



The Mathematics Department

10

Data Handling

and Statistics 4

Higher Tier

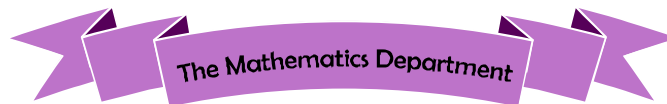
Version
3.0
08/2025

Name:

Contents

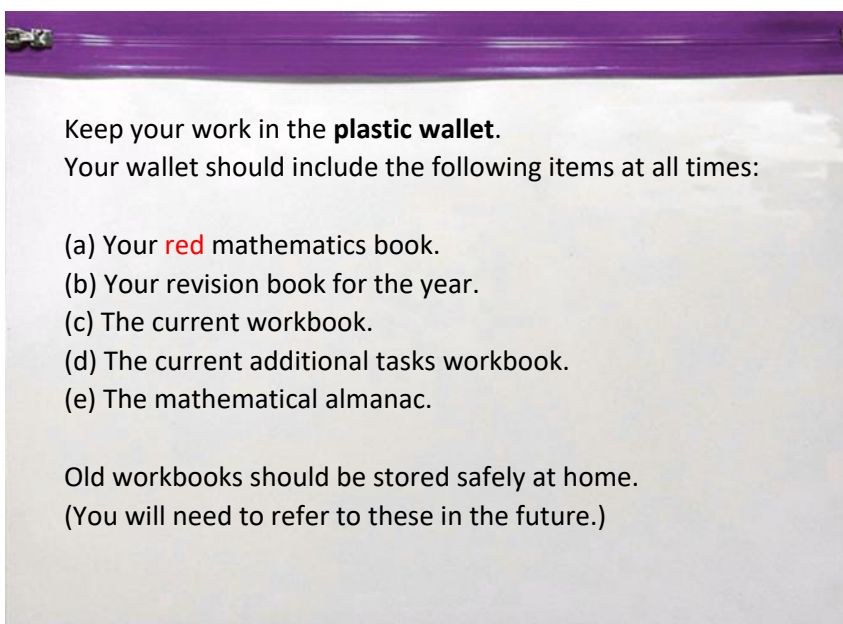
Chapter	Mathematics	Page Number
The Mathematics Department	Looking after your work. Content of the workbooks. Mathematics learning journey.	3
The Sequential Data Handling Cycle	An infographic explaining the sequential data handling cycle.	6
Comparing Averages	Revising averages. Choosing the most appropriate average. Comparing averages.	7
Questionnaires	Designing questionnaires. Criticising questions. Hypotheses.	12
Sampling	Simple random sampling. Systematic sampling. Stratified sampling.	15
Frequency Polygons	Drawing frequency polygons. Interpreting frequency polygons.	22
Box and Whisker Diagrams	Drawing box and whisker diagrams. The connection between box and whisker diagrams and cumulative frequency diagrams.	26
Histograms	Frequency density. Drawing a histogram. Interpreting a histogram. Estimating the median from a histogram. Estimating the quartiles from a histogram. Comparing histograms.	32





Welcome back to year 10! Here is a reminder of the guidelines for looking after your work.

- At the start of every lesson, write "Classwork", the date and a suitable title for the work. Each of these should be underlined.
- No blank pages should be left in the red mathematics book.
- Draw diagrams using a pencil and, where appropriate, using the correct equipment, e.g. ruler, protractor, compass.
- Show your method fully.
- Remember to include relevant units in your answers, e.g. cm, £, ml.



Equipment

- Black, red and blue biros.
- HB Pencil.
- Ruler (a 30 cm one is better).
- Eraser.
- Protractor.
- Compass.
- Scientific calculator (Casio fx-83GTCW).
- Highlighter.

The Revision Book

This will form the basis for your GCSE revision.

- Complete at least 4 pages in your revision book for each unit of work.
- You should include material you will need in the future for remembering the work quickly. This can include notes on the work; examples; important facts; and revision posters.

Workbooks

You will receive 1 copy of the workbook and 1 copy of the additional tasks workbook at the start of each new unit of work. (If you lose the workbook, a new one will cost 50p.)

A Welsh copy of the workbook, and additional supporting materials, can be found on the department's website, www.mathemateg.com



Content of the workbooks

When you see a QR code (like the one on the left), scan it using your mobile device in order to reach a Welsh YouTube video hosted on the following channel.

www.youtube.com/adolygumathemateg

The numbers in circles, for example **3**, show which GCSE unit the work appears in.

Circle	1	2	3	12	13	23	A
Units	1	2	3	1 and 2	1 and 3	2 and 3	1, 2 and 3

All the workbooks contain a variety of exercises, labelled as follows.



Exercises on new topics.



Answering a question in context, or solving a problem.



A more difficult question.



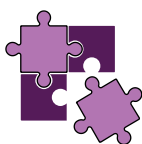
Revision of material from previous workbooks.



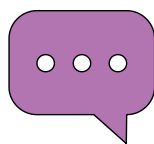
There are evaluation boxes at the end of each chapter to revise the completed work.

Key words	Corrections	I am happy with...	I need to revise...
Write down the new or important mathematical terms from the chapter.	What do you need to remember when completing this type of work in the future?	Write down the topics you had success with.	Write down the topics you need to look at again.

Curriculum for Wales Proficiencies



Conceptual understanding



Communication using symbols



Strategic competence



Logical reasoning



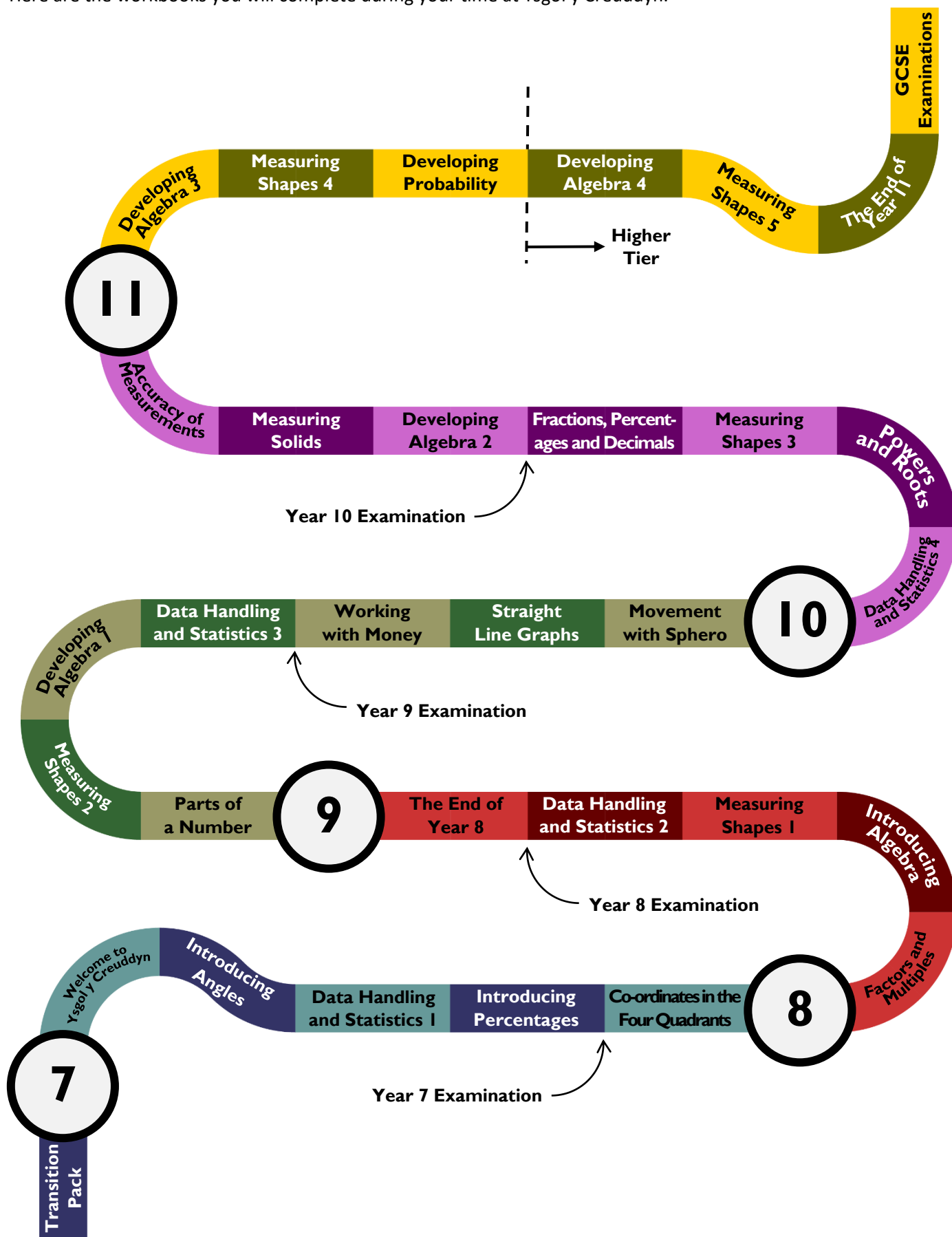
Fluency

Supporting Materials:

- Diagnostic Questions
 - A quiz for each workbook on the website www.diagnosticquestions.com.
- Reflection Sheet
 - An opportunity to assess your understanding of a workbook, and to see the test question order.
- Old WJEC examination questions; worksheets; investigations; puzzles.
 - Available for some topics.

Mathematics Learning Journey for Ysgol y Creuddyn

Here are the workbooks you will complete during your time at Ysgol y Creuddyn.



The Sequential Data Handling Cycle

- Designing questionnaires
- Criticising questions
- Hypotheses



- Types of data
- Simple random sampling
- Systematic sampling
- Stratified sampling



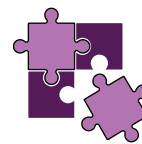
- The mean
- The range
- The mode
- The median
- The interquartile range
- Box and whisker diagrams
- Choosing the most appropriate average
- Comparing averages

- Bar charts
- Frequency diagrams
- Pie charts
- Pictograms
- Vertical line diagrams
- Line graphs
- Scatter diagrams
- Cumulative frequency diagrams
- Frequency polygons



Comparing Averages

Revising Averages



We have studied three different types of averages over the last few years.

Year 7	Year 8	Year 9	Year 10
Mean	Mode	Median	Comparing Averages

Exercise 1



Fill in the boxes below to explain how to calculate the mean, the mode and the median.

The Mean	The Mode	The Median

Exercise 2

Calculate the mean, the mode and the median of the following data.

12, 14, 14, 15, 16, 17, 17, 17, 19, 20.



3

Choosing the Most Appropriate Average

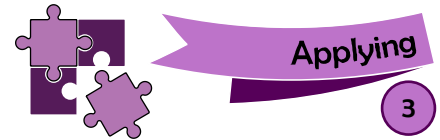
	The Mean	The Mode	The Median
Advantages	<ul style="list-style-type: none"> Uses all the data values. 	<ul style="list-style-type: none"> Not affected by outliers. Can be used for qualitative data. 	<ul style="list-style-type: none"> Not affected by outliers.
Disadvantages	<ul style="list-style-type: none"> Can be affected by outliers. Needs to be calculated. 	<ul style="list-style-type: none"> Doesn't use all the data values. There is no mode for some data sets. 	<ul style="list-style-type: none"> Doesn't use all the data values. Need to rearrange the data to find it.
Used for	<ul style="list-style-type: none"> Data which doesn't include outliers. 	<ul style="list-style-type: none"> Qualitative data. Data which includes outliers. 	<ul style="list-style-type: none"> Data which includes outliers.

Challenge!

Use the internet to investigate the term *skewness*.
Which average is the best to use when the data has a skewed distribution?



Exercise 3



Which average is most appropriate for the following data sets?

- (a) Favourite football team:
Liverpool, Chelsea, Man City, Everton, Liverpool, Man Utd.
- (b) Times in a 100 m race (in seconds):
9.81, 9.89, 9.91, 9.93, 9.94, 9.96, 10.04, 10.06.
- (c) Price of Heinz Baked Beans in different shops:
75p, 60p, 74p, 80p, 70p, 95p, 85p.
- (d) Age of players who start a football game:
28, 31, 19, 24, 25, 28, 30, 23, 20, 29, 26.
- (e) Year 10 pupil heights:
162 cm, 160 cm, 161 cm, 148 cm, 163 cm, 161 cm.
- (f) Favourite subject in school:
Science, Music, Drama, Mathematics, Music.
- (g) Spelling test scores (out of 10):
5, 7, 8, 4, 5, 3, 6, 4, 5, 4, 7.
- (h) Number of brothers:
0, 1, 2, 1, 0, 6, 1, 0, 1, 2.

Exercise 4

Are the following statements TRUE or FALSE?

- (a) The mode is the most popular data item in a set of data.
- (b) Half the values in a data set are greater than the mean.
- (c) Half the values in a data set are greater than the median.
- (d) When finding the median, it doesn't matter if you order the data from least to greatest or from greatest to least.
- (e) It is always possible to find the mean of a data set.



Exercise 5

The following table shows the percentages of 10 pupils in Welsh and Mathematics tests.

Welsh	57	63	91	58	56	75	59	76	91	54
Mathematics	67	68	66	68	68	66	70	69	68	70

(a) Complete the following table.

	Welsh	Mathematics
The mean		
The median		
The mode		
The range		

(b) Which statistics from the table support the following newspaper headlines?

- (i) **The Welsh results are very high this year.**
- (ii) **Pupils do not perform better in Mathematics than in Welsh.**
- (iii) **There was lots of copying during the Mathematics test.**
- (iv) **The Mathematics test was easier than the Welsh test.**
- (v) **The Welsh results show that some people tried harder than others.**

Exercise 6

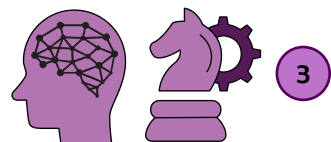
3

Three people were asked to estimate when 30 seconds had passed. This was done 5 times.

Osian	Siân	Shubnam
31	36	32
26	19	30
32	39	24
27	36	37
29	20	32

Who, in your opinion, is best at estimating the length of 30 seconds?

Exercise 7



Fill the following boxes with whole numbers.

<p>(a) 2 numbers mode = 5</p> <div style="display: flex; justify-content: space-around; width: 100px;"> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> </div>	<p>(b) 2 numbers mean = 5</p> <div style="display: flex; justify-content: space-around; width: 100px;"> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> </div>	<p>(c) 2 numbers mean = 5 range = 6</p> <div style="display: flex; justify-content: space-around; width: 100px;"> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> </div>
<p>(d) 2 numbers mean = 5 mode = no mode</p> <div style="display: flex; justify-content: space-around; width: 100px;"> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> </div>	<p>(e) 2 numbers mean = 6 range = 8</p> <div style="display: flex; justify-content: space-around; width: 100px;"> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> </div>	<p>(f) 2 numbers mean = 10 largest number = 12</p> <div style="display: flex; justify-content: space-around; width: 100px;"> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> </div>
<p>(g) 3 numbers mean = 3 mode = 2</p> <div style="display: flex; justify-content: space-around; width: 100px;"> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> </div>	<p>(h) 3 numbers mean = 7 mode = 10</p> <div style="display: flex; justify-content: space-around; width: 100px;"> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> </div>	<p>(i) 3 numbers median = 4 range = 0</p> <div style="display: flex; justify-content: space-around; width: 100px;"> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> </div>
<p>(j) 4 numbers mode = 5 range = 1</p> <div style="display: flex; justify-content: space-around; width: 100px;"> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> </div>	<p>(k) 4 numbers median = 3 range = 1</p> <div style="display: flex; justify-content: space-around; width: 100px;"> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> </div>	<p>(l) 4 numbers mean = 6 median = 6.5 range = 11</p> <div style="display: flex; justify-content: space-around; width: 100px;"> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> <div style="border: 1px solid black; width: 40px; height: 40px;"></div> </div>

Exercise 8

3

During a skiing trip, the PE department recorded the daily snowfall for 5 consecutive days. Here is some information about the daily snowfall.

Mean	Mode	Median	Range
5.8 cm	3 cm	5.6 cm	6.6 cm

(a) Use these statistics to calculate what was the daily snowfall for these 5 consecutive days.

(b) If it had snowed exactly 2 cm more each day, what would the new statistics be?

Mean	Mode	Median	Range

Exercise 9

In a game, it is possible for every player to score between 1 and 10 points. Lois and Beca play the game 5 times.



The table below shows the points scored by Lois in every game.

	Game 1	Game 2	Game 3	Game 4	Game 5
Lois	5	2	8	5	1
Beca					

- Beca had a higher mean score than Lois.
- Beca had a lower median score than Lois.
- Beca had a lower range of scores than Lois.

Complete the table above with a set of possible scores for Beca.

Exercise 10

Jim and Andy play for their local cricket team. They scored the following runs in their six most recent games.

Jim	42	71	39	62	70	40
Andy	115	6	84	36	10	85

- (a) Calculate Jim’s mean and Andy’s mean.
- (b) Calculate Jim’s median and Andy’s median.
- (c) There is no room for both Jim and Andy on the team for the next game. The management team need to choose either Jim or Andy to play for the team. Use your answers to parts (a) and (b) to give advice to the management team.



Exercise 11

The table below shows the number of ice hockey season tickets sold by a team last season, together with the prices for the tickets.

Ticket Price (£)	Number of tickets sold
250	180
300	230
350	230
500	150

For the cost of the tickets sold last season, calculate

- (a) The mode;
- (b) The median;
- (c) The range.
- (d) The ice hockey team owner says that more than half of the tickets sold were greater in value than £300. Explain why the team owner is incorrect.

Exercise 12

50 people took part in a charity walk. The table below shows the grouped frequency distribution of the sums of money raised, to the nearest £.

Sum, s , in £	Number of people
$10 \leq s < 19$	2
$20 \leq s < 29$	18
$30 \leq s < 39$	29
$40 \leq s < 49$	1

- (a) Find the modal class of the data.
- (b) Find the median class of the data.
- (c) Calculate an estimate for the mean sum of money raised per person.
- (d) Another 50 people took part in the same charity walk. The total of the money collected by these 50 extra people was £1,600. Is it possible to say that these 50 people raised more money than the original 50 people?



Evaluation

Key words	Corrections	I am happy with...	I need to revise...

Questionnaires

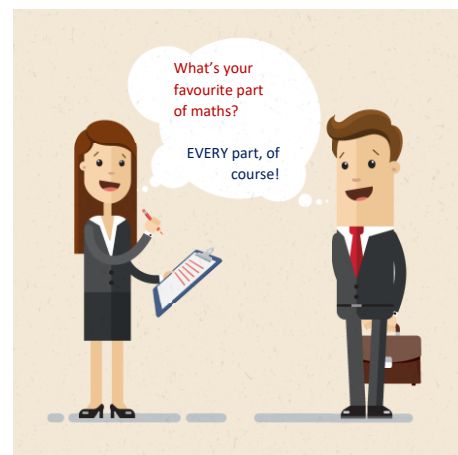
A **questionnaire** is a good way to collect data, but we must be careful when designing the questions.

- (1) We must avoid asking **leading** questions which favour one answer over another. For example, the question “Do you agree that eating ice cream is bad for you?” leads people to agree with the statement. (Why?)
- (2) We must avoid using answer boxes where the **options overlap**. For example, in the following question, people who are 20 years old can choose two different answer boxes.

How old are you?

Under 10 10–20 20–30 Over 30

- (3) We must be careful **where, when and how** a questionnaire is conducted. For example, the following questionnaires would not be appropriate.
 - a. Conducting a sports questionnaire outside of a football stadium.
 - b. Conducting a questionnaire about night shift workers at mid-day.
 - c. Conducting a questionnaire on mobile phone usage through a phone survey.
- (4) We must use **clear** and **concise** questions. For example, the question “How often do you go to the gym?” is not appropriate without explaining the meaning of the word “often”, i.e. daily, weekly, monthly...?
- (5) We must use **appropriate** and **relevant** questions. For example,
 - a. Some people would refuse to answer “How old are you?”, however they might answer if options showing different age ranges were included.
 - b. The question “What is your eye colour?” would not be appropriate in a questionnaire asking opinions about recycling.



Exercise 13

Write a criticism of the following questions.

(a) Do you read books? Circle your answer.

Yes No Sometimes

(b) Do you agree that the cruel sport of fox hunting should be made illegal?

(c) How often do you use the gym in a typical month? Circle your answer.

Never Once or Twice 2–5 Times More than 5 times

(d) How old are you? Circle your answer.

10–15 16–20 21–25 26–30 31–35

(e) In your current job, how much money do you earn?

(f) How often do you shop in a supermarket? Circle your answer.

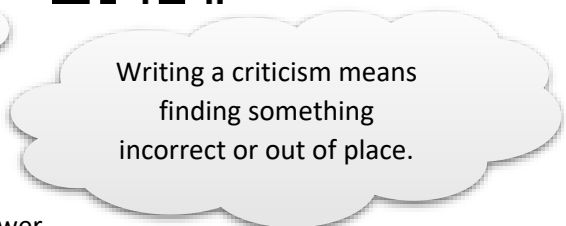
Three times a week Twice a week Once a week Once a month

(g) Do you agree that the amazing Liverpool are the best football team in the world? Circle your answer.

Yes Of course Absolutely



13



Exercise 14

A survey is conducted to see how often teenagers buy trainers.
The following questions are included in the questionnaire.



Question 1: Where do you live?

Question 2: How often do you buy trainers?

Never	1–10 times	10–15 times	More than 15 times
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



- (a) For each question state **one** reason why it is **not** appropriate.
- (b) The survey is conducted by leaving copies of the questionnaire on the seats in a sports clothing shop. Give one criticism of how the survey was conducted.

Exercise 15

A survey was carried out to discover whether people would rather watch sports or detective programs on the television.
The following three questions were included.

Question 1: What is your address?

Question 2: Which type of TV program do you prefer? Tick one box.

Sports	Detective
<input type="checkbox"/>	<input type="checkbox"/>

Question 3: How many hours do you spend watching TV?

Less than 1 hour	1–5 hours	More than 5 hours
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



- (a) Give **one** reason why question 1 is not appropriate.
- (b) Give **one** reason why question 3 is not appropriate.
- (c) The survey was conducted by asking people who were leaving a football stadium one Saturday afternoon. Give **one** criticism of how the survey was carried out.

Exercise 16

Elen is conducting a survey in her school about the standard of food in the canteen. She asks every 20th person that goes to get hot food. Explain what is wrong with Elen’s plan.



Hypotheses

A **hypothesis** is a statement like “children spend more time playing computer games than adults”. It is possible to create a questionnaire and collect data to test a hypothesis.



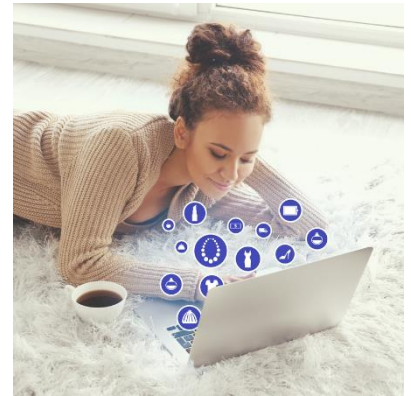
Exercise 17

Steffan wants to prove the following hypothesis.

‘Most people spend more than two hours on the internet every night.’

He intends to

- give a short questionnaire to people in the local fitness centre,
- ask the following questions:
 - In your opinion, do people spend too much time on the internet?
 - How much time do you spend on the internet?
- ask each participant to post their completed questionnaire in an envelope with a stamp on it.



Write **three** unfavourable comments about this plan.

Exercise 18

Mari wants to prove the following hypothesis.

‘Older pupils in a secondary school are better at remembering their times tables than younger pupils.’

She intends to

- give a short questionnaire to 50 randomly chosen pupils in each year,
- ask the same 5 multiplication questions to everyone, to complete in a maths lesson,
- ask the pupils to mark each other’s work and return the questionnaires to her through the maths teacher.

Write **three** unfavourable comments about this plan.

Exercise 19

Iwan wants to prove the following hypothesis.

‘The children in Llandudno spend more time playing computer games than the adults’.



Write a questionnaire that Iwan could use to prove or disprove this hypothesis.



Key words	Corrections	I am happy with...	I need to revise...

Sampling

In the previous chapter, we discussed how to design an appropriate questionnaire in order to prove a specific hypothesis. Often, it is not possible to ask the opinion of **all** members of a **population**. For example, it would be laborious to ask all pupils in a secondary school their opinion on a matter. Instead, we often use a **sample** of the population, which is a smaller group, and attempt to come to a conclusion about the opinion of the whole population based on the information from the sample.



We must be careful when selecting a sample. It must be **large enough**, and **representative** of the population. Asking only 5 pupils would not represent the opinion of a whole school, neither would asking only the pupils in year 8.

At GCSE level, we must be familiar with the following methods of selecting a sample.

Simple random sampling

Systematic sampling

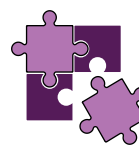
Stratified sampling

Simple Random Sampling

In a **simple random sample**, every member of a population has **the same chance** of being chosen.

There are two main methods of choosing a simple random sample:

- Using a table of random digits;
- Using the random number generator function on a calculator.



Example

A school hopes to change the start time for the school day and is eager to ask the opinion of the 600 pupils. The head teacher decides to choose a simple random sample of 10 pupils to question.

- The head teacher numbers all the pupils from 001 to 600.
- Starting from a random starting point in a table of random digits, the head teacher reads the numbers in groups of three.
- The head teacher accepts any number between 001 and 600, and rejects the others. The head teacher also ignores any repeated numbers.



Here is part of a table of random digits.

7087	0858	0164	1769	3218	1467	1938	8093	7918	2814
7796	7080	7227	3140	0933	0181	2013	7918	1177	4715
3830	9523	3653	8514	6061	0674	6025	9834	0499	3668
1347	1225	1910	3621	9722	8482	6298	1957	3507	7209

By starting at the digit in **red** (chosen at random), the head teacher chooses the following pupils: 218, 146, 388, 093, 147, 072, 273, 140, 301, 013.

Systematic Sampling

In a **systematic sample**, the sample is selected from the population in a **regular pattern**.



Example

A school hopes to change the start time for the school day and is eager to ask the opinion of the 600 pupils. The head teacher decides to choose a systematic sample of 10 pupils to question.



- The head teacher numbers all the pupils from 001 to 600.
- $600 \div 10 = 60$, therefore we work through the list of pupils in intervals of 60. (60 is the **sampling interval**.)
- To choose the starting number, we must use a table of random digits or the random number generator on a calculator to choose a number between 001 and 600.

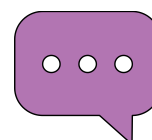
If we use the table of random digits from page 15, and read the digits in groups of three, the first number between 001 and 600 we see is 580. Therefore, the systematic sample is

580, 040, 100, 160, 220, 280, 340, 400, 460, 520.
 ↪ Circle around: $580 + 60 = 640$, $640 - 600 = 40$

Example

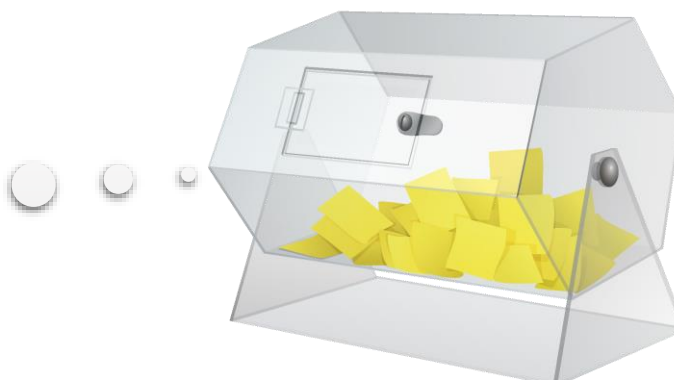
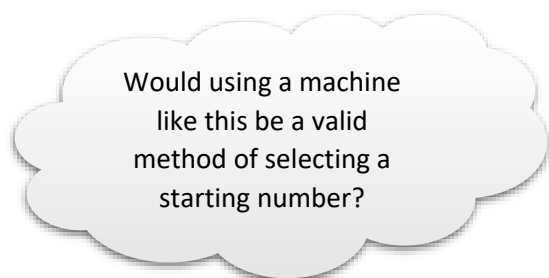
A young farmers' club is considering holding a fair and wishes to collect the opinion of the members on the content of the fair. The chairperson decides to choose a systematic sample of 10 members to question. The club has a total of 87 members.

- The chairperson numbers all the members of the club from 01 to 87.
- $87 \div 10 = 8.7$, therefore we need to work through the list of members in intervals of 9.
- To choose the starting number, we must use a table of random digits or the random number generator on a calculator to choose a number between 01 and 87.



The chairperson uses the random number generator function on a calculator to choose 03 as the starting number. Therefore, the systematic sample is

03, 12, 21, 30, 39, 48, 57, 66, 75, 84.



Exercise 24

Find the sampling interval in the following systematic samples.

- (a) Choosing 10 people out of 80. (b) Choosing 5 people out of 45.
 (c) Choosing 10 people out of 74. (d) Choosing 4 people out of 18.
 (e) Choosing 7 people out of 40. (f) Choosing 9 people out of 63.
 (g) Choosing 12 people out of 140. (h) Choosing 20 people out of 1,500.

**Exercise 25**

Choose a systematic sample of 10 people out of 70, by starting with

- (a) the second person (02);
 (b) the fourth person (04);
 (c) the seventh person (07).

Exercise 26

Choose a systematic sample of 10 people out of 140, by starting with

- (a) the first person;
 (b) the fifth person;
 (c) the twelfth person.

Exercise 27

Choose a systematic sample of 8 people out of 100, by starting with

- (a) the third person;
 (b) the eighth person;
 (c) the tenth person.

Exercise 28

Choose a systematic sample of 12 people out of 1,400, by starting with

- (a) the sixth person;
 (b) the twentieth person;
 (c) the 789th person.

Exercise 29

Choose a systematic sample of 9 people out of 50. Use the following table of random digits to decide where to start.

6841	4804	3748	9980	4225	5215	8258	3707	2575	8524
6966	5346	1628	1375	8214	8630	5766	5942	1463	2818
4049	7245	5872	1469	0956	9848	1042	0684	4823	1716
2041	3672	9958	9099	5660	9092	4286	7496	8092	1236

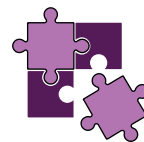
Exercise 30

Choose a systematic sample of 12 people out of 80. Use the random number generator function on your calculator to choose a starting point.



Stratified Sampling

Sometimes it is possible to split a population into **strata** or subgroups that reflect the composition of the population. For example, here are the details of the people working for a company creating computer games.



Role	Manager	Office staff	Programmers
Number of workers	3	7	27

An agency wants to sample the opinion of 10 people from the company. By looking at the table above, it would make sense to choose more programmers than managers to take part in the survey, since more programmers work for the company. Using a simple random sample or a systematic sample does not ensure that this will be the case, as it would be possible (for example) to choose all 3 managers – or none of them.

To reflect the strata (the subgroups) in the population, we use **stratified sampling** to ensure that every stratum of the population gets a fair representation in the sample.

Method:

$3 + 7 + 27 = 37$ people work for the company.

We choose $\frac{3}{37} \times 10 = 0.8108 \dots$ managers, which is 1 manager to the nearest whole number.

We choose $\frac{7}{37} \times 10 = 1.8918 \dots$ office staff, which is 2 office staff to the nearest whole number.

We choose $\frac{27}{37} \times 10 = 7.2972 \dots$ programmers, which is 7 programmers to the nearest whole number.

Multiply by 10 since we want a sample of 10.



Check: $1 + 2 + 7 = 10$, therefore we have selected the correct number of people to form the sample.

It would be possible to use a simple random sample or a systematic sample to choose which managers, which office staff and which programmers are questioned.

Example

An internet holiday club has members from 4 countries across the world. The number of members per country are given in the following table.

Country	Australia	China	Thailand	Mexico
Number of members	2,840	1,382	4,086	940

The company organises a meeting for 25 of the members to represent the opinions of all the members. Use a stratified sample to calculate how many members from each country should be invited to the meeting.

Answer:

There are $2,840 + 1,382 + 4,086 + 940 = 9,248$ members in total.

We choose $\frac{2840}{9248} \times 25 = 7.6773 \dots$ people from Australia, which is 8 people to the nearest whole number.

We choose $\frac{1382}{9248} \times 25 = 3.7359 \dots$ people from China, which is 4 people to the nearest whole number.

We choose $\frac{4086}{9248} \times 25 = 11.0456 \dots$ people from Thailand, which is 11 people to the nearest whole number.

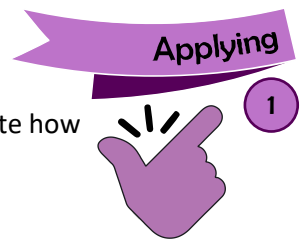
We choose $\frac{940}{9248} \times 25 = 2.5410 \dots$ people from Mexico, which is 3 people to the nearest whole number.

Check: $8 + 4 + 11 + 3 = 26$, therefore we have chosen one person too many. From the countries that have been rounded up (Australia, China, Mexico), we adjust the one that is originally closest to half a unit (Mexico, 2.5410...) So, our stratified sample is Australia 8, China 4, Thailand 11, Mexico 2.



Exercise 31

In a particular company, there are 359 Apple computers and 467 Microsoft computers. The technician has time to test 30 computers. Use the stratified sampling method to calculate how many Apple computers and how many Microsoft computers should be tested.

**Exercise 32**

A sports company employs people from a number of different countries. The following table shows the number of people employed by the company from each country.

Country	Canada	New Zealand	Turkey	China
Number of employees	2,785	804	1,207	8,763

The company is organising a promotional event and decides to invite a total of 45 employees to represent the opinion of all the employees. Use the stratified sampling method to calculate how many people from each country should be invited to the promotional event.

**Exercise 33**

A school in Wales has international links with schools in four countries across the world. The following table shows the number of pupils in each of the schools in the four countries.

Country	France	Australia	Canada	Brazil
Number of pupils	1,230	1,123	934	720

The school in Wales is organising a celebration and would like to invite a total of 35 pupils to represent the pupils from the four countries. Use the stratified sampling method to calculate the number of pupils who should be invited from each country.

Exercise 34

A movie society on the internet has members from four countries across the world. In the following table the number of members from each country is shown.

Country	USA	UK	France	The Netherlands
Number of members	12,637	8,382	4,010	720

The movie society organises a meeting for 30 members to represent the opinion of the whole society. Use the stratified sampling method to calculate how many members from each country should be invited to the meeting.

**Exercise 35**

In the following table the populations of 5 villages are given.

Village	Aberford	Bronglas	Carmel	Dunwern	Eiderfalls
Population	1,550	3,700	600	980	5,500

A committee of 20 people from the 5 villages needs to be chosen. Use the stratified sampling method to calculate how many people from each village should be invited to join the committee.

Exercise 36 (Revision)

1

(a) By starting at the first digit in the following part of a table of random digits, choose a random sample of 4 people from a list of 45 people.

06 56 06 14 27 93 24

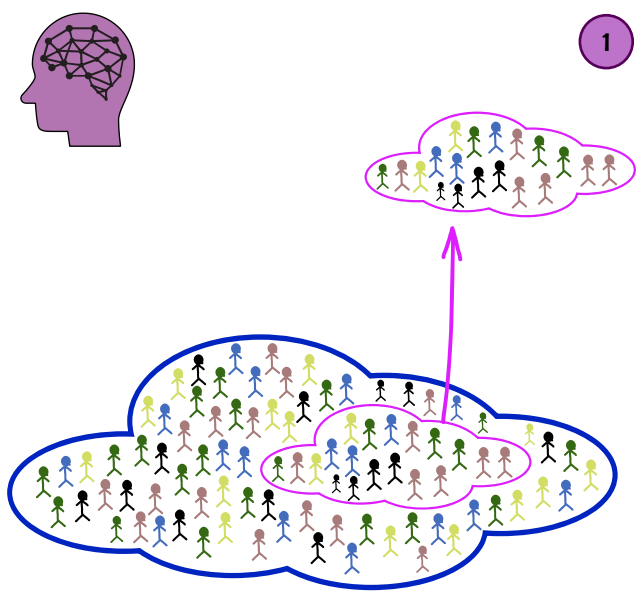
(b) The opinion of people in a queue is found by asking a number of people in the queue to answer a questionnaire. Explain why asking every tenth person in the queue is not a method of choosing a random sample to answer the questionnaire.

Exercise 37 (Revision)

1

(a) Are the following statements TRUE or FALSE?

- 1) Choosing the first name on the register of every class will give a random sample.
- 2) The ratio of boys to girls in a school is 2 : 3. The school council of 30 pupils is chosen by using a gender stratified sample. There are 10 boys and 20 girls on the school council.
- 3) A phone survey is conducted to discover which political party people support. The sample of people in the survey is **not** a random sample of the whole population.
- 4) Stratified sampling always considers the proportions according to specific criteria.
- 5) A random sample means that everybody has an equal chance of being selected.



(b) An international organisation employs people from Australia, Belgium, Canada, Denmark and Ecuador. The following table shows the number of people employed by the organisation in each country.

Country	Australia	Belgium	Canada	Denmark	Ecuador
Number of employees	5,243	1,004	8,745	545	762

The organisation is organising a charity event and chooses to invite 25 employees to represent the employees from all 5 countries. Use the stratified sampling method to calculate how many people from each country should be invited to the charity event.

Evaluation

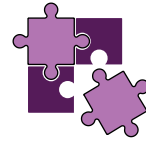
Key words	Corrections	I am happy with...	I need to revise...

Frequency Polygons

Drawing Frequency Polygons

We draw **frequency polygons** for the following types of data.

- Grouped discrete quantitative data.
- Continuous quantitative data.



A frequency polygon is a **line graph** where we plot the **midpoint** for every class against the **frequency**.

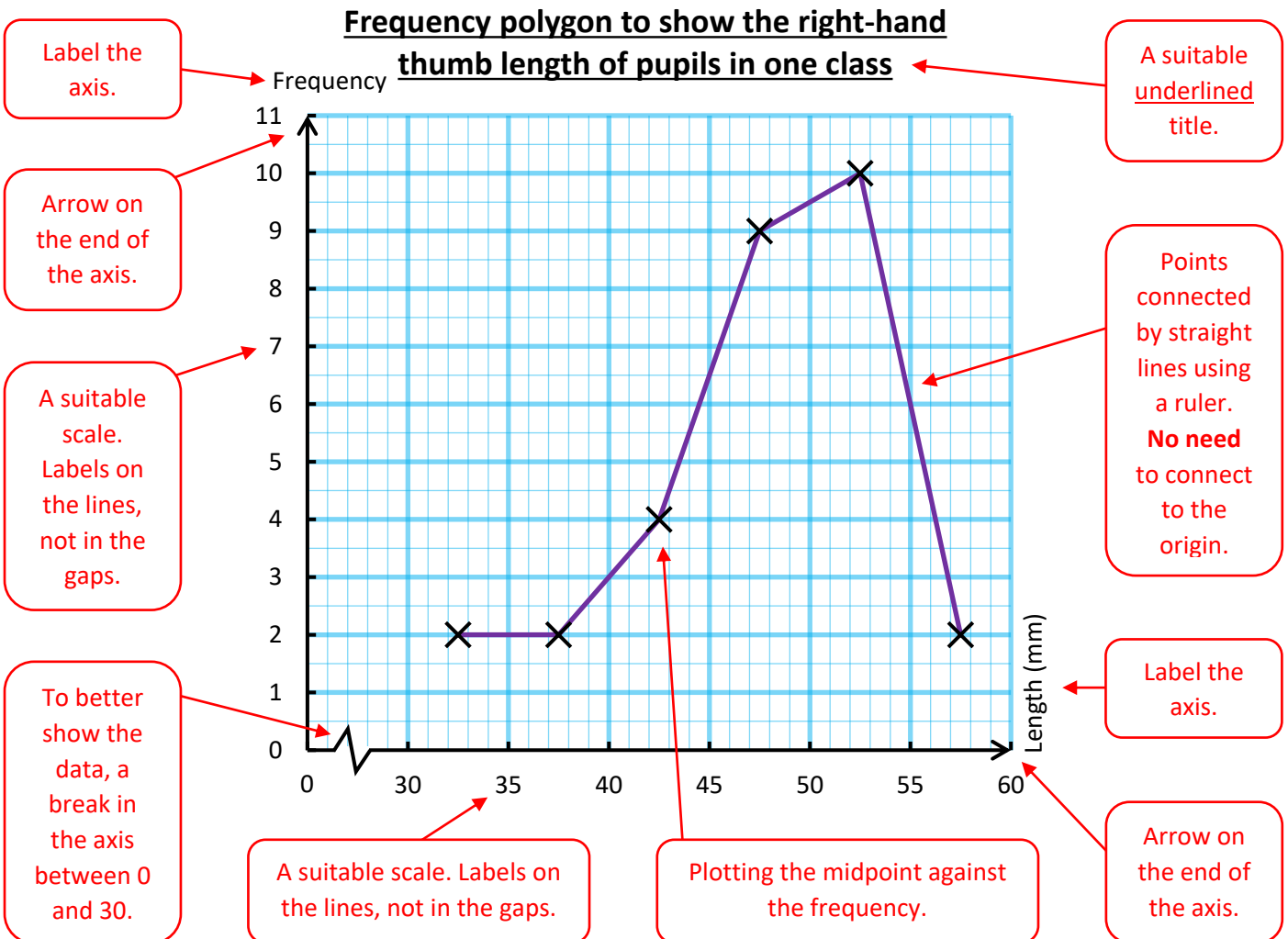
Example

The following frequency table shows the right-hand thumb length (*l*) for pupils in one class.

Class Interval (<i>l</i> mm)	Frequency
$30 \leq l < 35$	2
$35 \leq l < 40$	2
$40 \leq l < 45$	4
$45 \leq l < 50$	9
$50 \leq l < 55$	10
$55 \leq l < 60$	2

The midpoint for the class $30 \leq l < 35$ is 32.5.
Method 1: 32.5 is half way between 30 and 35.
Method 2: $30 + 35 = 65$. $65 \div 2 = 32.5$.

The frequency polygon below represents the data.



Interpreting Frequency Polygons



3

Exercise 41

The frequency polygon on the right shows the arm lengths of 100 females.

- (a) How many females have an arm length between 55 cm and 60 cm?
- (b) How many more females have an arm length between 65 cm and 70 cm, compared to females with arm length between 70 cm and 75 cm?
- (c) Complete the frequency table below, using the information from the frequency polygon.

Arm length, / cm	Frequency
$50 < l \leq 55$	
$55 < l \leq 60$	
$60 < l \leq 65$	
$65 < l \leq 70$	
$70 < l \leq 75$	

- (d) What is the modal class of the data?

Exercise 42

The daily rainfall for ten days was measured in Aberwen and Aberisel. The frequency polygon on the right shows the results. The purple line represents Aberwen, and the red line represents Aberisel.

- (a) For how many days was the rainfall between 0 mm and 1 mm of rain in Aberwen?
- (b) For how many days was the rainfall between 6 mm and 7 mm of rain in Aberisel?
- (c) Complete this sentence: Considering the days where the rainfall was between 5 mm and 6 mm of rain, Aberisel had this rainfall on _____ more days than Aberwen.
- (d) Over these 10 days, in your opinion where was the wettest place? Explain your answer.
- (e) Deiniol says “The frequency polygon shows that the rainfall in Aberwen and Aberisel is the same on the fourth day”. Is Deiniol telling the truth? Explain your answer.

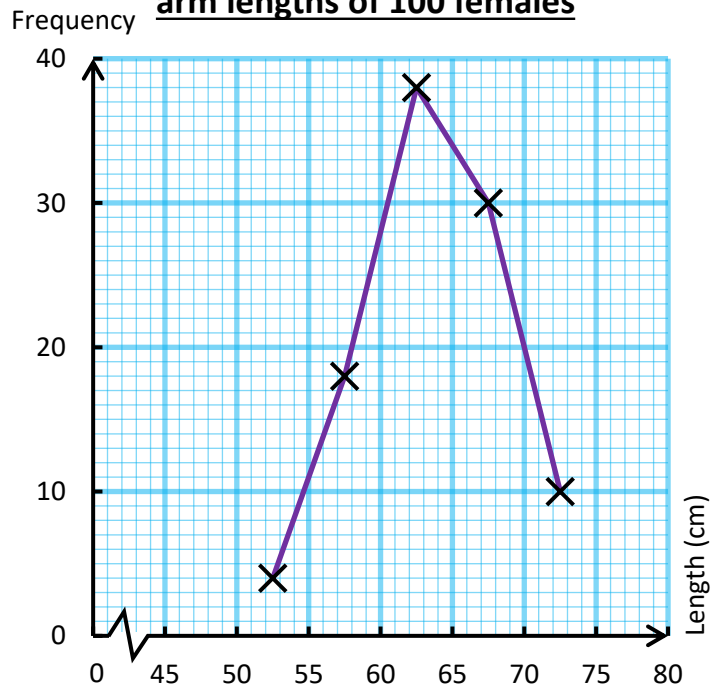
Challenge!

Use the frequency table in Exercise 41 to calculate an estimated mean arm length for the 100 females.

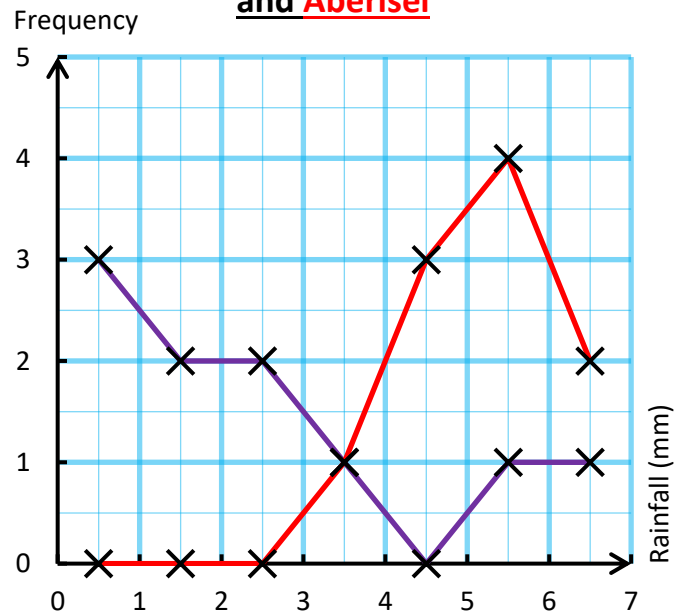
Challenge!

Use the frequency polygon from Exercise 42 to calculate the estimated mean rainfall in Aberwen and in Aberisel. Does your answer agree with your conclusion to part (d) of Exercise 42?

Frequency polygon to show the arm lengths of 100 females



Frequency polygon to show the rainfall over 10 days in Aberwen and Aberisel



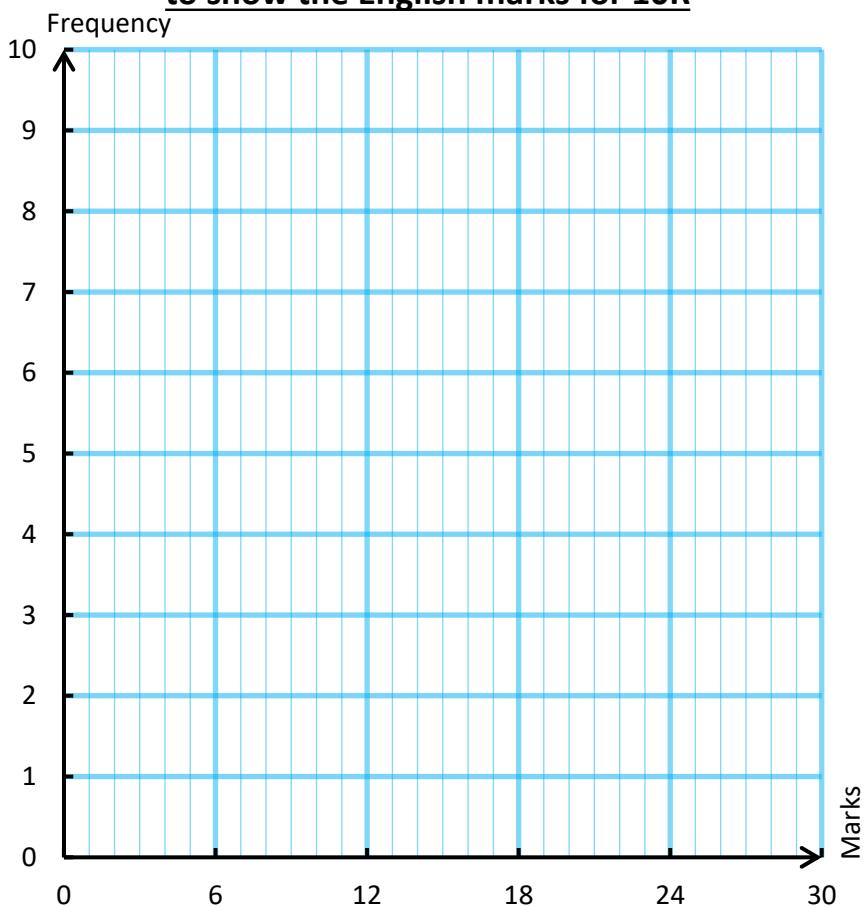
Exercise 43

3

The table below shows the marks for 10R in their English test. (The test was out of 30.)

Marks (m)	Frequency
$0 \leq m < 6$	3
$6 \leq m < 12$	8
$12 \leq m < 18$	7
$18 \leq m < 24$	6
$24 \leq m < 30$	2

Frequency polygon and frequency diagram to show the English marks for 10R



(a) Eric looks at the table and says: "Three people got 0 out of 30 in this test!". Is Eric telling the truth?

(b) Susan looks at the table and says: "Nobody got full marks in this test!". Is Susan telling the truth?

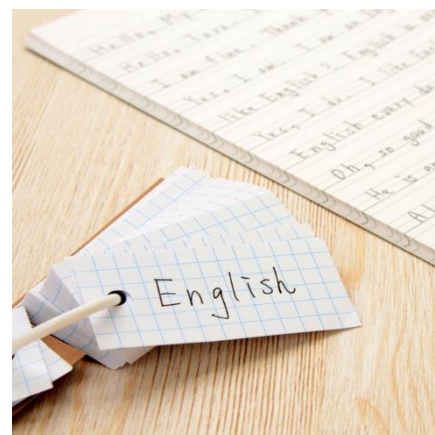
(c) On the graph paper on the right, draw a frequency polygon for the data.

(d) On the same graph paper, draw a frequency diagram for the data.

(e) What is the connection between any frequency polygon and frequency diagram drawn for the same data?

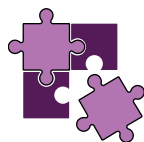
(f) Calculate the following for the data from the English test.

- (i) The modal class.
- (ii) The median class.
- (iii) The estimated mean.
- (iv) The estimated range.



Evaluation

Key words	Corrections	I am happy with...	I need to revise...

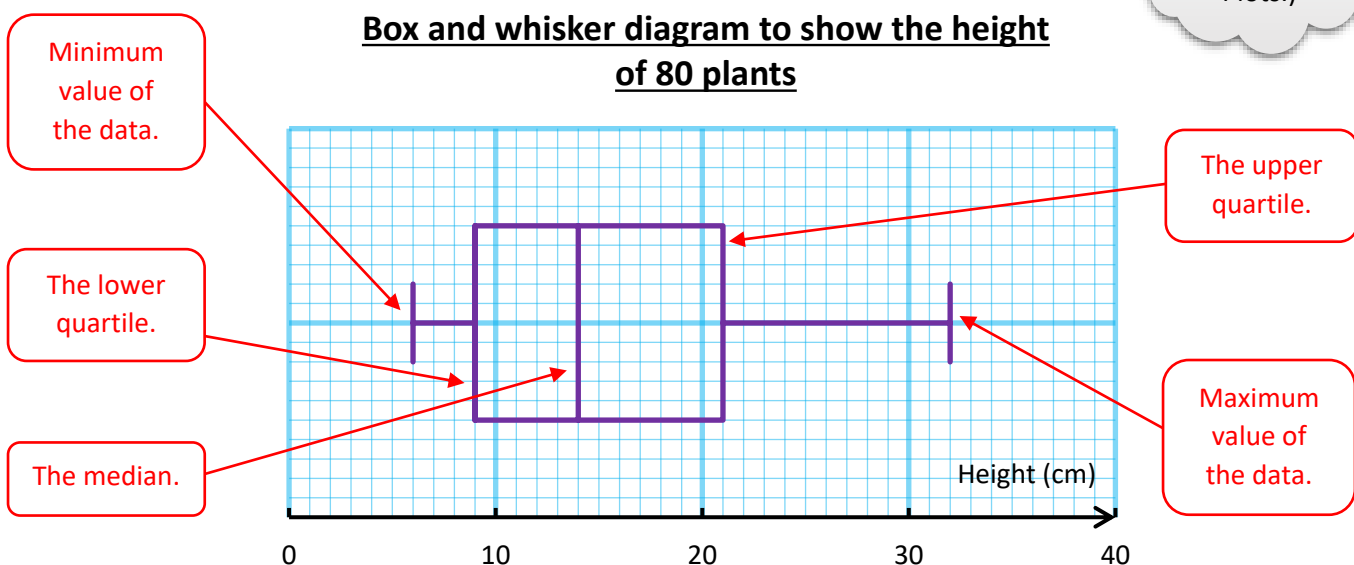


Box and Whisker Diagrams

(Or Box and Whisker Plots.)

A **box and whisker diagram** shows a number of statistics on one diagram.

Box and whisker diagram to show the height of 80 plants



Exercise 44

For the box and whisker diagram shown above, write

- (a) The smallest plant height.
- (b) The lower quartile.
- (c) The median plant height.
- (d) The upper quartile.
- (e) The largest plant height.
- (f) The range of plant heights.
- (g) The interquartile range of plant heights.



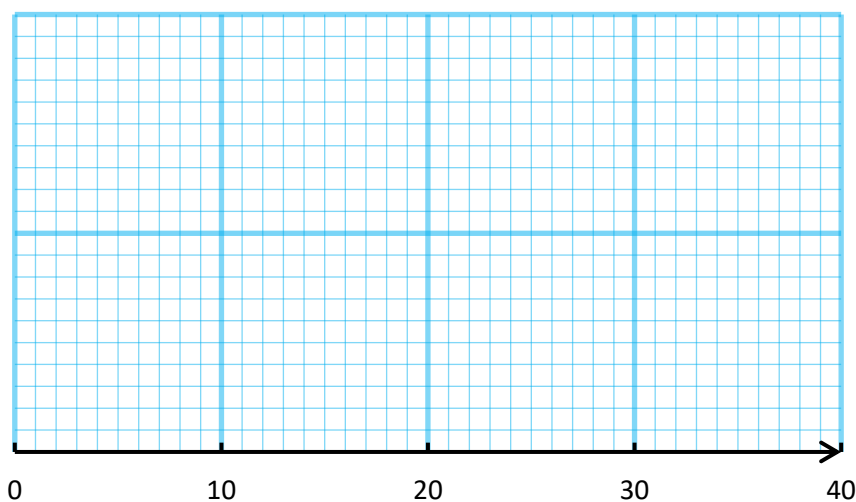
Exercise 45

Moli calculated the following statistics for 50 numbers.

Use the statistics to draw a box and whisker diagram on the graph paper below.

Smallest number = 5 Lower quartile = 12 Median = 19 Upper quartile = 24 Largest number = 34

Box and whisker diagram for Moli's 50 numbers



(You will not have to draw box and whisker diagrams in a GCSE examination, only interpret them.)

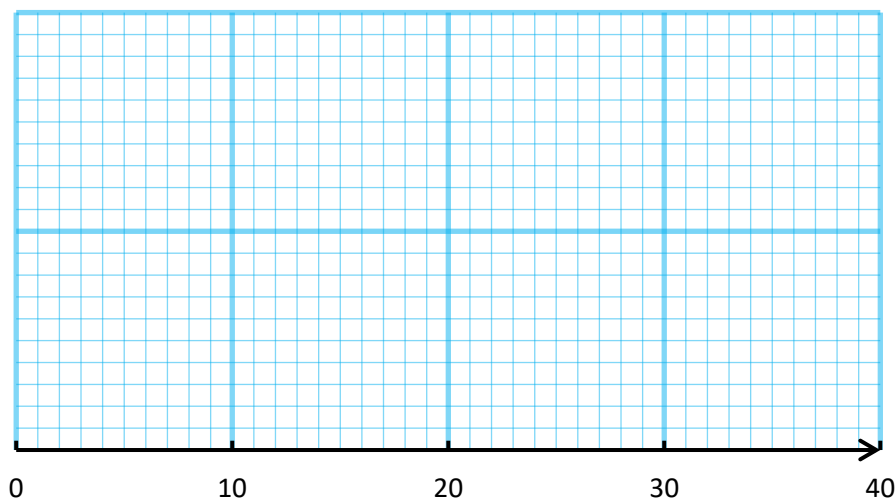
Exercise 46

Dafydd calculates the following statistics for 70 numbers.
Use the statistics to draw a box and whisker diagram on the graph paper below.



Smallest number = 8 Lower quartile = 15 Median = 20 Interquartile range = 12 Range = 27

Box and whisker diagram for Dafydd's 70 numbers



Exercise 47

Draw box and whisker diagrams for the following sets of data.

- (a) 4, 11, 14, 15, 17, 19, 20, 22, 22, 26, 29, 34, 35, 35, 38.
- (b) 24, 13, 9, 35, 3, 17, 21, 30, 12, 28.
- (c) 2, 6, 14, 18, 26, 27, 27, 30, 31.



The Connection between Box and Whisker diagrams and Cumulative Frequency Diagrams

There is a useful connection between cumulative frequency diagrams and box and whisker diagrams.

Example

Consider the following data which shows the width of books sitting on a shelf in a library.

Book width (w mm)	Frequency
$0 < w \leq 10$	3
$10 < w \leq 20$	14
$20 < w \leq 30$	35
$30 < w \leq 40$	8

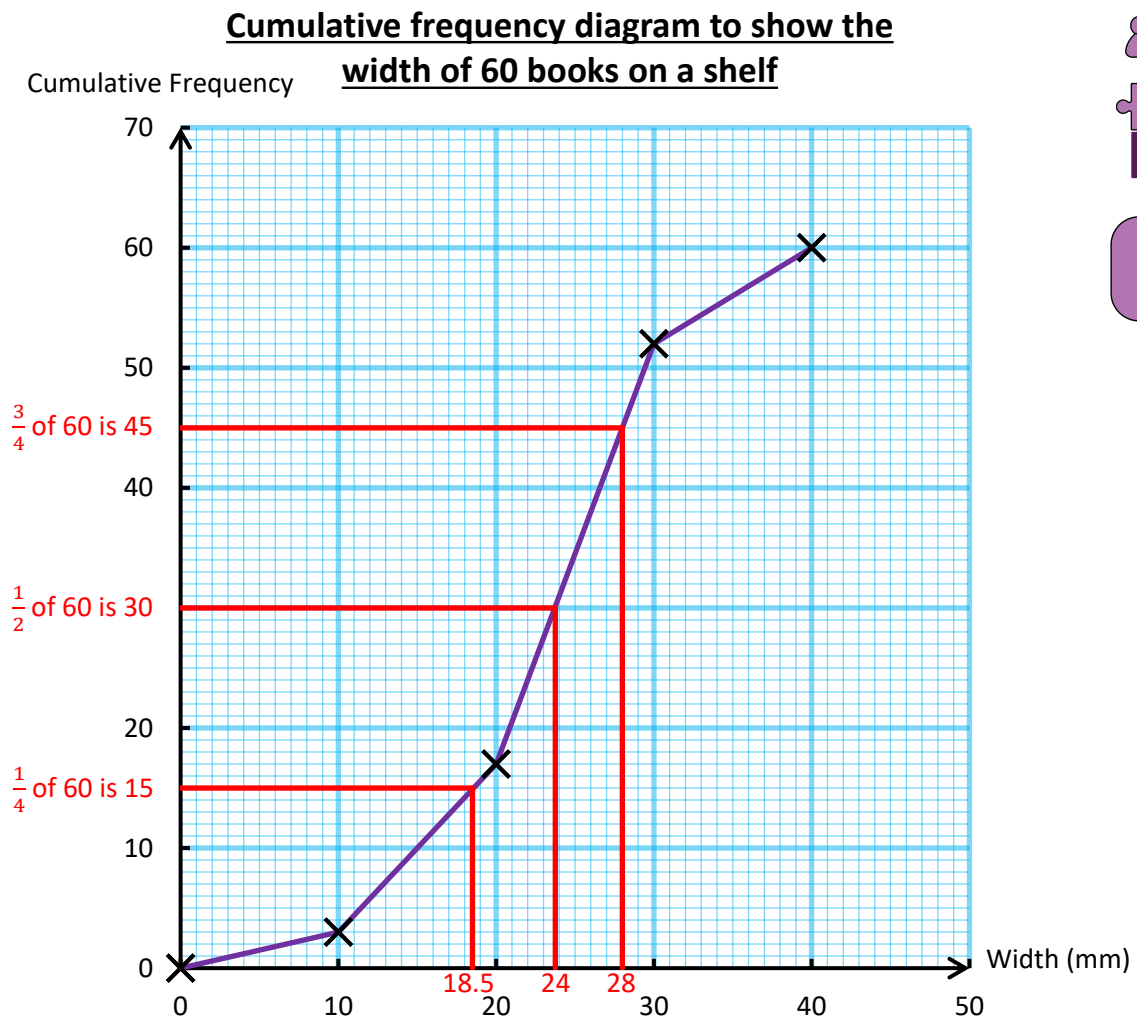
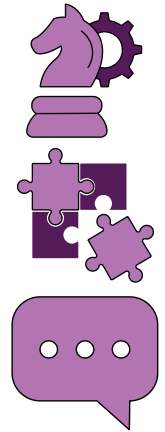
To draw a cumulative frequency diagram for the data we must first draw a cumulative frequency table.

Book width (w mm)	Frequency	Cumulative Frequency
$0 < w \leq 10$	3	3
$10 < w \leq 20$	14	17
$20 < w \leq 30$	35	52
$30 < w \leq 40$	8	60

$(3 + 14 = 17)$
 $(17 + 35 = 52)$
 $(52 + 8 = 60)$

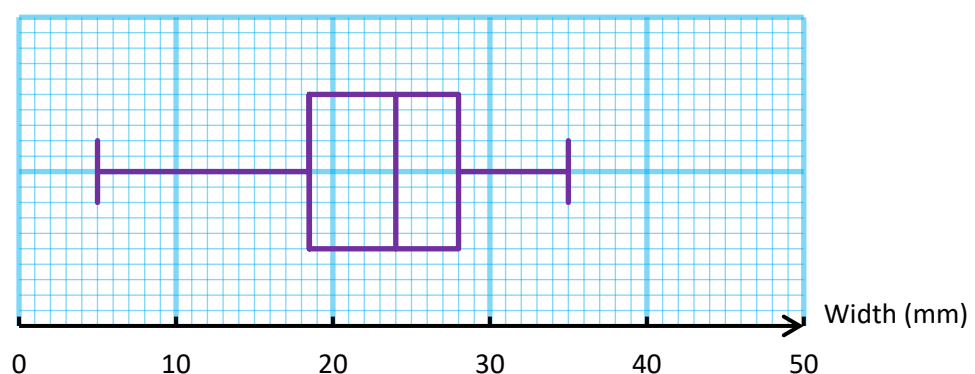


We can now draw a cumulative frequency diagram for the data.



To be able to draw the box and whisker diagram for the data, we need to find the quartiles. We can use the cumulative frequency diagram to estimate these. (These are the red lines on the above diagram). We also need to estimate the minimum value and the maximum value for the data. For the minimum value, we use the midpoint of the first class ($0 < w \leq 10$) which gives 5 mm. For the maximum value, we use the midpoint of the last class ($30 < w \leq 40$) which gives 35 mm. We can now draw the box and whisker diagram for the data.

Box and whisker diagram for the width of 60 books on a shelf



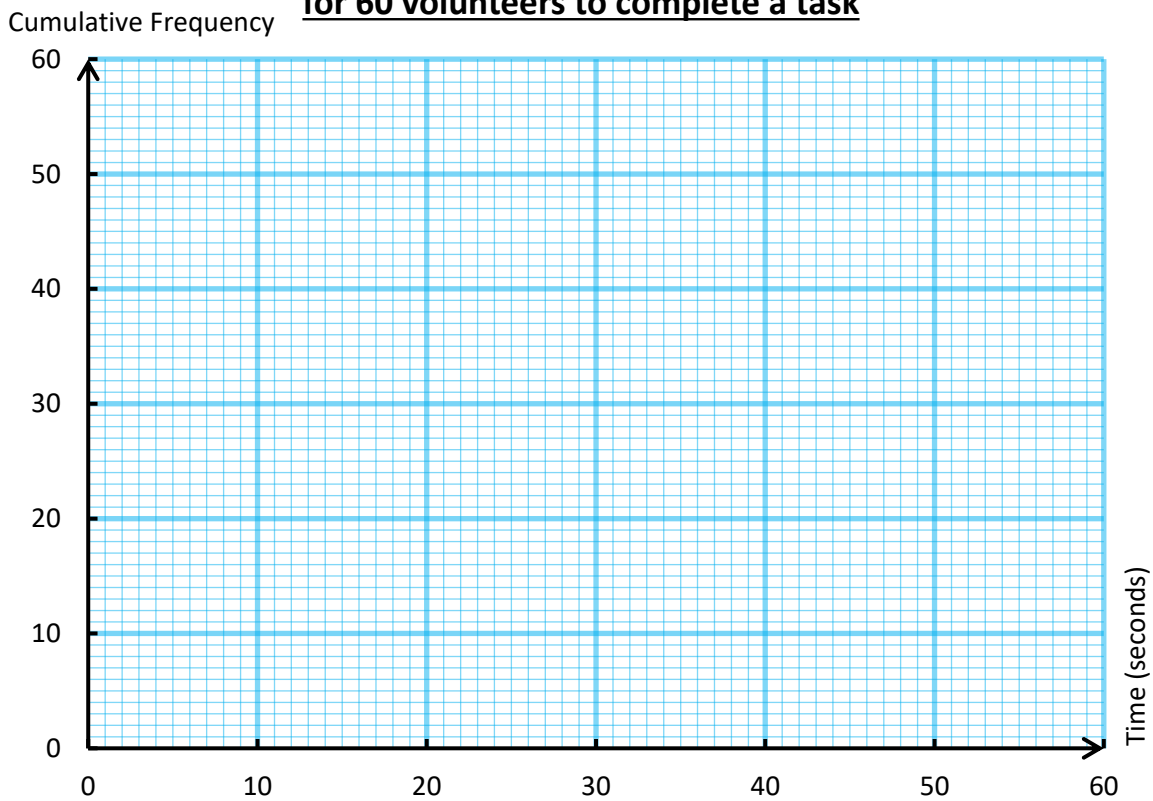
Exercise 48



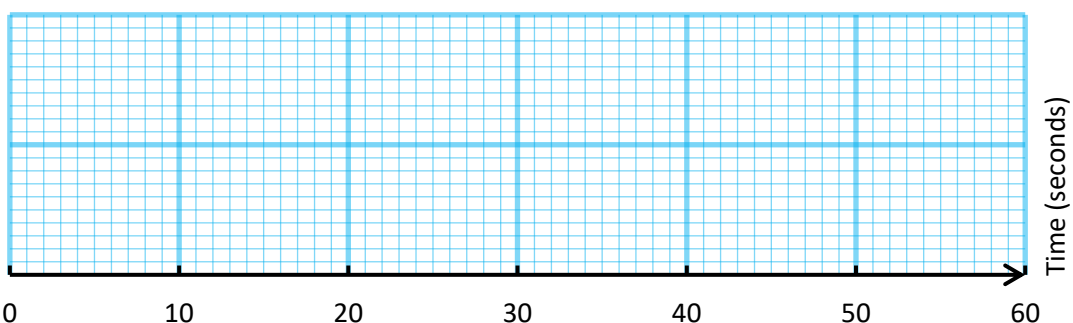
The times taken by 60 volunteers to complete a task were recorded, in seconds. The results are shown in the frequency table below. Complete the cumulative frequency table, the cumulative frequency diagram and the box and whisker diagram for the data.

Time, t , to finish task (seconds)	Frequency	Time, t , to finish task (seconds)	Cumulative Frequency
$15 < t \leq 20$	3	$t \leq 15$	
$20 < t \leq 25$	6	$t \leq 20$	
$25 < t \leq 30$	9	$t \leq 25$	
$30 < t \leq 35$	19	$t \leq 30$	
$35 < t \leq 40$	15	$t \leq 35$	
$40 < t \leq 45$	5	$t \leq 40$	
$45 < t \leq 50$	3	$t \leq 45$	
		$t \leq 50$	

Cumulative frequency diagram to show the times taken for 60 volunteers to complete a task



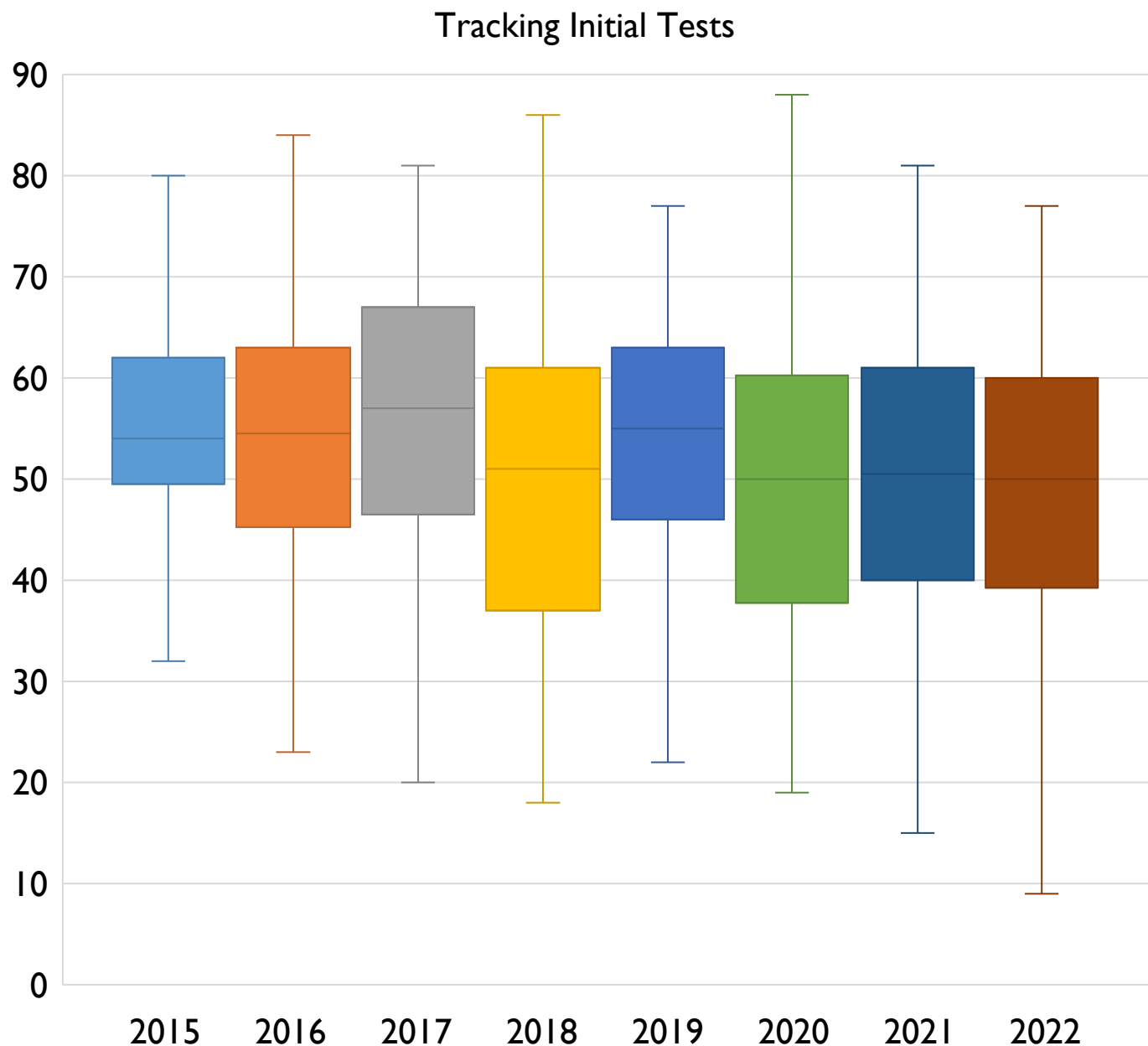
Box and whisker diagram to show the times taken for 60 volunteers to complete a task



Exercise 49

3

The following box and whisker diagrams show the marks (out of 90) gained by year 7 in their mathematics tests when starting school.



- What was the median mark in 2020?
- What was the lower quartile in 2021?
- What was the highest mark in 2015?
- In which year was the highest mark gained? What was this mark?
- What was the range in 2017?
- What was the interquartile range in 2021?
- In which year was the interquartile range at its lowest?
- In which year was the range at its lowest?
- Complete the following sentence: In 2016, 25% of the year gained a mark greater than _____.
- Complete the following sentence: In 2022, 75% of the year gained a mark greater than _____.
- Complete the following sentence: In _____, 50% of the year gained a mark greater than 55.
- In your opinion, which year performed best in the tests? Explain your answer.
- Between 2019 and 2022, the lowest mark decreases. What could be responsible for this?

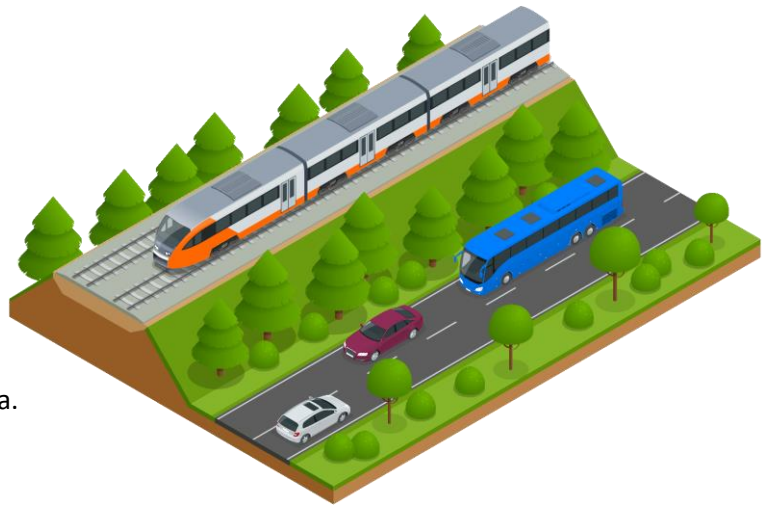


Exercise 50

3

The frequency table below shows the time taken, in minutes, for company workers to travel to work every morning.

Time, t , in minutes	Frequency
$0 < t \leq 10$	3
$10 < t \leq 20$	8
$20 < t \leq 30$	14
$30 < t \leq 40$	6
$40 < t \leq 50$	7
$50 < t \leq 60$	2



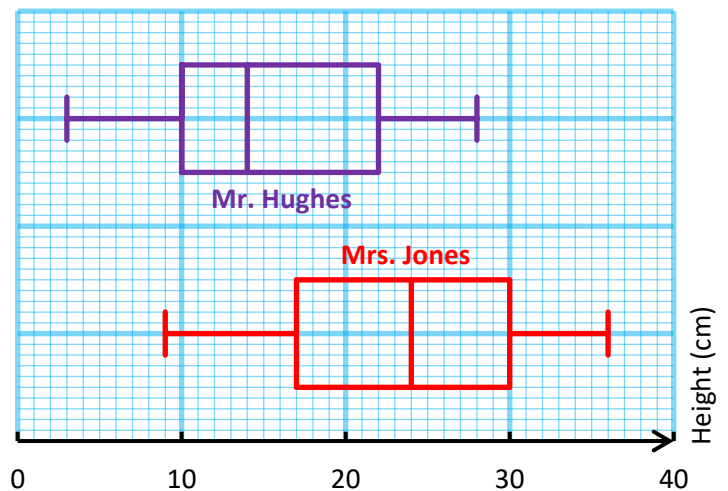
- (a) Draw a cumulative frequency table for the data.
- (b) Draw a cumulative frequency diagram for the data.
- (c) Draw a box and whisker diagram for the data.
- (d) 5 years ago the median time was 22 minutes. How has the journey changed? Give a possible explanation.
- (e) One day the median time was 24 minutes, the upper quartile was 50 minutes and the maximum time was 75 minutes. Explain what could have happened.

Exercise 51

Mr. Hughes and Mrs. Jones have bought the same type of seeds for planting a special plant. The two plant the seeds at the same time and measure the heights of the plants 6 months later. The box and whisker diagrams on the right show the results.

- (a) What is the median height of Mr. Hughes' plants?
- (b) What is the height of Mrs. Jones' tallest plant?
- (c) Calculate the interquartile range of Mr. Hughes' plants.
- (d) One of the people used fertiliser over the last 6 months. Who did this, in your opinion? Explain your answer.

Box and whisker diagrams for the heights of Mr. Hughes and Mrs. Jones' plants 6 months after planting



Evaluation

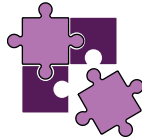
Key words	Corrections	I am happy with...	I need to revise...

Histograms



Consider the following data that shows the heights of parcels in an office one morning.

Height (h cm)	Frequency
$0 < h \leq 10$	3
$10 < h \leq 20$	2
$20 < h \leq 30$	2



It would be possible to draw a frequency diagram for this data; this is shown on the right.

Consider now that the same data is grouped as shown below.

Height (h cm)	Frequency
$0 < h \leq 10$	3
$10 < h \leq 30$	4

On drawing a frequency diagram for this data (shown on the right), the diagram does not give a fair reflection of the data – it is unfair to compare classes of different widths. To deal with this, we introduce the idea of drawing a **histogram** to show the data, where we plot not the height against the frequency, but the height against the **frequency density**.

$$\text{Frequency Density} = \frac{\text{Frequency}}{\text{Class width}}$$

For the second set of data above, we can calculate the frequency density using the table below.

Height (h cm)	Frequency	Frequency Density
$0 < h \leq 10$	3	$3 \div 10 = 0.3$
$10 < h \leq 30$	4	$4 \div 20 = 0.2$

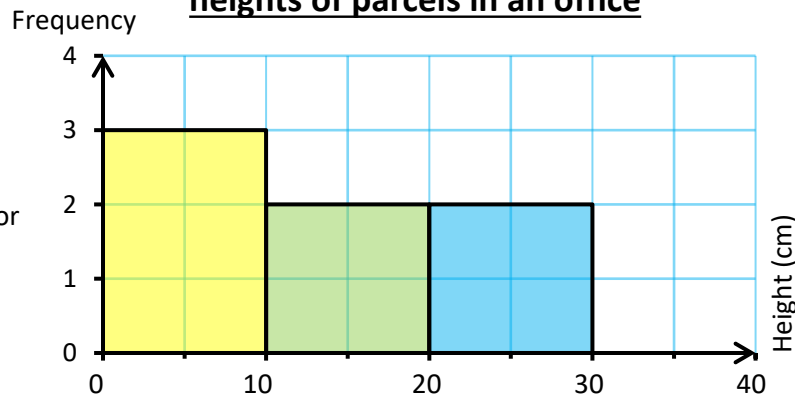
We can then draw the histogram on the right to illustrate the data. The histogram gives a fairer reflection of the data, as it takes into account that the second class is wider than the first class.

The **area** of a bar in a histogram gives the frequency of the class under consideration.

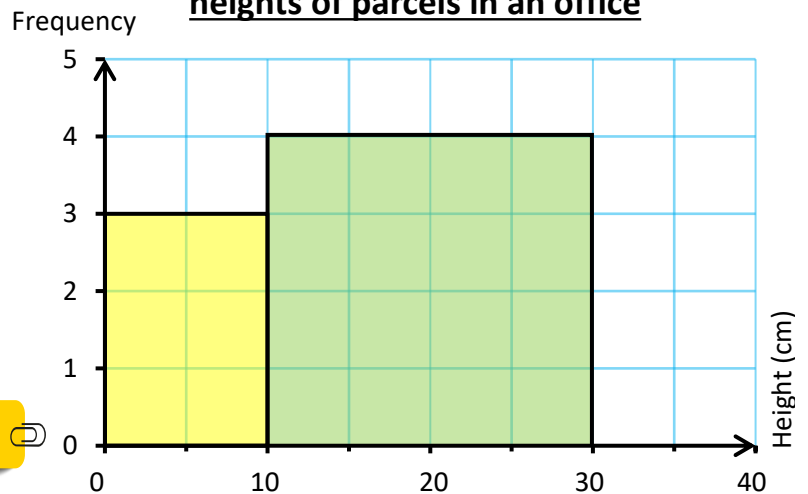
$$\text{Frequency} = \text{Class width} \times \text{Frequency Density}$$

For the histogram shown on the right, the frequency corresponding to the first class is $10 \times 0.3 = 3$, and the frequency corresponding to the second class is $20 \times 0.2 = 4$.

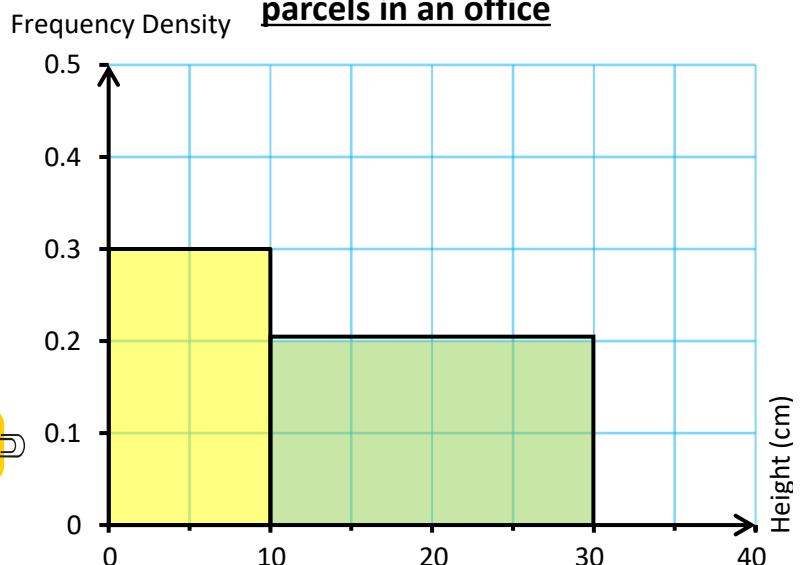
Frequency diagram to show the heights of parcels in an office



Frequency diagram to show the heights of parcels in an office



Histogram to show the heights of parcels in an office





3

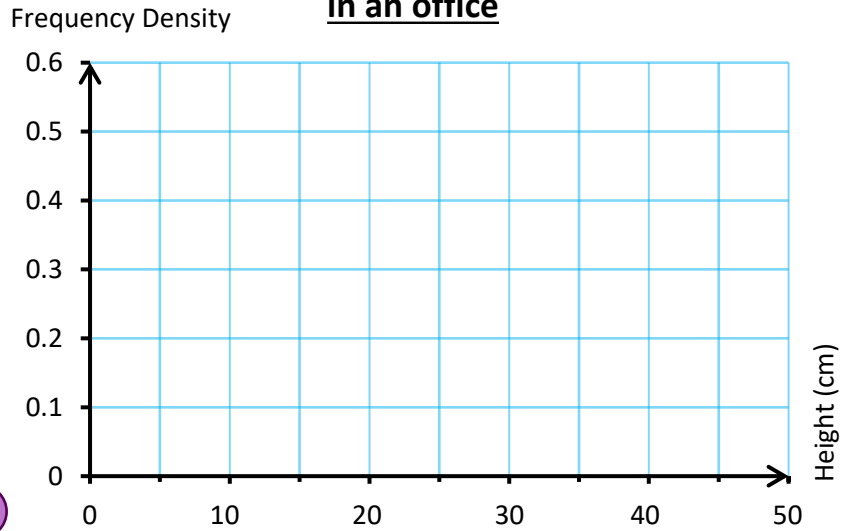
Exercise 52

The following data shows the heights of parcels in an office one morning.

Height (h cm)	Frequency	Frequency Density
$0 < h \leq 10$	5	
$10 < h \leq 30$	6	
$30 < h \leq 40$	2	

- (a) Complete the 'Frequency Density' column in the table.
- (b) Use the squared paper to draw a histogram for the data.

Histogram to show the heights of parcels in an office

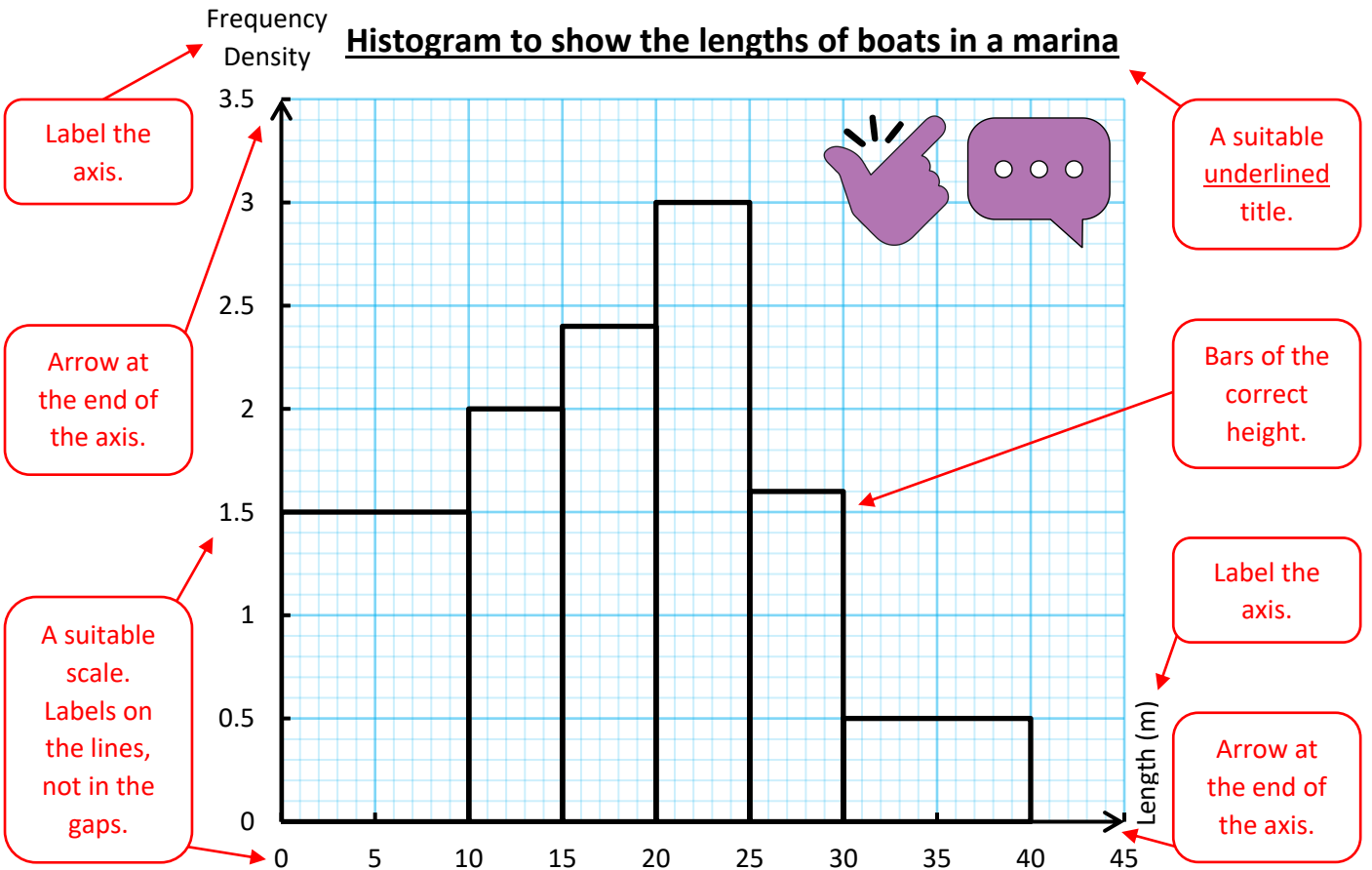


Exercise 53

3

The histogram below shows the length of boats in a marina.

Histogram to show the lengths of boats in a marina



(a) How many boats had a length between 0 m and 10 m?

(b) Complete the frequency table:

Length (l m)	Frequency
$0 < l \leq 10$	
$10 < l \leq 15$	
$15 < l \leq 20$	
$20 < l \leq 25$	
$25 < l \leq 30$	
$30 < l \leq 40$	

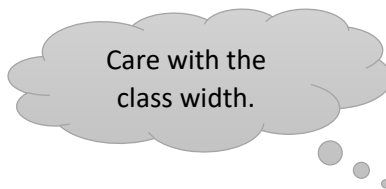
(c) How many boats had a length of less than or equal to 20 m?

(d) How many boats were measured in total?



Exercise 54

Draw histograms for the following sets of data.



(a) The sum raised by a group of people for a charity.

Sum raised (£s)	Frequency
$0 < s \leq 50$	6
$50 < s \leq 100$	22
$100 < s \leq 200$	31
$200 < s \leq 500$	42
$500 < s \leq 1,000$	15

(b) The age of people in a hockey club.

Age in years	Frequency
11–15	7
16–18	10
19–24	15
25–34	20
35–49	12
50–64	7

(c) The earnings of a group of students during one week.

Earnings (£e)	Frequency
$0 < e \leq 20$	5
$20 < e \leq 40$	15
$40 < e \leq 70$	27
$70 < e \leq 100$	30
$100 < e \leq 150$	6

(d) The weights of passengers' bags on an aircraft.

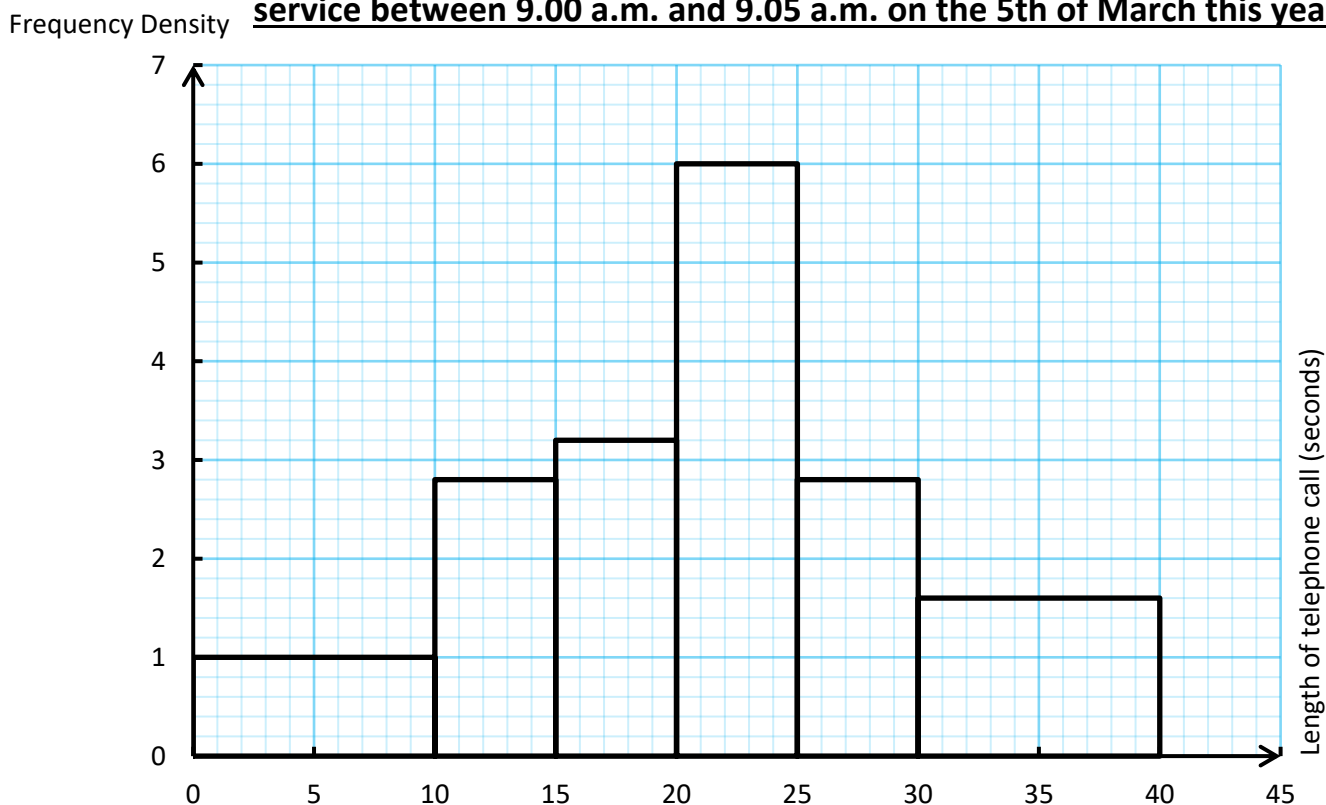
Weight (w kg)	Frequency
$0 < w \leq 5$	7
$5 < w \leq 10$	12
$10 < w \leq 20$	24
$20 < w \leq 40$	15
$40 < w \leq 50$	3

Exercise 55

The histogram below shows the lengths of telephone calls made to a directory enquiries service between 9.00 a.m. and 9.05 a.m. on the 5th of March this year.



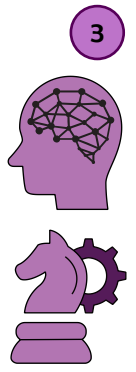
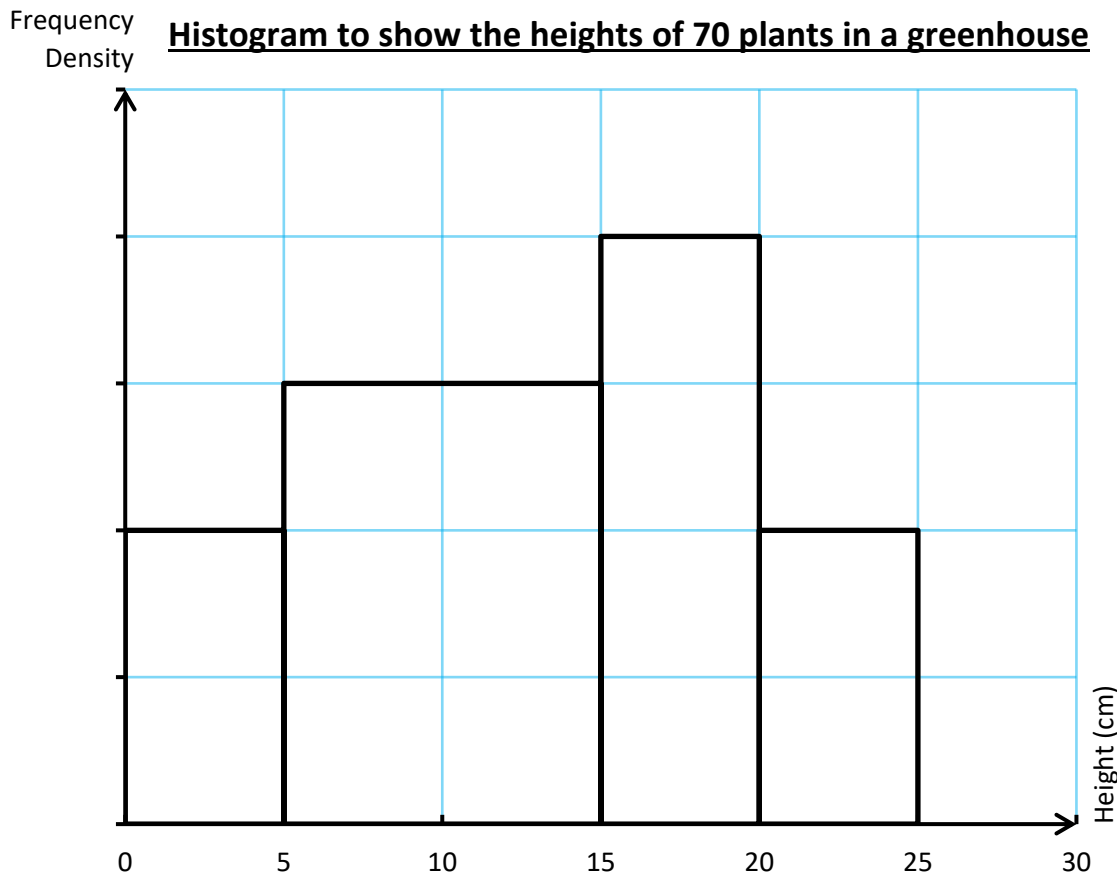
Histogram to show the lengths of telephone calls made to a directory enquiries service between 9.00 a.m. and 9.05 a.m. on the 5th of March this year



Use the histogram to calculate how many telephone calls were made to the directory enquiries service between 9.00 a.m. and 9.05 a.m. on the 5th of March this year.

Exercise 56

The following histogram shows the height distribution of 70 plants in a greenhouse.



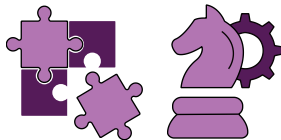
- (a) Complete the missing scale on the vertical axis.
- (b) How many plants had a height of between 15 cm and 20 cm?
- (c) Complete the following frequency table.

Height (h cm)	Frequency
$0 < h \leq 5$	
$5 < h \leq 15$	
$15 < h \leq 20$	
$20 < h \leq 25$	



- (d) Calculate an estimate of the amount of plants of height less than 10 cm.
- (e) What is the modal class of the data?
- (f) Calculate an estimate of the mean height of the plants.
- (g) Calculate an estimate of the range of the heights of the plants.
- (h) What percentage of all the plants have a height of more than 20 cm?
- (i) What fraction of all the plants have a height of more than 5 cm? Give your answer in its simplest form.
- (j) Which class is the median class of the data?

Estimating the median from a histogram



For any histogram,

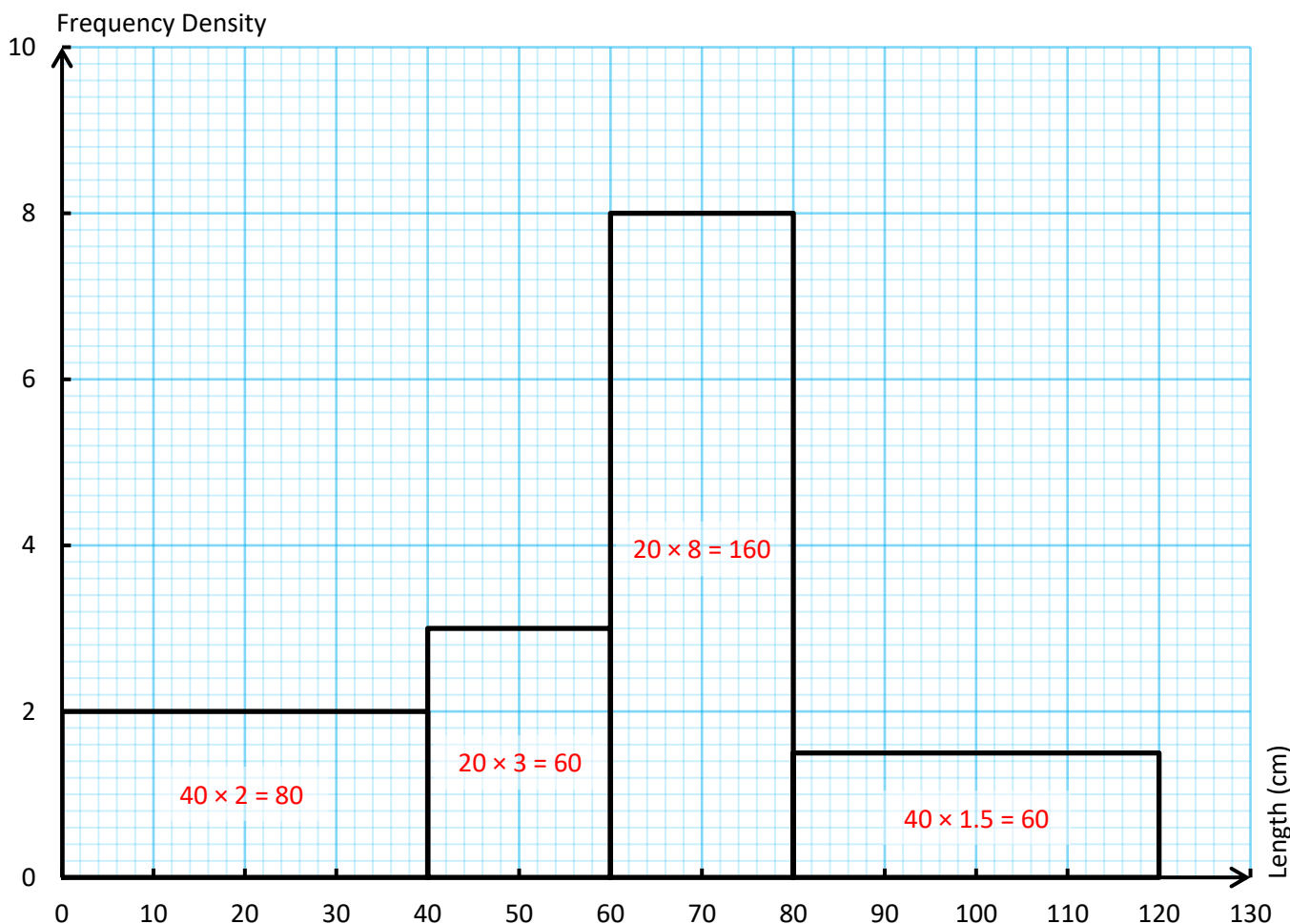


The estimate of the median is the vertical line in the histogram that halves the total area of the histogram.

Example

Let us consider the following histogram that represents the results of collecting and measuring the lengths of driftwood on a beach.

Histogram to show the lengths of driftwood on a beach



By calculating the area of each bar in the histogram (shown above in red), and adding the results, we see that a total of $80 + 60 + 160 + 60 = 360$ pieces of driftwood were collected and measured.

To estimate the median length of a piece of driftwood, we need to draw a vertical line in the histogram that halves the total area of the histogram. Because $360 \div 2 = 180$, we need to draw a vertical line in the histogram so that an area of 180 squared units is found on either side of the vertical line. This line must be in the third bar, as $80 + 60 = 140$ is less than 180, and $80 + 60 + 160 = 300$ is greater than 180.

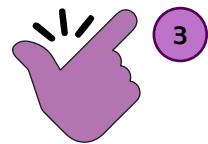
We need to travel across the third bar by the fraction $\frac{180-140}{160} = \frac{40}{160} = \frac{1}{4}$.

The width of the third bar is 20 cm, so we need to travel across the third bar by a distance of $20 \times \frac{1}{4} = 5$ cm.

So, the estimate of the median length of a piece of driftwood is $60 + 5 = 65$ cm.

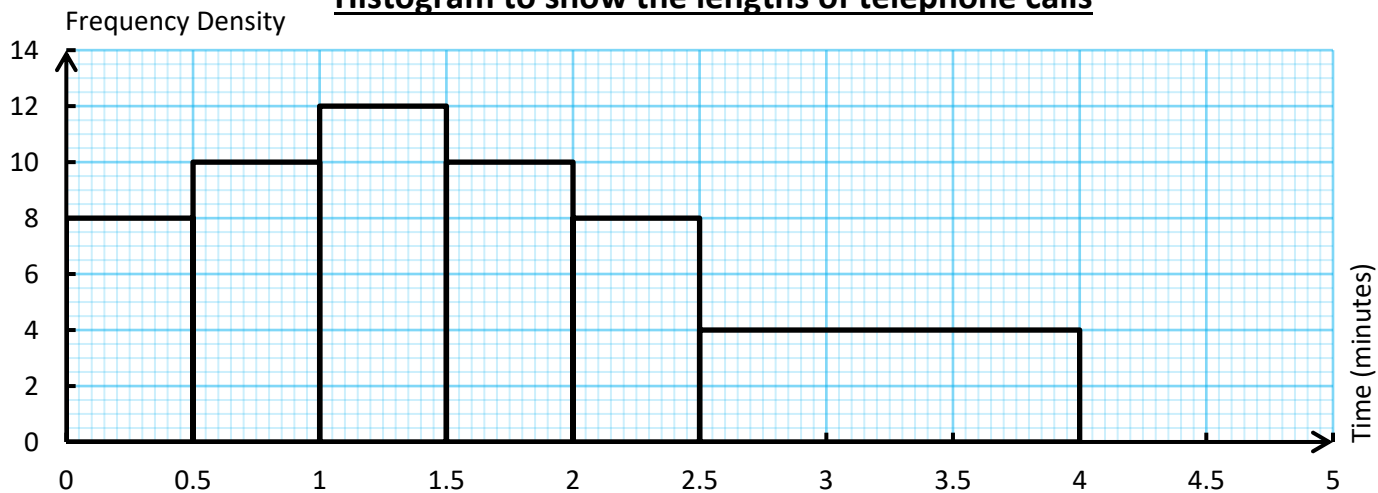


Exercise 57



The histogram below represents the results of recording the length of a number of telephone calls.

Histogram to show the lengths of telephone calls



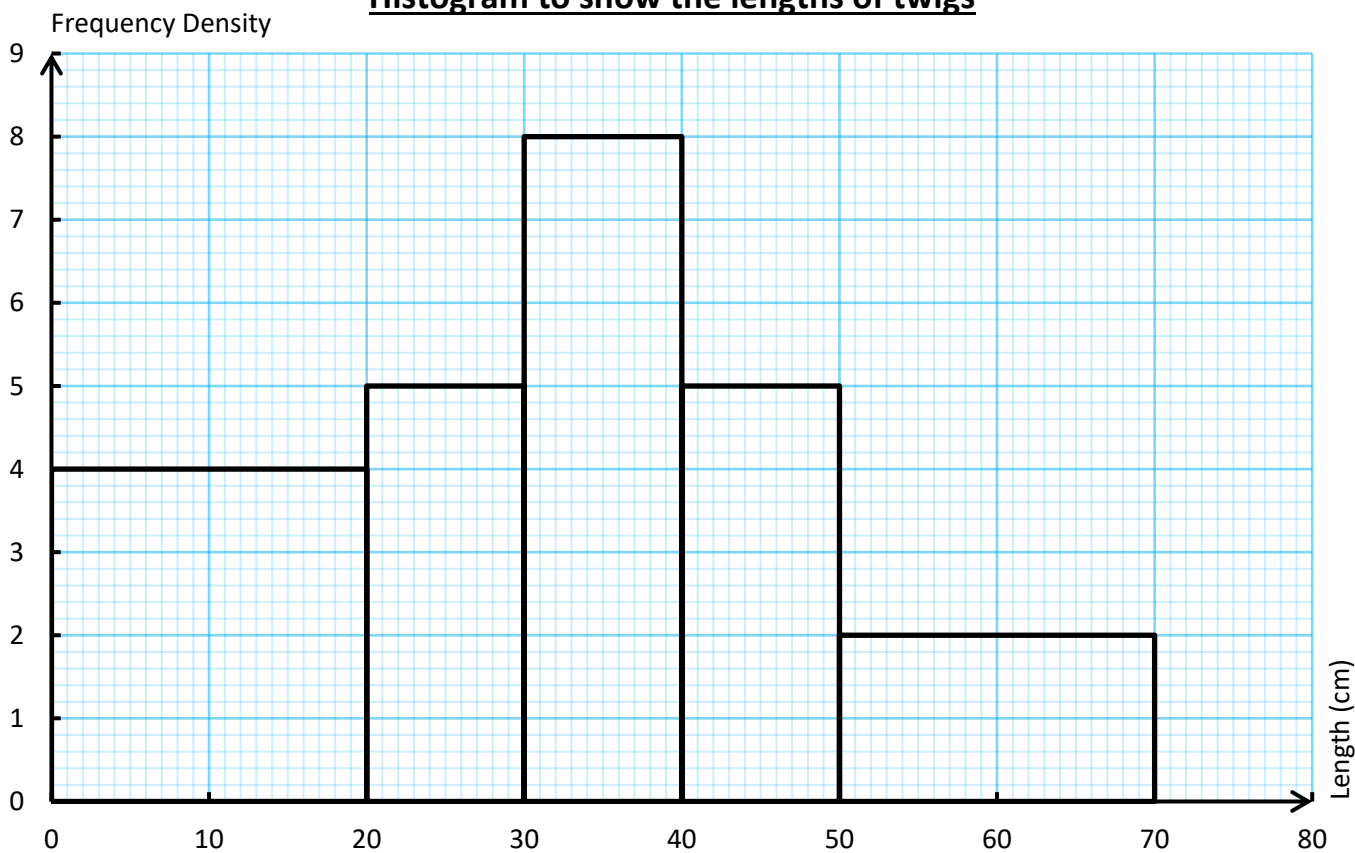
- (a) Use the histogram to calculate the total number of telephone calls.
- (b) Find an estimate for the median length of a telephone call, in minutes.

Exercise 58



The histogram below represents the results of recording the length of a number of twigs.

Histogram to show the lengths of twigs



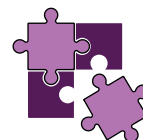
- (a) Use the histogram to calculate the total number of twigs.
- (b) Find an estimate for the median length of a twig, in cm.

Estimating the quartiles from a histogram

For any histogram,

The estimate of the lower quartile is the vertical line in the histogram that splits the histogram's area into the ratio 1 : 3.

The estimate of the upper quartile is the vertical line in the histogram that splits the histogram's area into the ratio 3 : 1.



Example

Let us again consider the histogram from page 36 that represents the results of collecting and measuring the lengths of driftwood on a beach.

By calculating the area of each bar in the histogram, and adding the results, we see that a total of $80 + 60 + 160 + 60 = 360$ pieces of driftwood were collected and measured.

To estimate the lower quartile, we need to draw a vertical line in the histogram that splits the histogram's area into the ratio 1 : 3. Because $360 \div 4 = 90$, we need to draw a vertical line in the histogram so that an area of 90 squared units lies to the left of the vertical line, and an area of $90 \times 3 = 270$ squared units lies to the right of the vertical line. This line must be in the second bar, as 80 is less than 90, and $80 + 60 = 140$ is greater than 90.

We need to travel across the second bar by the fraction $\frac{90-80}{60} = \frac{10}{60} = \frac{1}{6}$.

The width of the second bar is 20 cm, so we need to travel across the second bar by a distance of $20 \times \frac{1}{6} = \frac{10}{3} = 3\frac{1}{3}$ cm.

So, the estimate of the lower quartile is $40 + 3\frac{1}{3} = 43\frac{1}{3}$ cm.

To estimate the upper quartile, we need to draw a vertical line in the histogram that splits the histogram's area into the ratio 3 : 1. Because $360 \div 4 = 90$, and $90 \times 3 = 270$, we need to draw a vertical line in the histogram so that an area of 270 squared units lies to the left of the vertical line, and an area of 90 squared units lies to the right of the vertical line. This line must be in the third bar, as $80 + 60 = 140$ is less than 270, and $80 + 60 + 160 = 300$ is greater than 270.

We need to travel across the third bar by the fraction $\frac{270-140}{160} = \frac{130}{160} = \frac{13}{16}$.

The width of the third bar is 20 cm, so we need to travel across the third bar by a distance of $20 \times \frac{13}{16} = 16.25$ cm.

So, the estimate of the upper quartile is $60 + 16.25 = 76.25$ cm.

Exercise 59

3

For the histogram in Exercise 57,

- Find an estimate for the lower quartile;
- Find an estimate for the upper quartile.

Exercise 60

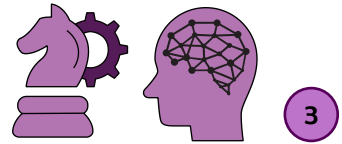
3

For the histogram in Exercise 58,

- Find an estimate for the lower quartile;
- Find an estimate for the upper quartile.

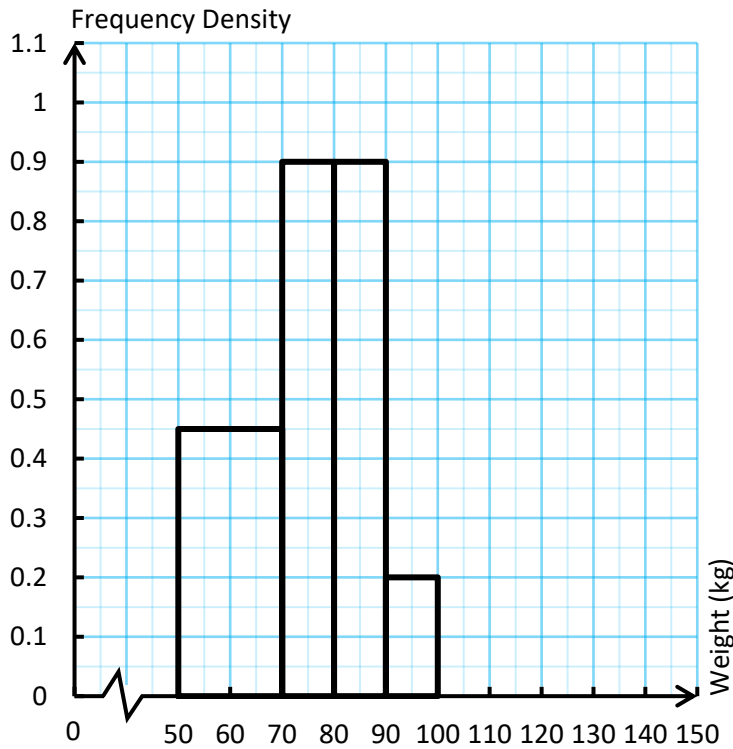


Comparing histograms

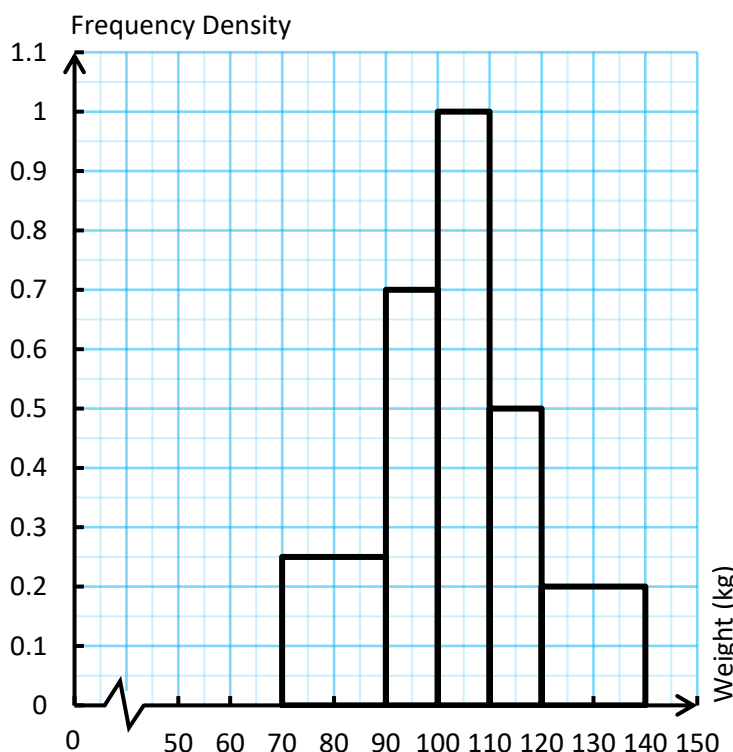


Exercise 61

A histogram to show the weight of Wales' rugby squad at the 2017 World Cup (women)



A histogram to show the weight of Wales' rugby squad at the 2019 World Cup (men)



The two histograms on the left show information about the weights of Wales' rugby squads at the 2017 World Cup (women) and the 2019 World Cup (men).

- (a) How many women weighed between 70 kg and 80 kg?
- (b) How many men weighed between 110 kg and 120 kg?
- (c) Complete the following frequency table for the women.

Weight (w kg)	Frequency
$50 < w \leq 70$	
$70 < w \leq 80$	
$80 < w \leq 90$	
$90 < w \leq 100$	

- (d) Complete the following frequency table for the men.

Weight (w kg)	Frequency
$70 < w \leq 90$	
$90 < w \leq 100$	
$100 < w \leq 110$	
$110 < w \leq 120$	
$120 < w \leq 140$	

- (e) How many women were in the squad in total?
- (f) How many men were in the squad in total?
- (g) Find an estimate for the median weight of a woman in the 2017 rugby squad.
- (h) Find an estimate for the median weight of a man in the 2019 rugby squad.
- (i) On average, which squad was heaviest?
- (j) What is the greatest possible range of the women's rugby squad?
- (k) What is the greatest possible range of the men's rugby squad?
- (l) Use your answers to (j) and (k) above to comment on which squad had the most consistent weight.

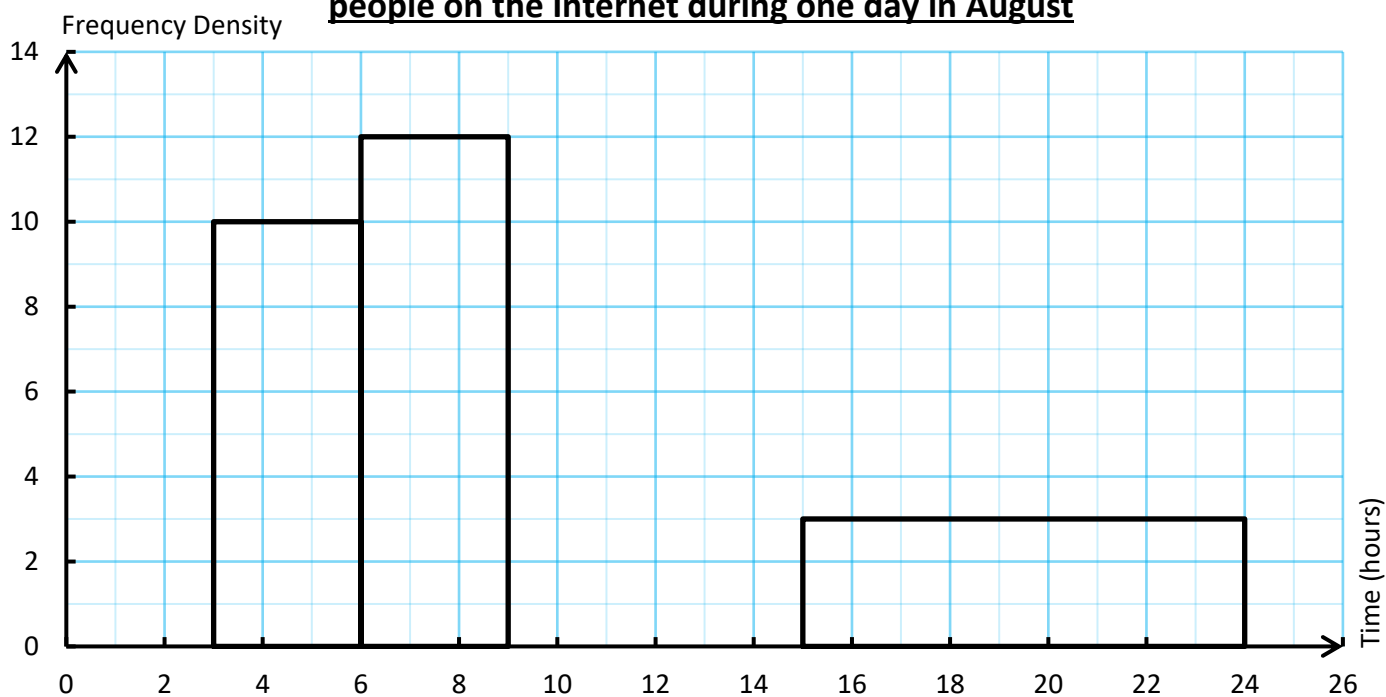
Exercise 62 (Revision)

3

The following histogram and frequency table shows some information about the time each person, in a group of people, spent on the Internet during one day in August.

Time (t hours)	Frequency
$0 < t \leq 3$	24
$3 < t \leq 6$	
$6 < t \leq 9$	36
$9 < t \leq 15$	30
$15 < t \leq 24$	

Histogram to show the time spent by a group of people on the Internet during one day in August



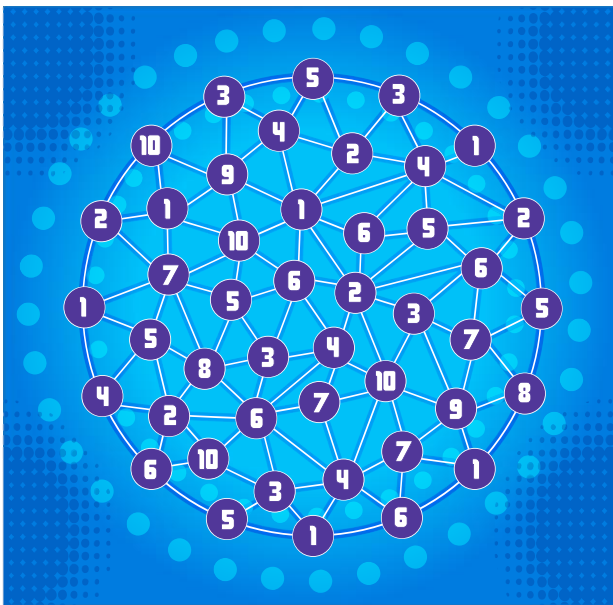
- (a) Complete the frequency table and histogram shown above.
- (b) Calculate an estimate for the median time spent on the Internet by the group of people during the day in August.

Evaluation

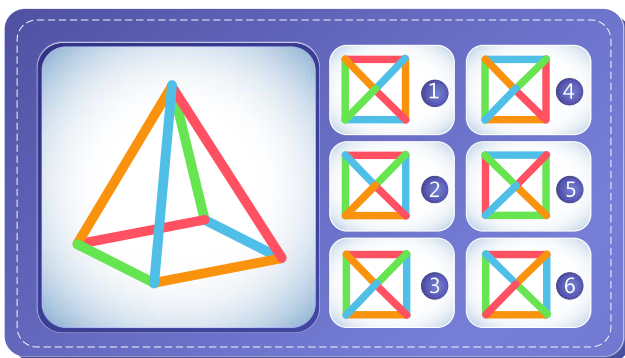
Key words	Corrections	I am happy with...	I need to revise...

Puzzles

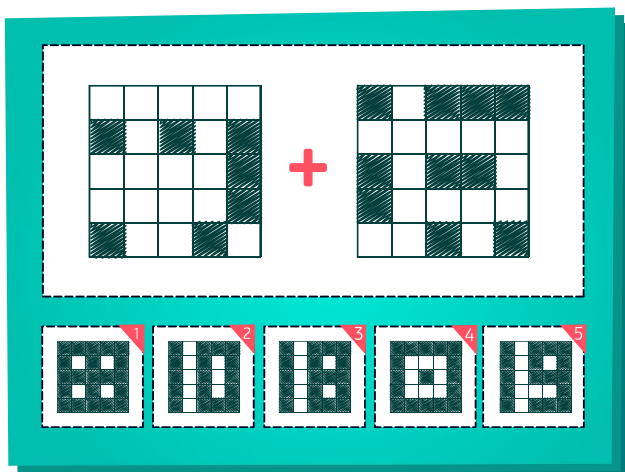
(a) Connect the numbers from 1 to 10.



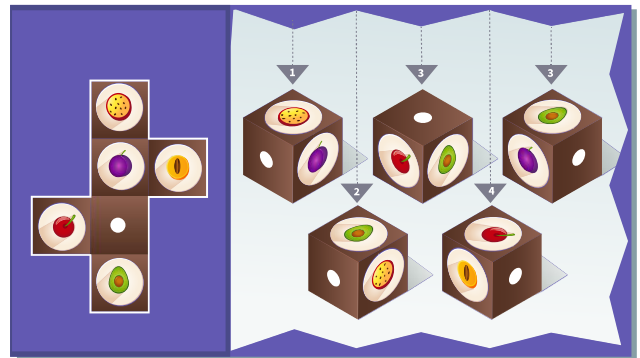
(b) Which is the correct plan view?



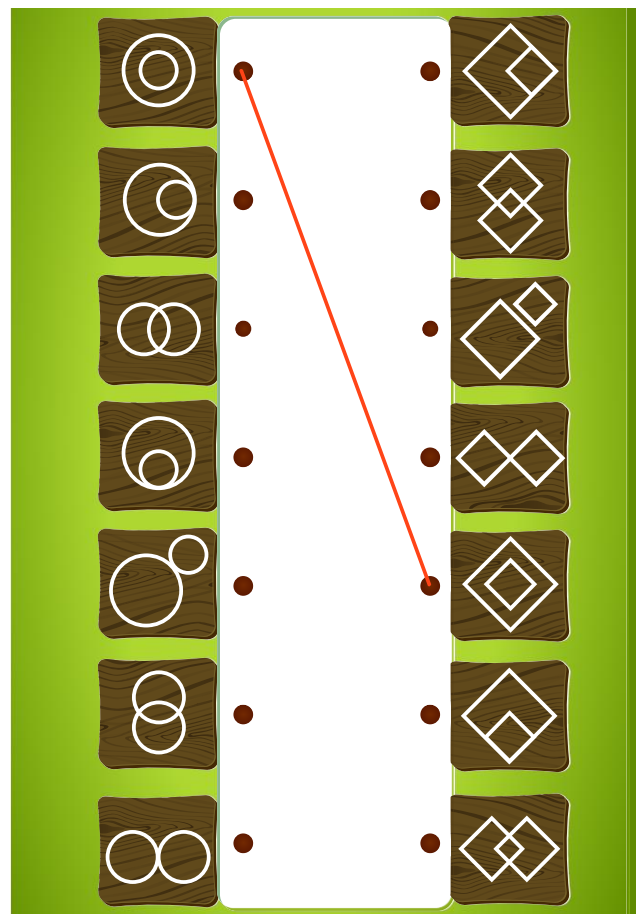
(c) Join the pictures to deduce which image is formed.



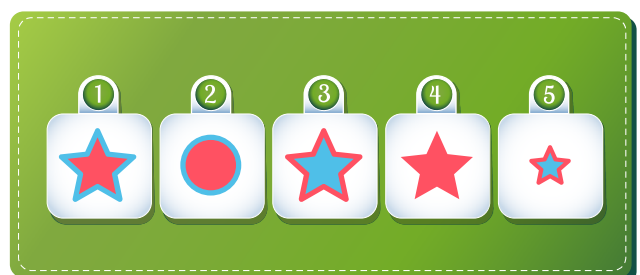
(d) Which cube is formed by folding the net?



(e) Pair the pictures.



(f) Which picture is the odd one out?







Reflection Sheet

Name:

Percentage in the test:

	I know this. 	I need to revise this. 	Question in the test:	Correct in the test?
I can calculate the mode , the median , the mean and the range for discrete data .				
I can calculate the modal class , the median class , an estimate of the mean and an estimate of the range for grouped data .			4	
I know how to decide which average is most suitable for a set of data.			9	
I know how to use averages and measures of spread to compare two data sets.			9	
I can find the original data set given information about averages and the range.			6	
I know how to criticise questions in a questionnaire .			1	
I know how to write unfavourable comments about a plan to prove a specific hypothesis .			1	
I know how to choose a simple random sample using a table of random digits or using the random number generator function on a calculator .			2	
I know how to calculate the sampling interval for a systematic sample.			3	
I know how to choose a systematic sample .			3	
I know how to choose a stratified sample .			10	
I know how to draw a frequency polygon .			4	
I know how to interpret a frequency polygon.				
I know how to draw a box and whisker diagram .			5	
I can use a cumulative frequency diagram in order to draw a box and whisker diagram.			7	
I can draw a histogram for a specific data set.			8	
I can re-create the frequency table for a specific histogram.			8	
I can estimate the median for a specific histogram.			8	
I can estimate the lower quartile and upper quartile for a specific histogram.				
I can compare two histograms using averages and measures of spread.				