18 PROCESSING, REPRESENTING AND INTERPRETING DATA

When you buy food, the packaging gives you information about the nutritional value of that food, but you will need to interpret it to understand what it means for your health. For example, a grilled salmon fillet gives you 30g of protein but unless you know that a woman needs approximately 46g of protein a day and a man approximately 56g, this is of little use. Now you can work out that for a woman, the salmon fillet gives her about 65% of her daily protein intake and for a man about 54%.

Objectives

In this chapter you will be able to produce and interpret the following, for various types of data:

- pie charts
- stem and leaf diagrams
- bar charts and composite bar charts
- frequency diagrams
- histograms for continuous data
- frequency polygons
- cumulative frequency graphs
- box plots.

Before you start

You need to be able to:

- measure and draw angles to the nearest degree
- measure and draw lines to the nearest mm
- understand grouped data.

18.1 Producing pie charts

Objective

- You can represent the proportions of different categories of data using a pie chart.

Why do this?

When a council collects council tax they like to show the taxpayers how they are spending their money. They might use a pie chart to show the proportions spent on different things.

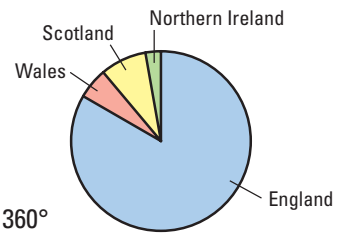
Get Ready

1. How many degrees are there in a circle?
2. How many degrees are there in a quarter-circle?
3. What is **a** $\frac{1}{3}$ of 360 **b** $\frac{2}{8}$ of 180 **c** $\frac{4}{6}$ of 90?

Key Points

- A **pie chart** is often used to show data. It shows how the total is split up between the different categories.
- In a pie chart the area of the whole circle represents the total number of items.
- The area of a **sector** represents the number of items in the category represented by that sector.

  This pie chart shows how the population of the United Kingdom is split between the different countries.
  It shows that the lowest number of people live in Northern Ireland.
  The greatest number of people live in England.



- The angles at the centre must add up to 360°.
- The angle for a particular sector is found as follows: sector angle $= \dfrac{\text{frequency} \times 360°}{\text{total frequency}}$

Example 1

The table shows the number of theatre-goers who attended each type of performance at least once in a 1-year period.

Performance	Musical	Play	Entertainment	Dance	Opera
Number	38	27	14	17	24

Draw a pie chart to represent this information.

Total frequency = 38 + 27 + 14 + 17 + 24 = 120
Use the formula to find each angle.

Musical $\dfrac{38}{120} \times 360° = 114°$

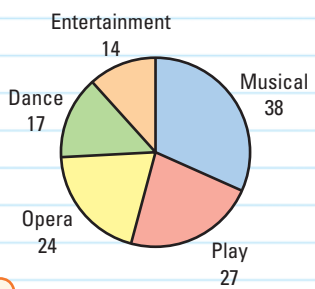
Play $\dfrac{27}{120} \times 360° = 81°$

Entertainment $\dfrac{14}{120} \times 360° = 42°$

Dance $\dfrac{17}{120} \times 360° = 51°$

Opera $\dfrac{24}{120} \times 360° = 72°$

Check: 114 + 81 + 42 + 51 + 72 = 360



ResultsPlus
Examiner's Tip

Add the angles together to make sure they add up to 360°.

Exercise 18A

Questions in this chapter are targeted at the grades indicated.

* 1 The numbers of drinks dispensed by a vending machine in one day are shown in the table.

Type of drink	Tea	Black coffee	Chocolate	Orange	Coke	Latte
Number of drinks	54	42	18	30	12	24

Draw a pie chart to represent these data. Use a radius of 4 cm.

* 2 The snack bar at a bus station sold 120 sandwiches one lunch time.
The table shows the number of each type of sandwich sold.

Type of sandwich	Cheese	BLT	Tuna	Prawn	Ham	Chicken
Number of drinks	10	35	20	30	15	10

Draw a pie chart to represent these data. Use a radius of 4 cm.

* 3 A factory manager asks the employees how they travel to work.
The table shows these data.

Method of getting to work	Walk	Car	Cycle	Train	Motorbike
Number of employees	14	32	12	10	4

Draw a pie chart to represent these data.

18.2 Interpreting pie charts

Objective

- You can interpret a pie chart.

Why do this?

Election results are often shown in a pie chart. You can interpret these graphs to see how many people voted for each party.

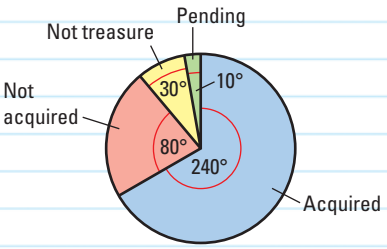
Key Points

- To read frequencies from a pie chart use the formula

$$\text{Frequency} = \frac{\text{sector angle} \times \text{total frequency}}{360°}$$

- The frequency represented by corresponding sectors in two pie charts is dependant upon the total populations represented by each of the pie charts.

Example 2 The pie chart shows the number of Bronze Age finds made with a metal detector and the outcomes when they were submitted to a local museum.
There were 36 finds altogether.



ResultsPlus
Watch Out!
In an exam 'work out' means calculate the frequency, so don't just measure the angle.

a Which type of outcome was most frequent?
b Work out the frequency for each outcome.

a 'Acquired' was the most common outcome.

b Acquired frequency $= \dfrac{\text{sector angle} \times \text{total frequency}}{360^\circ} = \dfrac{240^\circ \times 36}{360^\circ} = 24$

$$\text{Not acquired frequency} = \frac{80^\circ \times 36}{360^\circ} = 8$$

$$\text{Not treasure frequency} = \frac{30^\circ \times 36}{360} = 3$$

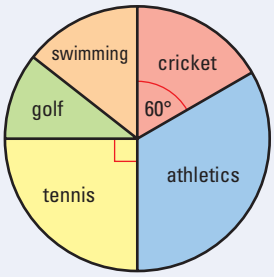
$$\text{Pending frequency} = \frac{10^\circ \times 36}{360} = 1$$

Check: $24 + 8 + 3 + 1 = 36$

ResultsPlus
Examiner's Tip
Always add up the frequencies for each sector to make sure they total to the right number.

Exercise 18B

1 The pie chart shows how the 180 boys in Year 11 at Windup Academy chose from five sports options.



a Write down the least popular option.
b Write down the most popular option.
c Work out how many boys chose tennis.
d Work out how many boys chose cricket.

* 2 A company owns two coffee shops in Twyfield.
They do a survey to find the number of each type of coffee they sell between 9 am and 10 am on one particular day.

	Frequency	
Coffee type	**Shop A**	**Shop B**
Espresso	5	5
Americano	15	12
Latte	10	24
Mocha	40	20
Cappuccino	20	11

Compare and contrast the information by drawing two pie charts.

18.3 Representing and interpreting data in a stem and leaf diagram

Objectives

- You can represent data as a stem and leaf diagram.
- You can use a stem and leaf diagram to find the mode, median, range and quartiles of a set of data.

Why do this?

If you surveyed the number of DVDs that your friends have, you could use a stem and leaf diagram to show the pattern of the results.

Get Ready

1. Write the numbers in each set in order of size, with smallest number first.

a 65, 54, 72, 50 **b** 4.3, 4.6, 4.0, 4.4 **c** 0.11, 0.1, 0.01, 0.12

Key Points

- A **stem and leaf diagram** is a way of presenting data that makes it easy to see the pattern without losing the actual data.
- A stem and leaf diagram should always have a key.
- From a stem and leaf diagram you can find statistics about the data. The lower quartile (Q_1) is the value a quarter of the way through the data, the second quartile (Q_2) or median is halfway through, and the upper quartile (Q_3) is three-quarters of the way through.
- The interquartile range (IQR) is the difference between the upper and lower quartiles $= Q_3 - Q_1$.

Example 3

Here are the numbers of cigarettes smoked per day by 15 people who are going to attempt to give up smoking:

20, 35, 40, 42, 32, 15, 22, 30, 28, 34, 40, 43, 28, 41, 25

a Write these data as an **ordered stem and leaf diagram**.

b Write down the mode of these data.

c Find the median of these data.

d Work out the range of these data.

e Find the lower and upper quartiles and interquartile range.

a

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

Key 1|5 stands for 15

The digit that each number begins with is called the stem.

The following digit is called the leaf.

Under stem, write the numbers 1 to 4.

Opposite each stem, write the leaves. Don't worry about the order. This gives you an unordered stem and leaf diagram.

1	5				
2	0	2	5	8	8
3	0	2	4	5	
4	0	0	1	2	3

Key 1|5 stands for 15

Next draw a stem and leaf with the leaves in order, starting with the smallest. This is an ordered stem and leaf diagram as asked for in the question.

b There are two modes: 28 and 40.

Each appears twice, the others only once.

c The median is 32.

32 is the middle value.

d The range is $43 - 15 = 28$

The range is the difference between the largest and smallest values. The largest and smallest values are the first leaf and the last leaf.

e $Q_1 = 25$ $Q_3 = 40$

$IQR = 40 - 25 = 15$

$Q_1 = \frac{16}{4}$th value = 4th value

$Q_3 = 3 \times \frac{16}{4}$th value = 12th value

You can find the values by counting in from each end.

Exercise 18C

C

1 Nassim records the number of emails he receives every day for 35 days.
The data he collects are shown in the stem and leaf diagram.

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

Key 3 | 1 stands for 31

a Write down the mode of these data.
b Find the median of these data.
c Work out the range of these data.
d Find Q_1 and Q_3 of these data.
e Work out the interquartile range for these data.

2 Here are the number of minutes a sample of 19 people had to wait to see a dentist.

10	12	8	9	21	24	17	4	28	30
5	7	9	15	7	9	14	9	6	

a Draw an ordered stem and leaf diagram for these data.
b Use your stem and leaf diagram to find the mode of these data.
c Use your stem and leaf diagram to find the median of these data.
d Work out the range of these data.
e Use your stem and leaf diagram to find Q_1 and Q_3 of these data.
f Work out the interquartile range for these data.

3 A delivery driver does a journey on 23 days every month.
Here are the distances, in kilometres, that he travelled in March.

56	74	83	74	65	92	52	59
64	68	72	94	82	63	74	65
88	69	68	85	68	74	63	

a Draw an ordered stem and leaf diagram for these data.
b Use your stem and leaf diagram to find the mode of these data.
c Use your stem and leaf diagram to find the median of these data.
d Work out the range of these data.
e Use your stem and leaf diagram to find Q_1 and Q_3 of these data.
f Work out the interquartile range for these data.

18.4 Interpreting comparative and composite bar charts

Objectives

- You can interpret comparative bar charts.
- You can interpret composite bar charts.

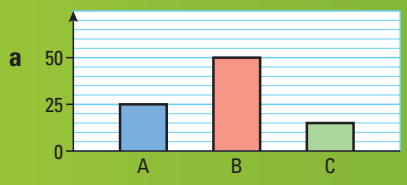
Why do this?

You may want to compare the sales of various categories of music in two shops. Composite bar charts would allow you to do this.

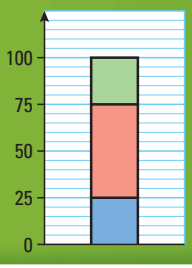
Get Ready

1. What can you say about the data in these two charts?

a



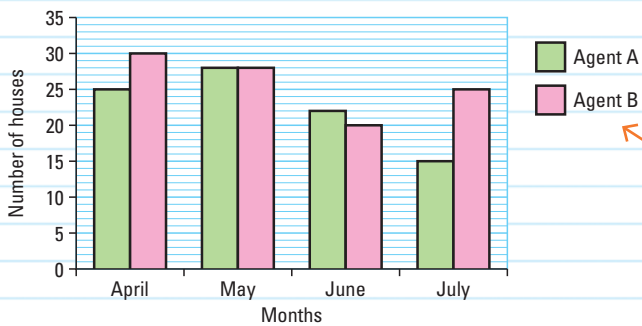
b



Key Points

- Composite bar charts (sometimes called compound or **component bar charts**) can be drawn to compare data.
- A composite bar chart shows the size of individual categories split into their separate parts.
- A **dual bar chart** is a type of comparative bar chart.
  In a comparative bar chart, two (or more) bars are drawn side-by-side for each category, and the heights of the bars can be compared category-by-category.

Example 4 The **dual bar chart** shows the number of houses sold by two agents in four months.



The key tells you which bars are agent A's and which are agent B's.

a In which month did A and B sell the same number of houses?
b Which agent sold the most houses in June?
c How many houses did B sell in April?
d How many houses did A sell in April?
e How many more houses than A did B sell in July?

a A and B sold the same number of houses in May.
b A sold more houses in June.
c B sold 30 houses in April.
d A sold 25 houses in April.
e 25 − 15 = 10
B sold 10 more houses than A in July.

The bars are the same height for May.

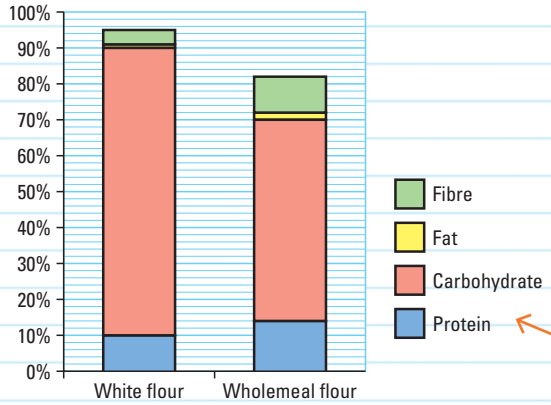
A's bar is higher than B's.

Use the key to identify B's colour. Find the month April and read off B's sales from the left-hand scale.

B sold 25 houses in July and A sold 15.

Example 5

This composite bar chart shows the amounts of protein, carbohydrate, fat and fibre in 100 g of white and wholemeal flour.



a How many grams of protein are there in 100 g of white flour?

b How many grams of carbohydrate are there in 100 g of wholemeal flour?

c Write down the flour which had the greater amount of fibre.

d Write down the flour with the smaller amount of fat.

e How many grams of wholemeal flour were not protein, carbohydrate, fat or fibre?

The key tells you the colour for each constituent.

a 10 g ← Identify the bar for white flour and use the colour key to find out which is protein.

b 70 − 14 = 56 g ← Read off from scale.

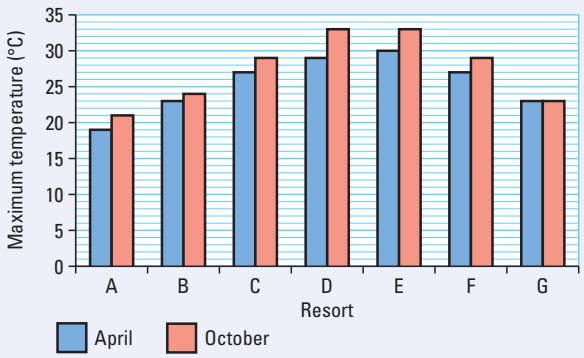
c Wholemeal ← The higher bar for fibre was for wholemeal flour (10 g).

d White ← The shorter bar for fat was for white flour (1 g).

e 100 − 82 = 18 g ← Read off the total height of the bar for wholemeal and take it away from 100 g (the figures are per 100 g).

Exercise 18D

1 The composite bar chart shows the temperature in a number of resorts in April and October.



a Write down the maximum temperature in April.

b Write down the maximum temperature in October.

c Write down the resort that had the same maximum temperature in both months.

d Write down the resorts in which the maximum temperature in October was 29°C.

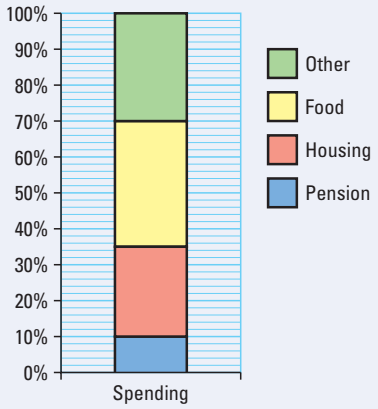
e Write down the resort in which the maximum temperature in April was 19°C.

2 The composite bar chart shows how David spends his money.

a What did David spend most on?

b What did David spend least on?

c What percentage of his income did he spend on housing?

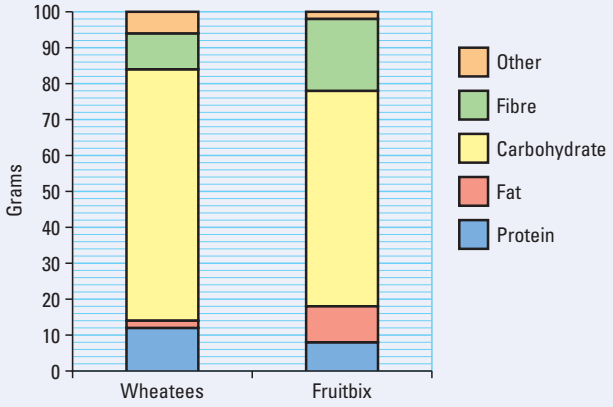


3 The composite bar charts show the make-up of 100 grams of each of two cereals: Wheatees and Fruitbix.

a How many grams of carbohydrate are there in 100 g of Wheatees?

b Estimate the number of grams of fat in 100 g of Fruitbix.

c Write down the name of the cereal that has more fibre.

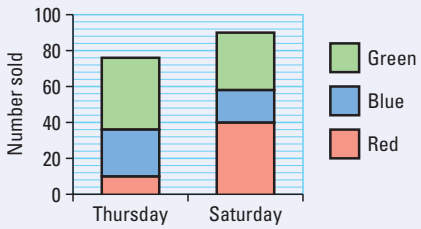


4 Market days in Ulvston are on Thursday and Saturday. Mattie runs a market stall that sells jumpers on each of these days. The composite bar chart shows his sales in one week in November.

a On which day were most jumpers sold overall?

b On which day were most green jumpers sold?

c How many red jumpers were sold on Saturday?



18.5 Drawing and interpreting frequency diagrams and histograms

Objectives

- You can draw a frequency diagram for grouped discrete data.
- You can draw a histogram for continuous data.

Why do this?

An exam board may choose to illustrate how the candidates in a particular year performed in one of its exams by creating a frequency diagram or histogram with the data.

Get Ready

1. What is the width of each class interval?
   a $0 \leqslant h < 3$ **b** $8 \leqslant h < 24$ **c** $75 \leqslant h < 100$

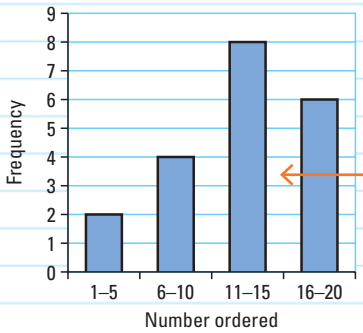
Key Points

- A **frequency diagram** can be drawn from grouped discrete data.
- A frequency diagram for grouped discrete data looks the same as a bar chart except that the label underneath each bar represents a group.
- A **histogram** can be drawn from grouped continuous data.
- A histogram is similar to a bar chart but represents continuous data so there is no gap between the bars.
- You can find information from a histogram, such as the median or the number of people in a given interval.

Example 6

The table shows the number of pizzas ordered in a restaurant from 7 pm to 8 pm on consecutive nights. Draw a frequency diagram for this information.

Number ordered	Frequency
1–5	2
6–10	4
11–15	8
16–20	6



There is a gap between the bars because, for example, there is no whole number between 15 and 16.

Example 7

The **grouped frequency table** shows information about the lengths of a series of roadworks.

Length (l metres)	Frequency
$160 \leqslant l < 165$	10
$165 \leqslant l < 170$	14
$170 \leqslant l < 175$	8
$175 \leqslant l < 180$	5
$180 \leqslant l < 185$	3
$185 \leqslant l < 190$	2

a Write down the modal class interval.

b The length of one set of roadworks is 177.2 m. In which class interval is this recorded?

c The length of another set is exactly 180 m. In which class interval is this length recorded?

d Draw a histogram for these data.

a The modal class is $165 \leqslant l < 170$.

This class interval has the highest frequency, 14.

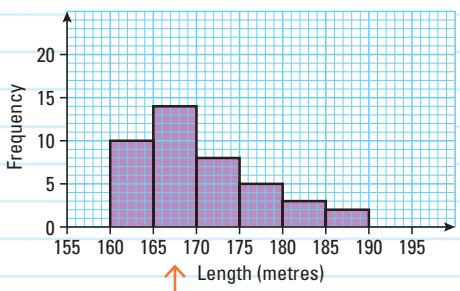
b This set is in the class interval $175 \leqslant l < 180$.

177.2 m is greater than 175 m but less than 180 m.

c This set is in the class interval $180 \leqslant l < 185$.

180 is shown at the end of one class interval and at the beginning of another. The sign for 'less than or equal to' ($\leqslant$) shows that 180 m should go in the class interval $180 \leqslant l < 185$.

d



In this histogram the area of the bars is proportional to the frequency.
In the class interval 160 to 165 there are 50 little squares representing a frequency of 10.
Each little square is equal to a frequency of $\frac{1}{5}$.

In the class interval 165 to 170 there are 70 little squares so it represents a frequency of $70 \times \frac{1}{5} = 14$.

Exercise 18E

D

1 The grouped frequency table shows information about the number of computer games owned by each of 35 college students.

Draw a frequency diagram for this information.

Number of games	Frequency
0 to 2	2
3 to 5	5
6 to 8	9
9 to 11	12
12 to 14	7

2 The grouped frequency table shows information about the wingspans of 36 snowy owls.

a Write down the modal class.

b The first snowy owl measured had a wingspan of 140 cm. In which class interval is this recorded?

c Draw a histogram for these data.

Wingspan (w cm)	Frequency
$125 \leqslant w < 130$	2
$130 \leqslant w < 135$	10
$135 \leqslant w < 140$	14
$140 \leqslant w < 145$	7
$145 \leqslant w < 150$	3

3 In a research project 40 young otters were weighed. Some information about their weights is shown in the table.

a Write down the modal class.

b In which class interval does the weight of 137 g fall?

c Draw a histogram for these data.

Weight (w g)	Frequency
$135 \leqslant w < 137$	3
$137 \leqslant w < 139$	10
$139 \leqslant w < 141$	14
$141 \leqslant w < 143$	8
$143 \leqslant w < 145$	5

18.6 Drawing and using frequency polygons

Objectives

- You can draw frequency polygons.
- You can recognise simple trends from a frequency polygon.
- You can use two polygons to make comparisons between two sets of data.

Why do this?

If you take a sample of your classmates' long-jump results, a frequency polygon would give you a good idea of how the lengths are distributed.

Get Ready

1. Which number is halfway between:

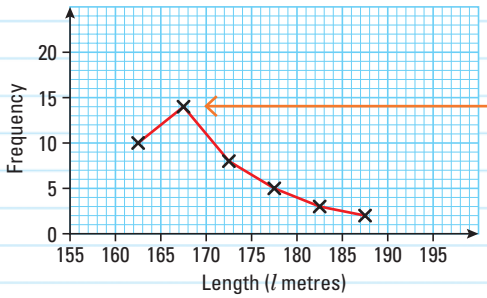
a 3 and 7 **b** 15 and 20 **c** 112 and 119?

Key Points

- A **frequency polygon** is another graph which shows data.
- When drawing a frequency polygon you draw a histogram then mark the midpoints of the tops of the bars and join these with straight lines.
- More than one frequency polygon can be drawn on the same grid to compare data.

Example 8

Draw a frequency polygon for the data in Example 7.



Plot the points at the midpoints of the class intervals.

Example 9

The frequency table gives information about the time waited, in seconds, at a set of traffic lights.

a Write down the modal class.

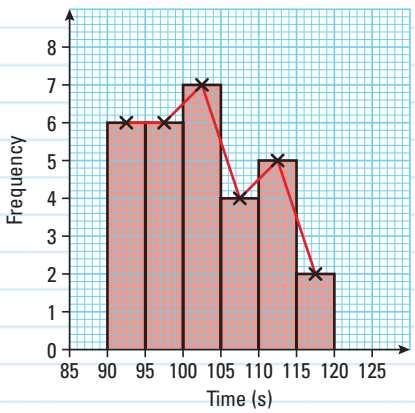
b Use the information to draw a histogram.

c Draw a frequency polygon to represent the information.

Time waited (t seconds)	Frequency
$90 \leqslant t < 95$	6
$95 \leqslant t < 100$	6
$100 \leqslant t < 105$	7
$105 \leqslant t < 110$	4
$110 \leqslant t < 115$	5
$115 \leqslant t < 120$	2

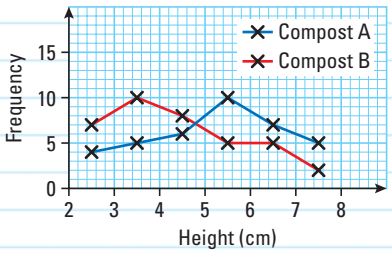
a *The modal class is* $100 \leqslant t < 105$.

b, c



As the question asks for both a histogram and a frequency polygon to be drawn, draw the histogram first.

Example 10 These two frequency polygons show the heights of seedlings growing in two different composts.
Compare the heights of the two groups.
Give reasons for your answers.



Compost A gives taller seedlings overall. ← Above 5 cm, the line showing the heights with compost A is above the line for compost B.

There are more very tall seedlings with compost A. ← There are five seedlings in the 7 – 8 cm class interval which were grown in compost A compared to two for compost B.

There are more very short seedlings with compost B. ← There are seven seedlings grown in compost B but only four for compost A in the 2 – 3 cm class interval.

Exercise 18F

1 A seed producer wants to know the numbers of peas in pods of a new variety of peas.
He records the number of peas in 60 pods. The table shows this information.

Number of peas	3	4	5	6	7	8
Frequency	2	4	7	10	22	15

Draw a frequency polygon for these data.

2 The noise levels at 40 locations near an airport were measured in decibels.
The data collected are shown in the grouped frequency table.

Noise level (d decibels)	$60 \leqslant d < 70$	$70 \leqslant d < 80$	$80 \leqslant d < 90$	$90 \leqslant d < 100$
Frequency	15	16	7	2

a Write down the modal class.

b Use the information in the table to draw a histogram.

c Use your answer to part **b** to draw a frequency polygon.

C

C

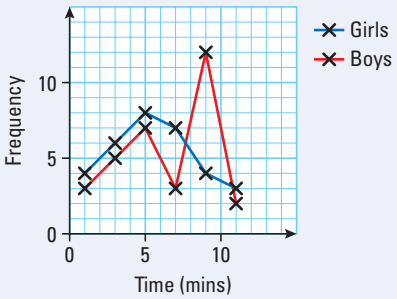
3 In a fishing competition the lengths, in centimetres, of all the trout caught were measured. The information collected is shown in the table.

Trout length (l cm)	Frequency
$24 \leqslant l < 25$	4
$25 \leqslant l < 26$	14
$26 \leqslant l < 27$	6
$27 \leqslant l < 28$	10
$28 \leqslant l < 29$	6

Draw a frequency polygon for these data.

AO3

* **4** The two frequency polygons show the amount of time it took a group of boys and a group of girls to do a crossword puzzle. Who were better at doing the puzzle, boys or girls? Give a reason for your answer.



18.7 Drawing and using histograms with unequal class intervals

Objectives

- You can draw a histogram with unequal class intervals.
- You understand frequency density.
- You can find the number of people in a given interval.

Why do this?

If you measure the heights of a number of people, they will cluster around a middle value. Adjusting the size of the class intervals makes these irregularities less noticeable.

Key Points

- In histograms, when there are unequal class intervals in a bar you adjust the height by using a scale of **frequency density** rather than width, where:

  $\text{frequency density} = \dfrac{\text{frequency}}{\text{class width}}$

  or $\text{frequency} = \text{frequency density} \times \text{class width}$.
- The area of each bar gives its frequency.

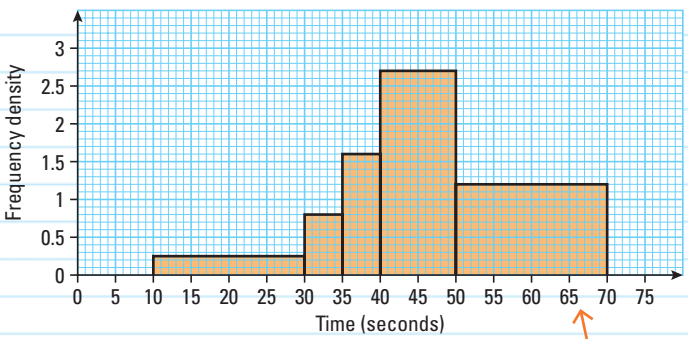
Example 11 The table gives information about the times taken, in seconds, by a number of workers to complete an operation in a factory.

Time taken (t seconds)	Frequency
$10 < t \leqslant 30$	5
$30 < t \leqslant 35$	4
$35 < t \leqslant 40$	8
$40 < t \leqslant 50$	27
$50 < t \leqslant 70$	24

Draw a histogram for these data.

Time taken (t seconds)	Frequency	Class width	Frequency density $= \frac{\text{frequency}}{\text{class width}}$
$10 < t \leqslant 30$	5	20	$\frac{5}{20} = 0.25$
$30 < t \leqslant 35$	4	5	$\frac{4}{5} = 0.8$
$35 < t \leqslant 40$	8	5	$\frac{8}{5} = 1.6$
$40 < t \leqslant 50$	27	10	$\frac{27}{10} = 2.7$
$50 < t \leqslant 70$	24	20	$\frac{24}{20} = 1.2$

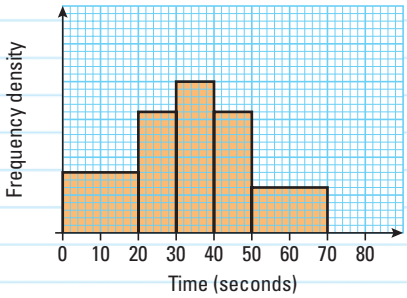
Work out the width of each class interval (the class width).
Divide the frequency by the class width to find the frequency density which gives the height of each bar.



On a grid label the horizontal axis 'Time (seconds)' and the vertical axis 'Frequency density'.
Scale the horizontal axis from 0 to 75 and the vertical axis from 0 to 3.
Draw the bars with no gaps between them.
The first bar goes from 10 to 30 and has a height of 0.25

Example 12 The histogram gives information about the time, in seconds, taken by students to solve a puzzle.

Time taken (t seconds)	Frequency
$0 < t \leqslant 20$	
$20 < t \leqslant 30$	8
$30 < t \leqslant 40$	
$40 < t \leqslant 50$	
$50 < t \leqslant 70$	

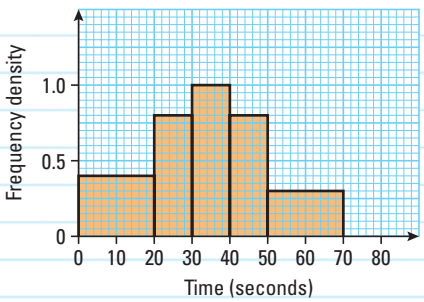


a Complete the frequency table.

b Use the histogram to estimate the number of people who took between 10 and 36 seconds to solve the puzzle.

a Frequency density for $20 < t \leqslant 30$ seconds $= \frac{8}{10} = 0.8$.

Frequency density $= \frac{\text{frequency}}{\text{class width}}$

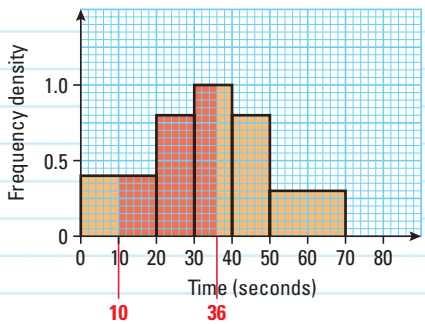


Now put a scale on the histogram.

Time taken (t seconds)	Frequency
$0 < t \leqslant 20$	$20 \times 0.4 = 8$
$20 < t \leqslant 30$	8
$30 < t \leqslant 40$	$10 \times 1.0 = 10$
$40 < t \leqslant 50$	$10 \times 0.8 = 8$
$50 < t \leqslant 70$	$20 \times 0.3 = 3$

Frequency = frequency density × class width

b



Frequency $= (10 \times 0.4) + (10 \times 0.8) + (6 \times 1.0)$
$= 4 + 8 + 6$
$= 18$ people

Work out the area between time = 10 and 36 seconds using frequency = frequency density × class width.

Exercise 18G

A

1 The table gives information about the lifetime of a certain make of torch battery.

Lifetime (l hours)	Frequency	Class width	Frequency density
$10 \leqslant l < 15$	4		
$15 \leqslant l < 20$	10		
$20 \leqslant l < 25$	20		
$25 \leqslant l < 30$	15		
$30 \leqslant l < 40$	6		

a Copy and complete the table.

b Draw a histogram for these data.

A02 A03

* **2** The table gives information about the distances a group of workers have to travel to work.

Distance (d kilometres)	Frequency
$0 < d \leqslant 5$	8
$5 < d \leqslant 10$	16
$10 < d \leqslant 20$	30
$20 < d \leqslant 30$	20
$30 < d \leqslant 40$	6

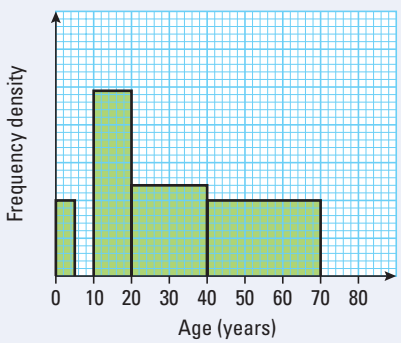
Draw a histogram for these data and find an estimate of the number of workers who travel between 15 and 25 minutes.

A02 A★

3 The table gives information about the age of people visiting a theme park one April morning.

a Copy and complete the table and histogram, table with a scale.

Age (y years)	Frequency
$0 < y \leqslant 5$	10
$5 < y \leqslant 10$	28
$10 < y \leqslant 20$	
$20 < y \leqslant 40$	
$40 < y \leqslant 70$	



b Find an estimate of how many people between 5 years and 30 years visited the theme park that morning.

A03

18.8 Drawing and using cumulative frequency graphs

Objectives

- You can construct a cumulative frequency table.
- You can draw a cumulative frequency graph.

Why do this?

Data are sometimes displayed in a cumulative frequency curve, for example, weights of babies as they get older.

Key Points

- The **cumulative frequency** of a value is the total number of observations that are less than or equal to that value.
- **Cumulative frequency diagrams (graphs)** can be used to find estimates for the number of items up to a certain value.

Example 13

The grouped frequency table shows information about the time, in minutes, taken by 40 runners who had competed in a cross-country race.

Time (t minutes)	Frequency
$t \leqslant 60$	0
$60 < t \leqslant 65$	2
$65 < t \leqslant 70$	12
$70 < t \leqslant 75$	21
$75 < t \leqslant 80$	5

a Draw up a **cumulative frequency table**.

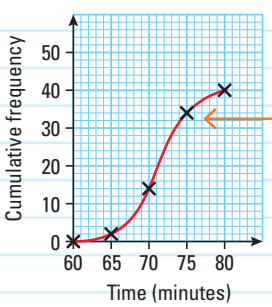
b Draw a cumulative frequency graph.

a

Time (t minutes)	Frequency	Cumulative frequency
$t \leqslant 60$	0	0
$60 < t \leqslant 65$	2	$0 + 2 = 2$
$65 < t \leqslant 70$	12	$2 + 12 = 14$
$70 < t \leqslant 75$	21	$14 + 21 = 35$
$75 < t \leqslant 80$	5	$35 + 5 = 40$

Each time add the frequency to the previous cumulative frequency. The previous frequency was 2 so add the frequency 12 to get the new cumulative frequency 14.

b



The cumulative frequency 35 for the interval $70 < t \leqslant 75$ is plotted at (75, 35).
The plotted points may be joined by a curve or by straight lines.

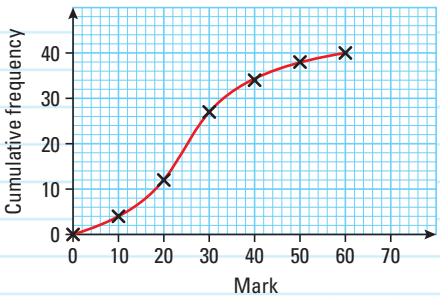
Example 14

Forty students took a test. The cumulative frequency graph gives information about their marks.

a Use the graph to estimate the number of students who had marks less than or equal to 26.

b Use the graph to work out an estimate for the number of students whose mark was greater than 44.

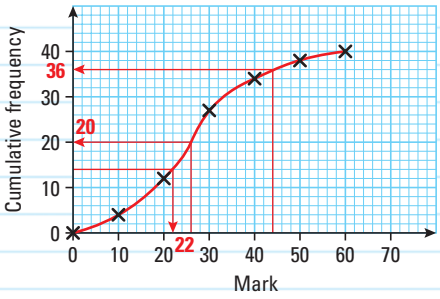
c 26 students passed the test. Work out the pass mark for the test.



a There are 20 students with a mark less than 26.

b There are 36 students with a mark less than or equal to 44 so there are $40 - 36 = 4$ with a mark greater than 44.

c If 26 pass there will be $40 - 26 = 14$ that fail.
From the graph the pass mark was 22.



Exercise 18H

1 The table shows the ages of people using a bowling alley.

Age (x years)	Frequency	Cumulative frequency
$x \leqslant 10$	3	
$10 < x \leqslant 15$	7	
$15 < x \leqslant 20$	10	
$20 < x \leqslant 25$	15	
$25 < x \leqslant 30$	8	
$30 < x \leqslant 35$	5	
$35 < x \leqslant 40$	2	

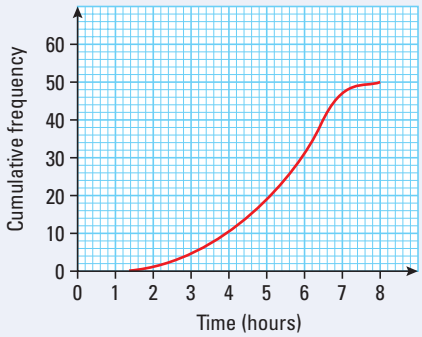
a Copy and complete the table.

b Draw a cumulative frequency graph for these data.

B

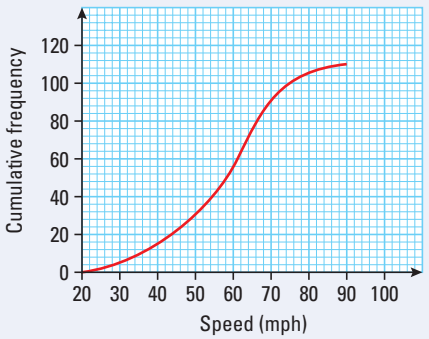
B

2 The cumulative frequency graph shows the time a group of girls spent on school computers.



- **a** Use the cumulative frequency graph to estimate the number of girls who spent up to 4 hours on the computer.
- **b** Use the cumulative frequency graph to estimate the number of girls who spent more than 6 hours on the computer.
- **c** Use the cumulative frequency graph to estimate the number of girls who spent between $3\frac{1}{2}$ and $6\frac{1}{2}$ hours on the computer.

3 The cumulative frequency graph shows the speeds of cars on a motorway.



- **a** Use the cumulative frequency graph to find an estimate for the number of motorists
  - **i** driving at 45 mph or less
  - **ii** driving at between 40 mph and 70 mph.
- **b** How many motorists' speeds were recorded altogether?
- **c** The speed limit on a motorway is 70 mph. Estimate the percentage of cars with a speed greater than 70 mph.

AO3

18.9 Finding quartiles from a cumulative frequency graph

Objective

- You can estimate the median and quartiles from a cumulative frequency graph.

Why do this?

Looking at the age of Britain's population in a frequency table, it is difficult to estimate the median and range. A cumulative frequency graph makes it easy to find the values.

Get Ready

1. Look at this list of numbers: 5, 5, 6, 7, 9, 9, 12, 13, 18, 20, 22, 23.
Which numbers are: **a** halfway along the list **b** three-quarters along the list?

Key Points

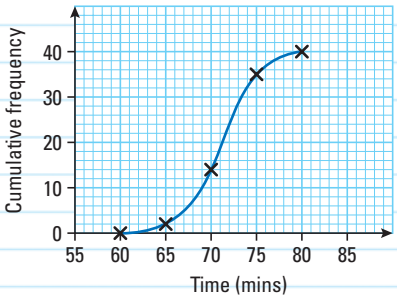
- The quartiles divide the frequency into four equal parts.
- If there are n values then the quartiles can be estimated from the cumulative frequency graph.
- The estimate for the lower quartile is the $\frac{n}{4}$th value.
- The estimate for the median is the $\frac{n}{2}$th value.
- The estimate for the upper quartile is the $\frac{3n}{4}$th value.
- You can compare measures of spread for two cumulative frequency graphs.

Example 15

The cumulative frequency graph shows information about the times, in minutes, taken by 40 runners who competed in a cross-country race.

a Find estimates for the median and quartiles.

b Find estimates for the range and interquartile range.



a $Q_1 = 69$ min

median $= Q_2 = 71.5$ min

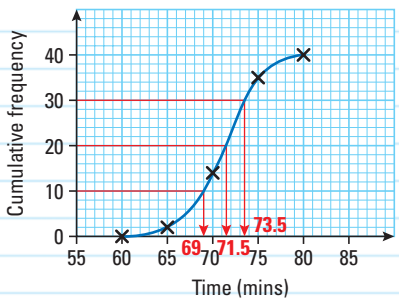
$Q_3 = 73.5$ min

Q_1 is the $\frac{40}{4} = 10$th value.
Q_2 is the $\frac{40}{2} = 20$th value.
Q_3 is the $3 \times \frac{40}{4} = 30$th value.

b Range $= 80 - 60 = 20$ min

IQR $= 73.5 - 69 = 4.5$ min

Range = highest − lowest values
IQR $= Q_3 - Q_1$



Exercise 18I

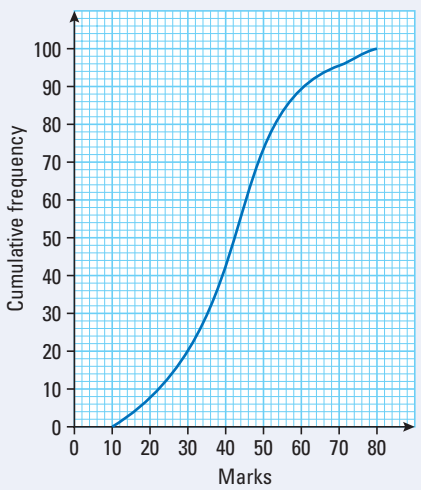
1 The cumulative frequency graph shows the scores a group of 100 apprentices got in an engineering examination.

a Find an estimate for the median (Q_2).

b Find an estimate for Q_1 and Q_3.

c Work out the interquartile range.

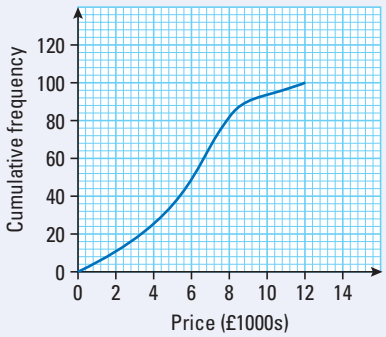
d Work out the range.



B

B

2 The cumulative frequency graph shows the prices of second-hand cars at a garage.

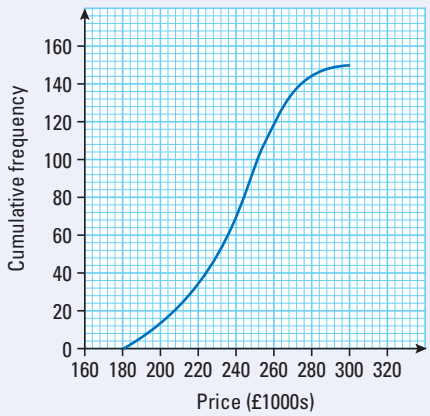


a Find an estimate for the median (Q_2).

b Find an estimate for Q_1 and Q_3.

c Work out the interquartile range.

3 The cumulative frequency graph shows the prices of detached houses on an estate agent's website.



a Find estimates for the median and quartiles.

b Find estimates for the range and the interquartile range.

18.10 Drawing and interpreting box plots

Objectives

- You can construct a box plot given the raw data.
- You can find the median, quartiles and interquartile range given a box plot.

Why do this?

You can easily show the median and range of data with a box plot. For example, speeds of cars on a section of motorway.

Get Ready

1. What are the median, lower and upper quartiles, and interquartile range of this list of numbers?

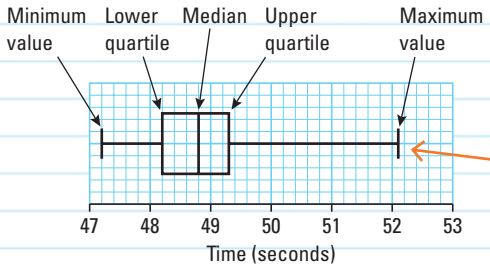
5 6 6 8 11 13 13 15 17 20 22 25 26 26 29

Key Points

- Box plots (sometimes called **box and whisker plots**) are diagrams that show the median, upper and lower quartiles and the maximum and minimum values of a set of data and are often used to compare distributions.

Example 16

The times run by an athlete had a maximum of 52.1 seconds, a minimum of 47.2 seconds, a median of 48.8 seconds and upper and lower quartiles of 49.3 seconds and 48.2 seconds. Draw a box plot for these data.



The box shows the spread over the middle 50% of the data (the interquartile range).

The whiskers show the lower 25% and the upper 25% of the data.

Example 17

The numbers of downloads from a music site during 15 time periods were as follows.

5 5 7 12 16 20 21 23
26 26 27 27 28 29 31

Draw a box plot for these data.

5 5 7 (12) 16 20 21 (23) 26 26
27 (27) 28 29 31

The lowest value is 5 and the highest is 31.

The lower quartile is the $\frac{1}{4}(15 + 1)$th
$= 4$th value $= 12$.

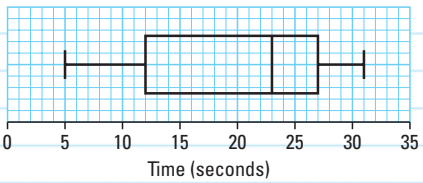
Lower quartile is the $\frac{1}{4}(n + 1)$th value.

The median is the $\frac{1}{2}(15 + 1)$th
$= 8$th value $= 23$.

Median is the $\frac{1}{2}(n + 1)$th value.

Upper quartile is the $\frac{3}{4} \times (15 + 1)$th
$= 12$th value $= 27$.

Upper quartile is the $\frac{3}{4}(n + 1)$th value.

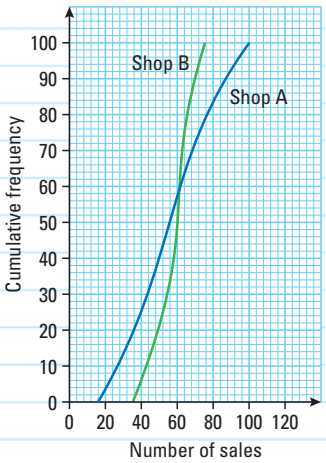


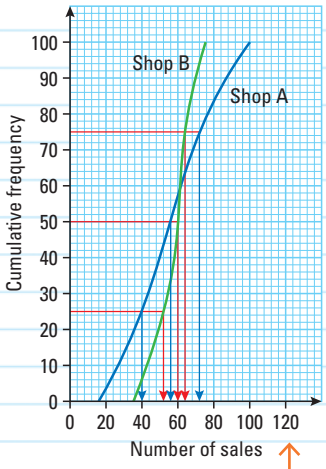
ResultsPlus
Examiner's Tip

These formulae are only for discrete data.

Example 18

The cumulative frequency graphs give information about the number of sales of mobile phones at two shops over 100 days.





A03

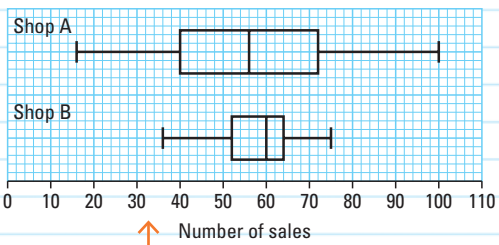
a Draw comparative box plots for these data.

b Compare the sales of the two shops.

Find the maximum and minimum values and the median and quartiles from your graph.

a

	Shop A	Shop B
Least number	16	36
Lower quartile	40	52
Median	56	60
Upper quartile	72	64
Greatest number	100	75



Draw your box plots to the same scale.

b *Shop B had a higher median so their sales are generally greater.*
Both the range and interquartile range of shop A were greater than those of shop B.
The sales of shop A are more variable from day to day.

Exercise 18J

B

1 A wildlife park ranger estimated the heights of all the adult giraffes in the park. The tallest was 5.8 metres tall and the shortest was 4.2 metres. The median height was 5 metres, the lower quartile 4.6 metres and the upper quartile 5.6 metres. Draw a box plot for these data.

2 The heights of the trees in a small piece of mature woodland were measured in metres. They were as follows.

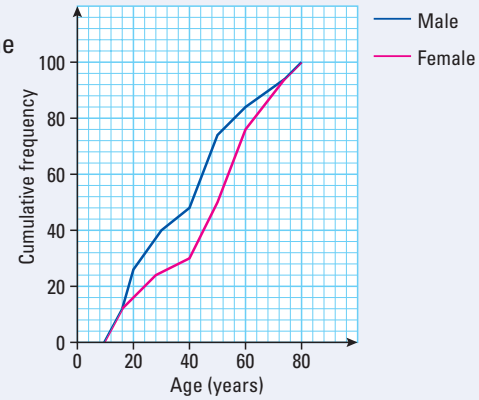
29 29.2 30.1 32 32.5 34.5 34.5 36.7 38 39.2 39.5 40.0 40.3 40.3 40.4

Draw a box plot for these data.

3 The cumulative frequency graph gives information about the ages of the male and female members of a cycling club.

a Use the cumulative frequency diagram to find the quartiles and the maximum and minimum values.

* b Draw two box plots on the same scale using these data and compare and contrast the data.



A

A03

Chapter review

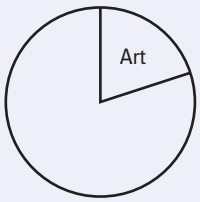
- In a **pie chart** the area of the whole circle represents the total number of items.
- The area of each **sector** represents the number of items in that category.
- The angles at the centre must add up to 360°.
  sector angle $= \dfrac{\text{frequency} \times 360°}{\text{total frequency}}$ or frequency $= \dfrac{\text{sector angle} \times \text{total frequency}}{360°}$
- The frequency represented by corresponding sectors in two pie charts is dependant upon the total populations represented by each of the pie charts.
- A **stem and leaf diagram** is a way of presenting data that makes it easy to see the pattern without losing the actual data.
- A stem and leaf diagram should always have a key.
- From a stem and leaf diagram you can find statistics about the data. The lower quartile (Q_1) is the value a quarter of the way through the data, the second quartile (Q_2) or median is halfway through, and the upper quartile (Q_3) is three-quarters of the way through.
- The interquartile range (IQR) is the difference between the upper and lower quartiles $= Q_3 - Q_1$.
- A composite bar chart shows the size of individual categories split into their separate parts.
- A comparative bar chart shows two or more bars side-by-side for each category.
- A **frequency diagram** for grouped discrete data looks the same as a bar chart except that the label underneath each bar represents a group.
- A **histogram** is similar to a bar chart but because it represents continuous data, no gap is left between the bars.
- You can find information from a histogram, such as the median or the number of people in a given interval.
- When drawing a **frequency polygon** you draw a histogram then mark the midpoints of the tops of the bars and join these with straight lines.
- More than one frequency polygon can be drawn on the same grid to compare data.
- In histograms the area of each bar is proportional to the frequency it represents.
  Frequency density $= \dfrac{\text{frequency}}{\text{class width}}$ or frequency = frequency density $\times$ class width.
- The **cumulative frequency** of a value is the total number of observations that are less than or equal to that value.
- The quartiles divide the frequency into four equal parts and can be estimated from the **cumulative frequency graph**.
- If there are n values, the estimates are:
  lower quartile $= \dfrac{n}{4}$th value, median $= \dfrac{n}{2}$th value, upper quartile $= \dfrac{3n}{4}$th value.
- You can compare measures of spread for two cumulative frequency graphs.
- Box plots (sometimes called **box and whisker plots**) are diagrams that show the median, upper and lower quartiles and the maximum and minimum values of a set of data and are often used to compare distributions.

Review exercise

1 60 students were asked to choose one of four subjects.
The table gives information about their choices.

Subject	Number of students	Angle
Art	12	72°
French	10	
History	20	
Music	18	

Copy and complete the pie chart to show this information.



Nov 2008

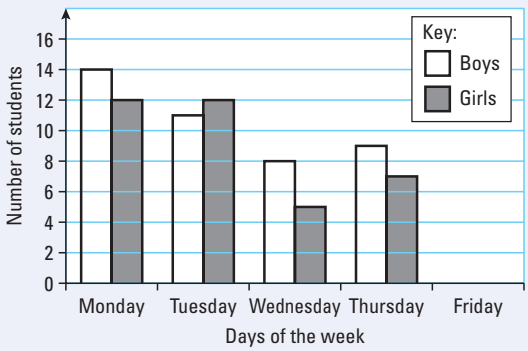
2 The table gives information about the drinks sold in a café one day.

Drink	Frequency	Size of angle
Hot chocolate	20	80°
Soup	15	
Coffee	25	
Tea	30	

Copy and complete the pie chart to show this information.

Nov 2008

3 Mr White recorded the number of students absent one week.
The dual bar chart shows this information for the first four days.



a How many boys were absent on Monday?
b How many girls were absent on Wednesday?

On Friday, 9 boys were absent and 6 girls were absent.
c Use this information to complete the bar chart.

On only one day more girls were absent than boys.
d Which day?

March 2008

D A02

4 Zoe recorded the weights, in kilograms, of 15 people. Here are her results.

87 51 46 77 74 58 68 78 48 63 52 64 79 60 66

a Draw a diagram to show these results.
b Write down the number of people with a weight of more than 70 kg.
c Work out the range of the weights.

March 2009, amended

D

C

5 Jason collected some information about the heights of 19 plants.
This information is shown in the stem and leaf diagram.

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

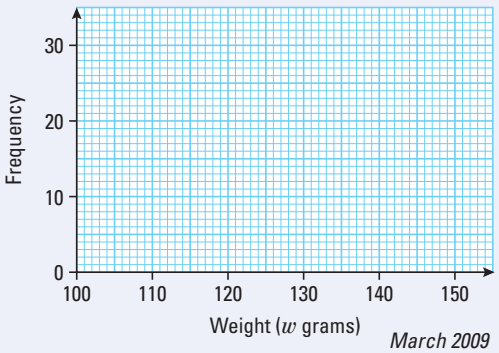
Key 4|8 means 48 mm

Find the median.

Nov 2008

6 The table shows some information about the weights (w grams) of 60 apples.
On a copy of the grid, draw a frequency polygon to show this information.

Weight (w grams)	Frequency
$100 \leqslant w < 110$	5
$110 \leqslant w < 120$	9
$120 \leqslant w < 130$	14
$130 \leqslant w < 140$	24
$140 \leqslant w < 150$	8

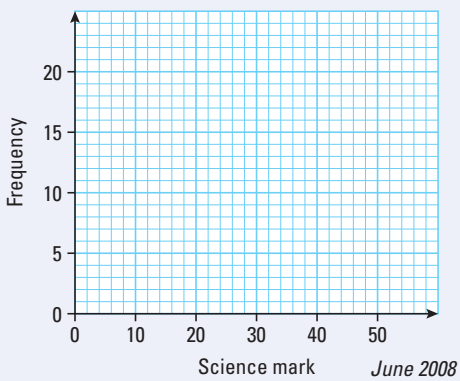


March 2009

7 60 students take a science test.
The test is marked out of 50.
This table shows information about the students' marks.

Science mark	0–10	11–20	21–30	31–40	41–50
Frequency	4	13	17	19	7

On a copy of the grid, draw a frequency polygon to show this information.

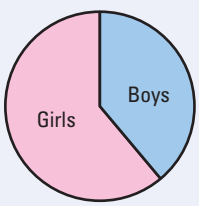


June 2008

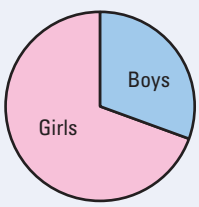
FS

A03

* **8**

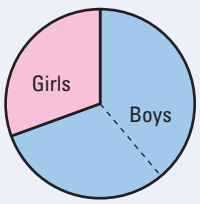


Pie chart showing proportion of boys and girls in Year 9



Pie chart showing proportion of boys and girls in Year 10

To draw the pie chart for boys and girls in Years 9 and 10 combined, Kimberly drew the pie chart on the right:

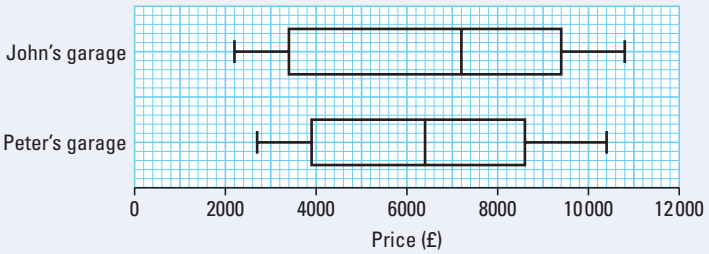


Pie chart showing proportion of boys and girls in Year 9 and Year 10

James said that this could not be correct.
Explain who is right.

B AO3

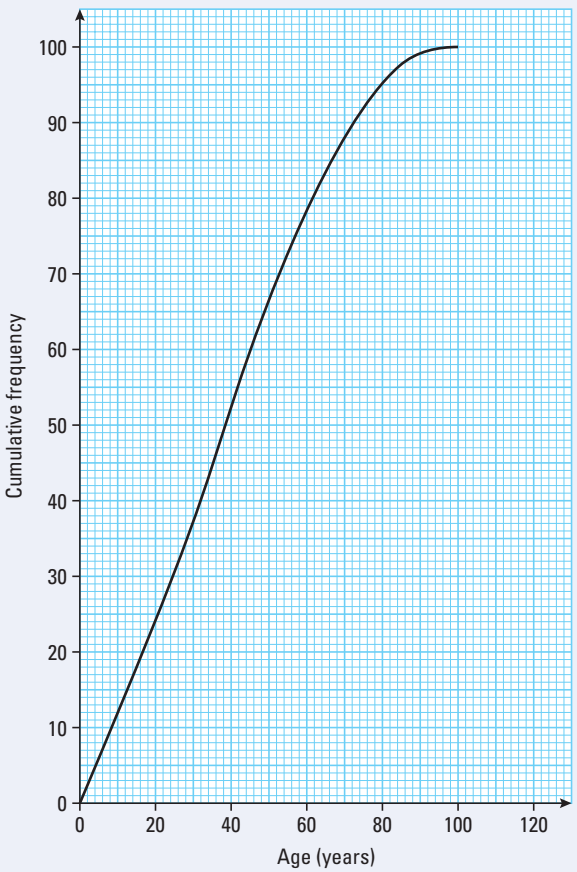
* 9 John and Peter each own a garage. They both sell used cars.
The box plots show some information about the prices of cars at their garages.



Compare the distribution of the prices of cars in these two garages.
Give **two** comparisons.

Nov 2008

10 The cumulative frequency graph shows some information about the ages of 100 people.



a Use the graph to find an estimate for the number of these people less than 70 years of age.

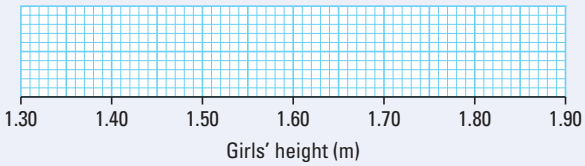
b Use the graph to find an estimate for the median age.

c Use the graph to find an estimate for the interquartile range of the ages.

Nov 2008

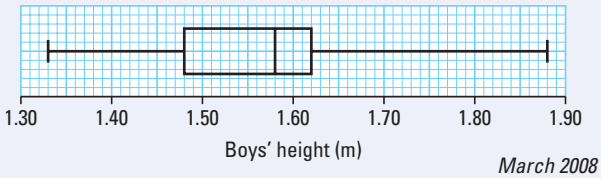
11 Verity records the heights of the girls in her class.
The height of the shortest girl is 1.38 m.
The height of the tallest girl is 1.81 m.
The median height is 1.63 m.
The lower quartile is 1.54 m.
The interquartile range is 0.14 m.

a Using this scale, draw a box plot for this information.



The box plot shows information about the heights of the boys in Verity's class.

b Compare the distributions of the boys' heights and the girls' heights.



A03

March 2008

* **12** Lucy did a survey about the amounts of money spent by 120 men during their summer holidays. The cumulative frequency table gives some information about the amounts of money spent by the 120 men.

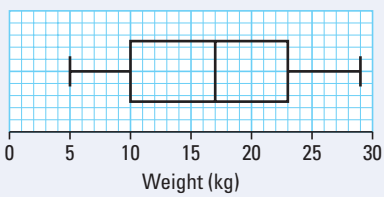
A survey of the amounts of money spent by 200 women during their summer holidays gave a median of £205.
Compare the amounts of money spent by the women with the amounts of money spent by the men.

Amount (£A) spent	Cumulative frequency
$0 < A \leqslant 100$	13
$0 < A \leqslant 150$	25
$0 < A \leqslant 200$	42
$0 < A \leqslant 250$	64
$0 < A \leqslant 300$	93
$0 < A \leqslant 350$	110
$0 < A \leqslant 400$	120

A02

May 2009

13 The box plot gives information about the distribution of the weights of bags on a plane.



A03

a Jean says the heaviest bag weighs 23 kg.
She is **wrong**. Explain why.

b Write down the median weight.

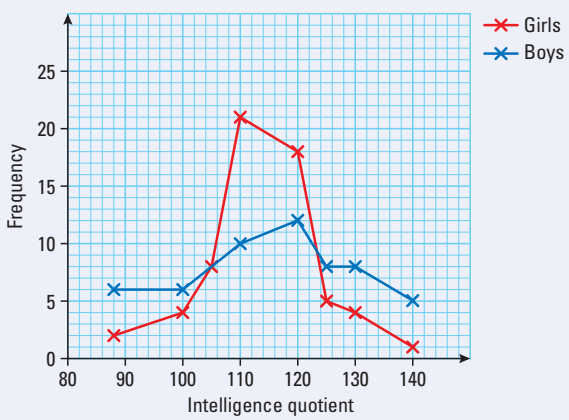
c Work out the interquartile range of the weights.

There are 240 bags on the plane.

d Work out the number of bags with a weight of 10 kg or less.

June 2009

14 The frequency polygons show information about the IQs of a group of boys and a group of girls.



A03

a Write down an estimate for the number of girls with an IQ of 110.

b Write down an estimate for the number of boys with an IQ of 110.

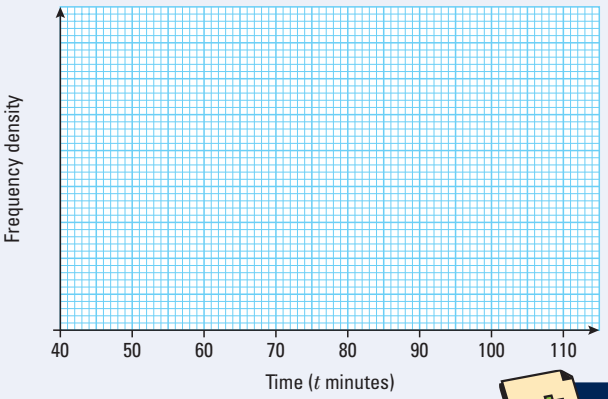
c Use the frequency polygon to compare the overall IQs of the boys and the girls.

A

15 The table gives some information about the lengths of time some boys took to run a race.

Draw a histogram for the information in the table.

Time (t minutes)	Frequency
$40 \leqslant t < 50$	16
$50 \leqslant t < 55$	18
$55 \leqslant t < 65$	32
$65 \leqslant t < 80$	30
$80 \leqslant t < 100$	24



ResultsPlus
Exam Question Report

73% of students answered this sort of question poorly.

March 2009

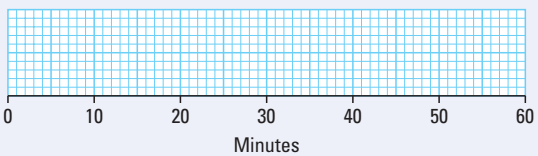
16 On Friday, Peter went to the airport.

He recorded the number of minutes that each plane was delayed.

He used his results to work out the information in this table.

	Minutes
Shortest delay	0
Lower quartile	2
Median	8
Upper quartile	18
Longest delay	41

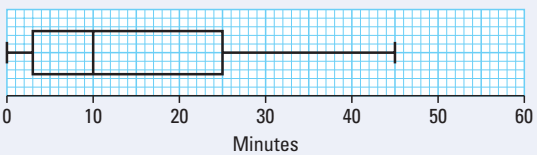
a Using this scale, draw a box plot to show the information in the table.



Peter also went to the airport on Saturday.

He recorded the number of minutes that each plane was delayed.

The box plot below was drawn using this information.



b Comment on the plane delays.

March 2009, adapted

AO3

AO2 A

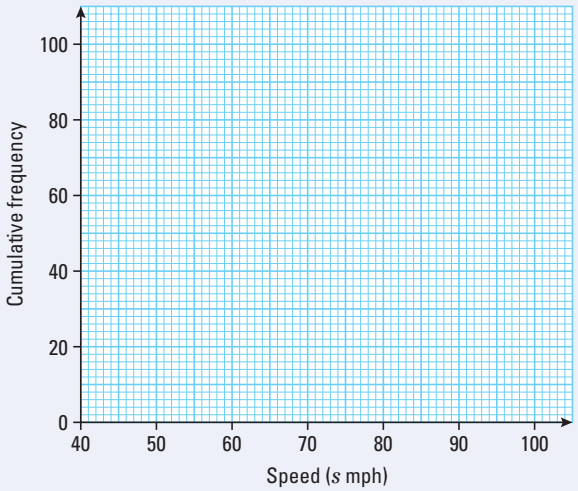
17 The speeds of 100 cars on a motorway were recorded.
The grouped frequency table shows some information about the speeds of these cars.

Speed (s mph)	Frequency
$40 < s \leqslant 50$	4
$50 < s \leqslant 60$	19
$60 < s \leqslant 70$	34
$70 < s \leqslant 80$	27
$80 < s \leqslant 90$	14
$90 < s \leqslant 100$	2

a On a copy of the grid, draw an appropriate graph for your table.

b Find an estimate for the median speed.

c Find an estimate for the interquartile range.

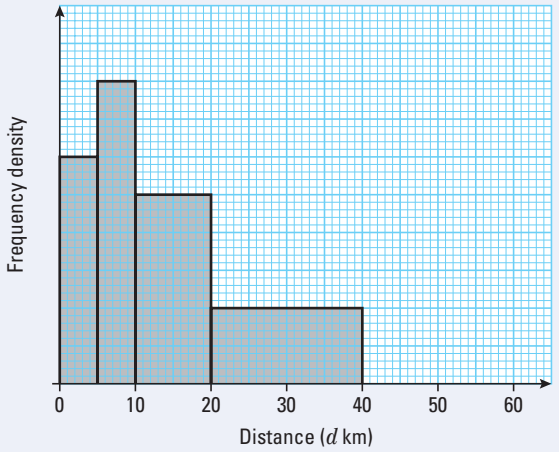


June 2008, adapted

18 The incomplete histogram and table give some information about the distances some teachers travel to school.

a Use the information in the histogram to complete the frequency table.

Distance (d km)	Frequency
$0 < d \leqslant 5$	15
$5 < d \leqslant 10$	20
$10 < d \leqslant 20$	
$20 < d \leqslant 40$	
$40 < d \leqslant 60$	10



b Use the information in the table to complete the histogram.

Nov 2008

A A02 A03

19 The table gives information about parcel sizes and their frequency.

Weight (w kg)	Frequency	Frequency density
$0 < w \leqslant 5$	20	
$5 < w \leqslant 15$	30	
$15 < w \leqslant 25$	15	
$25 < w \leqslant 35$	10	
$35 < w \leqslant 40$	5	

a Copy and complete the table.

b Draw a histogram for these data.

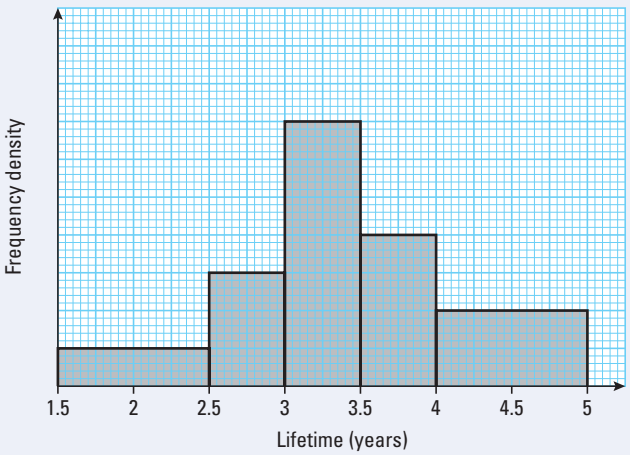
The weight limit for parcels going by Royal Mail is 20 kg.

c Work out an estimate for the number of parcels which will weigh 20 kg or less.

d Work out an estimate for the number of parcels weighing between 10 and 30 kg.

A*

20 The histogram shows information about the lifetime of some batteries.



Two of the batteries had a lifetime of between 1.5 and 2.5 years.
Find the total number of batteries.

June 2008