

 **SCAN ME**

CHECK YOUR ANSWERS

- (a) Identify the population. **(1)**
- (b) Explain what is meant by the term census. **(1)**
- (c) Explain what is meant by the term sample. **(1)**
- (d) Should the car manufacturer complete a census or a sample? **(1)**
Give a reason for your answer.



- The club chair wishes to sample the members of the club to see what their chess ratings are.

- The club chair decides to take a sample of 20 members of the chess club. They wish to ensure that the proportion of males and females in the same is the same as that in the entire club.

- (d) State the type of sampling method that the club chair should use. **(1)**
- (e) Work out how many male and female members should be included in the sample. **(1)**

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. In the bottom-left corner, there is a small, light blue sticker with a scalloped edge. The rest of the page is empty.

(Total for Question 2 is 5 marks)



3 A sports company produces trainers for athletes who compete in long distance running races.

The sports company produces 1000 units of a new type of trainer and wishes to know how many miles athletes can run in the trainer before it needs replacing. The company takes sample a of the 1000 units and gives them to athletes to test.

- (b) Describe how the sports company could produce a sample of size 20 using simple random sampling. (3)

Instead the sports company uses this following method to sample 20 of the units of trainers.

- Number of each of units from 1 to 1000.
- Use a random number generator to select a random number between 1 and 50.
- Include the unit corresponding to this number in the sample.
- Include every 50th unit after the unit sampled in the previous step.

- (c) Name the type of sampling method used by the sports company. (1)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, light blue decorative element resembling a scalloped-edge sticker or a piece of tape. The rest of the page is empty and white.

(Total for Question 3 is 5 marks)

- (d) Explain why this sampling technique would not be suitable for this investigation. (1)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, light blue sticker with a scalloped edge. The rest of the page is empty.

1st



- (d) Work out how many drivers and non-drivers should be in the sample using the council's estimate. **(1)**

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. In the bottom-left corner, there is a small, light blue decorative element resembling a scalloped-edge sticker or a piece of tape. The rest of the page is empty and ready for writing.

6

7 A headteacher wants to investigate how much revision students do to prepare for mock exams.

There are 750 students who attend the school that are split into 25 tutor groups.

The headteacher considers three sampling methods:

Method A:

- Whilst on duty at lunchtime, sample the first 150 students that they meet.

Method B:

- Number all students in the school from 1 to 750.
- Randomly generate 150 different numbers between 1 and 750.
- Sample all 150 students that correspond to the numbers generated.

Method C:

- Number each of the tutor groups in the school from 1 to 25.
- Randomly generate 5 different numbers between 1 to 25 and sample all students within the tutor groups that correspond to these numbers.

- (a) State the sampling technique used in Method A. (1)
- (b) State **two** reasons why Method A might be biased. (2)
- (c) State the sampling technique used in Method B. (1)
- (d) State one advantage of Method A over Method B. (1)
- (e) State the sampling technique used in Method C. (**AQA only**) (1)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, circular blue sticker with a scalloped edge. The rest of the page is empty.

(Total for Question 7 is 6 marks)



 SCAN ME

CHECK YOUR ANSWERS

- 2 The following table summarises heights, y cm to the nearest centimetre, of the female athletes at an athletics club.

Height (y cm)	$160 \leq y < 170$	$170 \leq y < 180$	$180 \leq y < 190$	$190 \leq y < 200$
Number of athletes (f)	3	7	7	8

- State the modal class for the heights of the female athletes. **(1)**
- Work out the class interval containing the median female height. **(1)**
- Work out an estimate for the mean height of the female athletes. **(2)**

For the heights, x cm to the nearest centimetre, of the male athletes at the club

$$\sum x = 6615 \quad n = 35$$

- (d) Use this data, and your answer to part (c) to estimate the mean height of all of the male and female athletes at the athletics club. (2)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, stylized blue icon resembling a gear or a sunburst. The rest of the page is empty.

(Total for Question 2 is 6 marks)



- 3 The following tables summarises the annual salaries, £S to the nearest pound, of staff at a multi-academy trust.

Annual Salary (£ S)	Frequency (f)
$30,000 \leq S < 50,000$	84
$50,000 \leq S < 70,000$	18
$70,000 \leq S < 90,000$	9
$90,000 \leq S < 110,000$	3
$110,000 \leq S < 200,000$	1

- Work out the class interval containing the median annual salary. **(1)**
- Work out an estimate, to the nearest pound, for the mean annual salary of the staff. **(1)**
- Explain why your answer to part (b) is an estimate. **(1)**

It is found that the CEO of the multi-academy trust has an annual salary of £195,000 (yes really)

- (d) Explain what effect, if any, will this information have on
- (i) your answer to part (a)
 - (ii) your answer to part (b)
- (2)**

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, light blue circular sticker with a scalloped edge. The rest of the page is empty.

(Total for Question 3 is 5 marks)

- The number of minutes late L , to the nearest minute, of the first 8 students are shown below.

Student	A	B	C	D	E	F	G	H
Lateness (L minutes)	1	2	3	3	4	6	8	9

- State the modal number of minutes late for the first 8 students. **(1)**
- Work out the median number of minutes late for the first 8 students. **(1)**
- Work out the mean number of minutes late for the first 8 students. **(1)**

On Monday 4 students received a detention for being late.

(d) Work out the mean of minutes late for the 4 students who received a detention on Monday. **(2)**

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. In the bottom-left corner, there is a small, light blue decorative sticker with a scalloped edge. The rest of the page is empty, providing space for writing or drawing.

(Total for Question 4 is 5 marks)



5 The table below summarises the number of fillings of 20 patients randomly sampled at a dentist.

Number of Fillings	0	1	2	3	4	5	6
Frequency (f)	7	4	3	3	0	2	1

- (a) State the modal number of fillings. **(1)**
- (b) Work out the median number of fillings. **(1)**
- (c) Work out the mean number of fillings. **(2)**

Another patient is added to the sample. They have 2 fillings.

- (d) Explain what affect, if any, this will have on
- (i) your answer to part (a)
 - (ii) your answer to part (b)
 - (iii) your answer to part (c)
- (3)**

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, light blue decorative element resembling a scalloped-edge sticker or a piece of tape. The rest of the page is empty and ready for writing.

(Total for Question 5 is 7 marks)

7 The speeds, S mph to the nearest mph, of some vehicles on a motorway are recorded.

Lane of Motorway	Lane 1	Lane 2	Lane 3
Number of Vehicles	124	148	165
Mean Speed (mph)	55.2	67.8	71.3

Work out the mean speed of all vehicles across all lanes of the motorway.
Give your answer to 3 significant figures. (2)

(Total for Question 7 is 2 marks)

8 The following tables summarises the ages, A years, of people taking a driving test at a test centre.

Age (A years)	10-19	20-29	30-39	40-49	50-59	60-69
Frequency (f)	21	4	1	2	1	1

- (a) Work out the class interval containing the median age. (1)
- (b) Work out an estimate for the mean age of the people taking a driving test at the test centre. (2)
- The minimum age for taking a driving test is 17 years old.
- (c) Comment on whether the true mean age of the drivers at the test centre likely to be different to your answer to part (b). Give a reason for your answer. (1)

(Total for Question 8 is 4 marks)





CHECK YOUR ANSWERS

- A sample of 11 teachers from school A are selected and their ages are shown below.

The interquartile range of the sample is 14.

- (b) Work out two possible values for x . (2)

A sample of 7 teachers from school B are selected and their ages are shown below.

(c) Compare the ages of teachers who work at school A with the ages of teachers working at school B. **(2)**

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, stylized blue icon resembling a gear or a sunburst. The rest of the page is empty.

1st

3

They recorded the speeds of cars as they passed under a bridge until they had 800 cars from each of the three lanes.

The table below shows information about the speeds of the vehicles in mph.

	Lane 1	Lane 2	Lane 3
Lowest Speed	55	57	63
Q_1	58	59	64
Q_2	59	62	65
Q_3	60	64	70
Greatest Speed	61	69	71

- (a) Name the type of sampling method used by the traffic officer. (1)
- (b) Work out the lane for which the speeds of the vehicles were most consistent.
Give a reason for your answer. (2)
- (c) Write down percentage of the vehicles in lane 1 were travelling below 58 mph. (1)
- (d) Work out how many of the vehicles were travelling more than 64 mph. (2)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small blue tab with a scalloped edge, resembling a binder ring or a piece of tape. The rest of the page is empty and white.

(Total for Question 3 is 6 marks)

4 The table below shows the GCSE grades of Year 12 students taking A-Level maths at a college.

GCSE Grade	Frequency (f)
6	8
7	17
8	14
9	12

- Work out the range of the GCSE grades. **(1)**
- Work out the median GCSE grade. **(1)**
- Work out the lowest quartile of the GCSE grades. **(1)**
- Work out the upper quartile of the GCSE grades. **(1)**

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. In the bottom-left corner, there is a small, circular blue sticker with a scalloped edge. The rest of the page is empty.

(Total for Question 4 is 4 marks)





6 The race times for runners of the Brighton Half Marathon are shown below.

Race Time (t minutes)	Frequency (f)
$60 \leq t < 90$	437
$90 \leq t < 120$	3657
$120 \leq t < 150$	3053
$150 \leq t < 180$	846
$180 \leq t < 210$	152
$210 \leq t < 240$	21

The winner completed the race in 1 hour, 7 minutes and 17 seconds.

- (a) Use linear interpolation to calculate an estimate for the median race time giving your answer in hours, minutes and seconds. (2)

The actual median time for the race was 1 hour, 59 minutes and 57 seconds.

- (b) Give a reason why your answer to part (a) is different to the true median time. (1)

- (c) Use linear interpolation to calculate an estimate for the 10th to 90th interpercentile range of the race times. Give your answer in hours, minutes and seconds. (3)

This image shows a full page of white paper with horizontal blue ruling lines. The lines are evenly spaced and run across the width of the page. In the bottom-left corner, there is a small, partially visible blue circular sticker or logo with some white text that appears to say "1st".

(Total for Question 6 is 6 marks)



- 7 The following tables summarises the maximum daily temperatures in a city, T °C to the nearest degree, recorded for the months May to October in a particular year.

Temperature ($T^{\circ}\text{C}$)	13 – 17	18 – 22	23 – 27	28 – 32
Frequency (f)	25	32	70	57

- (a) Use linear interpolation to calculate an estimate for the interquartile range for the temperatures.
Give your answer to 2 decimal places. **(3)**

- (b) Use linear interpolation to calculate an estimate for the 38th percentile for the temperatures.
Give your answer to 2 decimal places. (2)

Sunny correctly calculates an estimate for the 85th percentile for the temperatures as 30.1 °C

Sunny claims “This means that 15% of the temperatures were greater than 30.1°C”

- (c) Explain why Sunny's claim could be incorrect. **(1)**

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, light blue decorative element resembling a scalloped-edge sticker or a piece of tape. The rest of the page is empty and white.

(Total for Question 7 is 6 marks)



- 8** The following tables summarises the annual salaries, £*S* to the nearest £1000, of employees at a company in the year 2023.

Salary (£<i>S</i>)	20,000 – 29,000	30,000 – 39,000	40,000 – 49,000	50,000 – 59,000
Frequency (f)	25	15	6	2

- (a) Use linear interpolation to calculate an estimate for the 20th to 80th interpercentile range of the annual salaries in 2023. Give your answer to the nearest pound. (3)

The company claims that between 2023 and 2024 they increased annual salaries by 10%.

For the 2024 salaries: $Q_1 = \text{£}27,000$ $Q_2 = \text{£}31,000$ $Q_3 = \text{£}41,000$

- (b) Use linear interpolation to calculate Q_1 , Q_2 and Q_3 for the salaries in 2023. (3)

- (c) Comment on the company's claim. (1)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, blue, circular sticker with a scalloped edge. On the sticker, the text "1st" is written in a white, sans-serif font. The rest of the page is empty.

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and run across the width of the page. In the bottom-left corner, there is a small, light blue decorative sticker with a scalloped edge. The rest of the page is empty and ready for writing.





REVISE THIS TOPIC



CHECK YOUR ANSWERS





3 A sample of 8 students from Year 7 and 8 students from Year 10 were taken.

The heights of the Year 7 students (x) and Year 10 students (y), in cm, are summarised below.

$$\Sigma y^2 = 235032$$

- (d) Calculate the mean and standard deviation of the heights of all 16 students. **(3)**

[illegible]

(Total for Question 3 is 11 marks)



- 4 8 different teams were timed to escape from an escape room.
Their times to escape x , to the nearest minute, are summarised below.

$$\sum (x - \bar{x})^2 = 187.5$$

- (b) Calculate the standard deviation of the times taken to escape. (2)

This extra team escapes in 55 minutes.

(c) Without further calculation, state what effect, if any, including the extra time of 55 minutes will have on

- (i) the mean
(ii) the standard deviation

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, light blue sticker with a scalloped edge. The rest of the page is empty.

(Total for Question 4 is 5 marks)



The table below summarises the headteacher's calculations.

Day	Mon	Tues	Weds	Thurs	Fri
Mean	12	12.5	12.6	10.25	23
Standard Deviation	0	1.66	3.04	-2.28	2.55

As maths teacher checks the data and realises that two of the values were calculated incorrectly.

- (a) State which two values were calculated incorrectly. Give reasons for your answers. (2)

The headteacher also wants to know about the number of detentions being set.

The number of detentions received in a one-month period for 30 Year 10 students (x) and 30 Year 11 students (y) are summarised below.

$$\sum x = 126 \qquad \sum x^2 = 828 \qquad \sum y = 93 \qquad \sum (y - \bar{y})^2 = 392.7$$

- (b) By working out the standard deviation determine which of the year groups had more variation in their number of detentions. (5)



- 7 A lottery draws 6 winning numbers. Players win a prize if their ticket matches 2 or more of the winning numbers.

The prize breakdown for a lottery draw is shown below.

Numbers Matched	Prize Value (£)	Number of Prizes
6	11 449 068	1
5	1750	186
4	140	12 142
3	30	222 468
2	2	1 678 011

- Calculate the mean prize value to the nearest pound. **(1)**
- Calculate the percentage of the prizes awarded that were below the mean prize value. **(2)**
- Calculate the standard deviation of the prize values to the nearest pound. **(1)**
- State a reason why the standard deviation may not be a suitable measure of spread to describe these data. **(1)**

The prize values (£ x) for a different lottery draw are summarised below.

$$n = 1\,800\,000$$

$$\sum x = 22\,752\,692$$

$$S_{xx} = 96\,868\,911\,115\,690$$

- (e) Calculate, to the nearest pound, the mean and standard deviation of the prize values for this lottery draw. (3)

[illegible]

(Total for Question 7 is 8 marks)



8 The speeds S , to the nearest mph, of 40 cars travelling through a village are shown below.

Vehicle speed (S mph)	Frequency (f)	Midpoint (x mph)
$0 \leq S < 10$	1	5
$10 \leq S < 20$	3	15
$20 \leq S < 30$	31	25
$30 \leq S < 40$	4	35
$40 \leq S < 50$	1	45

You may use $\sum fx = 1010$ and $\sum fx^2 = 27\,000$

- Calculate, to 2 decimal places, an estimate for the mean vehicle speed. **(1)**
- Calculate, to 2 decimal places, an estimate for the standard deviation of the vehicle speeds. **(2)**
- Explain why your answers to parts (a) and (b) are only estimates. **(1)**

It is found that one of the vehicle speeds was incorrectly recorded as 2 mph but was in fact 25 mph.

- (d) Without calculate a new estimates state what effect, if any, using the correct speed will have on
- (i) your estimate for the mean
 - (ii) your estimate for the standard deviation

Give reasons for your answers. (2)

[illegible]

(Total for Question 8 is 6 marks)

- There are 50 entrants, some of which are professionals, and the rest are amateurs.

The times of the amateurs (x) and the professionals (y), to the nearest minute are summarised below.

$$\sigma_y = 3$$

The total time for all 50 entrants was 3578 minutes.

Work out the standard deviation of the times all 50 of the entrants.

(6)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small portion of a blue spiral binding. The rest of the page is empty, providing space for writing or drawing.

(Total for Question 9 is 6 marks)





SCAN ME

NOT ON AQA **Linear Coding** NOT ON AQA



SCAN ME

REVISE THIS TOPIC

CHECK YOUR ANSWERS

1 The race times, x , to the nearest minute, of the first 10 finishers in the London marathon are shown.

124 124 127 127 128 129 130 130 131 131

The data is coded using the formula $y = \frac{x - 120}{4}$

(a) Work out the coded value for the runner who finished in 3rd place. (1)

One of the coded race times is 2.5 minutes.

(b) Work out the original race time for this coded value. (1)

$\bar{x} = 128.1$ $\sigma_x = 2.468$ (to 3 decimal places).

(c) Using these values, or otherwise, work out the mean ($\bar{y}$) and standard deviation (σ_y) of the coded values. (2)

(Total for Question 1 is 4 marks)



2 The heights, h , to the nearest centimetre, of 10 students are summarised below.

$$\sum h = 1776 \qquad S_{hh} = 236.4$$

(a) Work out the mean ($\bar{h}$) and standard deviation (σ_h) (3)

The data is coded using the formula $c = h - 165$

(b) Work out the mean ($\bar{c}$) and standard deviation (σ_c) for the coded values. (2)

(Total for Question 2 is 5 marks)

3 The masses, x , to the nearest kg, of 8 cows are summarised below.

$$\sum x = 4120 \qquad \sum x^2 = 2\,121\,856$$

(a) Work out the mean ($\bar{x}$) and standard deviation (σ_x) (3)

The data is coded using the formula $y = \frac{x}{5} - 100$

(b) Work out the mean ($\bar{y}$) and standard deviation (σ_y) for the coded values. (2)

(Total for Question 2 is 5 marks)





6 The weekly wages, w , to the nearest pound, are recorded for 6 employees at a company.

The data is coded using the formula $x = \frac{w - 800}{5}$

$$\sum x = 66 \qquad \sum x^2 = 4356$$

- of the 6 employees.

(Total for Question 6 is 5 marks)

7 The masses, m , of ten 1p coins are recorded in grams.

The data is coded using the formula $x = 1000m - 3500$

$$\sum x = 640 \qquad S_{xx} = 17\,284$$

- (b) Work out the standard deviation, in grams, of the masses of the ten 1p coins (3)

(Total for Question 7 is 5 marks)



SCAN ME

Outliers, Box Plots and Cumulative Frequency Diagrams

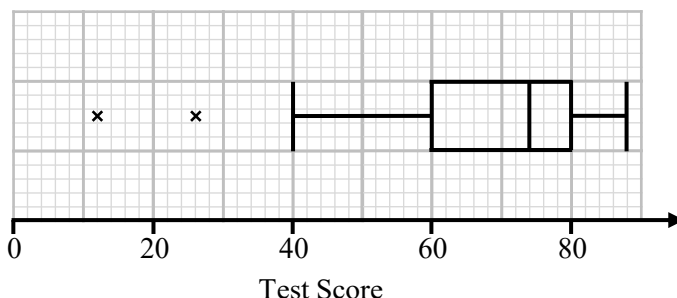


SCAN ME

REVISE THIS TOPIC

CHECK YOUR ANSWERS

- 1 180 students completed a maths assessment. The maximum mark for the assessment was 100. The results are summarised in the box plot below.



- (a) Write down the median test score. (1)
- (b) Find the range of the test scores. (1)
- (c) Find the interquartile range of the test scores. (1)
- (d) State the meaning of the $\times$ symbols shown on the box plot. (1)

One of the students is selected at random. Eric says

“the probability that the student scored between 40 and 60 marks is the same the probability that the student scored between 60 and 80 marks.”

- (e) Explain how you know that Eric must be incorrect. (1)

(Total for Question 1 is 5 marks)



2 The table below shows the race times for the 100 metre final at the 2009 World Championships.

Race Position	1	2	3	4	5	6	7	8
Time (seconds)	9.58	9.71	9.84	9.93	9.93	10.00	10.00	10.34

- Calculate the interquartile range of the race times. **(1)**
- Calculate, to 3 decimal places, the mean race time. **(1)**
- Calculate, to 3 decimal places, the standard deviation of the race times. **(1)**

Hannah defines a race time an outlier if it falls either

more than $1.5 \times$ (interquartile range) above the upper quartile or
more than $1.5 \times$ (interquartile range) below the lower quartile.

- (d) Determine if any of the race times are considered outliers using Hannah's definition. (2)

Ross defines a race time an outlier if it falls either

more than $2 \times$ (standard deviation) above the mean or
more than $2 \times$ (standard deviation) below the mean.

- (e) Determine if any of the race times are considered outliers using Ross' definition. (2)

[illegible]

(Total for Question 2 is 7 marks)



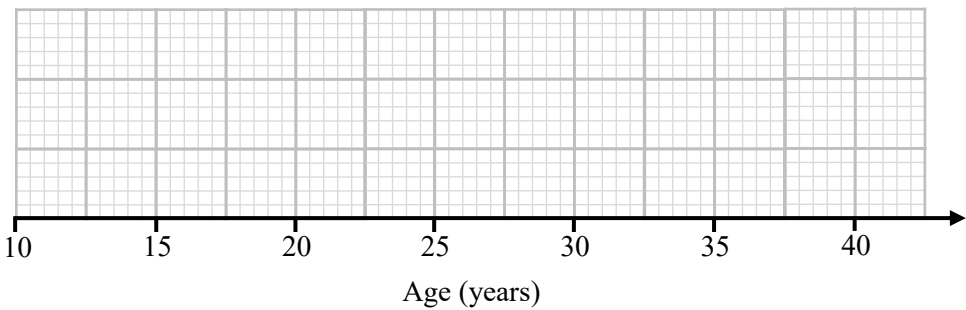
3

15	17	19	19	20	21	21	22	23	23	23	23	23	23	24
24	24	24	24	25	25	26	27	28	29	29	31	33	35	36

An outlier is any value that falls either

more than $1.5 \times$ (interquartile range) above the upper quartile or
more than $1.5 \times$ (interquartile range) below the lower quartile.

- (a) Show that the ages two of the players in the football squad are considered outliers. (3)
- (b) Draw a box plot for the ages of the players. (2)
You should indicate outliers with a ✕



(Total for Question 3 is 5 marks)

- 4 A college tracks lateness to school to the nearest minute.
Students that are early or on time are considered to have 0 minutes of lateness.

A sample of 20 students from Year 12 and 20 students from Year 13 were taken on Monday.

The number of minutes lateness to school for the Year 12 students (x) is summarised below.

$$\sum x = 72 \qquad \sum x^2 = 1048$$

- (a) Calculate, to 1 decimal place, the mean and standard deviation of the lateness for Year 12 students. (2)

An outlier is any value that falls either

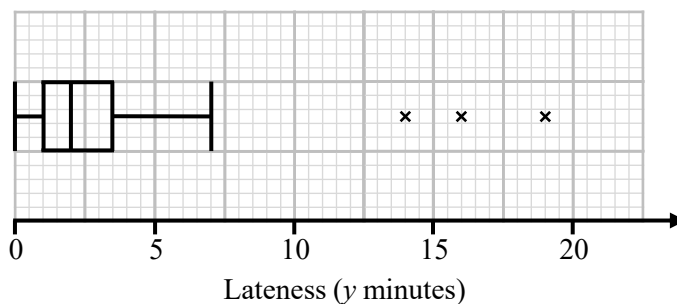
more than $2 \times$ (standard deviation) above the mean or
more than $2 \times$ (standard deviation) below the mean.

The students in Year 12 who were the latest to school are

Merry (13 minutes late) and Perry (26 minutes late)

- (b) Show that Perry's lateness is the **only** outlier from the sample of 20 students in Year 12. (3)

The box plot below shows the lateness to school for the Year 13 students (y)



- (c) State the meaning of the $\times$ symbols shown on the box plot. (1)
- (d) Find the range of the lateness to school for the Year 13 students. (1)

Students who arrive to school 6 or more minutes late receive a detention.
A teacher claims that 5 of the Year 13 students received a detention on Monday for lateness.

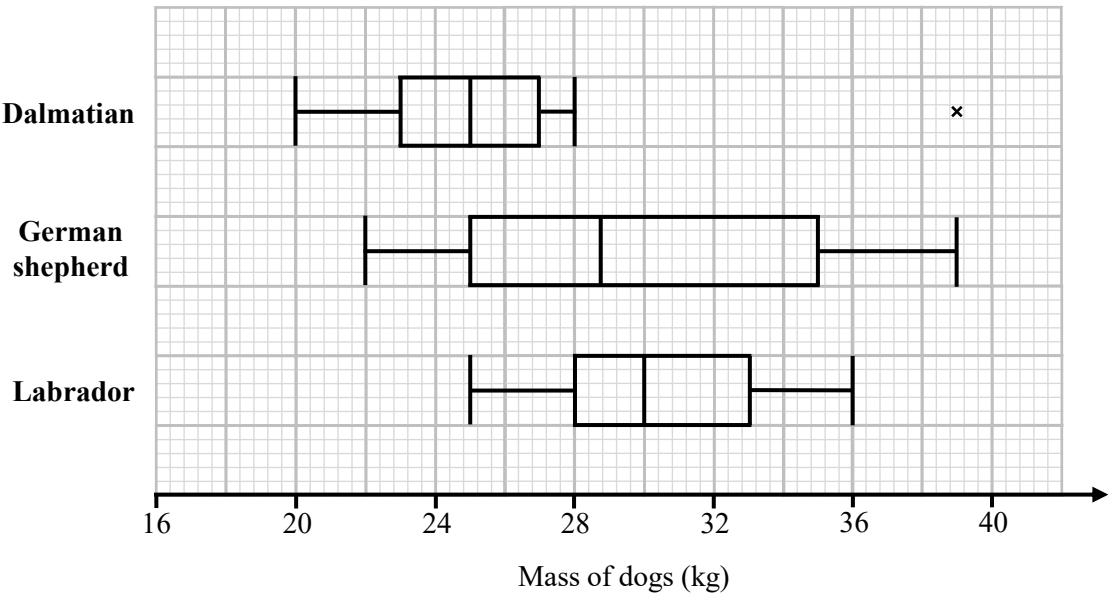
- (e) Show that the teacher must be incorrect. (3)



(Total for Question 4 is 10 marks)



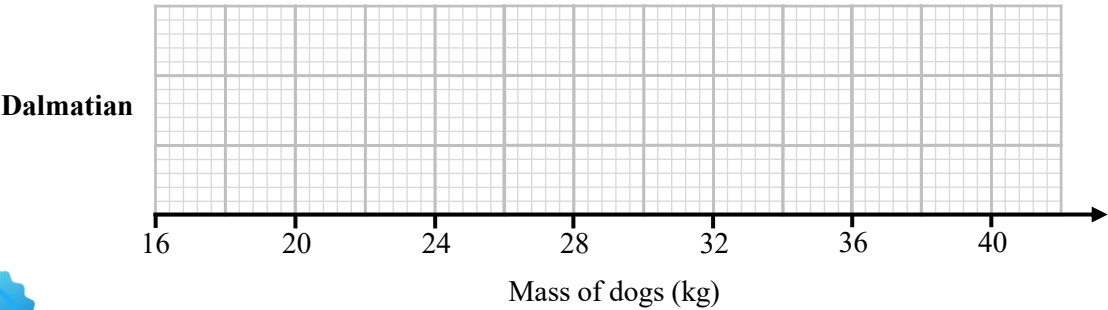
5 A vet measured the masses of 40 dogs of three different breeds. The results are summarised in the box plots below.



(a) Compare the masses of the German shepherds to the labradors. (2)

The heaviest dalmatian had a mass of 29 kg but was incorrectly logged at 39 kg.

(b) Given that the mass of 29 kg dalmatian is not considered an outlier, draw a corrected box plot for the masses of the 40 dalmatians on the grid below. (2)

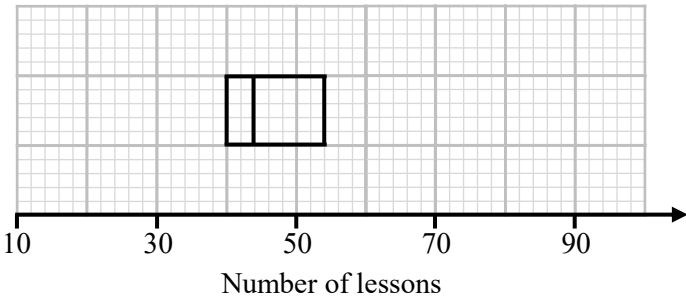


(Total for Question 5 is 4 marks)



6 A driving instructor recorded the number of lessons before each of their learners were ready to take their driving test in 2023.

The box plot below shows the lower quartile, median and upper quartile for this data.



An outlier is any value that falls either
more than $1.5 \times (\text{interquartile range})$ above the upper quartile or
more than $1.5 \times (\text{interquartile range})$ below the lower quartile.

The learners with the greatest and least number of lessons are shown below.

Learner	A	B	C	D	E	F
Number of Lessons	14	24	25	72	80	85

(a) (i) Determine which of the values above are outliers. (2)

(ii) Hence, complete the box plot on the next page for the number of lessons. (2)

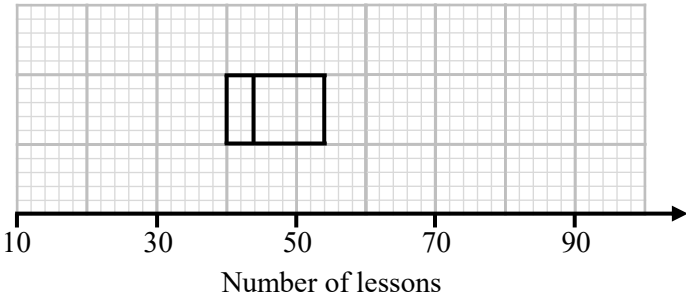
The driving instructor also recorded the number of lessons their learners had before they were ready to take their test in 2024. The 2024 data is summarised below.

Q_1	Q_2	Q_3
38	41	50

(b) Compare the number of lessons before learners were ready to take their driving test in 2023 and 2024. (2)



Handwriting practice lines (10 lines).



Handwriting practice lines (10 lines).



- 7 A student recorded the sound level (S , dB) in a town centre every minute between 7am and 7pm. The results are summarised in the table below.

Sound level (S , dB)	Frequency (f)	Midpoint (S , dB)
$40 \leq S < 50$	99	45
$50 \leq S < 60$	162	55
$60 \leq S < 70$	132	65
$70 \leq S < 80$	129	75
$80 \leq S < 90$	120	85
$90 \leq S < 100$	78	95

You may use $\sum fS = 49230$ and $\sum fS^2 = 3544800$

- (a) Calculate, to 3 decimal places, an estimate for the mean sound level. (1)
- (b) Calculate, to 3 decimal places, an estimate for the standard deviation of the sound levels. (2)

An outlier is any value that falls either

more than $2 \times$ (standard deviation) above the mean or
 more than $2 \times$ (standard deviation) below the mean.

The minimum sound level recorded was 41 dB.
 The maximum sound level recorded was 98 dB.

- (c) Use your answers to parts (a) and (b) to show that none of the sound levels are outliers. (2)
- (d) Use linear interpolation to calculate estimates for the lower quartile, median and upper quartile for the sound levels. (4)

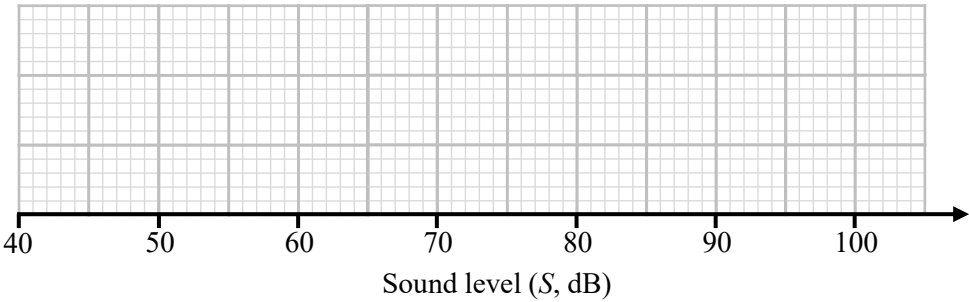
Another way to define an outlier is any value that falls either

more than $1.5 \times$ (interquartile range) above the upper quartile or
 more than $1.5 \times$ (interquartile range) below the lower quartile.

- (e) (i) Show that none of the sound levels were outliers using this definition. (2)
- (ii) Draw a box plot for the sound levels. (2)



Handwriting practice lines consisting of 20 horizontal lines.



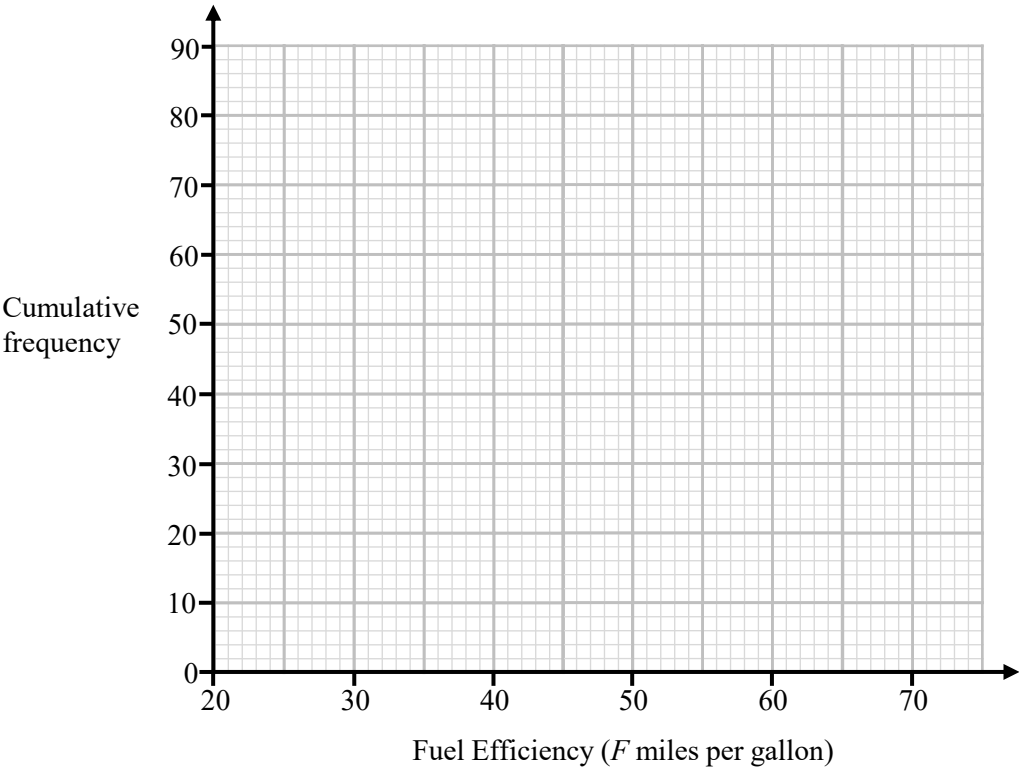
(Total for Question 7 is 13 marks)



8 The grouped frequency table gives information about the fuel efficiency (F , miles per gallon) of 80 different vehicles.

Fuel Efficiency (F miles per gallon)	$20 \leq F < 30$	$30 \leq F < 40$	$40 \leq F < 50$	$50 \leq F < 60$	$60 \leq F < 70$
Frequency	15	21	19	17	8

(a) On the grid, draw the cumulative frequency graph for this information.
(2)



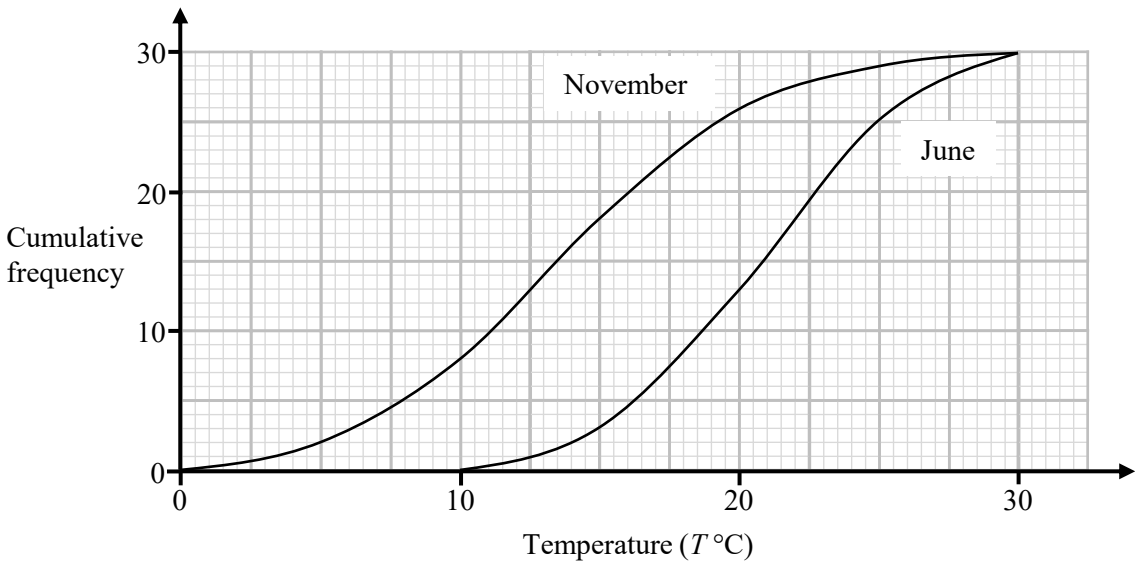
(b) Use your cumulative frequency diagram to estimate

(i) the median
(ii) the upper quartile
(iii) the 80th percentile
(iv) the 15th percentile

(4)



- 9 The cumulative frequency diagram below shows the maximum temperatures ($T^{\circ}\text{C}$) for a city for each of the days during June and November.



- Work out estimates for the median and interquartile range of the temperatures in June. (2)
- Work out estimates for the median and interquartile range of the temperatures in November. (2)
- Compare the maximum temperatures in June and November. (2)

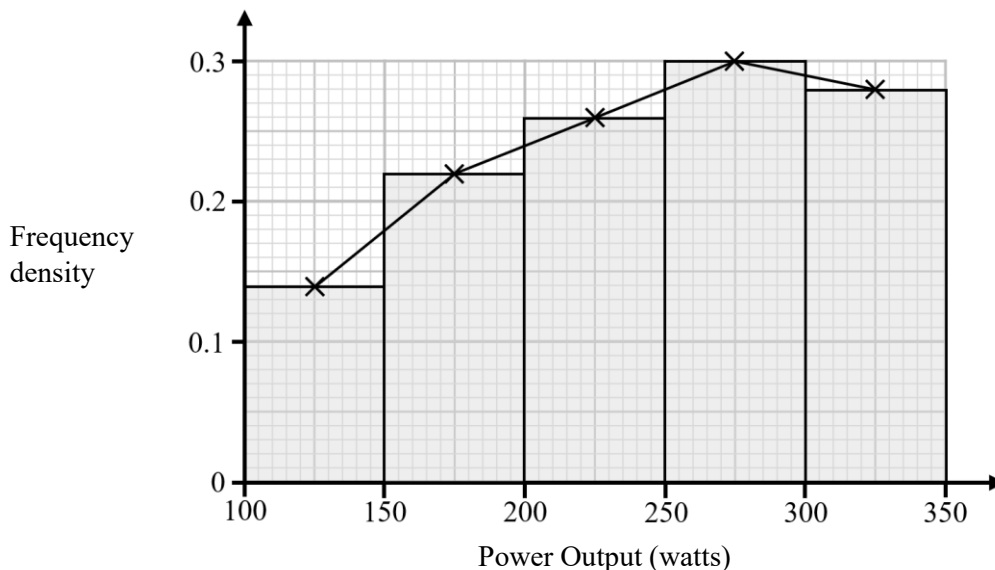
This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small blue circular icon with a gear-like or cogwheel design. The rest of the page is empty.

(Total for Question 9 is 6 marks)



CHECK YOUR ANSWERS

- 2 The histogram and its frequency polygon below give information about the power outputs, in watts, of 60 cyclists from a cycling club.



- (a) Calculate an estimate for the mean power output of the 50 cyclists. (2)
- (b) Calculate an estimate for the standard deviation of the power outputs of the 50 cyclists. (2)

An outlier is any value that falls either

more than $2 \times (\text{standard deviation})$ above the mean or
 more than $2 \times (\text{standard deviation})$ below the mean.

The lowest power output of the 50 cyclists was 115 watts.

The highest power output of the 50 cyclists was 339 watts.

- (c) Use your answers to parts (a) and (b) to show that, based on the estimates, none of the power outputs are outliers. (2)
- (d) Use linear interpolation to estimate the median power output. Give your answer to 1 decimal place. (4)

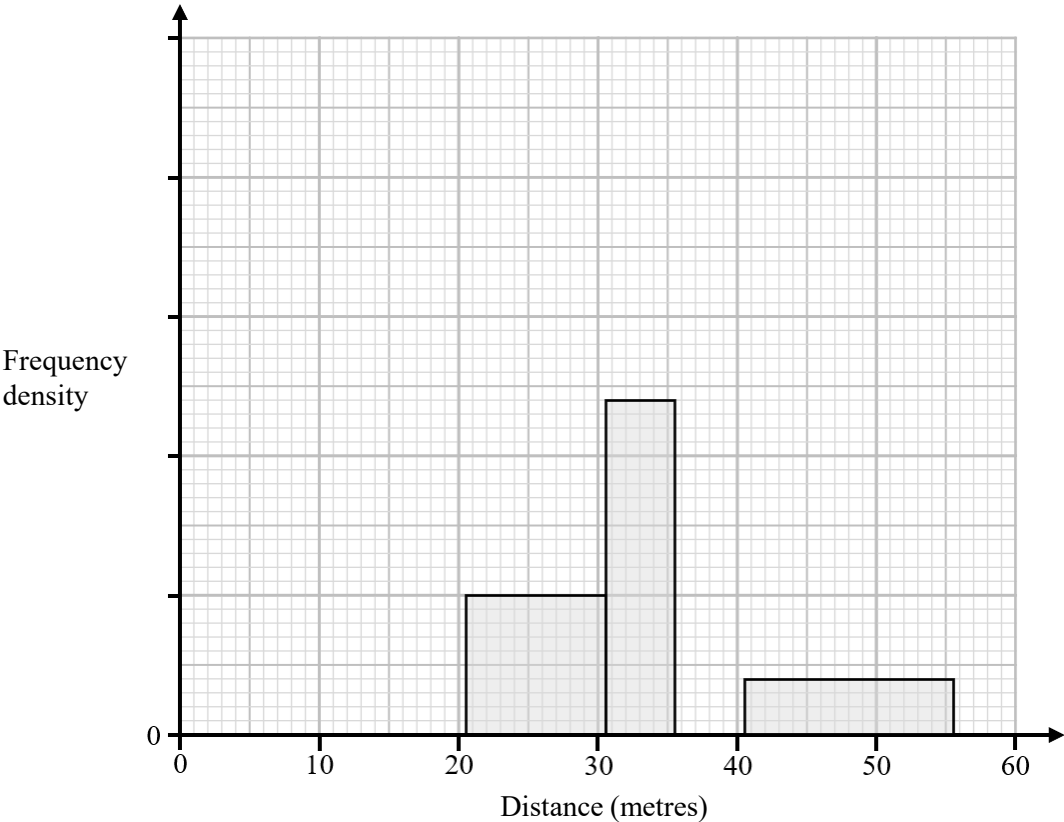


(Total for Question 2 is 10 marks)



3 The partially completed table and partially completed histogram below show the distances, to the nearest metre, of javelin throws by athletes in a competition.

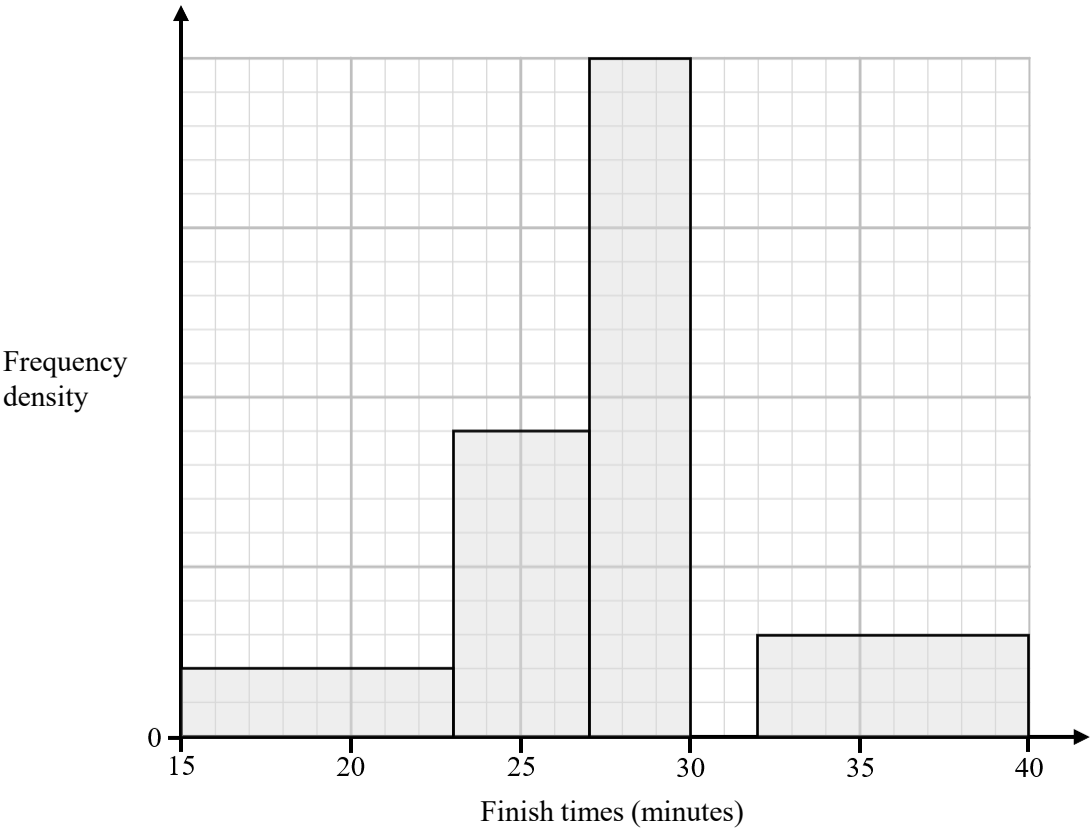
Distance (metres)	5 – 20	21 – 30	31 – 35	36 – 40	41 – 55
Frequency	4	5		12	



- (a) Complete the table. (2)
- (b) Complete the histogram. (2)



4 The partially completed histogram below gives information about the finish times, in minutes, of runners at a fun run.



All the runners finished with times between 15 and 40 minutes.

10% of the runners finished with a time less than 23 minutes.

(a) Complete the histogram (3)

One of the runners is selected at random.

(b) Estimate the probability that this runner had a finish times less than 18 minutes 30 seconds. (2)

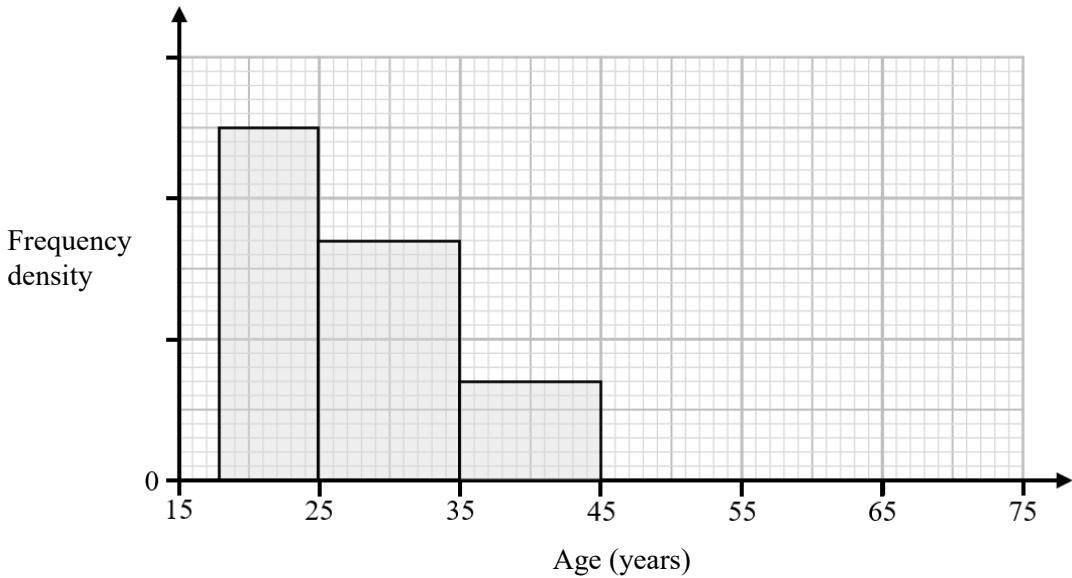
(c) Use linear interpolation to estimate the interquartile range of the finish times. (5)



(Total for Question 4 is 10 marks)



- 5 The partially completed histogram below gives information about the ages of employees at a large company.



The number of employees aged between 25 and 35 is 80 more than employees aged between 35 and 45.

- (a) Work out an estimate for the number of employees aged between 18 and 20. (3)

There are a total of 416 employees at the company.
The age of the oldest employee is 75

- (b) Complete the histogram by adding a bar for ages between 45 and 75. (3)

An outlier is any value that falls either

more than $1.5 \times (\text{interquartile range})$ above the upper quartile or
 more than $1.5 \times (\text{interquartile range})$ below the lower quartile.

Given that $Q_1 = 23$ and $Q_3 = 36$

- (c) Show that at least one employees age is considered an outlier. (2)

The mean age of the employees is 31.7 years, and the standard deviation is 11.8 years.

The CEO of the company assumes that in 1 year's time, the same 416 employees will be working at the company and that there will be no new employees.

- (d) Using the CEO's assumption, write down
 (i) the mean age of employees at the company in 1 year's time.
 (ii) the standard deviation of the ages of the employees in 1 year's time. (2)



(Total for Question 5 is 10 marks)



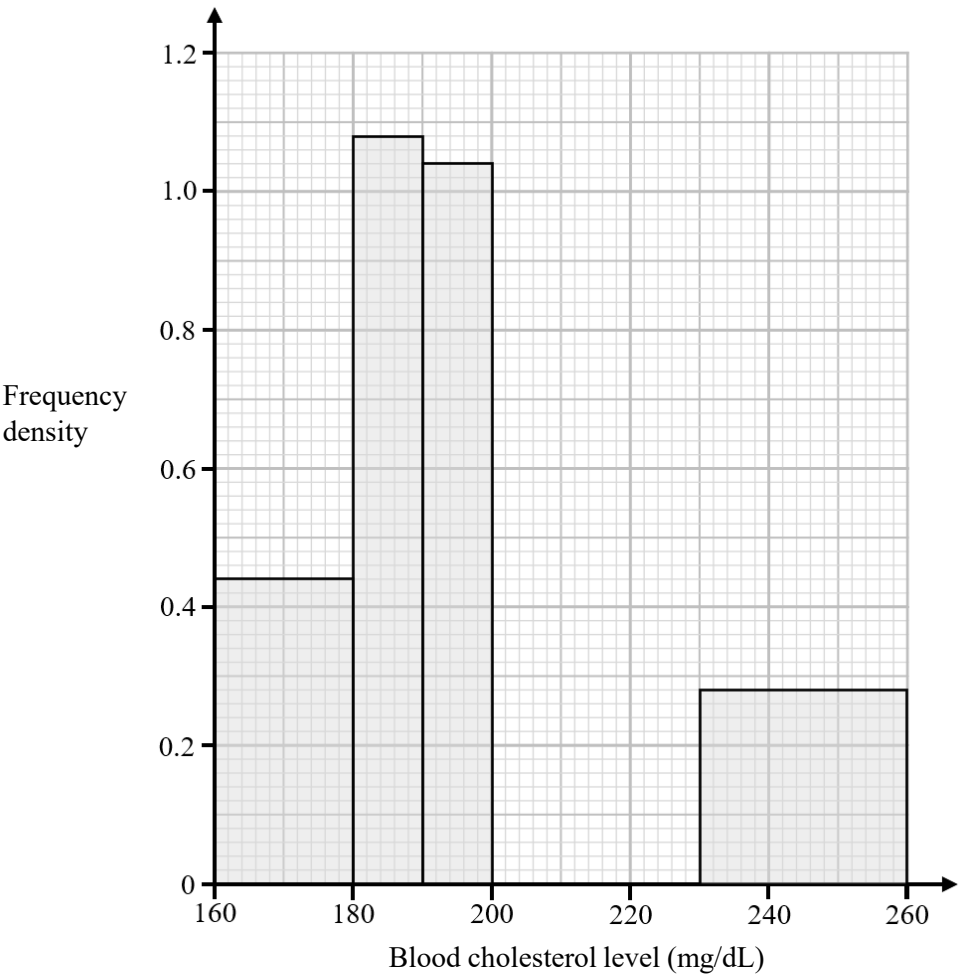
6 A doctor wanted to investigate the blood cholesterol levels of patients at their surgery. They decide to take a sample of 200 patients.

(a) Identify the population. (1)

The doctor asks the next 200 patients who book appointments at the surgery if they will take a blood cholesterol level test.

(b) State the type of sampling technique used by the doctor. (1)

The partially completed histogram below shows the blood cholesterol levels of the patients who agreed to do the test.



22 of the patients had a blood cholesterol level between 160 mg/dL and 180 mg/dL.

(c) Work out the number of patients with a blood cholesterol level between 230 mg/dL and 260 mg/dL. (3)

60% of the 200 patients who were asked, agreed to take the test.
Of these patients, all recorded a blood cholesterol level between 160 mg/dL and 260 mg/dL.

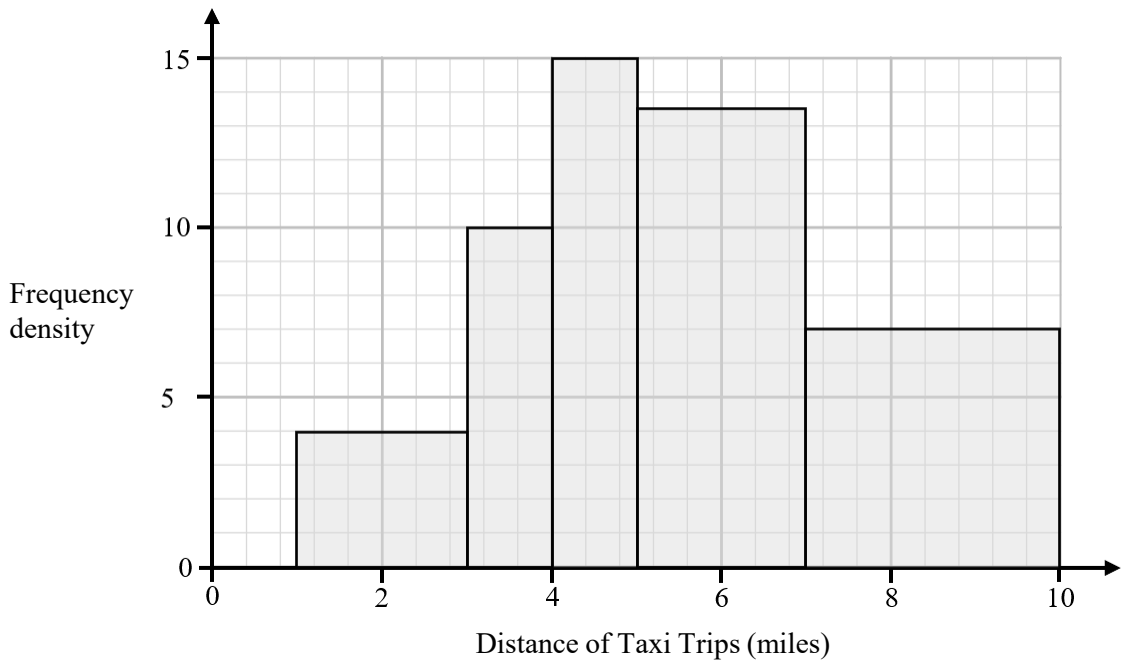
(d) Complete the histogram by adding the bar for blood cholesterol levels between 200 mg/dL and 230 mg/dL (3)



(Total for Question 6 is 8 marks)



7 The histogram below shows the distances, in miles, of 81 taxi trips provided by a taxi driver.



The taxi driver only provides trips if the journey distance is between 1 mile and 10 miles.

(a) Calculate the percentage of the taxi trips that were between 3 and 5 miles. (2)

The mean distance of the taxi trips is 5.6 miles.

The standard deviation of the distances of the taxi trips is 2.35 miles.

An outlier is any value that falls either

more than $2 \times$ (standard deviation) above the mean or
 more than $2 \times$ (standard deviation) below the mean.

(b) Show that none of the distances of the taxi trips were considered outliers. (2)

The taxi driver decides to model the **frequency density** for the 81 taxi trips by a curve with equation

$$y = k(1 - x)(x - 10) \quad 1 \leq x \leq 10$$

where k is a constant.

(c) Find the value of k . (4)

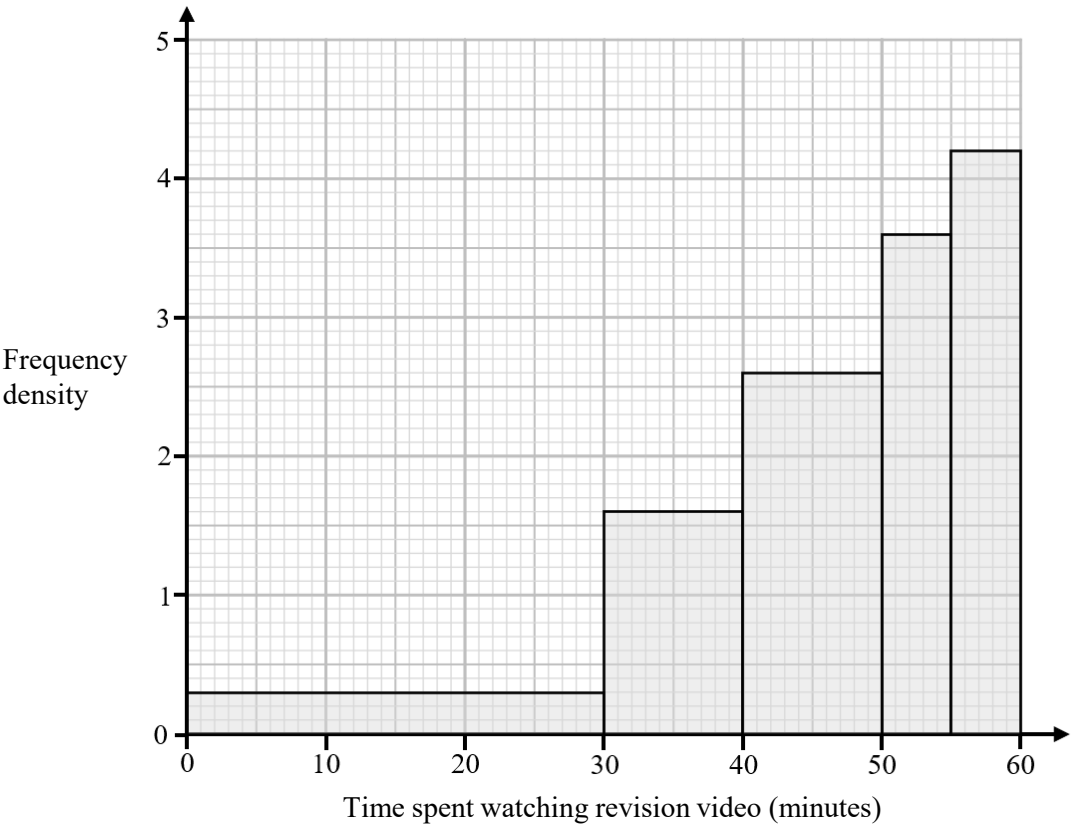


(Total for Question 7 is 8 marks)



8 90 students were asked to watch a 60-minute revision video to prepare for their exams.

The histogram below shows the amount of time each student spent watching the video, up to a maximum of 60 minutes.



- (a) Use linear interpolation to work out an estimate for the median time spent watching the video. Give your answer in minutes and seconds, to the nearest second. (4)

The class teacher decides to model the **frequency density** for the 90 times by a curve with equation

$$y = \frac{1}{k}x^2 \quad 0 \leq x \leq 60$$

where k is a constant.

- (b) Find the value of k . (4)



(Total for Question 8 is 8 marks)



 **SCAN ME**

CHECK YOUR ANSWERS

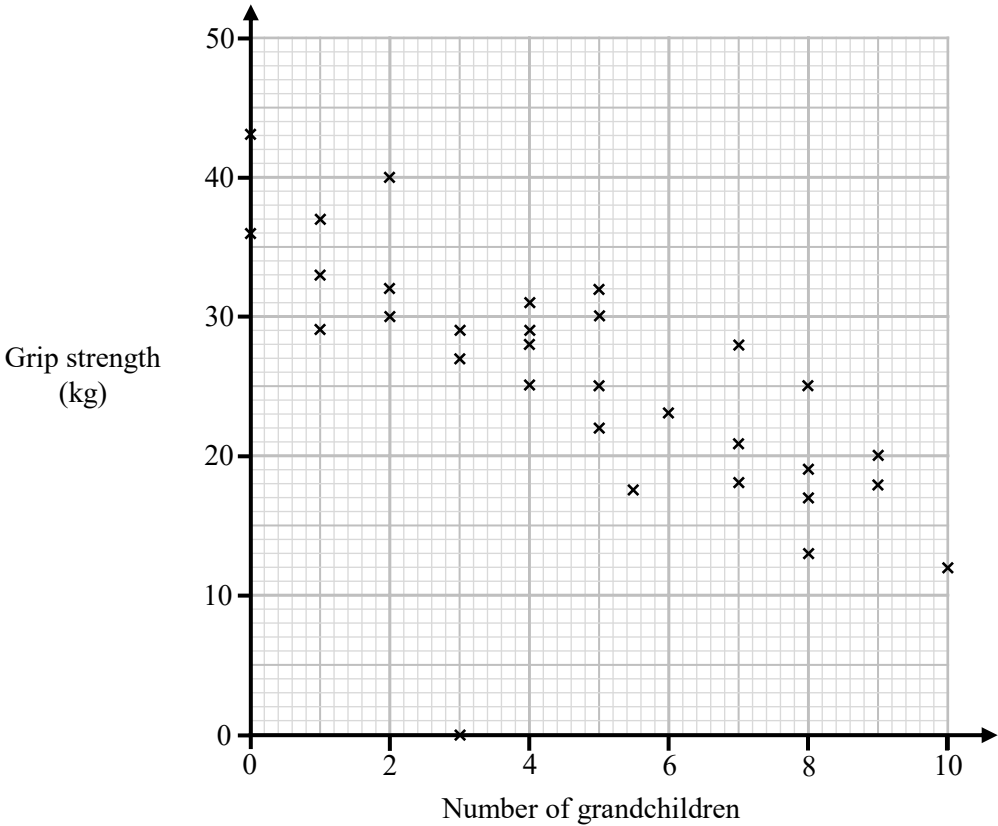
- (b) Describe the correlation between the age of the car and its value. **(1)**
- (c) Give an interpretation of the gradient of this regression line. **(1)**
- (d) Describe one limitation of this regression model. **(1)**

[illegible]



3 Kareem is investigating whether there is a linear relationship between grip strength, in kilograms and the number of grandchildren that a person has.

Kareem selects a sample of 30 people and draws the scatter diagram below using the data.



Kareem says: “Having more grandchildren causes a reduction in grip strength”.

(a) Comment on Kareem’s claim. (1)

One of the points on the scatter diagram is an anomalous result.

(b) Circle the point and explain how you know that it is anomalous. (1)

Before doing further calculation Kareem decides to clean the data.

(c) Explain what is meant by cleaning the data. (1)

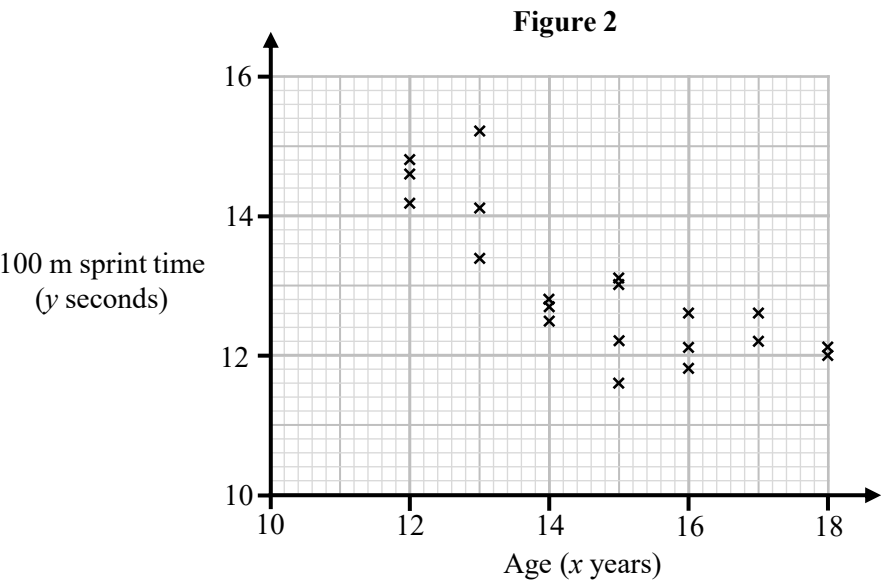
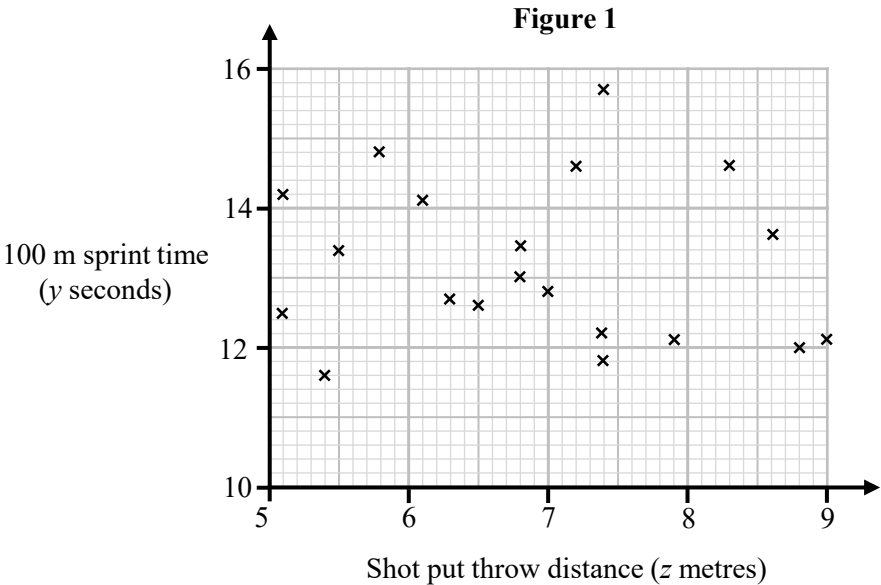
(Total for Question 3 is 3 marks)





5 A athletics coach records the age (x years), 100-metre sprint time (y seconds), and shot-put throw distance (z metres) for a sample of 20 athletes at their athletics club.

The coach uses the data to draw the two scatter diagrams shown below in **Figure 1** and **Figure 2**.







 **SCAN ME**

CHECK YOUR ANSWERS

A Venn diagram with two overlapping circles, labeled A and B . The region inside circle A but outside circle B is labeled 0.12 . The region inside circle B but outside circle A is labeled 0.42 . The intersection of the two circles is labeled 0.18 . The number 0.28 is written outside the circles, representing the total number of elements in the universal set.

- (a) Find
 - (i) $P(A)$ (1)
 - (ii) $P(B)$ (1)
 - (iii) $P(A \text{ or } B \text{ or both})$ (1)
- (b) State, giving a reason, whether events A and B are mutually exclusive. (1)
- (c) State, giving a reason, whether events A and B are independent. (1)

[illegible]

2

The probabilities for each fish caught are:

$$P(\text{large fish}) = 0.24$$

$$P(\text{medium fish}) = 0.72$$

$$P(\text{small fish}) = 0.04$$

Finn goes fishing four times in May, catching one fish each time.

For the four fish caught by Finn in May, find, to 4 decimal places, the probability that:

- (c) exactly one fish is large and the rest are medium. (3)

This image shows a full page of white paper with horizontal blue ruling lines. The lines are evenly spaced and run across the width of the page. There is a small blue decorative element in the bottom left corner, which appears to be a stylized wave or cloud shape. The rest of the page is completely blank.

(Total for Question 2 is 7 marks)



3 A college has 160 students. Of these students

60 study maths

50 study physics

64 were late to school

24 of the students who were late to school also study maths.

20 of the students who were late to school also study physics.

91 study maths or physics or both

One student from the college is selected at random.

M is the event that the student selected studies maths

P is the event that the student selected studies physics

L is the event that the student selected was late to school

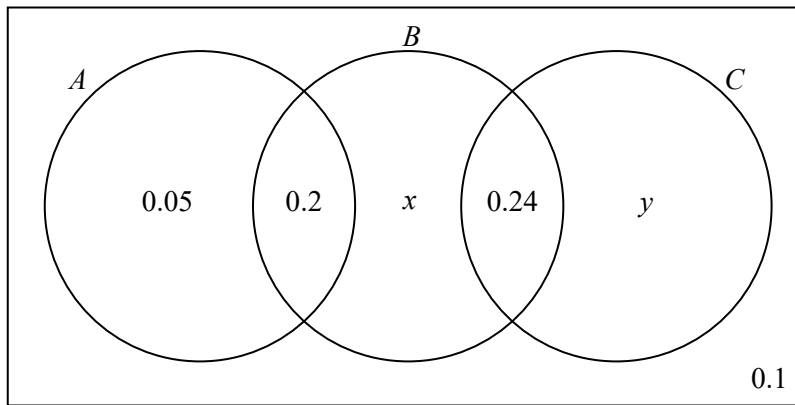
- Determine, by calculation, if the events M and L are independent. (2)
- Determine, by calculation, if the events P and L are independent. (2)
- Determine, by calculation, if the events M and P are independent. (3)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small blue gear-shaped icon. The rest of the page is empty.

(Total for Question 3 is 7 marks)



- 4 The Venn diagram shows the events A , B and C and their associated probabilities.



Events A and B are independent.

- (a) Find the value of x **(2)**
- (b) Find the value of y **(1)**
- (c) Determine, by calculation, if the events B and C are independent. **(1)**

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, stylized blue gear or cogwheel icon. The rest of the page is empty.

(Total for Question 4 is 4 marks)

5 Jeremy and Doug both play a computer game that has three levels.

If a player completes level 1 they progress to level 2.

If a player completes level 2 they progress to level 3.

If a player completes level 3 they complete the game.

If players fail to complete any of the levels the game is over and they stop playing.

The probabilities that Jeremy and Doug complete each of the levels of the game are shown below.

	Level 1	Level 2	Level 3
Jeremy	0.9	0.7	0.3
Doug	0.8	0.6	0.4

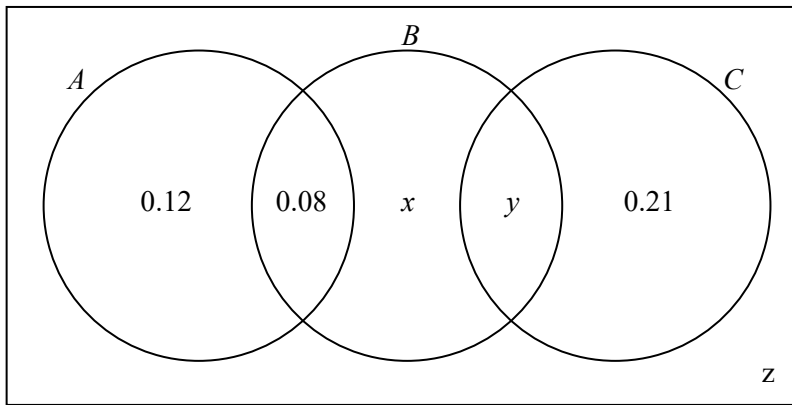
Both Jeremy and Doug make one attempt to complete the game.

- Work out who is more likely to complete the game. **(1)**
- Find, to 4 decimal places, the probability that exactly one of them completes the game. **(2)**
- Find, to 4 decimal places, the probability that at least one of them completes the game. **(2)**
- Find, to 4 decimal places, the probability that Jeremy fails on level 3 and Doug fails on level 2. **(2)**

This image shows a single sheet of white paper with horizontal blue or grey ruling lines. The lines are evenly spaced and run across the width of the page. There are approximately 20 lines visible. The paper has a slight shadow on the right side, suggesting it's resting on a surface.

(Total for Question 5 is 7 marks)

6 The Venn diagram shows the events A , B and C and their associated probabilities.



Events A and B are independent.

Events B and C are independent.

Find the values of x , y and z

(5)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, stylized blue gear or cogwheel icon. The rest of the page is empty.

(Total for Question 6 is 5 marks)

9 A bag contains 16 red dice, 5 blue dice and 3 green dice.

Each red die is fair and 4-sided, with faces numbered 1, 2, 5 and 6

Each blue die is fair and 5-sided, with faces numbered 1, 2, 3, 4 and 5

Each green die is fair and 3-sided, with faces numbered 1, 2 and 3

One of the dice is selected at random and then rolled.

R is the event the dice selected is red.

X is the event the number rolled is 2.

(a) Show that events R and X are independent. (3)

The first die is not returned to the bag.

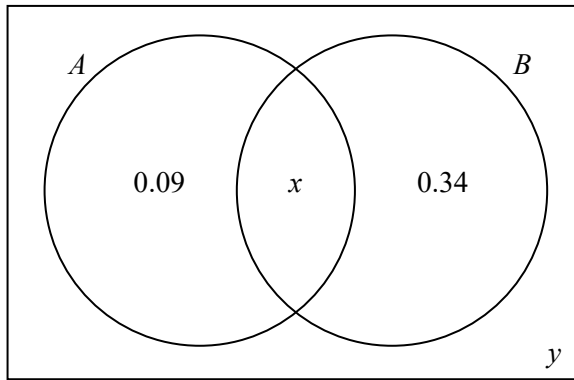
A second die is selected at random and then rolled.

(b) Work out the probability that the total of both rolls is 11. (3)
Give your answer as a fraction in its simplest form.

Handwritten solution area with horizontal lines.



10 The Venn diagram shows the events A and B and their associated probabilities.



Events A and B are independent.

Give that $x > y$, find the value of x and y .

(4)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, stylized blue wave icon. The rest of the page is empty.

(Total for Question 10 is 4 marks)



(5)

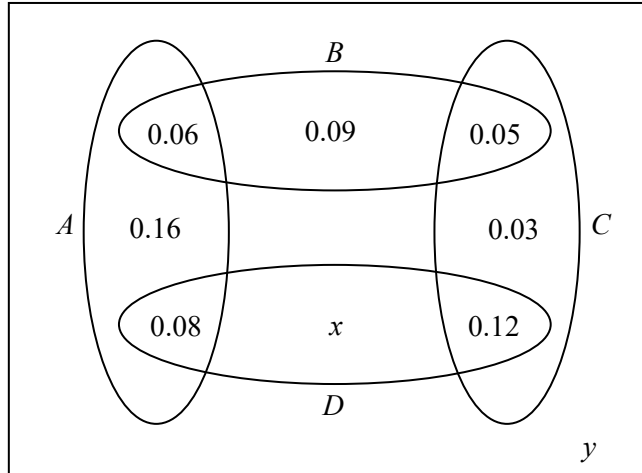
This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, stylized blue gear or cogwheel icon. The rest of the page is empty.

(Total for Question 11 is 5 marks)





13 The Venn diagram shows the events A , B , C and D and their associated probabilities.



- (a) Write down all pairs of mutually exclusive events from A , B , C and D . **(1)**
- (b) Show that events A and B are independent. **(1)**
- (c) Show that events B and C are not independent. **(1)**
- Events C and D are independent.
- (d) Work out the values of x and y . **(3)**

Events C and D are independent.

[illegible]

(Total for Question 13 is 6 marks)



CHECK YOUR ANSWERS

y	2	4	6	8
$P(Y=y)$	0.2	0.3	0.4	q

- (c) Find
- (i) $P(Z = 3)$ **(1)**
- (ii) $P(Z > 1)$ **(1)**

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, stylized blue gear or cogwheel icon. The rest of the page is empty.

2

x	0	1	2	3	4	5	6
$P(X=x)$	p	q	r	0.14	0.16	0.19	0.21

(a) Find $P(4 \leq X < 6)$

$$P(X > 1) = 0.88$$

$$P(X=0) = 2 \times P(X=1)$$

(b) Find the values of p , q and r .

The random variable $Y = X^2$

(c) Find

(i) $P(Y = 3)$

(ii) $P(Y > 4)$

(iii) $P(X < Y)$

3 X is a random variable with the following probability distribution.

x	5	6	7	8	9
$P(X=x)$	k	k	$3k$	$5k$	$5k$

- (a) Find the value of k (2)

- (b) Find
- (i) $P(X \leq 6)$ (1)
- (ii) $P(5 < X < 9)$ (1)

The random variable $Y = 4X$

n is an integer with $P(X + Y < n) > 0.5$

- (c) Find the minimum possible value for n . (3)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, stylized blue gear or cogwheel icon. The rest of the page is empty.

(Total for Question 3 is 7 marks)



5 X , Y and Z are independent discrete random variables with the following probability distributions.

x	1	2	3	4	5
$P(X=x)$	0.2	0.1	0.05	0.25	0.4

y	0	2	4	6	8
$P(Y=y)$	0.24	0.18	0.11	0.15	0.32

z	5	6	7	8
$P(Z = z)$	0.4	0.07	0.13	0.4

(a) Find

(i) $P(2 < X < 5)$

(1)

(ii) $P(Y > 2)$

(1)

(b) Find

(i) $P(X + Y > 11)$

(2)

(ii) $P(X + Z = 11)$

(2)

(iii) $P(Y > Z)$

(3)

(Total for Question 5 is 9 marks)





7

The probability that Hannah passes each of the subjects is shown below.

Subject	Maths	Chemistry	History
Probability of passing	0.9	p	0.6

The outcome of each exam is independent of the others.

The probability that Hannah passes all three exams is 0.459

(a) Find the value of p . **(1)**

The discrete random variable X represents the number of exams that Hannah passes.

(b) Find the complete probability distribution of X . (4)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small blue circular icon with a gear-like or scalloped edge. The rest of the page is empty.

(Total for Question 7 is 5 marks)



8 Kat plays a game at a fair that costs £5 per play.

The probability that Kat wins the game is 0.3

Kat has £20 and plays the game either until she has won the game twice or until she has no money left.

The discrete random variable X represents the number of times that Kat plays the game.

(a) Find the complete probability distribution of X . (4)

The discrete random variable Y represents the number of times that Kat wins the game.

(b) Find the complete probability distribution of Y . (4)

[illegible]

(Total for Question 8 is 8 marks)



9 A biased dice can land on the numbers 1, 2, 3, 4, 5, or 6.

The random variable X represents the number that the dice lands on.

$$P(X=r) = P(X=7-r) \text{ for } r = 1, 2, 3$$

Given that $P(X=3)=0.05$ and $P(X=2)=2 \times P(X=1)$

(a) Find the complete probability distribution of X . (4)

The dice is rolled 3 times.

The random variable Y represents the number of times that the dice lands on the number 2.

(b) Find the complete probability distribution of Y . (4)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small blue circular icon with a gear-like or scalloped edge. The rest of the page is empty.

(Total for Question 9 is 8 marks)



10 X is a random variable with the following probability distribution.

x	1	2	3	4	5
$P(X=x)$	0.12	0.16	0.2	0.24	0.28

$$Y = X^2$$

$$Z = 2X$$

(a) Find

(i) $P(X < 3)$

(1)

(ii) $P(Y < 3)$

(1)

(iii) $P(2 < Z \leq 10)$

(1)

(b) Find

(i) $P(Y = Z)$

(1)

(ii) $P(Y - 4X = -3)$

(3)

(iii) $P(Y + 8 \leq 3Z)$

(3)

11 The random variable X has a probability function

$$x = 1, 2, 3$$

where k is a constant.

- (2)**

The random variable Y has a probability function

$$y = 1, 2, 3, 4$$

where c is a constant.

- (2)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, stylized blue circular logo with a scalloped edge. The logo contains some faint, illegible text or a symbol. The rest of the page is empty.

(Total for Question 11 is 8 marks)

12 The random variable X has a probability function

$$P(X=x) = (k+1)x \quad x=1, 2, 3$$

where k is a constant.

(a) By working out the value of k , find the complete probability distribution of X . (4)

The random variable Y has a probability function

$$P(Y=y) = \frac{c-y^2}{100} \quad y = 1, 2, 3, 4$$

where c is a constant.

(b) By working out the value of c , find the complete probability distribution of Y . (4)

[illegible]

(Total for Question 12 is 8 marks)



13 The random variable X has a probability function

$$P(X=x) = \frac{1}{k-x} \quad x=1,3$$

where k is a constant.

Find the value of k giving your answer in the form $a + \sqrt{b}$, where a and b are integers. (5)

[illegible]

(Total for Question 13 is 5 marks)



14 The random variable X has a probability function

$$x = 5, 6, 7, 8$$

where k and c are constants.

Given that $P(X=5) = 0.01$ find the values of k and c .

(4)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small blue circular icon with a gear-like or sunburst pattern. The rest of the page is empty.

(Total for Question 14 is 4 marks)



15 The random variable X has a probability function

$$P(X=x) = \begin{cases} kx & x=1, 2 \\ (k-0.2)x & x=3, 4 \end{cases}$$

where k is a constant.

(a) By working out the value of k , find the complete probability distribution of X . (5)

The random variable Y has a probability function

$$P(Y=y) = \begin{cases} c(y+0.5) & y=0, 1 \\ cy^2 & y=2, 3, 4 \end{cases}$$

where c is a constant.

(b) By working out the value of c , find the complete probability distribution of Y . (5)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small blue gear-shaped icon. The entire page is otherwise empty.

(Total for Question 15 is 10 marks)





SCAN ME

The Binomial Distribution



SCAN ME

REVISE THIS TOPIC

CHECK YOUR ANSWERS

1 The random variable $X \sim B(30, 0.4)$

Find

- | | |
|---------------------|-----|
| (i) $P(X = 18)$ | (1) |
| (ii) $P(X \leq 10)$ | (1) |
| (iii) $P(X < 10)$ | (1) |
| (iv) $P(X \geq 13)$ | (2) |
| (v) $P(X > 13)$ | (2) |

(Total for Question 1 is 7 marks)

2 The random variable $Y \sim B(12, 0.65)$

Find

- | | |
|--------------------------|-----|
| (i) $P(Y = 7.8)$ | (1) |
| (ii) $P(Y < 7.8)$ | (1) |
| (iii) $P(5 < Y < 11)$ | (2) |
| (iv) $P(3 \leq Y < 9)$ | (2) |
| (v) $P(4 \leq Y \leq 6)$ | (2) |

(Total for Question 2 is 8 marks)





5

For Penny's coin, $P(\text{Heads}) = 0.78$

Penny flips the coin 15 times and records the number of times it lands on heads.

- (2)

Franc also has a biased coin that can land on either heads or tails.

For Franc's coin, $P(\text{Heads}) = 0.6$

Franc flips the coin n times and records the number of times it lands on heads.

$$0.1 < P(\text{Franc throws less than 10 heads}) < 0.4$$

- (4)

[illegible]

(Total for Question 5 is 8 marks)

6

A

$$P(\text{Skye scores a penalty kick}) = 0.85$$

$$P(\text{Ava scores a penalty kick}) = 0.81$$

I

- (

I

- (

- (

2

“When I miss a few penalties in a row, I feel pressure and I am more likely to miss the next one”

- (

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, stylized blue gear or cogwheel icon. The rest of the page is empty.

(Total for Question 6 is 11 marks)





8 A supermarket runs a promotion for its customers.

When reaching the checkout, customers are asked to pick an integer from 1 to 100.

The cashier then generates a random integer between 1 and 100 using a random number generator.

If the numbers match, then the customer wins and does not have to pay for their shopping.

Between 3pm and 4pm, 350 customers take part in the promotion.

- (a) State a suitable binomial distribution to model the number of customers that win and do not have to pay for their shopping between 3pm and 4pm. **(1)**
- (b) Find the probability that between 3pm and 4pm
 - (i) exactly 3 customers win **(1)**
 - (ii) no more than 5 customers win **(1)**
 - (iii) more than 7 customers win **(2)**

Between the hours of 7pm and 8pm the store manager changes the promotion, so customers only need to select a number between 1 and 50.

Between 7pm and 8pm 630 customers take part in the promotion.

The store manager says “*I am 99% sure that no more than 20 customers will win*”

- (c) Comment, with calculation, on this claim. (3)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, stylized blue gear or cogwheel icon. The rest of the page is empty.

(Total for Question 8 is 8 marks)

10 To be able to drive, learners need to pass a theory test and then a practical test.

Data shows that 46% of attempts at the theory test result in pass and that 49% of attempts at the practical test result in a pass.

On Monday 60 learners take a theory test, and 28 learners take a practical test.

- (a) State suitable binomial distributions to model the number of learners who pass each of the tests. **(2)**
- (b) Use your models to find
 - (i) the probability that the number passing the theory test is exactly 26 **(1)**
 - (ii) the probability that the number passing the theory test is less than 30 **(1)**
 - (iii) the probability that the pass rate for the practical test was more than 75% **(2)**

On Tuesday 52 learners take the theory test.

The random variable A represents the number of learners who pass the theory test on Tuesday.

$$P(A \leq k) < 0.15$$

$$P(A \geq k) < 0.99$$

- (c) Find a set of possible values for k . (4)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, stylized blue gear or cogwheel icon. The rest of the page is empty.

(Total for Question 10 is 8 marks)



===== **BONUS AQA ONLY QUESTIONS** =====

- 12

Allan is attempting to complete a level in a computer game as fast as possible.

One level has a difficult jump to perform, so Allan practices this jump everyday. During a daily practice session, Allan always attempts this jump 80 times.

On each attempt, Allan successfully performs the jump with probability 0.72

The number of successful jumps in a daily practice session can be modelled using a binomial distribution.
- 12 (a)

Find the mean of the number of times that Allan successfully performs the jump in a daily practice session. [1 mark]

- 12 (b)

Find the standard deviation of the number of times that Allan successfully performs the jump in a daily practice session. [2 marks]

- 12 (c)

Alan practices the jump 80 times per day, for 5 days in a row.

Find the probability that on each of these 5 days Allan performs the jump successfully at least 55 times. [3 marks]



BONUS AQA ONLY QUESTIONS

13

$$Y \sim B(b, 0.81)$$

13 (a)

13 (b)



BONUS AQA ONLY QUESTIONS

14

$$X \sim B(n, 0.48)$$

where $n > 0$

$$Y \sim \text{B}(120, p)$$

where $p > 0$

mean of $X = 3 \times$ mean of Y

variance of $X = 2 \times$ variance of Y

Find the values of n and p .

[6 marks]



 **SCAN ME**

CHECK YOUR ANSWERS

- (b) For the test in part (a)
- (i) write down the p -value **(1)**
 - (ii) find the critical region **(2)**
 - (iii) find the actual level of significance **(1)**

This image shows a full page of white paper with horizontal blue ruling lines. The lines are evenly spaced and run across the width of the page. In the bottom-left corner, there is a small, stylized blue gear-like icon.



- (d) Find the actual find the actual level of significance for the test. (1)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small blue circular icon with a gear-like or cogwheel design. The rest of the page is empty.

5 Data from a school shows that each week, 8% of the students receive a detention.

One week, the headteacher selects a sample of 25 students.

- (i) no students receive a detention **(1)**
(ii) less than 3 students receive a detention. **(1)**

The headteacher believes that with a new behaviour policy the proportion of students receiving a detention in a week will decrease.

After the introduction of the new policy, a new sample of 50 students is taken and the headteacher finds that 1 student received a detention.

- (c) Carry out a suitable test to determine whether the proportion of students receiving a detention has decreased. You should state your hypothesis clearly and use a 10% level of significance. (4)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small blue circular icon with a gear-like or cogwheel design. The rest of the page is empty.

(Total for Question 5 is 7 marks)



-
- This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. In the bottom-left corner, there is a small, stylized blue gear or cogwheel icon. The rest of the page is empty.

1st



8

A random sample of 1000 people from the UK is taken and they are asked if they are left-handed.

- (ii) at least 110 are left-handed (2)

Jeremy wants to test if the proportion of people in his city that are left-handed is different to that of the UK.

Jeremy randomly samples 90 people and finds that 4 of them are left-handed.

- You should state your hypothesis clearly and use a 5% level of significance. (4)

[illegible]

(Total for Question 8 is 8 marks)



10 A skincare company claims that 4 in 5 users of their product reported “visibly better skin”.

An advertising standards organisation decided to conduct a hypothesis test, at the 1% level of significance, to investigate if the true proportion differed from 4 in 5 users.

A random sample of 45 users of the product was taken and they were asked if they had “visibly better skin”.

- (a) State suitable hypotheses that the advertising standards organisation could use for this test. **(1)**

- (b) Find the critical region for this test.

You should state the probability associated with each tail, which should be as close to 0.5% as possible.

- (c) Find the actual find the actual level of significance for the test. (1)

The advertising standards organisation observed that 27 of the 45 users of the product reported “visibly better skin”.

- (d) State the conclusion of the hypothesis test. (1)

[illegible]

(Total for Question 10 is 7 marks)

