investigating best buys, finding measures of central tendency and calculating areas of shapes and volumes of prisms

Problem Solving includes formulating and solving authentic problems using numbers and measurements, working with transformations and identifying symmetry, calculating angles and interpreting sets of data collected through chance experiments

Reasoning includes applying the number laws to calculations, applying known geometric facts to draw conclusions about shapes, applying an understanding of ratio and interpreting data displays

General capabilities

General capabilities	1	2	3	4	5	6	7	R
Literacy								
Numeracy								
Information and communication technology (ICT) capability								
Critical and creative thinking								
Personal and social capability								
Ethical behaviour								
Intercultural understanding								



Indicates general capabilities are explicitly covered in that section of the resource.

Cross-curriculum priorities

Cross-curriculum priorities	1	2	3	4	5	6	7	R
Aboriginal and Torres Strait Islander histories and cultures								
Asia and Australia's engagement with Asia								
Sustainability								



Indicates cross-curriculum priorities are explicitly covered in that section of the resource.

This resource contains extracts from The Australian Curriculum Version 3.0 © Australian Curriculum, Assessment and Reporting Authority 2012.

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

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

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



creativecommons.org/licenses/by-nc-sa/3.0/au/

1. Finding the range and mode

When you complete this section you should be able to:

- calculate the range and mode for numerical data.

Key words

- range
- mode
- bimodal
- score

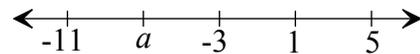
Warm-up 1

1. List the factors of 20. _____

2. $8 + 7 =$ _____

3. What is the missing number?

$a =$ _____



4. Circle the greater fraction. $\frac{5}{4}$ or $\frac{9}{8}$

5. Find a half of 15. _____

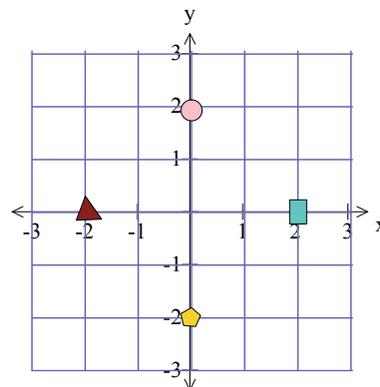
6. $7.45 - 2.5 =$ _____

7. $6.3 \times 5 =$ _____

8. Write 2.7 as a mixed fraction. _____

9. Complete by writing the next number: 1692, 1695, 1698, _____

10. Which shape is at $(-2, 0)$?



Review 1

Example

The ages (last birthday) of a group of kids playing basketball are recorded below.

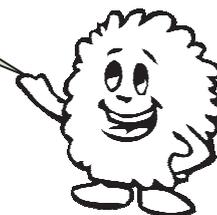
6, 8, 8, 11, 9, 12, 8, 8, 10, 11, 9

1. How many kids were playing basketball?
2. What is the age of the youngest player?
3. What is the age of the oldest player?
4. What is the difference in age between the oldest and the youngest player?
5. What is the most common age?
6. The kids line up in ages from the youngest to the oldest. How old is the middle kid in the line?
7. What is the total of all the ages of the kids?
8. Divide the total of all the ages by the number of kids to find the average age. (Use a calculator.)

Solution

1. 11 kids were playing.
2. The youngest player is six.
3. The oldest player is 12.
4. There is a difference in age of six years.
5. Eight is the most common age.
6. The middle kid in the line is nine years of age.
7. The total of the ages is 100 years.
8. The average age is 9.09 years of age, by dividing 100 by 11.

The values in a list of numerical data are often called **scores**.



1. For each of the following sets of **scores**

- (a) 5, 8, 5 (b) 3, 3, 6, 7, 4, 9, 10 (c) 10, 10, 15, 7, 5, 5, 9, 11, 9, 10, 8

Find:

- (i) the difference between the highest and lowest score
 (ii) the most common score
 (iii) the middle score when they are in order from smallest to largest
 (iv) the average of the scores.

Answers to be set out as shown below.

- (a) (i) _____ (ii) _____
 (iii) _____ (iv) _____
 (b) (i) _____ (ii) _____
 (iii) _____ (iv) _____
 (c) (i) _____ (ii) _____
 (iii) _____ (iv) _____

2. The winning margins in the weekend football games were

2, 53, 25, 37, 10, 6, 34, 21 and 2.

Find:

- (a) the difference between the highest and lowest margin _____
 (b) the most common winning margin _____
 (c) the middle winning margin when they are in order from smallest to largest

- (d) the average of the winning margins. _____



Check your work before continuing.

Focus problem 1



Tori had just got a part-time job at a shoe shop.

The manager had started to organise the records of the last week's sales. She didn't have time to complete them so Tori was given the task to finish them.

This is what was given to her.

Day	Sizes sold	Mode
Sunday	$6\frac{1}{2}$, 5, 6, 6, $9\frac{1}{2}$, 6, 5, 6, $7\frac{1}{2}$, 6, 6, 9, 5, 6	6
Monday	$4\frac{1}{2}$, 5, 11, 5, $9\frac{1}{2}$, 5, 6, 5, 6, $7\frac{1}{2}$, 5, 8, 9, 5, 10, 5, 11, 5	5
Tuesday	5, 8, 8, $5\frac{1}{2}$, 7, $5\frac{1}{2}$, $5\frac{1}{2}$, $5\frac{1}{2}$, $7\frac{1}{2}$, 8, 9, $5\frac{1}{2}$, 10, $5\frac{1}{2}$	$5\frac{1}{2}$
Wednesday	$5\frac{1}{2}$, 5, 10, 6, 5, 6, $7\frac{1}{2}$, 5, 8, 8, $9\frac{1}{2}$, 10, 5, 5	5
Thursday	8, 11, 5, $5\frac{1}{2}$, 7, $6\frac{1}{2}$, 8, 6, $7\frac{1}{2}$, 8, 8, 9, 10, 5, $7\frac{1}{2}$, 8, 8	
Friday	$6\frac{1}{2}$, 5, 6, 5, $6\frac{1}{2}$, 7, 6, 5, 6, $6\frac{1}{2}$, 8, $6\frac{1}{2}$, 6, $6\frac{1}{2}$, 10, $6\frac{1}{2}$	

- (a) Check the values in the table and decide how the number in the **mode** column was chosen. Write in your own words how the mode is decided.

Hint: Look at how many of each size was sold on a particular day.

- (b) Complete the values missing in the mode column for Thursday and Friday.
- (c) The sizes were recorded by each sales person writing down the size of the shoes they sold. Tori decided that there was a better way to record the sales to make it easier to make it easier to find the mode. How would you record the sales?

- (d) The manager asked Tori how to decide which sizes of a new style of shoe should be ordered. Tori said the **mode** for the week could be used as a guide. Why would Tori suggest the mode?

- (e) Tori was also interested in the different prices of the shoes for sale. The manager said the price **range** was \$230 as the cheapest pair of shoes was \$12 and the most expensive sold for \$242. Tori noticed that the most expensive shoes were now on sale for \$210. What is the range of the prices now?

- (f) Write instructions for another student telling them how to find the mode and the range of a set of numbers. Use an example to help.

In a lot of the answers a sentence will be needed. Don't forget that just because it is maths it doesn't mean that sentences and spelling are not important. Write carefully and clearly.



Check your work before continuing.

Shoe bits and pieces

Some scientists have indicated that humans were probably wearing some sort of foot covering as far back as 40 000 years ago. This was decided by looking at the state of foot bones in ancient skeletons. The earliest shoes that have been found are around 10 000 years old. The original shoes and sandals were usually made from animal hides and held on by string made from plant fibres.

Until the recent advent of cheap mass production of shoes the majority of people in the world had no shoes because they couldn't afford them. Shoes protect the feet from the elements and sharp objects when we walk so wearing shoes is a health and safety priority.

Skills development 1

Example

Find the **mode** and the **range** of:

- (a) 2, 4, 5, 2, 6, 7, 8, 4, 5, 7, 3, 6, 8, 11, 7, 11
- (b) \$23, \$47, \$29, \$18, \$23, \$47, \$28, \$94
- (c) 1.3, 2.7, 0.2, 1.3, 0.2, 1.6, 2.1, 2.1, 0.4

Solution

- (a) The mode (the most common value) is 7. The highest value is 11 and the lowest is 2 so the range is 9 (11 minus 2).
- (b) There are two modes, \$23 and \$47, so the data is **bimodal**. The range is \$76.
- (c) There are more than two modes so it means that no value really occurs more often than any other so there is no mode. The range is 2.5.

1. Find the range and the mode of:

- (a) 56, 43, 37, 85, 34, 43, 40

- (b) 15 cm, 11 cm, 12 cm, 15 cm, 25 cm, 25 cm, 11 cm, 15 cm, 12 cm

- (c) 2.9 kg, 2.8 kg, 2.8 kg, 3.6 kg, 2.9 kg, 2.1 kg, 7.6 kg

2. Find the value of n if the mode of the scores 3, n , 5, 7 and 8 is 7. Explain how you found the value.

3. The maximum temperatures were recorded for a Western Australian town on the first ten days of February. They were measured in degrees Celsius and recorded to the nearest degree, as given below.

41, 39, 37, 38, 32, 35, 31, 33, 32, 36

- (a) What was the range of the maximum temperatures for those ten days? _____
- (b) What is the mode of the temperatures? _____
- (c) If the maximum temperature on the 11th day was 33 degrees would the **range** and **mode** change? If so, write down the new values.

- 4 Why would the following set of class marks, out of ten, be considered as having no mode?

5, 6, 7, 8, 9, 5, 6, 8, 9, 4, 5, 6, 8

5. Find the two possible values of n if the range of the **scores** 5, 8, 17, 4, 10, 6, n and 7 is 16. Explain how you found the values.



Check your work before continuing.

2. Finding the mean and median

When you complete this section you should be able to:

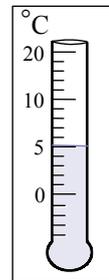
- find the mean and median of numerical data.

Key words

- mean
- median
- measure of central tendency
- measure of spread

Warm-up 2

1. Circle the largest common factor of 18 and 15. 2, 3, 4, 5, 6, 9
2. $13 - 5 =$ _____
3. The temperature was 5 degrees but it dropped 8 degrees.
What is the new temperature? _____
4. Insert $<$, $>$ or $=$ to make the following sentence true.
 $\frac{15}{5} \square 3$
5. $\frac{3}{2} \times 10 =$ _____
6. Round 10.491 to a whole number. _____
7. $7 \overline{)42.14}$
8. Write 6% as a fraction. _____
9. Complete: 9.93, 9.96, 9.99, _____
10. Determine the size of the missing angle.



Review 2

Example

For the nine scores in a test, 5, 7, 9, 9, 8, 6, 7, 3 and 9:

- find the middle score when they are in order from lowest to highest
- find the total of all the scores and divide by 9 to calculate the average score
- write down the range and the mode.

Solution

- The numbers in order of size are 3, 5, 6, 7, 7, 8, 9, 9, 9.
The middle score is a 7.
- The total of the scores is 63 so the average is $63 \div 9 = 7$.
- The range is $9 - 3 = 6$ and the mode is 9.

1. Find the middle score, when the scores are in order, and the average of the following sets of data.

- The hourly pay rate for a group of friends from their part-time jobs

\$9.50, \$8.50, \$8.50, \$9.10, \$8.30, \$8.20, \$9.50

- The length of the long jumps in the final at the school athletics carnival

3.43 m, 2.56 m, 2.87 m, 3.05 m, 1.99 m

Don't forget to include units in your answer if appropriate to the question.



2. Find the range and the mode for each of the following sets of data.

- The ages of the children at a birthday party

5, 5, 6, 7, 4, 5, 6, 6, 7, 8, 7, 6, 3 and 4

- The maximum temperatures, in degrees Celsius, for the first two weeks of autumn

27, 31, 29, 32, 28, 27, 31, 31, 31, 27, 30, 28, 26, 33

Focus problem 2

Where is the **median** strip on a busy road?

If you don't know the answer you may need to ask someone or do some quick research.



Mathematicians also use the term median when doing calculations with sets of numerical data. They also use another term called the '**mean**'. You have already completed calculations that gave you the median and the mean although those terms weren't used to describe them.

- (a) For the set of scores 15, 16, 16, 18, 21, 23 and 24 the calculations have been done below but you have to match them to their names. Hint: Do the ones you know from the last section first, and then think of the term 'median'.

Draw a line between the value and the correct name.

19	median
16	mean
18	range
9	mode

Using the descriptions you have already come across in earlier parts of this book explain what you think the median and mean represent.

Refer back to some of the wording used in Review 2 if you are stuck.



- (b) The set of scores 3, 5, 10 has a **median** of 5 but the set of scores 3, 5, 7 and 10 has a median of 6.

In the first case the median is one of the scores and in the second case it isn't one of them.

What is the difference between the two sets of scores?

How do you think the 6 is calculated from the second set of scores?

- (c) Finding the median should be quite easy when you have a small set of data but what would be the problem if the list is long?

- (d) Explain what is being done in each of the following.

(i) $\cancel{1}, 2, 2, 4, 7, 9, 11, 15, \cancel{24}$ _____

$\cancel{1}, \cancel{2}, 2, 4, 7, 9, 11, \cancel{15}, \cancel{24}$ _____

$\cancel{1}, \cancel{2}, \cancel{2}, 4, 7, 9, \cancel{11}, \cancel{15}, \cancel{24}$ _____

$\cancel{1}, \cancel{2}, \cancel{2}, \cancel{4}, 7, \cancel{9}, \cancel{11}, \cancel{15}, \cancel{24}$ _____

↑

7

(ii) $\cancel{1}, 2, 2, 4, 7, 9, 11, 11, 15, \cancel{24}$ _____

$\cancel{1}, \cancel{2}, 2, 4, 7, 9, 11, 11, \cancel{15}, \cancel{24}$ _____

$\cancel{1}, \cancel{2}, \cancel{2}, 4, 7, 9, 11, \cancel{11}, \cancel{15}, \cancel{24}$ _____

$\cancel{1}, \cancel{2}, \cancel{2}, \cancel{4}, 7, 9, \cancel{11}, \cancel{11}, \cancel{15}, \cancel{24}$ _____

↑

8

- (e) Is there anything wrong with each of the following diagrams used to calculate the **median** of the list of numbers? If so, explain what is wrong.

(i) $2, 3, 4, 2, 3, 3, 10$
 \uparrow
 2

(ii) $15, 12, 12, 11, 3, 3, 4$
 \uparrow
 11

(iii) $4, 3, 7, 8, 11, 12, 12, 13$
 \uparrow
 9

- (f) If you wanted to find the median height of a group of nine people at a picnic what instruction could you give to them to find it easily?

- (g) What connection does the term 'median' in mathematics have to the median strip on a road?

The **mean** of a set of numerical data is the sum of all the scores divided by the number of scores in the set.

You may already know this as the average.

It can also be referred to as the arithmetic mean, but we will stick to the use of the simple term 'mean'.

A calculator is a useful tool for calculating the mean of a set of numbers.

For example, the mean of the numbers 5, 6, 8 and 10 is $\frac{5+6+8+10}{4} = \frac{29}{4} = 7.25$.

The mean, **median** and mode for a set of numerical data are all considered a type of average. The type of information require determines which one is the best average to use.

For example:

If we asked about the average shoe size of a 13 year old it would be the mode that was used.

The average house values sold in a month uses the median price of all those houses sold so that very expensive houses don't have too much influence on the average.

The average maximum temperature in Perth over summer is found by using the mean.

In general, we still use the term average when we need the mean.

The three averages are called **measures of central tendency** as they tend to show how the scores are grouped around a 'middle'.

The range is called a **measure of spread** as it tells you about how far the scores are spread.



Check your work before continuing.

Skills development 2

Example

- (a) Find the **median**, **mean**, mode and range for the following amounts of pocket money a group of teenage friends received last week.

\$20, \$16, \$17, \$5, \$10, \$10, \$25, \$34, \$10, \$8, \$10

Explain the calculations you made.

- (b) One of the friends, Freddy, was late giving his amount for the calculation. If he added the \$15 he got to the list, calculate the median, mean, mode and range for the new list.

Solution

- (a) In order of size the pocket money is:

\$5, \$8, \$10, \$10, \$10, \$10, \$16, \$17, \$20, \$25, \$34.

The median is the sixth score which is \$10. There are five scores less than it and five more than it.

The mean is \$15 because the total was \$165 and divided by 11, which was the number of values, the answer is \$15.

The mode is \$10 because it is the most common score.

The range is \$29 because $\$34 - \$5 = \$29$.

- (b) In order of size the pocket money will now be:

\$5, \$8, \$10, \$10, \$10, \$10, \$15, \$16, \$17, \$20, \$25, \$34.

The median is the average of the two middle scores. $\$ \frac{10+15}{2} = \12.50 which means the median is \$12.50.

The mean is \$15 because the total was \$180 and divided by 12, which was the number of values, the answer is \$15.

The mode is \$10 because it is the most common score.

The range is \$29 because $\$34 - \$5 = \$29$.

1. When calculating the median, mean, mode and range for a set of numerical data you may need a calculator to help with the calculation of two of the measures. One of the other two can usually be completed without a calculator.

Indicate which two may need a calculator and which one doesn't need a calculator at all.

2. At an aquarium in the city the lengths of the mature male sharks were recorded.

White sharks	5.5 m, 6.7 m
Tiger sharks	3.4 m, 4.3 m, 3.6 m
Bull sharks	1.6 m
Dusky sharks	2.4 m, 2.8 m, 3.6m
Grey nurse sharks	2.2 m

Find the range, **mean**, **median** and mode of the shark lengths.

Range: _____

Mean: _____

Median: _____

Mode: _____

3. Mrs Shoparound was concerned about what she thought were the varying prices of her dog's favourite food. She decided to survey a number of shops in her area and find the cost of a 700 g can of the food. The results she collected are given below.

\$3.20, \$4.50, \$3.30, \$3.75, \$5.02, \$3.07, \$2.99, \$3.20, \$3.20

- (a) Calculate the mean, median, mode and range of the prices.

Mean: _____

Median: _____

Mode: _____

Range: _____

- (b) Which of the four values would be best for her to use as a guide to the usual price of the can of food?

- (c) What does the range tell you in this problem?

- (d) Are the values for the **mean**, **median** and mode close to each other? Why, or why not?



Check your work before continuing.

3. The effect of outliers

When you complete this section you should be able to:

- understand how outliers may affect the calculations involving numerical data.

Key words

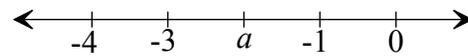
- outlier

Warm-up 3

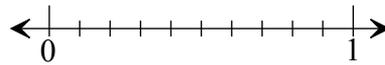
1. Circle the composite numbers. 4, 5, 6, 7, 8

2. $7 \times 12 =$ _____

3. What is the missing number?



4. Locate $\frac{3}{6}$ on the number line.



5. Find five-quarters of 12. _____

6. Estimate the sum by first rounding to whole numbers

$$271.9 + 19.9 \approx \text{_____}$$

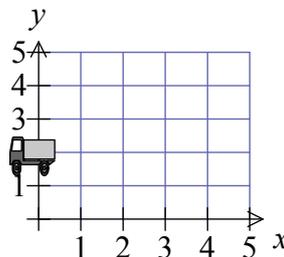
$$\begin{array}{r} 5.37 \\ \times \quad 8 \\ \hline \end{array}$$

8. Write $5\frac{1}{4}$ as a percentage. _____

9. Complete by adding the next number: $16\frac{3}{4}, 14\frac{2}{4}, 12\frac{1}{4},$ _____

10.

At what point is the truck?



Review 3

Example

Find the value of the number n if the set of scores 4, 5, 8, 9, n has:

- (a) a mean of 7
- (b) a median of 6.

Show the reasoning for your answers

Solution

- (a) If five numbers have a mean of 7 then their total is $5 \times 7 = 35$.
 $4 + 5 + 8 + 9 = 26$ and $35 - 26 = 9$. The value of n is 9.
- (b) If the median is 6 then it must be the middle number. There are five numbers so n will have to be 6 as there is no 6 in the list otherwise.

1. Calculate the mean of: 345, 567, 344, 212, 7.

2. A group of seven numbers is known to have a mean of 10. What is the total of all the numbers?

3. The median of a group of numbers is 25.

- (a) It is known that there are five numbers in the group. Will 25 be a number in the group or not? Explain your answer.

- (b) It is known that there are eight numbers in the group. Will 25 be a number in the group or not? Explain your answer.

Focus problem 3

A number of people in a fitness class were having a friendly competition as to who was the fittest. They decided that one of the ways they could compare themselves was to record how far in kilometres they could run in fifteen minutes on the treadmill.



The group then recorded their results as:

3.4, 2.4, 3.1, 2.7, 2.1, 3.0, 2.8,
2.7, 3.5, 3.3, 3.2, 170, 1.7, 2.8.

- (a) Calculate the mean and median of the distances.

- (b) Comment about any error that you see in the recorded results.

- (c) Why do you think the error has occurred?

- (d) Fix the error if possible, explaining what you did.

- (e) Re-calculate the median and the mean with the error fixed.

- (f) How did the error in the distances affect the median and mean?

- (g) One of the results in the original list was called on **outlier**. Which one is it and why does it have that name?

- (h) How can you be sure that there was an error in the original results?

An **outlier** is a score that lies outside the general values of the other scores.



Check your work before continuing.

Skills development 3

Example

The weights of bags of apples for sale at the market were listed on a sales sign as follows.

1.5 kg, 1.5 kg, 1.6 kg, 1.5 kg, 1.4 kg, 15 kg,

1.5 kg, 1.5 kg, 1.6 kg, 1.5 kg, 1.4 kg

- One of the customers pointed out a possible error in the weights. Which one would it be?
- What might indicate that the **outlier** is an error rather than a correct weight?
- Which of the calculations, mean, mode, median and range would be most affected if the outlier was to be deleted? Explain your answers.

Solution

- The possible error is the one listed as 15 kg.
- All the other weights are 1.4, 1.5 or 1.6 kg so 15 kg is much bigger. It possibly should have been 1.5 kg.
- The median and the mode would not be affected by removing the outlier. The range would be much smaller and the mean would be close to 1.5 kg if the outlier is taken out.

- A junior school basketball squad had their heights measured to include in their profiles. The heights of the team were as follows.

156 cm, 165 cm, 160 cm, 170 cm, 208 cm, 166 cm,

170 cm, 161 cm, 164 cm, 151 cm, 169 cm, 161 cm

- Is one of the heights an outlier? Explain the reasoning for your answer.

- Calculate the mean and median of the heights:

- as they are
- without the outlier.

(c) Did the results vary greatly?

2. The weekly wages of the staff at a small office are:

\$350, \$465, \$854, \$3200, \$535, \$350.

(a) One of the results is an **outlier**. Which one is it? _____

(b) Find the mean and median of the wages as they are listed.

(c) Remove the outlier then calculate the mean and median of the remaining wages.

(d) Which of the two results, mean or median, did the outlier most affect?

(e) The outlier wage is that of the boss. A new position is being advertised at the office. Does it make sense to leave the wage of the boss out when calculating the mean and median for the advertisement? Would it be honest to leave it in?

3. Jared got 45% in the class test which looked pretty bad compared to the top score of 80%. However the mean of the scores was 54% and the median was 43%.

(a) How could Jared argue that he was better than at least half of the class?

(b) If there were 23 students in the class how many got a mark above 43%?



Check your work before continuing.

4. Measures of central tendency

When you complete this section you should be able to:

- find the measures of central tendency from frequency tables.
- compare measures of central tendency for different sets of data.

Key words

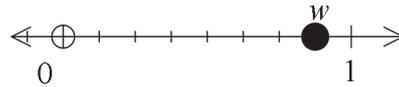
- frequency table

Warm-up 4

1. Circle the triangular numbers. 5, 10, 15, 20, 25
2. $66 \div 11 =$ _____
3. The temperature was minus 3 degrees but it went up 5 degrees.

What is the new temperature? _____

4. Express the value of w as a fraction.



5. $\frac{7}{4} \times 20 =$ _____

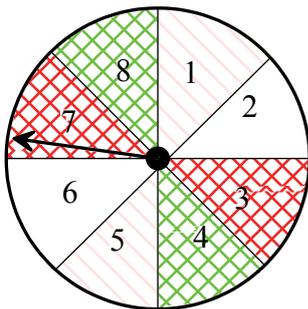
6. $506 \div 10\,000 =$ _____

7. $7.254 \div 2 =$ _____

8. Write 230% as a fraction. _____

9. Complete by writing the next number: 1031, 1016, 1001, _____

- 10.



Determine the probability the spinner will land on a factor of 7.

Express your answer as a fraction.

Review 4

Example

The number of mistakes by various students on a mental test is given below.

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

- (a) Find the mode, range, mean and median of the scores.
 (b) Complete the following frequency table.

Number of mistakes	Frequency

- (c) Comment on what the difficulties might be in finding the mode, range, mean and median of the scores if only the table was given.

Solution

- (a) Mode = 3, range = 4, mean = 3, median = 3

(b)

Number of mistakes	Frequency

- (c) The mode and range should be easy to find but the mean and median would be harder to find. I would wonder how we get the total of all the scores and how we find the middle score.

1. If possible, find the outlier in each of the following sets of data.

(a) 0.1, 83.2, 54.3, 62.2, 70.8 _____

(b) 9, 8, 7, 7, 6, 8, 5, 8, 78, 7, 7 _____

(c) 15 g, 23 g, 17 g, 19 kg, 12 g, 18 g _____

2. Write down the mode and range for each set of scores in the following frequency tables.

(a)

Score	Frequency
10	5
11	8
12	19
13	17
14	12

(b)

Score	Frequency
25	23
26	7
27	6
28	23
29	8
30	6

Focus problem 4



The physical education teacher set up a target with six sections and the students had to kick a ball at it. The students got the score indicated on the section the ball first hit. The scores were recorded in a **frequency table**.

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

1. (a) How many students were in the class? Explain how you found the answer from the frequency table.

- (b) What was the range of the scores? Explain how it can be found from the table.

- (c) What was the mode of the scores? Explain how it can be found from the table.

- (d) What does the mode tell you about the results of the kicks?

2. You already know that the median is the middle score when they are in order of size.

(a) If all the scores were written out in a list what would be the median?

(b) How could you find the median score in this **frequency table**?

Hint: Think of these questions to help you.

What shows that the scores in the table are already organised in order of size?

How many scores are less than four?

(c) Explain why the following reasoning to find the median is incorrect.

‘There are six numbers in the score column so the median will be the middle of the numbers from 1 to 6. This will be 3.5 because 3 and 4 are the two middle numbers.’

3. (a) What is the total of all the scores? _____

(b) How could the numbers in the frequency table be used to quickly find the total?

(c) Find the mean of all the scores, correct to one decimal place. _____

4. Summarise how the mean, median, mode and range can be found for numerical data that is given in a **frequency table**.

Mean:

Median:

Mode:

Range:

5. Use your methods for finding the mean, median, mode and range for numerical data with the following frequency table.

Score	Frequency	
1	2	
2	6	
3	14	
4	4	
5	3	

6. Explain why the table is called a **frequency table**.



Check your work before continuing.

You can easily find which position the median (the middle number) is in a list.

If there are five numbers the middle number will be the third number. $\frac{5+1}{2} = 3$

If there are six numbers the median will be the average of the third and fourth numbers. $\frac{6+1}{2} = 3.5$

If there are n scores in a list then the score in the position given by $\frac{n+1}{2}$ is the median.

Skills development 4

Example

Find the mode, range, mean and median of the scores shown in this **frequency table**. Complete any extra information that may help with the calculations. Provide working to explain your answers.

Score	Frequency	Score \times frequency
5	4	
6	2	
7	5	
8	1	
9	0	
10	2	

Solution

The mode is 7 because it has the largest entry in the frequency column.

The range is 5 because the largest number in the score column is 10 and the smallest is 5.
 $10 - 5 = 5$

To calculate the mean an extra column has been completed to give the total of all the scores.

Score	Frequency	Score \times frequency
5	4	20
6	2	12
7	5	35
8	1	8
9	0	0
10	2	20
Total = 14		Total = 95

The mean is $95 \div 14 = 6.8$ (rounded to one decimal place).

The median will be the score in the position given by $\frac{14+1}{2} = 7.5$. This means it will be the average of the 7th and 8th numbers. The table shows that there are four lots of five then two lots of six, which accounts for the first six scores. The 7th and 8th scores will both be 7 so the median is 7.

1. The following **frequency table** is being used to find the median of the scores in the table. Some working has been started for you.

Score	Frequency	Listed scores
2	2	2 , 2
3	4	3, 3, 3, 3
4	5	4, 4, 4, 4, 4
5	3	5, 5 , 5

- (a) After a quick look at the table what do you think the median might be? _____
- (b) Continue the crossing out of scores from the bottom and the top until you get to the middle.
- (c) What is the value of the median? _____
- (d) Which position in the list of scores is the median? _____
- (e) How many scores are represented in the table? _____
- (f) Show that the position of the median is given by $\frac{n+1}{2}$ where n is the number of scores in the table.

- (g) Calculate the mean of the scores, correct to one decimal place. Explain two ways you could find the mean from the table.

2. Determine the position in a list of ordered numbers for the median if there are:

- (a) 15 numbers _____
- (b) 25 numbers _____
- (c) 8 numbers _____
- (d) 100 numbers _____

3. Find the mean and median of the numbers in these **frequency tables**. Extra columns have been added if you need to use them.

(a)

Score	Frequency	
1	3	
2	5	
3	4	
4	3	
5	1	

(b)

Score	Frequency	
21	1	
22	3	
23	5	
24	0	
25	2	

4. If Joel said that the mean of the scores, displayed in the table below, is 25, how would you know for sure that he had made a mistake?

Score	Frequency
12	2
13	7
14	5
15	0
16	1



Check your work before continuing.

5. Calculations from graphs

When you complete this section you should be able to:

- find the measures of central tendency and outliers from data presented in graphs.

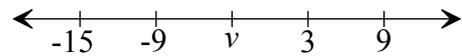
Warm-up 5

1. Express 24 as product of primes. _____

2. $38 + 346 =$ _____

3. What is the missing number?

$v =$ _____



4. $\frac{3}{6} + \frac{5}{6} =$ _____

5. Find six-fifths of 25. _____

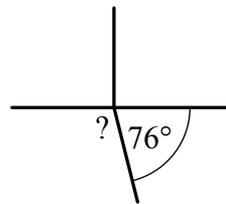
6. $400.2 \text{ mm} =$ _____ m

7. $7 + 8 \times 5 =$ _____

8. Write $4\frac{3}{8}$ as decimal. _____

9. Complete by adding the next number: 109.8, 105.5, 101.2, _____

10. Determine the size of the missing angle.



Review 5

Example

The areas, in hectares, of blocks of land for sale in a new housing area have been advertised as follows:

0.5, 0.6, 1.1, 1.3, 0.7, 0.7, 5.2, 1.2, 0.5, 0.5, 0.6

- Find the outlier in the areas.
- Does the outlier appear to be a genuine area or an error in the list?
- Calculate the mean and median with the outlier.
- Calculate the mean and median without the outlier.
- Comment on whether the outlier should be eliminated or not, when using the mean or median as a selling point.

Solution

- The outlier is 5.2 hectares.
- It is reasonable to believe that there is one very large block along with the smaller ones so it may be a genuine area.
- With outlier: median = 0.7 hectares, mean = 1.17 hectares (2dp).
- Without outlier: median = 0.65 hectares, mean = 0.77 hectares.
- The outlier is not really typical of the other areas. It would not really be representative to use it in the calculations of the mean of the block sizes. The median could be used with or without the outlier as it doesn't have a big effect on the answer.

- The latest power bill was delivered to the people living in a block of flats. The totals of the bill, to the nearest dollar, are shown below.

435, 324, 415, 260, 24, 356, 288, 420, 407, 378, 299, 400

- Find the outlier in the bills. _____
- Does the outlier appear to be a genuine amount or an error in the bill?

- Calculate the mean and median with the outlier included.

- (d) Calculate the mean and median without the outlier included.

- (e) Comment on whether the outlier should be eliminated or not, when using the mean or median to discuss the typical power bill for the people in the flats.

2. A smaller block of flats had the following power bills delivered:

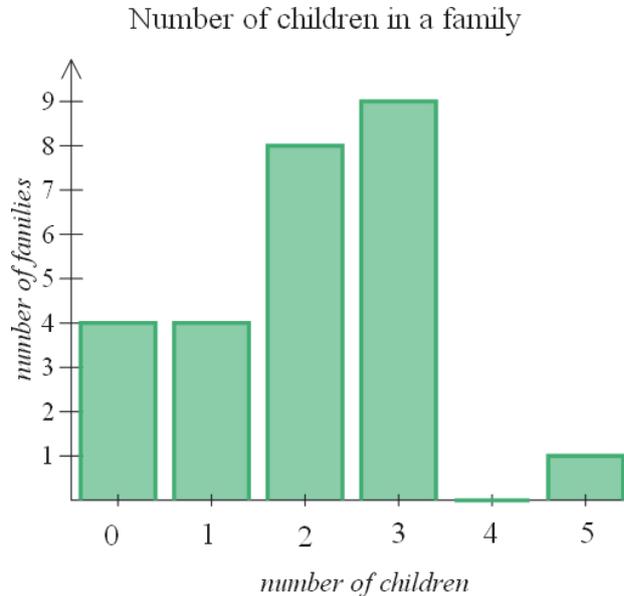
\$467, \$388, \$19, \$296.

- (a) Find the outlier. _____
- (b) Calculate the mean and median with and without the outlier.

- (c) Explain why the outlier had a bigger influence on the different calculations in this case than on the bills in question 1.

Focus problem 5

Becky and Derain completed a survey to find the number of children in each family in the street where they live. They then used their computer to draw the column graph for the data.



First they present their results to their class. Their next task was to tell the class how the four measures of mode, range, median and mean could be found.

Becky explained that the mode and the range are pretty easy to find from the graph.

- (a) What explanation would you give the class to find the mode and the range from the graph?

- (b) What are the mode and the range?

- (e) Write down any other information found in the survey that can be concluded from the graph.

- (f) Would the family with five children be considered an outlier in this case? Explain your answer.

The median of a set of results is used by the Australian Bureau of Statistics when they report comparisons between different areas of the state. Here is an example.

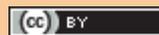
MEDIAN AGE in Western Australia 2010

WA had a median age (the age at which half the population is older and half is younger) of 36.2 years at June 2010. Females had a higher median age (37.0 years) than males (35.4 years). The median age of people living in Perth SD (Statistical Division) (35.8 years) was lower than for people living in the remainder of WA (37.3 years). For the LGA (Local government area) of Perth the median age was younger again at 32.8 years.

The Local government areas with the highest median ages in WA were the popular retirement and sea-change/tree-change locations of Denmark (47.1 years) and Bridgetown-Greenbushes (46.0) in South West SD, and York (45.2) and Toodyay (44.4) in the western regions of Midlands SD. Murray and Mandurah, other popular retirement destinations to the south of Perth SD, also had high median ages (43.6 and 43.5 respectively).

Halls Creek (27.0 years), Derby-West Kimberley (29.5) and Wyndham-East Kimberley (30.0), all located in the state's far north, had the lowest median ages in the state.

Based on Australian Bureau of Statistics data.



Summary

- From the focus problem you should have found that the mode is easy to find from a column graph because it is shown by the tallest column.
- The range is also easy because it is the highest value on the horizontal axis minus the lowest value.
- The median and mean can't be found without some calculation.
- Looking at the graph gives you a good idea of how well the scores are clustered around the centre of the data. However it would be necessary to put the information from the graph into a table or a list to determine the mean and median.
- To calculate the median you will need to know how many scores there are altogether. You do this by adding the heights of the columns.
- To find the position of the median you will use $\frac{\text{number of scores} + 1}{2}$.
- After you have the position you have to count through the scores until you get to that position in the list of scores.
- To calculate the mean you will also need to know the total of all the scores. This is found by multiplying each score by the height of the column then adding up the results. This total has to be divided by the number of individual scores to get the mean.

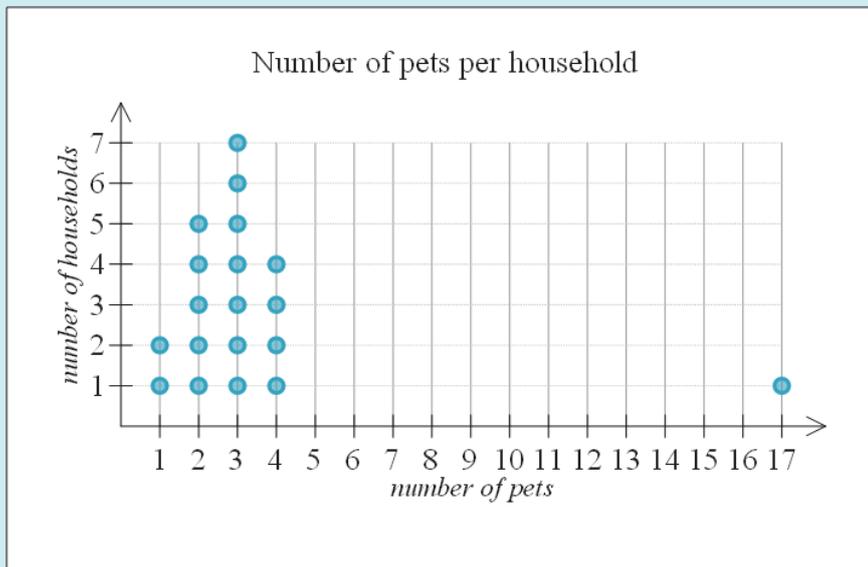


Check your work before continuing.

Skills development 5

Example

The number of pets per household for a class is shown in the graph below.



Without using a calculator, write as much as you can about the range, mode, mean, outliers and median of the data shown in the graph. Give reasons for your answers.

Solution

The range of the number of pets is 16, by taking 1 from 17.

The possible outlier is 17. It is vastly different to the other numbers of pets for this class.

The mode, or most common number of pets, is 3 because there are more dots for 3 than for the other numbers.

There are 19 households represented by the data. $\frac{19+1}{2} = 10$ so the median will be the 10th value if the outlier is used. The 10th number is a 3 so the median is 3.

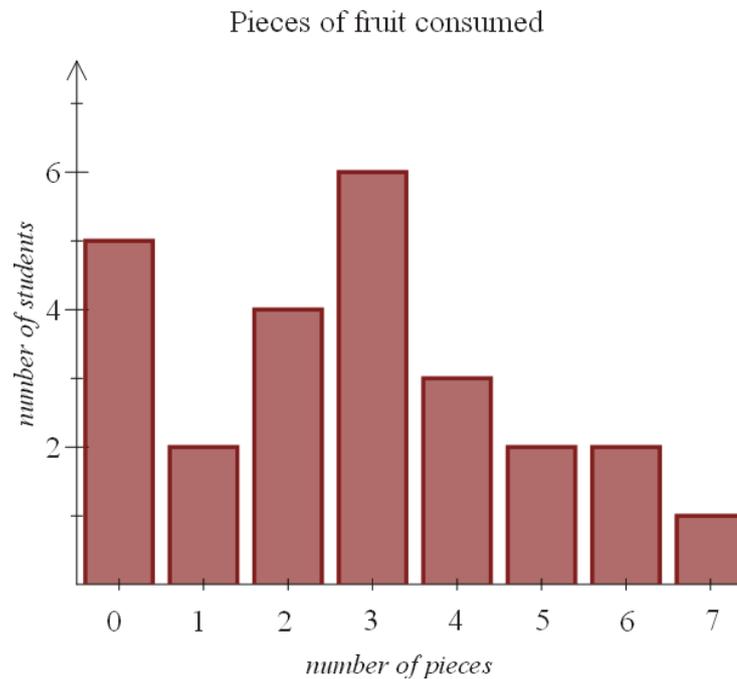
The median would still be the same even if the value of 17 was not considered. It would be the average of the 9th and 10th number, which is still 3.

The median does give a good idea of the typical number of pets.

The mean would be the total of the number of pets divided by 19 if the outlier is used. I would expect the mean to be a bit bigger than three.

The total would be divided by 18 to find the mean if the outlier is not used. It should be a little bit less than three.

1. The graph below shows the number of pieces of fruit each student ate over the five days of a normal school week.



- (a) What is the mode and range for the data?

- (b) How many students were in the class?

- (c) What is the position of the median of the data?

- (d) What is the median number of pieces of fruit eaten?

- (e) What is the total number of pieces of fruit that were eaten? _____

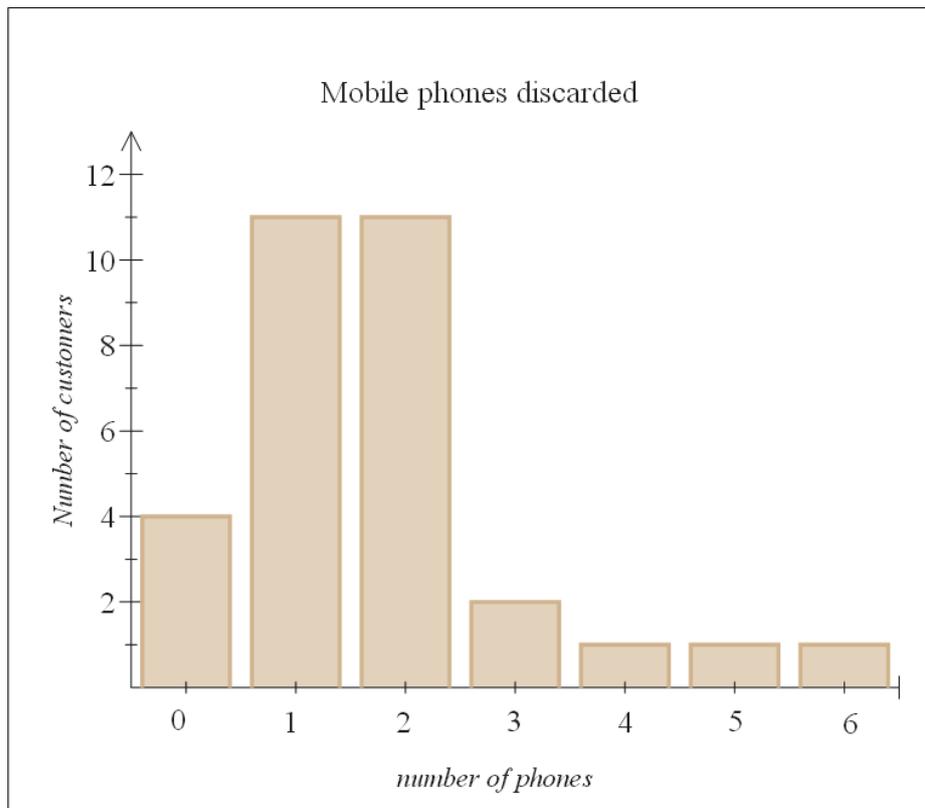
- (f) Determine the mean number of pieces of fruit eaten in the week. Show your working.

- (g) Dietary guidelines indicate that children should eat some fruit every day. How many students appear to have met the guidelines?
-
- (h) The ideal situation is that at least two pieces of fruit are eaten each day. Have any of the students met this target?
-
- (i) If you were in the class where would your consumption of fruit fall? Would you be above or below the mean?
-

Would you meet the dietary guidelines?

2. Australia, along with many countries, has a huge problem with e-waste. E-waste is short for electronic waste which is caused by unwanted electronic goods. Electronic goods are hard to recycle but they do contain heavy metals and toxins that should not be added to landfill. 90% of the components of a mobile phone could be recycled.

A survey was carried out of the 31 people who were looking at new mobile phones in a dealer's shop. They were asked to record the number of mobile phones they had discarded in the last five years. The graph below shows the results.



- (a) How many mobile phones were discarded by this group of people? _____

- (b) What are the mode and the range of the number of discarded phones?
-
- (c) Calculate the mean and the median of the number of phones that were discarded.
-
-
- (d) About 21% of a mobile phone is made from recyclable metal. How many phones from those discarded would that be equivalent to?
-
-
- (e) In Australia it is approximated that 7% of old mobile phones are thrown in the landfill waste. How many of the phones in this survey would that be?
-
-

3. The ages of the people waiting at a doctor's surgery are shown in the following stem-and-leaf plot.

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

- (a) The process of finding the median has been started by crossing out entries. Complete the crossing out to find the median.
-
- (b) Explain why it is useful to have the stem-and-leaf plot ordered to find the median.
-
-
- (c) Calculate the mean of the ages. _____

- (d) Explain how the range of the scores could be easily found from a stem-and-leaf plot. What is the value in this example?

- (e) This set of data does not have a mode. Explain that in terms of the ages of the people in the waiting room.



Check your work before continuing.

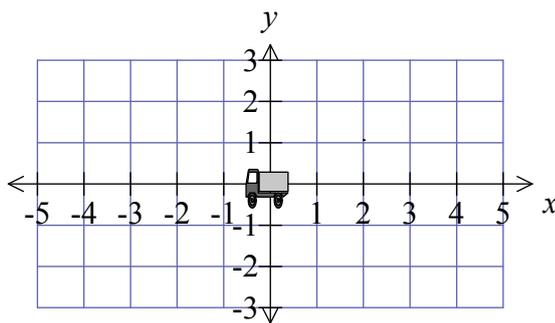
6. Use in everyday life

When you complete this section you should be able to:

- use the measures of central tendency and spread to answer problems related to everyday life.

Warm-up 6

- $90.9 \times 10 =$ _____
- $929 - 419 =$ _____
- The temperature is minus 5 degrees.
How much will it need to increase to get to zero degrees? _____
- $2\frac{1}{5} - \frac{3}{5} =$ _____
- $\frac{7}{5} \times 15 =$ _____
- $1.39 \text{ g} =$ _____ mg
- $30 \div 6 + 3 =$ _____
- Write 2.4 as a percentage. _____
- $10\frac{5}{10}, 9, 7\frac{5}{10},$ _____
- The truck is shown at $(0, 0)$.
If the truck moves 7 units left and 8 units up, where will it then be? _____



Review 6

Example

Marks out of 100 in a test

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

Find:

- the mode and the mean, two of the measures of central tendency
- the median and explain two different ways it could be found
- the measure of spread, the range
- any outliers. Explain how the calculations above are affected by the outlier.

Solution

- The mode is 66 and the mean is 71.
- The median is 72.5 because it is the average of the two middle numbers.

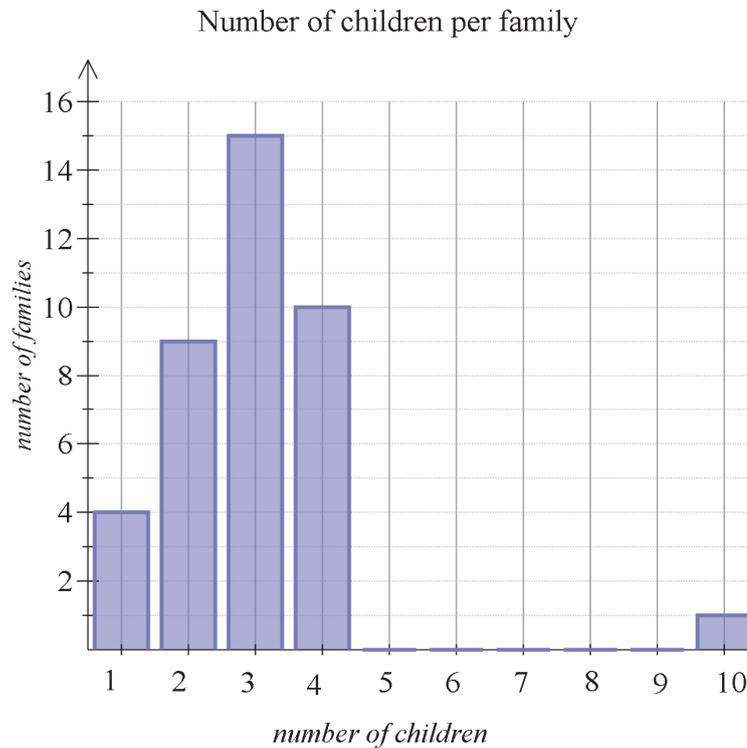
Method 1: Cross out scores from the top and bottom in pairs until you get to the middle.

Method 2: Find the position of the median by using $\frac{\text{number of scores} + 1}{2}$. In this case it is $\frac{12 + 1}{2} = 6.5$ so the median is the average of the 6th and 7th mark when in order.

- The range is 59 because $89 - 30 = 59$.
- The mark of 30 is an outlier because it is way below the other marks.

The outlier doesn't affect the mode and only makes a difference of 0.5 in the median. The range is most affected by the outlier and the mean is lower than it would be without the outlier.

1. In the example the mean is almost 75 if the outlier is left out. Explain why leaving the outlier out of the calculations gives a better idea of the class results.



- 2.
- How many families are represented in the graph? _____
 - What is the total number of children? _____
 - What is the most common number of children? What do we call this value?

 - What is the position of the median? _____
 - Which number of children is the median? _____
 - Which value is an outlier? _____
 - Calculate the mean of the number of children with the outlier removed.

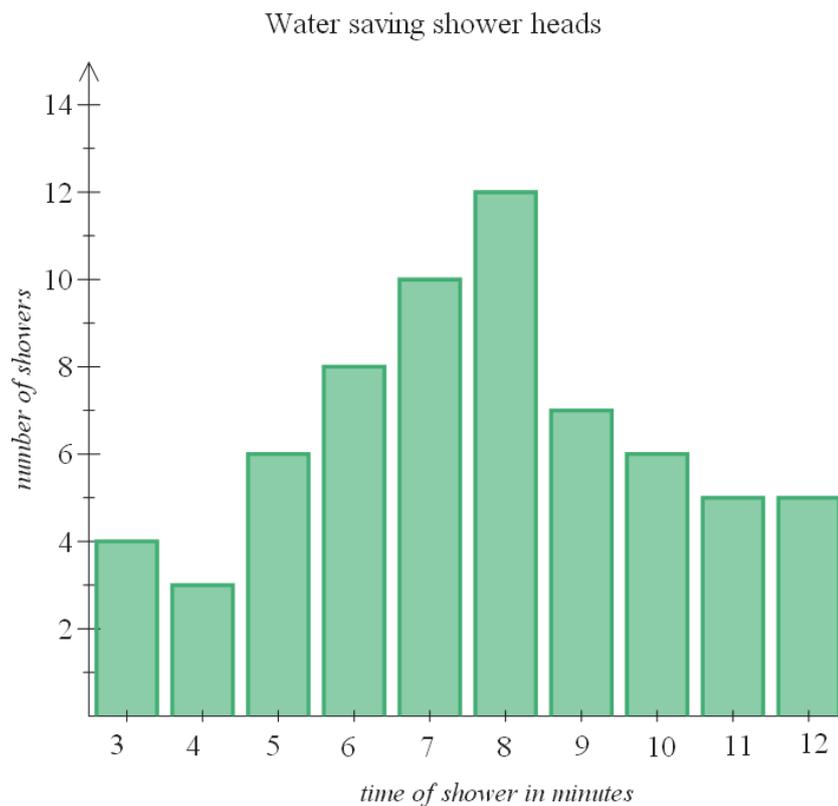
 - Without doing any more calculations a comment on whether the outlier would make a small or big difference to the mean if it was included.

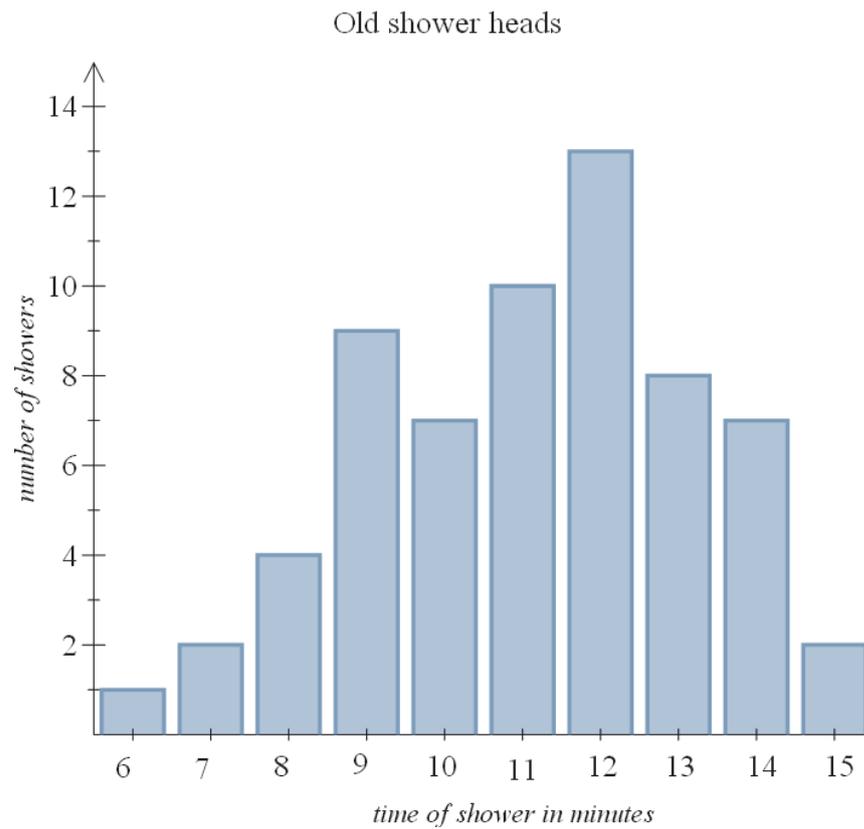
Focus problem 6

Saving water



The students in a class were asked to record how long they took in the shower each day. The results were then divided between those that had water saving shower heads and those that had old style ones. The results are graphed below.





- (a) What do you notice about the length of showers of people with water saving shower heads compared to those without?

- (b) What is the most common length of shower in each group and overall?

The average flow for water saving shower heads is eight litres per minute and the old style shower heads have a flow of 15 litres per minute.

- (c) How many litres of water are used for a 12 minute shower using each type of shower head?

- (d) A normal plastic bucket holds nine litres of water. How many buckets of water would be used for each of the showers in (c)?
-
-
- (e) What was the total number of showers in each set of data?
-
-
- (f) How many fewer litres of water per minute were used by the group that used water saving shower heads? Show your working
-
-

Another way to save water in a shower is to catch, in a bucket, the water that normally goes down the drain whilst you are waiting for the water to heat up. This water could be used to water plants in the garden or for other simple uses of cold water around the house.

- (g) Record how long you take to shower over three days.

If you took longer than 10 minutes and were able to reduce the shower by three minutes how much water would you save?

- (h) Check how long it takes for the water to get hot enough to use. How many nine litre buckets of water could you collect? How could this waste water be used in your household?
-
-
-
-
-

This problem illustrates how you can apply the knowledge you have about the measures of central tendency and spread to an everyday situation.

It shows ways that water, which is a scarce resource in many areas, could be saved. This could also lead to reduced household water costs.



Check your work before continuing.



Interesting water fact

Did you know that Australia is the driest inhabited continent on Earth?

It also has one of the highest consumption rates of water per head of population in the world.

Skills development 6

1. Leaving lights on when not in the room

A number of households were surveyed about the number of hours that lights were left on in empty rooms. The total was calculated over a day by adding up the time each light bulb was left on, and then rounding to the nearest hour.

Each family had already changed from the old incandescent bulbs to the new low energy bulbs to save power.

Number of hours	Number of households	
5	2	
6	7	
7	12	
8	11	
9	19	
10	16	
11	11	
12	0	
13	0	
14	2	

(a) What is the total number of hours the lights were left on?

(b) Write down the range, mode, median and mean of the data.

(c) If each light bulb uses 25 watts (W) of power per hour what is the total power, in watts, used by these lights in one day?

2. Greywater is the waste water from water use in a household. It may come from the washing machine, showers, baths and kitchens. Using greywater reduces the demand on the dwindling water resources. Two of the most common uses of greywater are to flush toilets or to water lawns and gardens.

The following stem-and-leaf plot shows the average amount of greywater, in litres, produced per person per day in a number of households.

Greywater production per person (litres)

Stem	Leaf			
11	4	5	5	
12	0	2	2	7
13	3	5	5	5
14	2	8		

- (a) Why is a stem-and-leaf plot more suitable for this data than a column graph?
- _____
- _____
- (b) What is the range of greywater production per person? _____
- (c) Find the mode of the data. _____
- (d) Find the median amount of greywater production per person.
- _____
- _____
- (e) Find the mean amount of greywater production per person.
- _____
- _____
- (f) If each household could reduce their greywater production by 20 litres per person per day how much water would be saved?
- _____
- _____
- (g) How much greywater would you estimate your household produces per person per day?
- _____
- _____

3. More and more households are turning to solar power to help reduce their power bills.

A solar photovoltaic panel system can be installed on the roof of the house. One of the most popular sizes is the 2.0 kW system.

One householder from each capital city was asked to give their average daily production of power. The amount produced depended on the length of the day and the amount of sunlight.

The results are shown in the table below.

City	Production in kW
Adelaide	8.5
Brisbane	8.4
Darwin	8.9
Hobart	7.2
Melbourne	7.3
Perth	8.6
Sydney	7.8

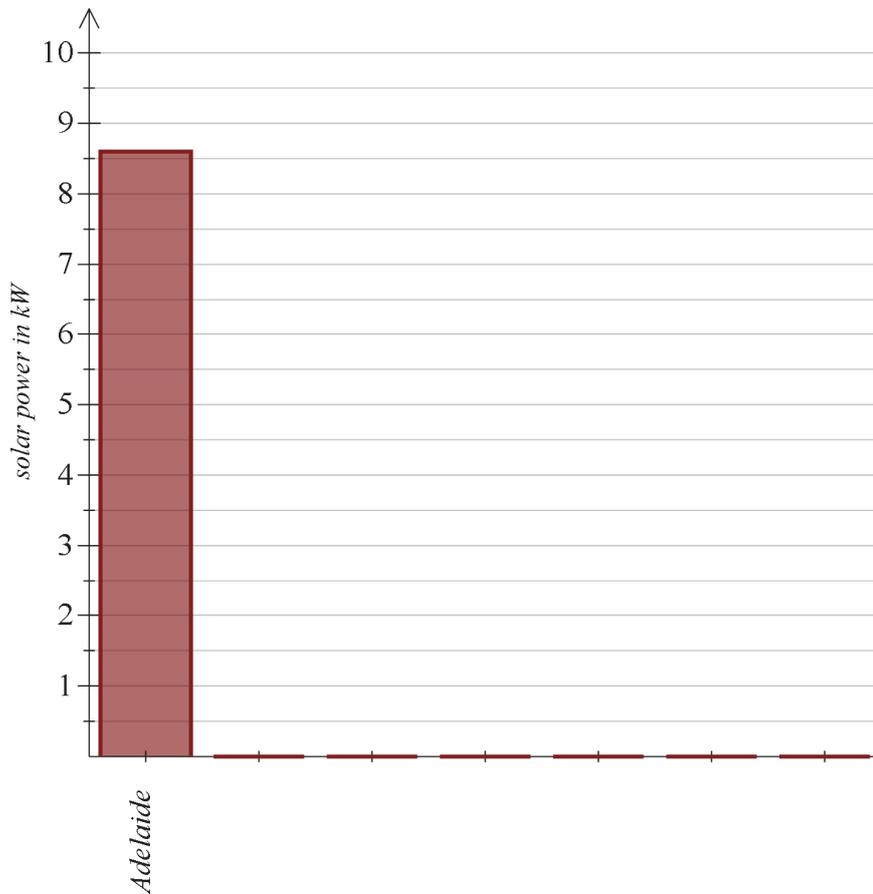
- (a) Calculate the range of the production of power. _____

- (b) Find the mean and the median of the production.

- (c) Are there any outliers in this set of data? _____

- (d) Put this data into a stem-and-leaf plot.

(e) Complete the column graph below for the data



(f) (i) What are the advantages of using a column graph for this data?

(ii) What are the advantages of using a stem-and-leaf plot for this data?



Check your work before continuing.

7. Calculations with provided data

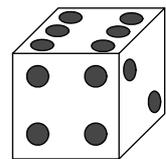
When you complete this section you should be able to:

- use the measures of central tendency and spread to interpret statistics provided by government and other agencies.

Warm-up 7

1. $21.07 \div 10 =$ _____
2. $642 \div 6 =$ _____
3. The temperature is 2 degrees. How much will it need to decrease to get to minus 3 degrees? _____
4. $\frac{1}{2} + \frac{1}{6} =$ _____
5. $\frac{2}{7} \times 21 =$ _____
6. $0.0003 \text{ L} =$ _____ mL
7. $21 \div 7 \times 3 =$ _____
8. Find 30% of \$75 _____
9. Describe the rule for the following pattern.
110, 106, 102, 98, 94, ...

10. A six-sided die is rolled.
Express as a fraction, the probability that it lands on a prime number.



Focus problem 7

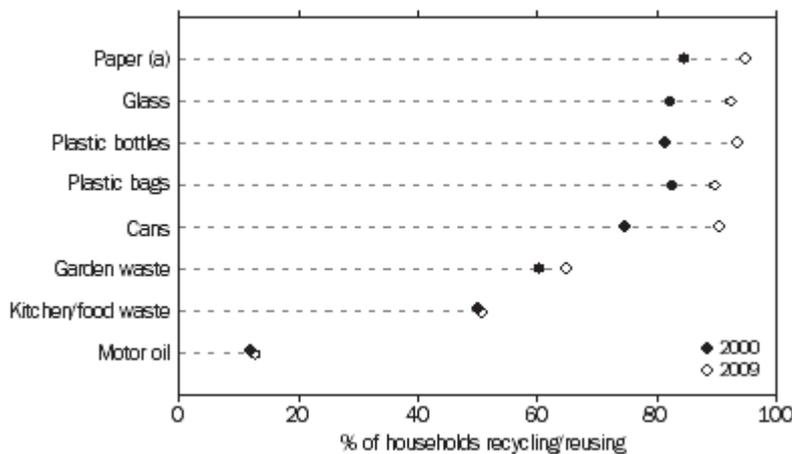
The following information about household recycling is based on information from the Australian Bureau of Statistics.

The recycling activities of households grew extensively between 1996 and 2009.

In March 1996, 91% of Australian households said they practised some form of waste recycling and/or reuse activity.

By March 2009, almost all households (99%) reported that they recycled and/or reused.

WASTE ITEMS RECYCLED AND/OR REUSED BY HOUSEHOLDS



(a) Includes cardboard and newspapers.

Source: ABS, 2009, *Environmental Issues: Waste Management and Transport Use March 2009* (cat. no. 4602.0.55.002).

Based on Australian bureau of Statistics data



1. What is the heading of the graph?

2. How many different types of waste items are shown in the graph?

3. What does the symbol \blacklozenge represent on the graph?

4. From the graph which type of waste items has the mode for the year 2000?

5. The table below shows the approximate percentages given on the graph for 2009.

Waste item	Percentage of households recycling/reusing
Paper	95
Glass	93
Plastic bottles	94
Plastic bags	90
Cans	91
Garden waste	66
Kitchen/food waste	50
Motor oil	13

- (a) What is the mean of the percentages?
- _____
- (b) What is the median of the percentages?
- _____
6. Look at the graph and compare the results for the two years. What would be the approximate mean and median for the year 2000 based on the results for 2009?
- _____
- _____



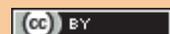
Check your work before continuing.

Plastic bag facts (2005)

Most of the supermarket plastic bags are made of high density polyethylene (HDPE).

- * Australians use 3.92 billion plastic shopping bags per year.
- * Nearly half a million plastic bags are collected on Clean Up Australia Day each year.
- * It takes only four grocery shopping trips for an average Australian family to accumulate 60 plastic shopping bags.
- * Plastic bags are produced from polymers derived from petroleum. The amount of petroleum used to make a plastic bag would drive a car about 11 metres.
- * In 2005, Australians used 192 HDPE bags per capita.
- * Only 14% of HDPE plastic carry bags are returned to major supermarkets for recycling.

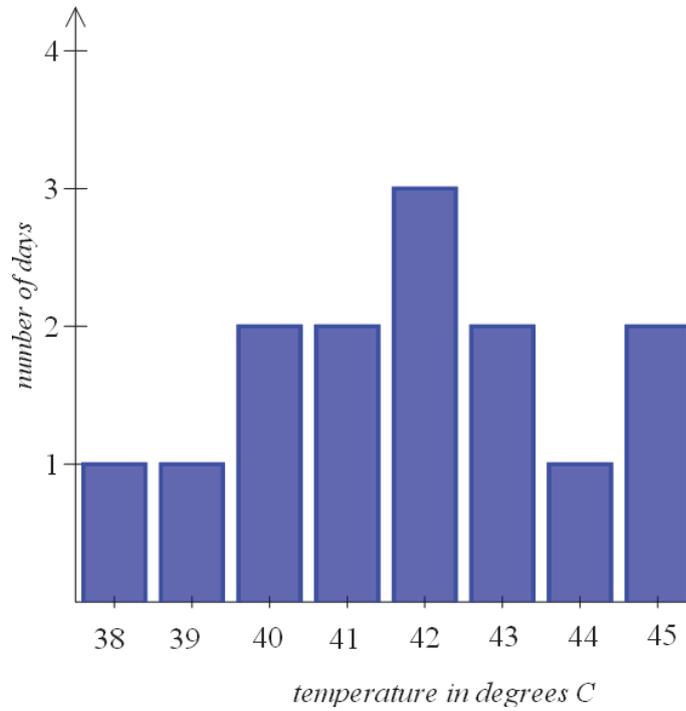
Based on Australian Bureau of Statistics data.



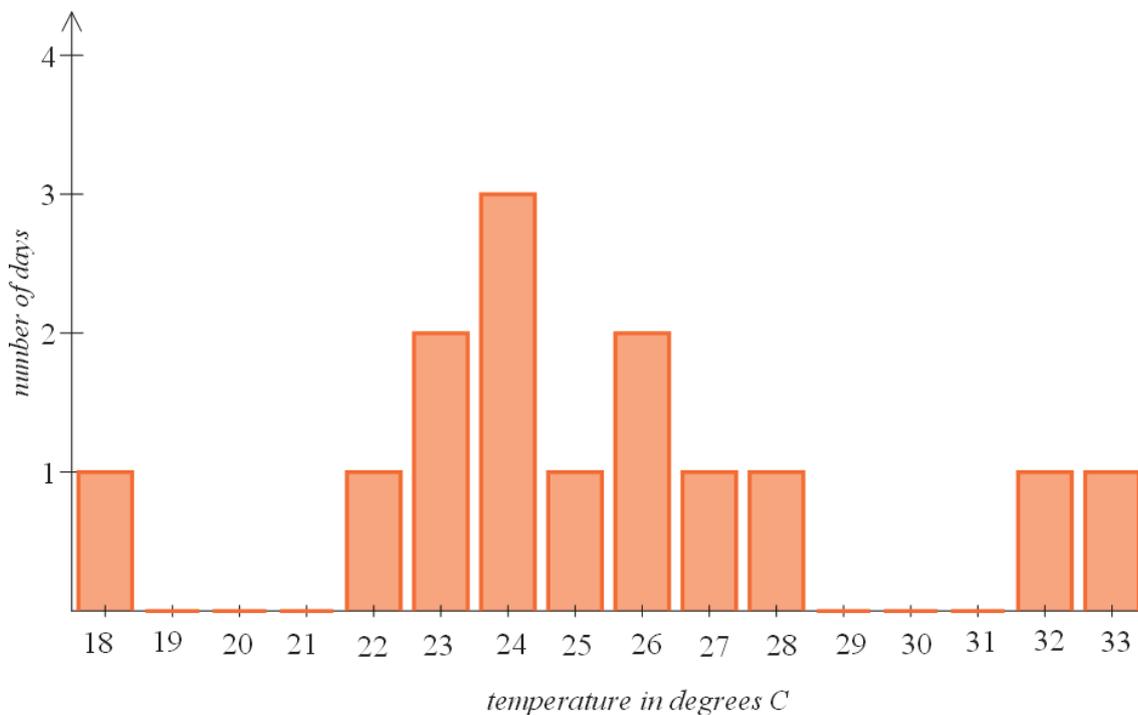
Skills development 7

- The temperatures in two West Australian towns were recorded and graphed for a two week period one February. The graphs are shown below.

Maximum temperatures at Marble Bar



Maximum temperatures at Albany



(a) What was the range of the temperatures for each town?

(b) Which town had the greatest range? How can this be seen on the graph?

(c) How is the mode of the temperatures found from the graphs?

(d) What is the mode of the temperature for each town?

(e) Are there any outliers in either graph? Give a reason for your answer.

(f) Calculate the mean maximum temperature for Marble Bar.

(g) Find the median maximum temperature for Albany.

2.

Median age at marriage, Australia.

Year	1988	1998	2003	2004	2005	2006	2007
Males	27.8	29.8	31.2	31.5	31.5	31.6	31.6
Females	25.4	27.7	29.1	29.2	29.3	29.3	29.3

Based on Australian bureau of Statistics data



(a) Why is the median used to compare the age at marriage for the various years?

(b) Look at the top row of the table. What do you notice about the values in there?

(c) From this table what change has occurred in the median age over the years given for:

(i) males _____

(i) females _____

8. Summary

- The mode of a set of data is the most common entry in the set.
- If two scores are the mode then the set of data is said to be bimodal.
- In a frequency table the mode is the score that has the highest frequency.
- On a column graph the mode is the score that has the highest column.
- The range of numerical data is the difference between the highest score and the lowest score.
- In a frequency table the range is the highest score minus the lowest score.
- On a column graph the range is the highest score minus the lowest score that have columns.
- The median of a set of numerical data is the middle number when the scores are in order of size.
- If there is an odd number of scores the median will be one of the scores.
- If there is an even number of scores the median will be the average of the two middle scores.
- For a frequency table the position of the median can be found by using $\frac{\text{number of scores} + 1}{2}$.
- The mean of a set of numerical data is the total of the scores divided by the number of scores. It is usually called the average in everyday life.
- In a frequency table, to find the mean add another column for the frequency multiplied by the score. The total of this column divided by the total of the frequency column is the mean.

$$\text{mean} = \frac{\text{total of (frequency} \times \text{score)}}{\text{total of frequency}}$$

- On a column graph the mean will have to be calculated by using the same method as for a frequency table.
- A score is an outlier if it lies outside the general group of the other scores. It can be a higher or a lower score.
- An outlier can be the result of an error in recording or an unusual value.
- An outlier will have a large effect on the range and can significantly change the mean.
- The mode is not affected by an outlier and the median can only be slightly affected by the outlier.

9. Review tasks

The following tasks will assist you to consolidate your learning and understanding of the concepts introduced in this resource, and assist you to prepare for assessments.

Task A

Name: _____

Suggested time: 50 minutes

Actual time taken: _____

Instructions

Complete this work on your own.

You may use a calculator, but show how you got your answer.

Attempt every question. Take as long as you need and record the time in the space provided above after you have finished.

1. Match the term to the description by drawing a line between those that match.

Term
outlier
mode
mean
range
median

Description
the average of the scores
the difference between the highest and the lowest score
a score that is vastly different to the other scores
the middle score when they are in order of size
the most common score

2. Calculate the range, mode, mean and median for each of the following sets of numerical data.

- (a) The heights, in metres, of a stand of Mountain Ash trees in the Kinglake National Park, Victoria

90.6, 91.4, 90.7, 88.8, 87.5

86, 87.5, 72, 93.1, 87.2

Range: _____

Mode: _____

Mean: _____

Median: _____

(b) Results in a class test, out of 10

Score	Frequency	
4	2	
5	6	
6	4	
7	2	
8	1	

Range: _____

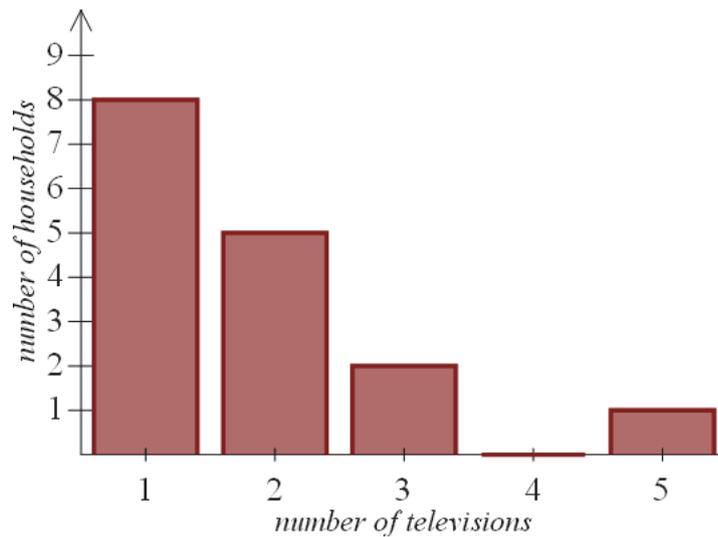
Mode: _____

Mean: _____

Median: _____

(c)

Numbers of televisions per household



Range: _____

Mode: _____

Mean: _____

Median: _____

3. Two suburbs in Perth have the median price of the houses sold in a three month period listed. One is listed as \$340 000 and the other as \$1 170 000. What conclusions could you make about the two suburbs given that information?

4. The median sale price of houses in a town was listed as \$585 000 one month and six months later it was listed as \$375 000. Could you come to the conclusion that the value of houses in the town had dropped by over \$200 000 in the two months? Give a reason for your answer.

5. In a lift in a large office block there is a notice that says 'Maximum load, 21 persons, 1430 kg'. If the lift was full and the 21 people had a combined weight of 1430 kg what would the mean weight of the 21 people be?

6. On a small plane the weight of the passengers is important as it has a maximum load it can carry. There are 11 passengers waiting for a plane. The median weight is 67 kg and the mean weight is 72 kg. If the maximum load of passengers is 760 kg will all the passengers be able to travel on this plane? Explain your answer.

7. Medical researchers carried out a set of trials of a new drug to kill skin cancer cells. The percentage decrease in the size of the cancer cells after one month on the trial drug treatment was recorded for a number of patients. The results were as follows.

11, 15, 12, 16, 11, 73, 18, 17, 15, 18, 16, 14, 11, 13, 13

- (a) What is the range of the percentage decreases? _____
- (b) Calculate the mode, median and mean of the percentage decreases.

Mode: _____

Median: _____

Mean: _____

- (c) Which of the measures of central tendency from (b) gives the best indication of the results? Explain your choice.

- (d) Which result is an outlier? _____

- (e) Comment on whether you think the outlier should be eliminated from the results as an error.

8. Earthquakes are recorded by a seismometer. The scale of their magnitude used to be measured by the Richter Scale. The scale used now is called the Moment Magnitude Scale (MMS) and is only slightly different to the Richter Scale.

A sample of the scales and their possible effect is found below.

9.0+	Devastating to a very large area
5.0	Can cause a lot of damage to buildings that are not well constructed
3.0	Can often be felt as a very slight shaking
1.0	Usually underground and cannot be felt

The magnitudes of 20 recent earthquakes around the world are listed below.

2.7, 3.1, 4.5, 2.6, 4.6, 4.1, 4.1, 4.7, 5.2, 4.6

2.8, 5.0, 3.7, 6.2, 4.7, 4.8, 2.6, 4.5, 6.8, 6.7

- (a) Display the data on a stem-and-leaf plot.

- (b) Why is this type of plot better for this data than a column graph?

- (c) Find the median magnitude of the data.

(d) Calculate the mean of the magnitudes.

(e) The earthquake that devastated the town of Meckering, 130 km east of Perth, in 1968 had a magnitude of 6.9. Comment the magnitude of that earthquake compared to the mean and median of the earthquakes in the data.

9. What are the errors in the following statements? How can they be corrected?

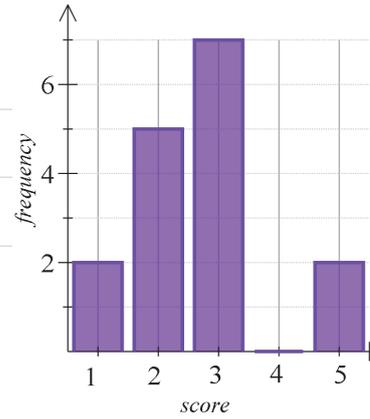
(a) The median of the set of scores 5, 7, 3, 8, 4 is 3.

(b) The mean of the set of scores 1.1, 1.2, 1.3, 1.3, 1.4, 1.5 is 13.

(c) There is no mode for the data in the plot below.

Stem	Leaf			
12	0	3		
13	0	3	4	0
14	1	6	7	
15	4			

(d) The outlier in the data shown in the graph is 5.



10. A mining company advertised positions in their company on their website.

Median salary for positions	
Operations manager	\$107 000
Mechanical engineer	\$80 000
Project manager	\$118 000
Geologist	\$90 000
IT staff	\$61 000

(a) Angela decided she had the qualifications to apply for one of the IT staff positions and said to her friends 'I will be earning at least \$61 000 per year'. What is wrong with Angela's statement?

(b) If the company only had five positions for geologists what conclusion can be made about the advertised salary?

Task B

Name: _____

Suggested time: 30 minutes

Actual time taken: _____

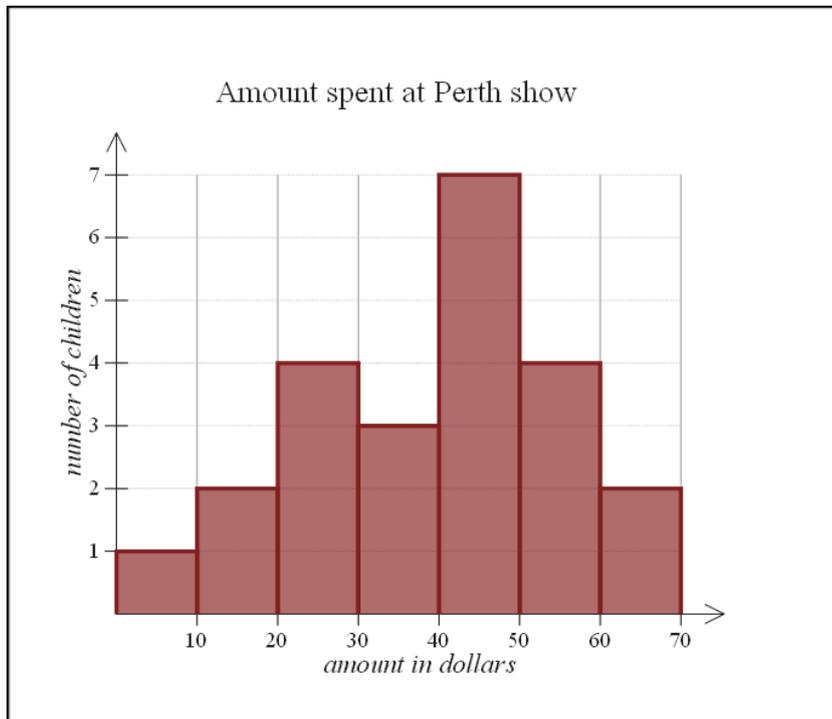
Instructions

Complete this work on your own.

You may use a calculator, but show how you got your answer.

Attempt every question. Take as long as you need and record the time in the space provided above after you have finished.

A group of children were asked how much money, apart from the entrance fee, they spent at the Perth show. The results are shown in the histogram below.



The calculations you have completed so far for the range, mode, median and mean have all been for non-continuous data.

Your task is to decide how those four values can be found, or approximated, for this data. You don't have to find the actual values although an approximation would be a good idea.

Write clear instructions for another student on how you think the four values could be found.

Self-evaluation task

Please complete the following.

How well did you manage your own learning using this resource?

	Always	Usually	Rarely	Not sure
Each section took approximately 45 minutes to complete.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I needed extra help.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I marked and corrected my work at the end of each section.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I made the journal entries and summaries when asked.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I have kept to my work schedule.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

How much mathematics have you learnt using this resource?

	Always	Usually	Rarely	Not sure
<i>Understanding</i>				
I understand how to find the range and mode from a list, table or graph of data.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I understand the difference between the mean and the median.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I understand which score is an outlier.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Fluency</i>				
I can calculate the mode, mean and median accurately.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I know how to find the position of the median.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I correctly named the outlier for a set of scores.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I included correct units in my answers where appropriate.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>Problem Solving</i>				
I am able to solve problems involving mean, median and mode.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I am able to solve problems involving outliers.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I can solve problems relating to data presented in graphs.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Reasoning

I can explain whether a mean or median should be used to solve a problem.

I can explain whether an outlier should be included in data calculations or not.

Write a list of topics for which you need additional assistance.

Solutions

1. Finding the range and mode

Solutions to Warm-up 1

- 1, 2, 4, 5, 10, 20
- 15
- $a = (-7)$ because the numbers are 4 units apart.
- $\frac{5}{4}$ should be circled because $\frac{5}{4} = \frac{10}{8}$ and $\frac{10}{8} > \frac{9}{8}$.
- 7.5 or $7\frac{1}{2}$
- 4.95
- 31.5
- $2\frac{7}{10}$
- 1701
- The triangle

Solutions to Review 1

- (a) (i) 3 (ii) 5 (iii) 5 (iv) 6

(b) (i) 8 (ii) 3 (iii) 6 (iv) 6

(c) (i) 10 (ii) 10 (iii) 9 (iv) $\frac{99}{11} = 9$
- (a) $53 - 2 = 51$

(b) 2

(c) ~~2~~, ~~2~~, ~~6~~, ~~10~~, 21, ~~25~~, ~~34~~, ~~37~~, ~~53~~

Don't forget that you need the middle number when they are in order of size.

The average is the total of the scores divided by the number of scores.



Values crossed off in pairs from top and bottom to get the middle number of 21

- (d) 21.1, rounded to one decimal place. This is usually written as 21.1 (1dp).

Solution to Focus problem 1

What you were asked was to investigate the term 'mode' for numerical data.

The wording of your answers may vary slightly from those given below.

- The mode is the size that was sold the most.
- For Thursday the mode is 8 and for Friday the mode is $6\frac{1}{2}$.

- (c) A good suggestion would be to set up a table with values from the smallest size to the largest size in the first column. There would be a tally column next to it for each day. The sales person could just put a tally mark next to the size they sold.
- (d) The mode would be a good guide as it shows which is the most popular size of shoe.
- (e) New range of prices = $\$210 - \$12 = \$198$.

I remember that this is called a frequency table.



Range = highest score - lowest score

- (f) The instructions may be like those below.

To find the mode you look at the list of numbers and find which number appears more often than any other number. That is the mode.

For the range you find the largest number and the smallest number and subtract the smallest from the largest.

For example, for the numbers 2, 3, 3, 4, 5, 6, 3 the mode is 3 because there are more of them than any other and the range is 4 because $6 - 2 = 4$.

If there are two modes the set of data is said to be bimodal.

Solutions to Skills development 1

1. (a) Range = $85 - 34 = 51$ Mode = 43
 (b) Range = $25 \text{ cm} - 11 \text{ cm} = 14 \text{ cm}$ Mode = 15 cm
 (c) Range = $7.6 \text{ kg} - 2.1 \text{ kg} = 5.5 \text{ kg}$ Mode = 2.9 kg and 2.8 kg

Include the units if they are given in the question.



2. For 3, n , 5, 7, 8 if the mode is 7 then it means there are more sevens than any other number. The value of n must be 7 to make the mode 7.
3. (a) Range = $41^\circ\text{C} - 31^\circ\text{C} = 10^\circ\text{C}$
 (b) Mode = 32°C
 (c) The range would not change with the extra piece of data. The set of temperatures would now have two modes, 32°C and 33°C so it is bimodal.
4. The numbers 5, 6 and 8 all occur three times so no number really occurs more often than any other number because this is just a small set of numbers.
5. At the moment the highest score is 17 and the lowest score is 4. For the range to be 16, n could be 20 since $20 - 4 = 16$, or n could be 1 since $17 - 1 = 16$

2. Finding the mean and median

Solutions to Warm-up 2

1. 3 should be circled.
2. 8
3. $(-3)^\circ$
4. $\frac{15}{5} = 3$
5. 15
6. 10
7. 6.02
8. $\frac{6}{100} = \frac{3}{50}$
9. 10.02
10. 36°

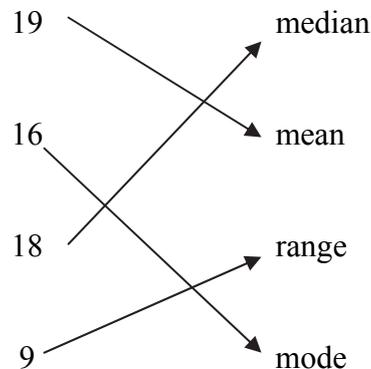
Solutions to Review 2

1. (a) In order the values are: \$8.20, \$8.30, \$8.50, \$8.50, \$9.10, \$9.50, \$9.50.
The middle score is \$8.50.
Both \$8.50 and \$9.50 occur twice so they are both the mode. The set is bimodal.
The average = $\$61.60 \div 7 = \8.80 .
- (b) The distances in order are: 1.99 m, 2.56 m, 2.87 m, 3.05 m, 3.43 m.
The middle score is 2.87 m but there is no mode because they are all individual scores.
The average = $13.9 \text{ m} \div 5 = 2.78 \text{ m}$
2. (a) Range = 7 years – 3 years = 4 years Mode = 6
(b) Range = $33^\circ\text{C} - 26^\circ\text{C} = 7^\circ\text{C}$ Mode = 31°C

Solution to Focus problem 2

You were asked to consider two new terms used in calculations with sets of data. The median strip is in the middle of the road.

(a)



The median is the middle score when they are in order of size.

The mean is the average of the numbers.

- (b) The first set of scores has an odd number of scores so the middle number is easy to find, it is the 5.

The second set of scores has an even number of scores so there are two middle numbers 5 and 7. 6 is the average of 5 and 7. $\frac{5+7}{2} = 6$

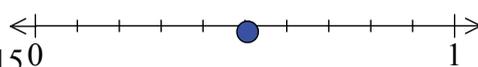
- (c) If the list is long it could be difficult to put the numbers in order first. Finding the middle number would also take time.
- (d) The diagrams show that if you cross out the same number of scores at each end you find the median. The arrows show the position of the median.
- (e) (i) The numbers were not in order of size so 2 is not the median.
 (ii) There is nothing wrong with finding the median here. However we usually order from smallest to largest.
 (iii) 9 is not the average of 8 and 11. The median is $\frac{8+11}{2} = 9.5$.
- (f) I would ask them to line up in order of size from the shortest to the tallest. Then one person from each end of the line can leave until the middle person is left. The height of this middle person will be the median.
- (g) The median is in the middle of a set of ordered numbers and the median strip is in the middle of the road.

Solutions to Skills development 2

- The mean and range may need a calculator but the mode doesn't need a calculator.
- $$\begin{aligned} \text{Range} &= 6.7 - 1.6 \text{ m} & \text{Median} &= \text{average of } 3.6 \text{ and } 3.4 \\ &= 5.1 \text{ m} & &= 3.5 \text{ m} \\ \text{Mean} &= \frac{\text{total length}}{10} & \text{Mode} &= 3.6 \text{ m} \\ &= \frac{36.1}{10} \\ &= 3.61 \text{ m} \end{aligned}$$
- Range = \$2.03 Mode = \$3.20 Median = \$3.20 Mean = \$3.58
(amount has been rounded to the nearest cent)
 - In this case the mode and median are the same but the mode would be used to find the most usual price.
 - The range tells you the difference between the price of the most expensive and the cheapest can.
 - All the values are close to each other although the mean is a bit larger because of the two expensive cans.

3. The effect of outliers

Solutions to Warm-up 3

- The numbers circled should be 4, 6 and 8.
- 84
- The missing number is (-2).
- 
- $272 + 20 = 292$
- 42.96
- 525%
- 10
- (0, 2)

Solutions to Review 3

- $\text{Mean} = \frac{1475}{5} = 295$
- Total = $7 \times 10 = 70$
- 25 will be a number in the group because there are an odd number of scores in the group. This means that 25 is the middle number.
 - There are an even numbers of scores so the median will be the average of the two middle numbers. If the two middle numbers are both 25 then the median will be in the group, otherwise it won't be.

Solution to Focus problem 3

- (a) Mean = $\frac{206.7}{14} = 14.76$ (2dp) Median = 2.9 (halfway between 2.8 and 3)
- (b) All the results except for one are quite small and seem to be in kilometres. The 170 is recorded wrongly.
- (c) Someone has written the result down wrongly. Maybe they left out the decimal point.
- (d) There are two possible ways to solve the problem.
- Rewrite the 170 as 1.7 and use that.
 - Discard the 170 as it looks like the person made a mistake, corrected it by writing the 1.7 after it but then forgot to remove the 170.
- (e) With the 170 left out: Mean = 2.82 (2dp) Median = 2.8
 With 170 changed to 1.7: Mean = 2.74 (2dp) Median = 2.8
- (f) The median was not affected very much in this case but the mean was much more sensible with the error changed.
- (g) The 170 was an outlier. All the other scores were between 1.7 and 3.5 so 170 was a long way outside the group of the other scores.
- (h) No-one could run 170 km in 15 minutes because people usually run only about 12 kph even at a fast pace. The 170 would have to be an error.

Solutions to Skills development 3

1. (a) 208 cm is quite a bit taller than the other heights so it could be considered an outlier.
 - (b) (i) Mean = 166.75 cm Median = $\frac{164+165}{2} = 164.5$ cm.
 (ii) Mean = 163 cm Median = 164 cm
 - (c) The median was almost the same but the mean was reduced by 3.75 cm without the outlier. While this is not a great change it is a significant one.
2. (a) \$3200
 - (b) Mean = $\$(5754 \div 6) = \959 Median = $\$ \frac{465+535}{2} = \500
 - (c) Mean = $\$(2554 \div 5) = \510.80 Median = \$465
 - (d) The outlier had the biggest effect on the mean.
 - (e) The wage of the boss does not reflect what the others in the office get paid. It would make sense to leave the wage of the boss out of the calculations. It would not be very honest to leave it in.
3. (a) 43% is the middle of the scores so half the class got less than 43%. This means that Jared must be in the top half of the class marks with his 45%.

- (b) The median of 43% will be the 12th score. This means that 11 students got a mark better than 43%.

4. Measures of central tendency

Solutions to Warm-up 4

- 10 and 15 should be circled.
- 6
- 2 degrees
- $w = \frac{7}{8}$
- 35
- 0.0506
- 3.627
- $\frac{230}{100}$ or $\frac{23}{10}$
- 986
- $\frac{2}{8}$ or $\frac{1}{4}$

Solutions to Review 4

- (a) 0.1
(b) 78
(c) no outlier
- (a) Mode = 12 Range = 14 – 10 = 4
(b) Mode = 25 and 28 Range = 30 – 25 = 5

Solution to Focus problem 4

- (a) If you add the values in the frequency column you get the total number of kicks. There are 21 kicks so there are 21 students in the class.
(b) The range is 5. To find the range just take the smallest value in the score column away from the largest value.
(c) The mode is 4. To find the mode look down the frequency column for the biggest number then the score that is connected to it is the mode.
(d) The mode tells you that more balls hit the four on the target than any other number.

2. (a) The list of scores in order is shown below with the median highlighted.
 1, 1, 2, 2, 2, 3, 3, 3, 3, 3, **4**, 4, 4, 4, 4, 4, 5, 5, 6, 6, 6
 The median is 4.
- (b) Work out the position of the median. The median will be the 11th score because $\frac{21+1}{2} = 11$. Now go through the frequency column until you reach the 11th score. There are two lots of 1, three lots of 2 and five lots of 3. This accounts for the first 10 scores so the 11th score is 4. The median is 4.
- (c) The numbers in the score column are not all the scores just a list of the possible scores so you don't find the middle of them for the median.
3. (a) Total = 75
- (b) Add another column to the table where you multiply each score by the frequency and record the result. Then add up that column of numbers. It will give you the total of all the scores as shown below.

Score	Frequency	Score x frequency
1	2	2
2	3	6
3	5	15
4	6	24
5	2	10
6	3	18
Total	21	75

- (c) mean = 3.6 (1dp)
4. Summary of how to find the range, mode, mean, median from a frequency table.
 (wording may vary)
- Mean: Add another column and in it put the results of when you multiply the score by the frequency. Divide the total of this column by the total of the frequency column.
- Median: Add 1 to the total of the frequency column then divide by 2. This will give the position of the median. Work down the frequency column until you get to this position then the score in that position is the median.
- Mode: Find the largest number in the frequency column then the score that has that frequency is the mode.
- Range: Using the score column, subtract the smallest value from the largest value.

5.

Score	Frequency	Score x frequency
1	2	2
2	6	12
3	14	42
4	4	16
5	3	15
Total	29	87

Mean: Multiply the scores by the frequency. The total of this column is 87 and the total of the frequency column is 29. Mean = $87 \div 29 = 3$.

Median: The position of the median is $\frac{29+1}{2} = 15$. Working down the frequency column until you get to the 15th number gives the median of 3. (The first 8 numbers are 1 or 2.)

Mode: The largest number in the frequency column is 14, so the mode is 3.

Range: Subtract the smallest value from the largest value. Range = $5 - 1 = 4$.

6. It is a frequency table because the table shows how many of each score there is.

Solutions to Skills development 4

- The median will probably be 4.
 - Crossing out to be completed from either end to get to the middle number.
 - The median is 4.
 - The median is the position halfway between the 7th and 8th score.
 - 14 scores are represented in the table.
 - $\frac{n+1}{2} = \frac{14+1}{2} = 7.5$ The score in the 7.5th position is halfway between the 7th and 8th scores.
 - Mean = 3.6 (1dp)
Method 1: You could use the list with all the 14 numbers written out and add them up then divide by 14.
Method 2: Include a frequency \times score column and divide the total of that column by 14.
- $\frac{15+1}{2} = 8$ Median is in the 8th position.
 - $\frac{25+1}{2} = 13$ Median is in the 13th position.

- (c) $\frac{8+1}{2} = 4.5$ Median is the average of the 4th and 5th numbers.
- (d) $\frac{100+1}{2} = 50.5$ Median is the average of the 50th and 51st numbers.
3. (a) Mean = 2.625 Median = 2.5 (the average of 2 and 3)
 (b) Mean = 22.9 (1dp) Median = 23
4. Joel must be wrong because the numbers in the table are all between 12 and 16 (inclusive) so the mean would have to be between 12 and 16.

5. Calculations from graphs

Solutions to Warm-up 5

- $24 = 8 \times 3 = 2 \times 2 \times 2 \times 3$
- 384
- $v = (-3)$
- $\frac{8}{6} = 1\frac{1}{3}$
- 30
- 0.4002
- 47
- 4.375
- 96.9
- 104°

Solutions to Review 5

- (a) The outlier is \$24.
 (b) The outlier looks like a mistake. If the people in the flat were on holidays and didn't use power it may be correct.
- (c) Mean = $\$ \frac{4006}{12} = \334 (to the nearest dollar) Median = $\$ \frac{356+378}{2} = \367
- (d) Mean = $\$ \frac{3982}{11} = \362 Median = \$378
- (e) The value of \$24 is not typical of the other bills which are all between \$260 and \$435. It would be a good idea to leave it out.

2. (a) Outlier is \$19

(b) With outlier: Mean = $\$ \frac{1170}{4} = \292.50 Median = \$342

Without outlier: Mean = $\$ \frac{1151}{3} = \384 (to the nearest dollar) Median = \$388

(c) The outlier had a bigger effect, especially on the mean because there were such a small number of scores.

Solution to Focus problem 5

What you were asked to find was the measures of central tendency and spread from data presented in a graph.

(a) Explanations may vary but a possible explanation is as follows. The mode is easy to find because you just look for the tallest column. The score along the bottom axis under that column is the mode.

To find the range you look across the bottom (horizontal) axis and take the smallest number from the biggest number that both have columns.

(b) Mode = 3 children Range = $5 - 0 = 5$ children

(c)

Score (Number of children)	Frequency (Number of families)	Score \times frequency
0	4	0
1	4	4
2	8	16
3	9	27
4	0	0
5	1	5
	Total = 26	Total = 52

$$\text{Position of median} = \frac{26+1}{2} = 13.5$$

The median is halfway between the 13th and the 14th number.

The 13th and 14th numbers are both 2 so the median is 2.

$$\text{Mean} = \frac{52}{26} = 2$$

(d) The conclusions will vary.

Possible conclusion: You would still have to do the same calculations so putting the data into a table would be easy.

- (e) Answers will vary but some possible conclusions are as follows.
- All except one family had three children or less.
 - Most families had two or three children.
 - There were 26 families in the street and a total of 52 children.
- (f) Five is not an outlier as it is not far enough from the other results. It will have little effect on the calculations, except the range, because it is only one score out of 26.

Solutions to Skills development 5

1. (a) Mode = 3 Range = 7
- (b) Number of students = $5 + 2 + 4 + 6 + 3 + 2 + 2 + 1 = 25$
- (c) Median position = $\frac{25+1}{2} = 13$
- (d) Median = 3
- (e) Total number of pieces of fruit = $0 + 2 + 8 + 18 + 12 + 10 + 12 + 7 = 69$
- (f) Mean = $\frac{69}{25} = 2.76$ pieces of fruit
- (g) Five students ate five or more pieces of fruit. If they spread it out over the five days they may have met the guidelines.
- (h) No students have met the ideal. No-one ate 10 pieces of fruit over the week.
- (i) Answers will be individual. Get your teacher to check your answer.
2. (a) $0 + 11 + 22 + 6 + 4 + 5 + 6 = 54$ 54 phones have been discarded.
- (b) Mode = 1 and 2 (bimodal) Range = 6
- (c) Mean = $\frac{54}{31} = 1.7$ (1dp) Median = 16th score = 2
- (d) 21% of 54 = $0.21 \times 54 = 11.34$ It is the equivalent of about 11 phones.
- (e) 7% of 54 = 3.78 Close to four phones are thrown in landfill.
3. (a)

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

The median is 37.

- (b) If the stem-and-leaf plot was not ordered it would be harder to find the middle number by crossing out.

- (c) Mean = total of ages \div 11
 $= 396 \div 11$
 $= 36$
- (d) Range = highest age - lowest age. Because the plot is ordered the highest age is the last entry on the last line and the lowest age is the first entry on the first line.
 Range = $48 - 22 = 26$ years
- (e) No mode means there are no two people in the waiting room with the same age.

6. Use in everyday life

Solutions to Warm-up 6

1. 909
2. 510
3. 5 degrees
4. $1\frac{3}{5}$
5. 21
6. 1390 mg
7. 8
8. 240%
9. 6
10. (-7, 8)

Solutions to Review 6

1. All the other scores are 65 or over and half the class got over 72 so leaving out the 30 gives a better idea of the class average.
2. (a) Number of families = $4 + 9 + 15 + 10 + 1 = 39$
 (b) Total number of children = $4 + 18 + 45 + 40 + 10 = 117$
 (c) Three is the most common number of children. It is called the mode.
 (d) The median is in the 20th position using $\frac{39+1}{2}$.
 (e) Median = 3
 (f) 10 is the outlier.
 (g) The mean without the outlier = $\frac{107}{38} = 2.8$.
 (h) The outlier won't make a big difference because there are a large number of families and 10 is not too far above the other values.

Solution to Focus problem 6

- (a) The people with the water saving shower heads actually tend to have shorter showers than the others.
- (b) The most common shower length is eight minutes for water saving shower heads, 12 minutes for the old style shower heads. The mode overall is 12 minutes.
- (c) Water saving shower heads: $12 \times 8 \text{ L} = 96 \text{ L}$
 Old style shower heads: $12 \times 15 \text{ L} = 180 \text{ L}$
- (d) $96 \div 9 = 10.7$ (1dp) It would take about 11 buckets of water.
 $180 \div 9 = 20$ It would take 20 buckets of water.
- (e) Water saving shower heads: 66 showers in total
 Old style shower heads: 63 showers in total
- (f) $66 \times 8 = 528 \text{ L}$ $63 \times 15 = 945 \text{ L}$
 The amount of water saved is $945 - 528 = 417 \text{ L}$ per minute.
- (g) Answers will be individual. Have them checked by your teacher.
- (h) Answers will be individual. Have them checked by your teacher.

Solutions to Skills development 6

1.

Number of hours	Number of households	Number of hours \times Number of households
5	2	10
6	7	42
7	12	84
8	11	88
9	19	171
10	16	160
11	11	121
12	0	0
13	0	0
14	2	28
Total	80	704

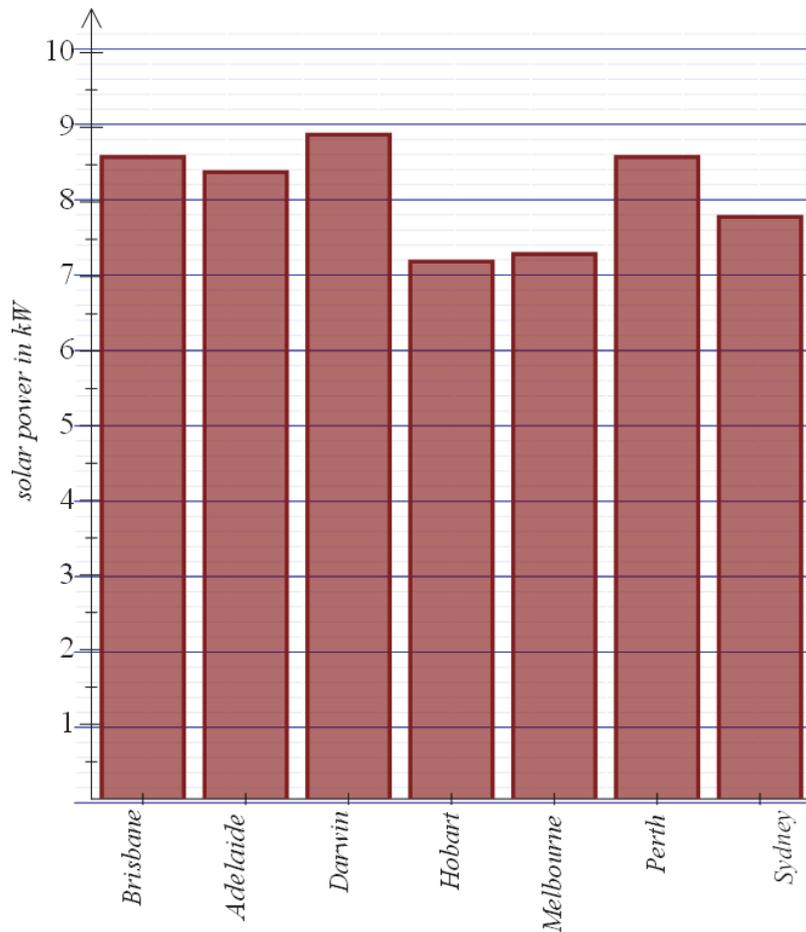
- (a) Total number of hours = 704.
- (b) Range = 9 hours Mode = 9 hours
 Mean = $704 \div 80 = 8.8$ hours Median = average of 40th and 41st value = 9 hours
- (c) $704 \times 25 \text{ W} = 17\,600 \text{ W}$

- (d) $17\,600\text{ W} \div 1000 = 17.6\text{ kWh}$
- (e) $17.6 \times \$0.30 = \5.28 per day
Extra cost per year = $\$5.28 \times 365 = \1927.20
- (f) The average household uses about 8.8 hr/day of extra electricity.
Using old globes:
 $8.8 \times 100 = 880\text{ W} = 0.88\text{ kWh}$ extra power per day
 $0.88 \times 365 = 321.2\text{ kWh}$ extra power per year
Using efficient globes:
 $8.8 \times 25 = 220\text{ W} = 0.22\text{ kWh}$ extra power per day
 $0.22 \times 365 = 80.3\text{ kWh}$ extra power per year
Difference in power use = $(321.2 - 80.3)\text{ kWh}$
 $= 240.9\text{ kWh}$
Savings per year = $240.9 \times 0.3 = \$72.27$
2. (a) A stem-and-leaf plot is more suitable because there are only a small number of scores but they range from 114 to 148.
- (b) Range = $148 - 114 = 34\text{ L}$ per person
- (c) Mode = 135 L
- (d) There are 13 scores so the median is the 7th score. Median = 127 L
- (e) Mean = total $\div 13 = 1663 \div 13 = 127.9\text{ L}$ (1dp)
- (f) $13 \times 20 = 260$ litres of water would be saved.
- (g) Answers will be individual. Discuss with your teacher. The calculations should be taken over a week and an average per day found.
3. (a) Range = $8.9 - 7.2 = 1.7\text{ kW}$
- (b) Mean = $56.7 \div 7 = 8.1\text{ kW}$ Median = 8.4 kW (the 4th value in order)
- (c) There are no outliers.
- (d) Stem-and-leaf plot:

Average daily power production in kW

Stem	Leaf			
7	2	3	8	
8	4	5	6	9

(e)



- (i) The column graph is good to find which are the highest and lowest productions of power.
- (ii) The stem-and-leaf plot organises the values in order and it makes finding the median easy.

7. Calculations with provided data

Solutions to Warm-up 7

1. 2.107
2. 107
3. 5 degrees
4. $\frac{4}{6} = \frac{2}{3}$
5. 6
6. 0.3 mL
7. 9
8. \$22.50
9. Reduce each number by 4 to get the next number.
10. The prime numbers are 2, 3 and 5. Probability = $\frac{3}{6} = \frac{1}{2}$

Solution to Focus problem 7

You were asked to examine and answer questions relating to statistics provided by the Australian Bureau of Statistics.

1. Waste items recycled and/or reused by households
2. There are eight different types of waste items.
3. The ♦ represents the percentage of households recycling/reusing each item in the year 2000.
4. There is no real mode for 2000 because the values are too close.
5. (a) Mean = $(592 \div 8)\% = 74\%$
(b) Median = 90.5% (the average of the 4th and 5th percentage in order)
6. The high values are about 10% different but the last two are about the same. The mean might be expected to be approximately 64% and the median about 80%.

Solutions to Skills development 7

1.
 - (a) Range for Marble Bar = $45\text{ }^{\circ}\text{C} - 38\text{ }^{\circ}\text{C} = 7\text{ }^{\circ}\text{C}$
Range for Albany = $33\text{ }^{\circ}\text{C} - 18\text{ }^{\circ}\text{C} = 15\text{ }^{\circ}\text{C}$
 - (b) Albany has the greatest range of temperatures. You can see this by how far spread out the graph is.
 - (c) The modes are found by deciding which temperature has the highest column.
 - (d) Mode for Marble Bar = $42\text{ }^{\circ}\text{C}$ Mode for Albany = $24\text{ }^{\circ}\text{C}$
 - (e) There are no outliers because there are no gaps in the Marble Bar temperatures and the gaps in the Albany temperatures are very small.
 - (f) Mean for Marble Bar = $585\text{ }^{\circ}\text{C} \div 14 = 41.8\text{ }^{\circ}\text{C}$ (1dp)
 - (g) Median for Albany = $24.5\text{ }^{\circ}\text{C}$ (average of the 7th and 8th temperatures)
2.
 - (a) The mode would not necessarily represent the average age. The median is used because it means that very high and very low ages don't affect the result.
 - (b) The last five columns go up one year at a time but the first gap is ten years then the next is five years. The increases are not all the same.
 - (c)
 - (i) The median age has increased by 3.8 years
 - (ii) The median age has increased by 3.9 years.
3. Answers will be individual. Check with your teacher.
Your answer should include:
 - information about why the median is used
 - sale prices for houses in two different suburbs or towns
 - the mean and median calculated for both sets of data
 - comments about the comparison of the results.

Solutions to Task A

1.

Term	Description
outlier	the average of the scores
mode	the difference between the highest and the lowest score
mean	a score that is vastly different to the other scores
range	the middle score when they are in order of size
median	the most common score

2.

	Range	Mode	Mean	Median
(a)	21.1 m	87.5 m	87.48 m	88.15 m
(b)	4	5	5.6	5
(c)	4 televisions	1 television	1.8 televisions (1dp)	1.5 televisions

3. In general the houses in the suburb with the higher median should have the more expensive houses but there is not enough information about how many houses were included to come to that conclusion. The second suburb may have had only that one house for sale. The first suburb may have had some expensive houses too.
4. No, because this is only the middle value of the houses sold in the town in that month.
5. Mean weight = $1430 \text{ kg} \div 21 = 68.1 \text{ kg}$ (1dp)
6. In this case you would need to use the mean weight, not the median weight.
 $72 \times 11 = 792$ This means that the total weight of the 11 passengers is 792 kg so they wouldn't all be able to go on the plane. It would be 32 kg overweight.
7. (a) Range = $73\% - 11\% = 62\%$
 (b) Mode = 11% Median = 8th score in order = 15%
 Mean = $\frac{273\%}{15} = 18.2\%$
- (c) The median is the best measure to use because the 73% doesn't affect it and the remainder of the scores are between 11 and 18. The mean is bigger than 18 so it doesn't really show what the scores are like.
- (d) 73% is the outlier.
- (e) Further study would be needed to decide if the 73% is an error or a real result. If it is a real result it may be very significant for the researcher.

8. (a) Recent world earthquake magnitudes

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

- (b) This plot is better because there are a lot of different values but not a lot of difference in the whole number parts. The actual values can still be seen and the calculations are easy to do.
- (c) The median is the average of the 10th and 11th scores which is 4.55.
- (d) $\text{Mean} = \frac{88}{20} = 4.4$
- (e) The Meckering earthquake was quite a lot stronger than the mean and median of these twenty earthquakes.
9. (a) The scores have not been put in order of size. They should read: 3, 4, 5, 7, 8 so the median is actually 5.
- (b) The mean cannot be 13 because the largest number is only 1.5. The decimal point was left out so the mean should be 1.3.
- (c) There is a mode it is 130. The leaf values were not all written in order to show there are two scores of 130.
- (d) Five is not an outlier because it is not far enough from the other data.
10. (a) The median is the middle score which means there are some values below it and some above. Angela may earn one of the lower salaries.
- (b) The advertised salary will be the middle one, and one of the actual salaries. There will be two salaries less than it and two more than it.

Solutions to Task B

Answers will vary. Check with your teacher.

One possible approach is to use the middle score in each column as the values and do the calculations from there.

The range will be \$70 or less.

The group representing the mode is the \$40 to \$50 group.

The median will be the 12th score in order of size so again in the \$40 to \$50 group.

The mean should also be somewhere between \$30 and \$50.



Department of
Education

MATHSAC030
ANALYSING DATA
ISBN: 9780730744504

