

Statistics GCSE**Paper 2**

Edexcel Foundation - 2026

Foundation Tier

Variant 3

1ST0/2F

Instructions

- Write all answers in the spaces provided.
- Answer all questions.
- You must show all your working.
- There may not be enough space to show all your working out.

Information

- This is a practise paper to aid your revision for your exams.
- This site, and all that work on it, have no affiliation or relationship with any exam board.
- This site is not endorsed by any company or charity, unless we state otherwise.

Copyright

- Sharing this PDF is strictly forbidden unless you have the author's permission.
- This paper is only authorised to be used by the person who has bought it.
- You may not store this online or any shared area such as an intranet.
- Please contact us if you have any queries.

Advice

- You can get support for all these questions at our website: www.statsgcse.com
- This paper and more are available on our site with questions that change subtly after each attempt.
- Good luck!

1



Noah has 6 number tiles shown above.

All of the tiles are placed inside a bag then a random tile is drawn from the bag.

- (a) Select the word describes the likelihood that the tile has a 3 on it.

(1 mark)

Impossible – This means it cannot happen at all.

Unlikely – This means it could happen, but it probably won't.

Evens – This means it has a 50/50 chance of happening – it's just as likely to happen as not.

Likely – This means it will probably happen, but it's not guaranteed.

Certain – This means it will definitely happen – there's no doubt at all.

Select **one** box.

impossible

likely

evens

unlikely

(b) Select the word describes the likelihood that the tile has a number more than 3 on it.

(1 mark)

Impossible – This means it cannot happen at all.

Unlikely – This means it could happen, but it probably won't.

Evens – This means it has a 50/50 chance of happening – it's just as likely to happen as not.

Likely – This means it will probably happen, but it's not guaranteed.

Certain – This means it will definitely happen – there's no doubt at all.

Select **one** box.

likely

impossible

unlikely

evens

(c)



Using the probability scale, write down the letter that shows probability that the tile has a 1 on it.

(1 mark)

Select **one** box.

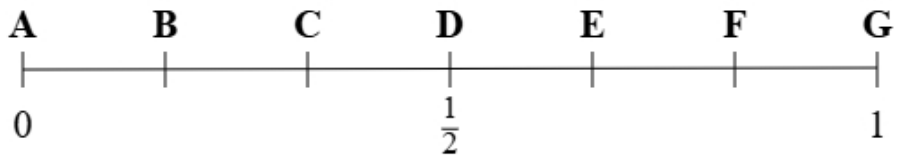
B

A

C

D

(d)



Using the probability scale, write down the letter that shows probability that the tile has a 2 or a 3 on it.

(1 mark)

Select *one* box.

F

D

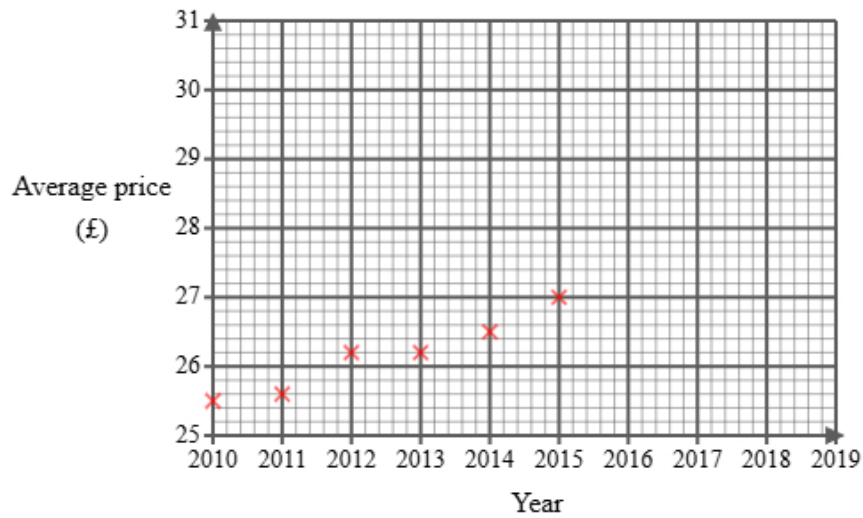
G

E

2 Ethan found the following information about the average price of a football match ticket in England.

| Year | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 |
|-------------------|-------|-------|-------|-------|-------|-------|------|-------|-------|-------|
| Average price (£) | 25.50 | 25.60 | 26.20 | 26.20 | 26.50 | 27.00 | | 28.00 | 28.20 | 28.50 |

He did not find the price for 2016 and has started to draw a graph for the data.

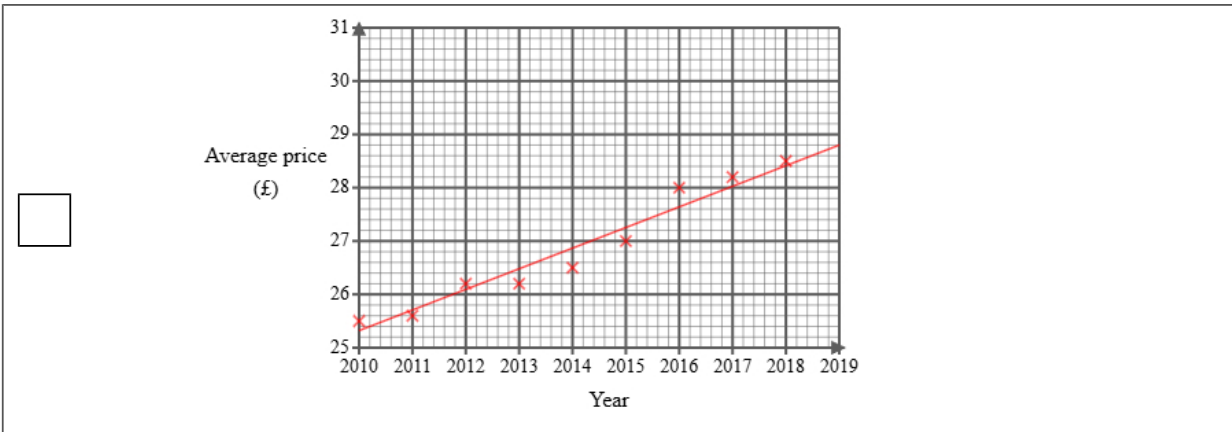
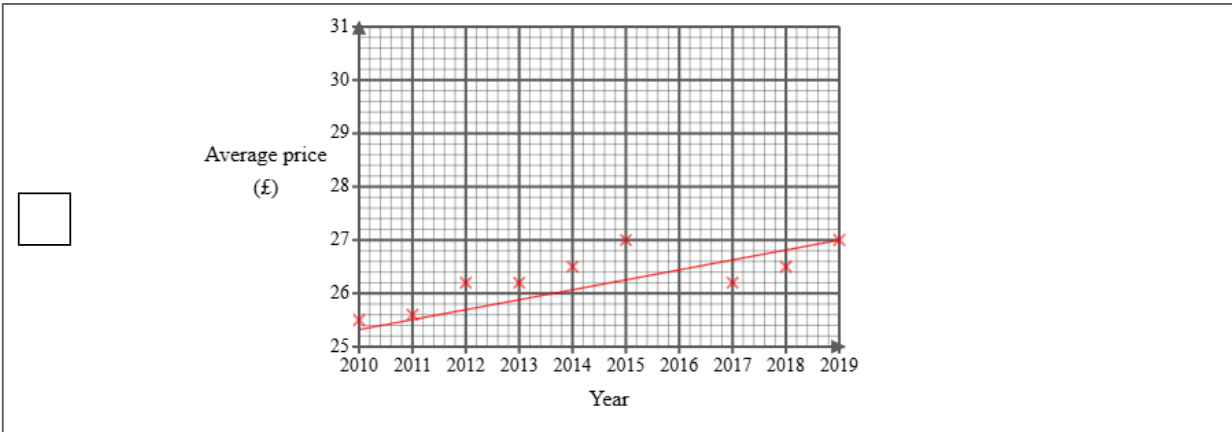
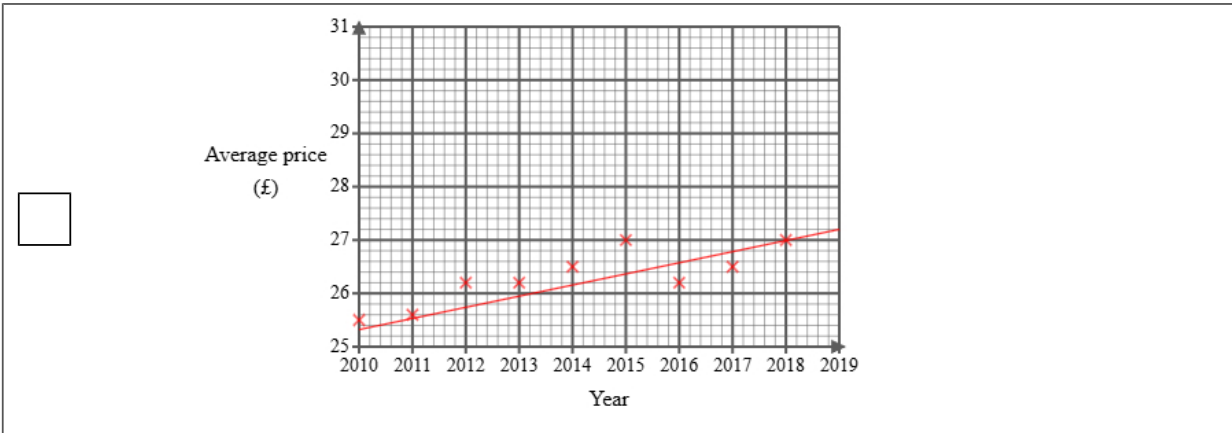
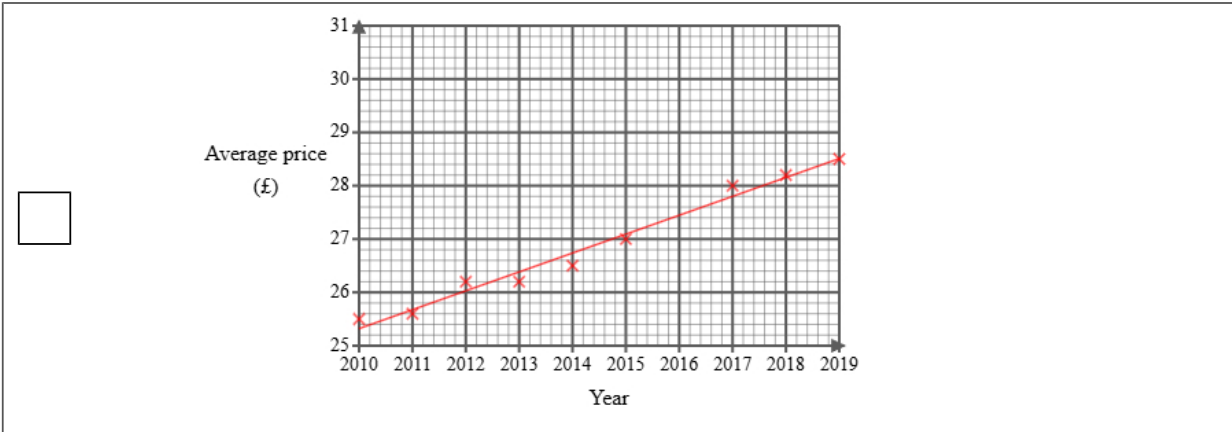


Ethan then used statistical software to calculate the equation for the trend line.

- (a) (i) Plot the average price for each of 2017, 2018 and 2019
(ii) Draw a trend line for Ethan's data
(iii) Describe the trend in the average price of a football match ticket in the UK from 2010 to 2019

(4 marks)

Select the correct answer.



Select the correct boxes.

- Decreasing
- Negative correlation
- Increasing
- Positive correlation

(b) The gradient of Ethan's trend line is 0.35

Interpret this gradient.

(1 mark)

Select **one** box.

- The average price increases per year.
- The average price increases.
- The average price started at 35 pence in 2010.
- The average price decreases.

(c) Explain whether or not the scale used on the 'average price' axis could make the graph misleading.

(2 marks)

Select the **two** correct statements (**two** statements are incorrect).

- The graph could be misleading.
- The 'average price' axis does not start from zero.
- The graph is not misleading.
- The graph is labelled correctly.

(d) Ethan draws the trend line onto the graph.

He suggests that the trend line could be used to estimate the average price for 2016 **and** 2020

Explain whether each of these estimates would be reliable.




You should **not** work them out.

(3 marks)


Select the **two** correct statements (**two** statements are incorrect).

- 2020 would not be reliable because it is outside the range of the data.
- 2020 would be reliable because it is within the data.
- 2016 would not be reliable because it is outside the range of the data.
- 2016 would be reliable because it is within the data.

- 3 A greengrocer keeps track of how many apples they sell every day of the week. This helps them ensure that they always have fresh stock.

| | |
|-----------|---|
| Monday |  |
| Tuesday |  |
| Wednesday |  |
| Thursday | |
| Friday | |





Key:

 represents 8 apples

- (a) On Thursday, the number of apples sold was 24.
Show this information on the pictogram.

(1 mark)

Select the correct answer.

| | | |
|--------------------------|----------|--|
| <input type="checkbox"/> | Thursday |  |
| <input type="checkbox"/> | Thursday |  |
| <input type="checkbox"/> | Thursday |  |
| <input type="checkbox"/> | Thursday |  |

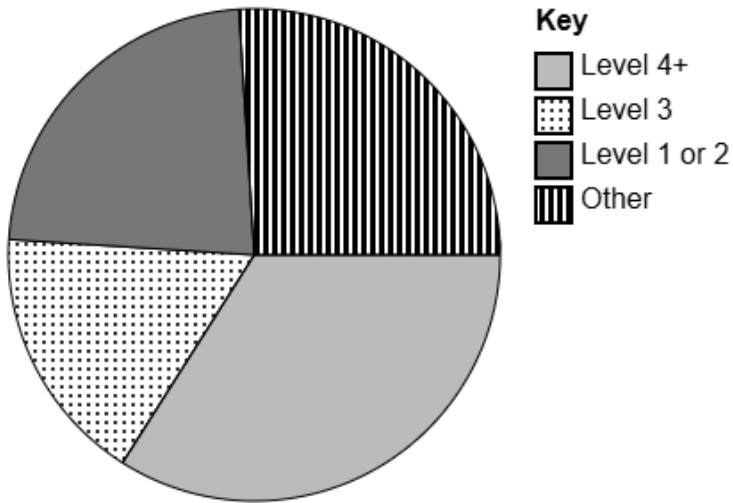
(b) Eduardo suggests redrawing the pictogram using a key with a whole-circle representing 5 apples. Explain why this key would **not** be suitable.

(1 mark)

Select **one** box.

- Tuesday shows 8 loaves of bread. This would be very difficult to show because 8 has a remainder 3 when divided by 5.
- If Eduardo uses 5 apples for the key this will lead to less sales being shown.
- The key must always be an even number.
- This would be much better because you can fit more circles on.

- 4 The accurately drawn pie chart shows information about the highest level of qualification in England and Wales in 2021.



- (a) Explain how you can tell that most people's highest level of qualification was level 4+ in England and Wales in 2021 using the pie chart.

(1 mark)

Select **one** box.

- 'Level 4+' is at the bottom of the pie chart.
- 'Level 4+' is the first value in the key.
- 'Level 4+' is the most positive response.
- 'Level 4+' has the largest sector.

(b) The population in the England and Wales in 2021 was estimated to be 60 million.

Calculate an estimate for the number of people in the UK in 2021 who's highest level of qualification was 'Level 4+'.

Round your answer to the nearest million.

(2 marks)

Start by measuring the angle with a protractor.

Next find the amount. Remember that there are 360 degrees in a circle.

_____ million

5 Emma owns a bookstore.

She wants to collect information about types of books liked by people in her city.

The following list gives the information she is going to collect about people's favourite books:

genre

average reading time

number of pages

(a) Select the information that is categorical data from the list.

(1 mark)

Categorical data can be grouped into non-overlapping categories.

Select **one** box.

average reading time

number of pages

genre

(b) Select the information that is discrete data from the list.

(1 mark)

Discrete data can take one of a set of certain values.

Select **one** box.

genre

number of pages

average reading time

(c) Emma would like to send a questionnaire to 100 of her customers.

She has a list of all 500 of her customers.

Explain how Emma can select a systematic sample of 100 people from her list of customers.

(2 marks)

Select the **two** correct statements (**two** statements are incorrect).

Calculate a random number between 1 and 5 to use a starting point.

Calculate a random number between 1 and 4 to use a starting point.

Select every 5th person.

Select every 4th person.

6 Chloe is a student and wants to study music preferences.
She would like to find out the most popular music genre in her school.

Chloe decides to do convenience sampling outside the school gate an hour after school and uses the data collection sheet shown here:

| Music genre | Tally |
|-------------|-------|
| Pop | |
| Rock | |
| Hip-hop | |
| Classical | |

(a) State the population for this study.

(1 mark)

Select **one** box.

- A selection of students in Chloe's school
- The students she asks
- All the students in Chloe's school
- All students in the UK

(b) Describe what is meant by a convenience sample.

(1 mark)

Select **one** box.

- Sampling people in proportion to a characteristic of the population
- Sampling people randomly
- Sampling every nth person from a list
- Sampling people who are easiest to reach

(c) Give **one** disadvantage of convenience sampling.

(1 mark)

Select **one** box.

- It takes too much time
- It may be biased
- It requires a large population
- It is expensive to do

(d) Discuss whether this data collection sheet is appropriate for Chloe to collect the data.

(2 marks)

Select the **two** correct statements (**two** statements are incorrect).

- There is no 'other' option.
- Chloe may not ask all the students.
- It will make collecting data very difficult.
- Chloe will be able to put the data into a graph very easily.

(e) After collecting the data, Chloe would like to display the data in a diagram.

Discuss whether or not a stem and leaf diagram would be suitable.

(2 marks)

Select the **two** correct statements (**two** statements are incorrect).

- because the data is quantitative.
- because the data is qualitative.
- It is not suitable
- It is suitable

7 A moderator wants to get feedback on a discussion on careers they ran last week.

289 people attended the discussion.

The moderator plans to give a questionnaire to a sample of 50 of the people who attended the discussion.

One of the questions on the questionnaire is:

To what extent do you agree with the statement, the principal communicated well?

Use a scale from 0 to 5, where 0 means **strongly disagree** and 5 means **strongly agree**.

(a) Select the word from the list that best describes the type of data collected by this question.

(1 mark)

Select **one** box.

bivariate

continuous

ordinal

(b) Explain how the moderator could use a list of random numbers to choose a simple random sample of 50 people who attended the discussion.

(3 marks)

Select the **three** correct statements (**three** statements are incorrect).

Select 50 numbers ignoring any numbers that have repeated or are out of range.

Select the people who were assigned the highest number.

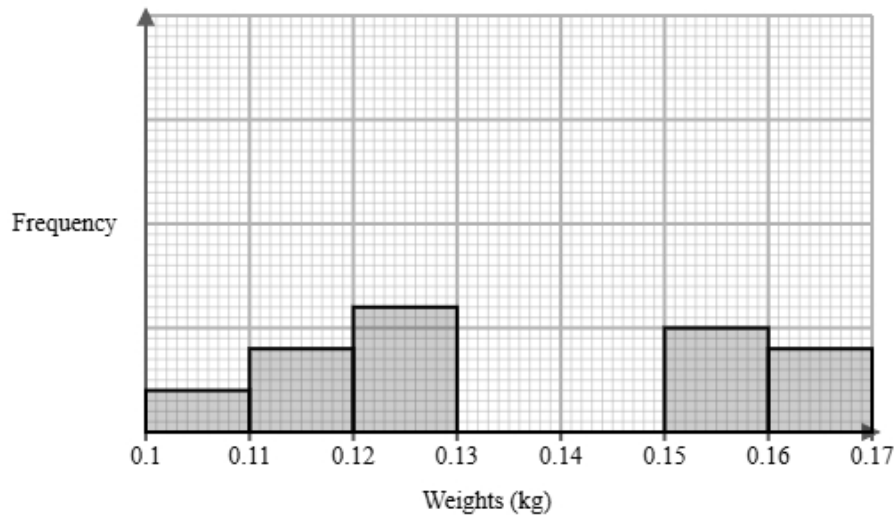
Select 50 people from a hat.

Select people with the corresponding number from random number list.

Assign a number to all the people.

List the people in alphabetical order based on their surname.

- 8 A greenhouse worker measured the weights of Roma tomatoes and plum tomatoes in the greenhouse. They recorded the weights after 3 months. The incomplete histogram and grouped frequency table give information about the weights of Roma tomatoes in the greenhouse.



| Weights w (kg) | Frequency |
|----------------------|-----------|
| $0.10 < w \leq 0.11$ | 2 |
| $0.11 < w \leq 0.12$ | 4 |
| $0.12 < w \leq 0.13$ | 6 |
| $0.13 < w \leq 0.14$ | 11 |
| $0.14 < w \leq 0.15$ | 9 |
| $0.15 < w \leq 0.16$ | |
| $0.16 < w \leq 0.17$ | |

- (a) Use the information in the histogram to complete the table.

(2 marks)

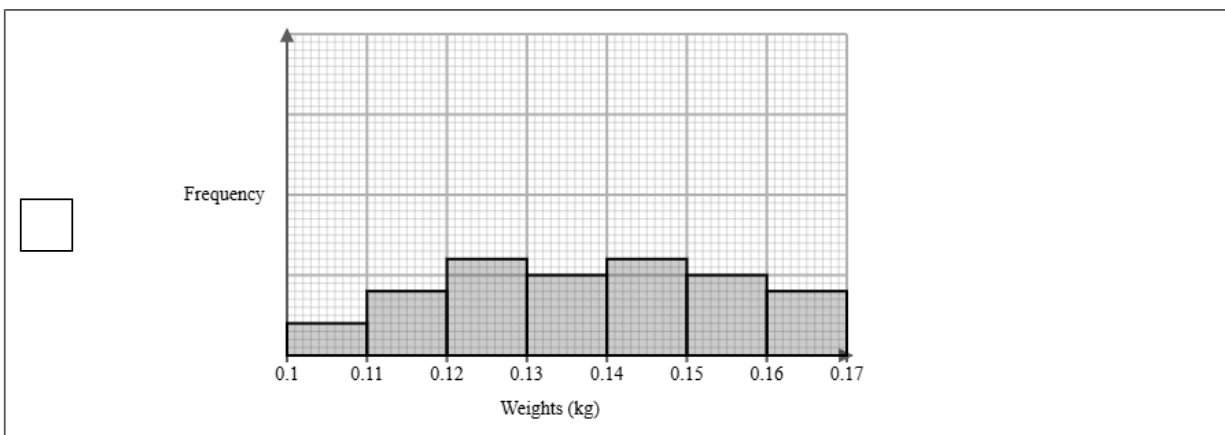
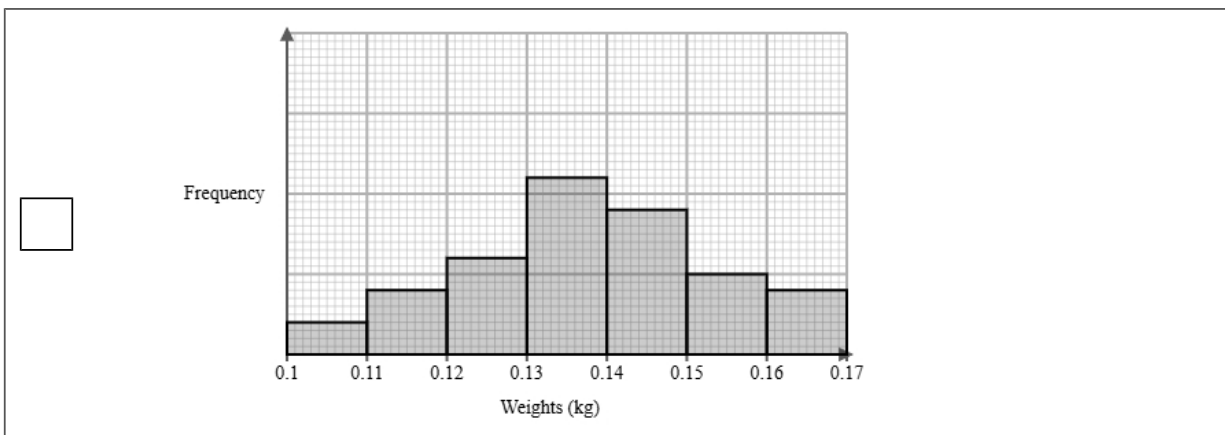
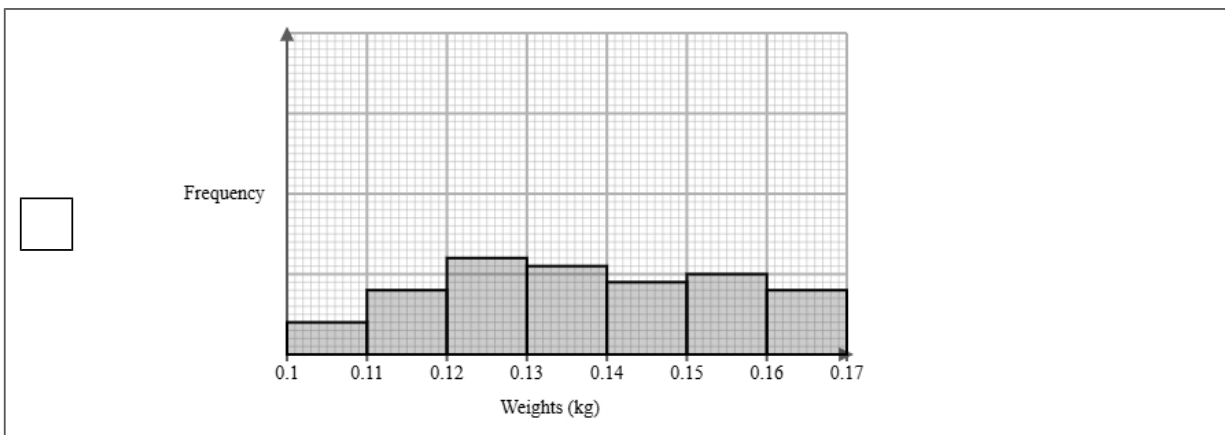
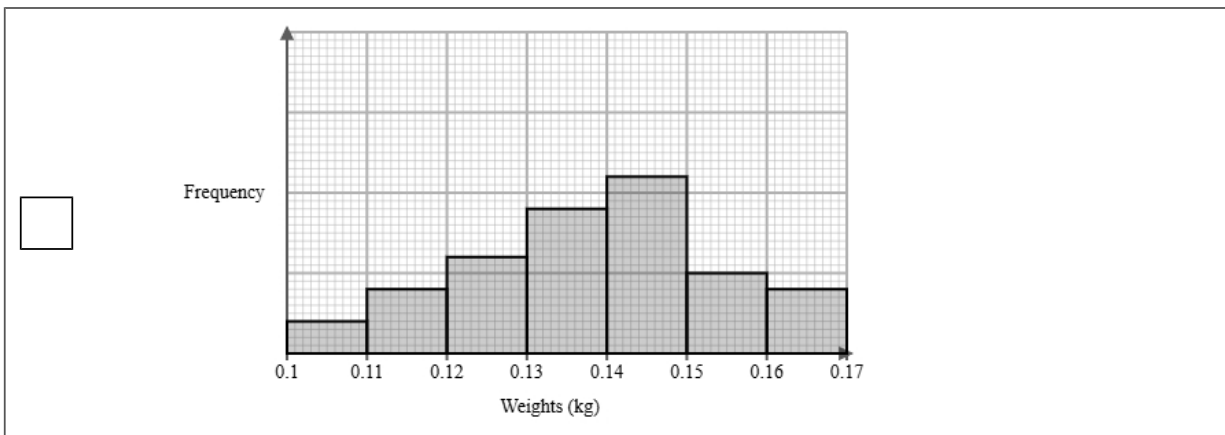
Find the height of $0.10 < w \leq 0.11$ on the graph and compare it with the frequency to find the scale
Use this scale to find the missing frequencies

| Weights w (kg) | Frequency |
|----------------------|-----------|
| $0.10 < w \leq 0.11$ | 2 |
| $0.11 < w \leq 0.12$ | 4 |
| $0.12 < w \leq 0.13$ | 6 |
| $0.13 < w \leq 0.14$ | 11 |
| $0.14 < w \leq 0.15$ | 9 |
| $0.15 < w \leq 0.16$ | _____ |
| $0.16 < w \leq 0.17$ | _____ |

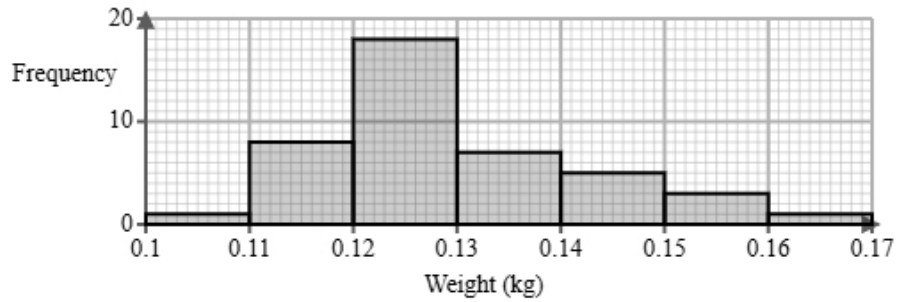
(b) Use the information in the table to complete the histogram.

(2 marks)

Select the correct answer.



(c) The histogram below shows data on the weights of plum tomatoes after 3 months.



Identify and interpret the type of skew shown in the histogram for plum tomatoes.

(2 marks)

Select the **two** correct statements (**two** statements are incorrect).

- The mean weight is smaller than the median.
- The mean weight is larger than the median.
- Negative skew.
- Positive skew.

- (d) The greenhouse worker also measured the weights of Roma tomatoes and plum tomatoes in the polytunnel. The grouped frequency table below gives information about the weights of Roma tomatoes and plum tomatoes in the polytunnel.

| Weights w (grams) | Frequency | |
|---------------------|---------------|---------------|
| | Roma tomatoes | plum tomatoes |
| $20 < w \leq 60$ | 4 | 5 |
| $60 < w \leq 100$ | 15 | 18 |
| $100 < w \leq 140$ | 13 | 6 |
| $140 < w \leq 180$ | 4 | 4 |
| Total | 36 | 33 |

The estimate of the mean for Roma tomatoes is calculated to be 98.9 g to 1 decimal place.

David uses the estimate of the means for Roma tomatoes and plum tomatoes to conclude that the Roma tomatoes weigh less than plum tomatoes.

Discuss whether or not David is correct and give **one** limitation of your conclusion.

You must show your working.

(4 marks)

Add a midpoint and fw column onto the table

Add up the frequencies and the fw columns

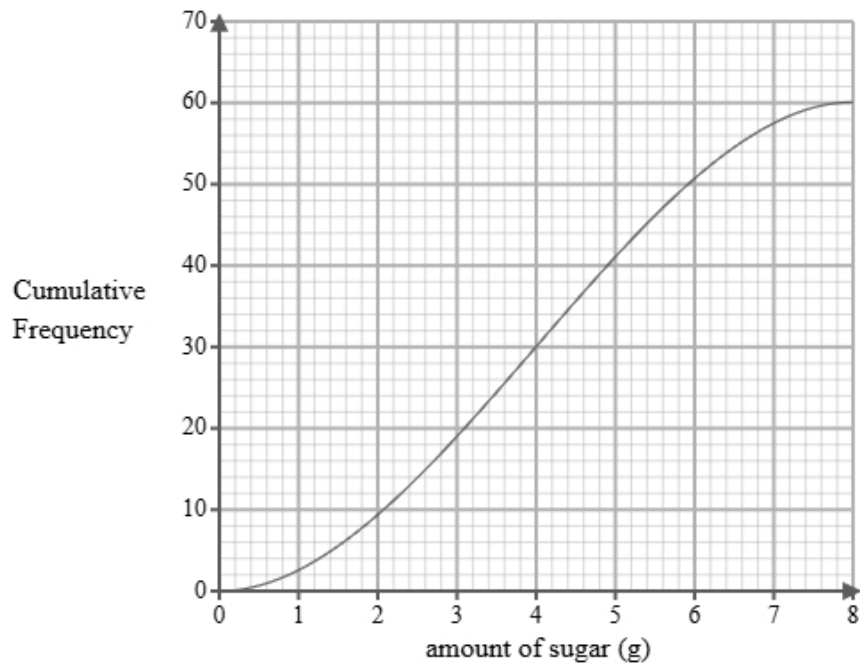
Find the mean by dividing $\sum fw$ and $\sum f$

Mean of plum tomatoes = _____

Select the **two** correct statements (**two** statements are incorrect).

- David is incorrect.
- David is correct.
- We cannot be sure because both means are in the same class interval.
- These are only estimates.

- 9 A researcher measures the amount of sugar, in grams, in 60 different breakfast cereals.
A cumulative frequency diagram is drawn from the data.



Complete the table below from the cumulative frequency diagram.

| Lower quartile | Median | Upper quartile |
|----------------|--------|----------------|
| | | |

(2 marks)

Select the correct answer.

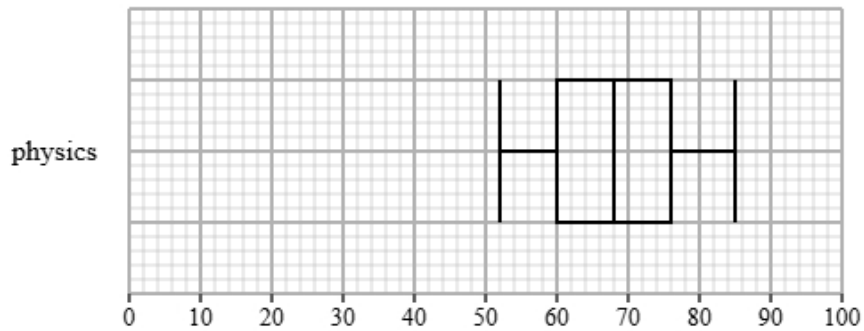
| | | | |
|--------------------------|----------------|--------|----------------|
| <input type="checkbox"/> | Lower quartile | Median | Upper quartile |
| | 1.4 | 4.7 | 6.8 |

| | | | |
|--------------------------|----------------|--------|----------------|
| <input type="checkbox"/> | Lower quartile | Median | Upper quartile |
| | 3.8 | 4 | 4.7 |

| | | | |
|--------------------------|----------------|--------|----------------|
| <input type="checkbox"/> | Lower quartile | Median | Upper quartile |
| | 1.9 | 4.7 | 6.1 |

| | | | |
|--------------------------|----------------|--------|----------------|
| <input type="checkbox"/> | Lower quartile | Median | Upper quartile |
| | 2.6 | 4 | 5.4 |

- 10** Olivia collected the marks for physics and business students from a statistics exam.
Both groups took the same exam.
The box plot presents data on the marks for the physics students.



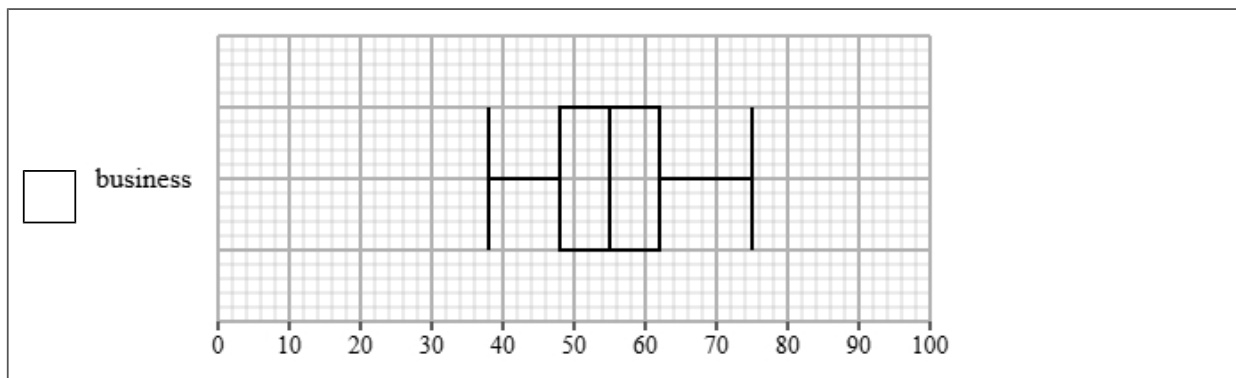
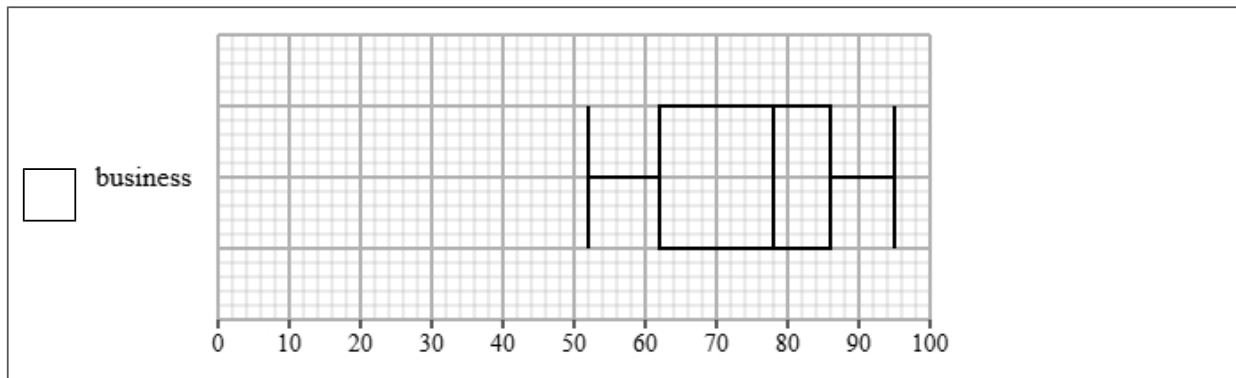
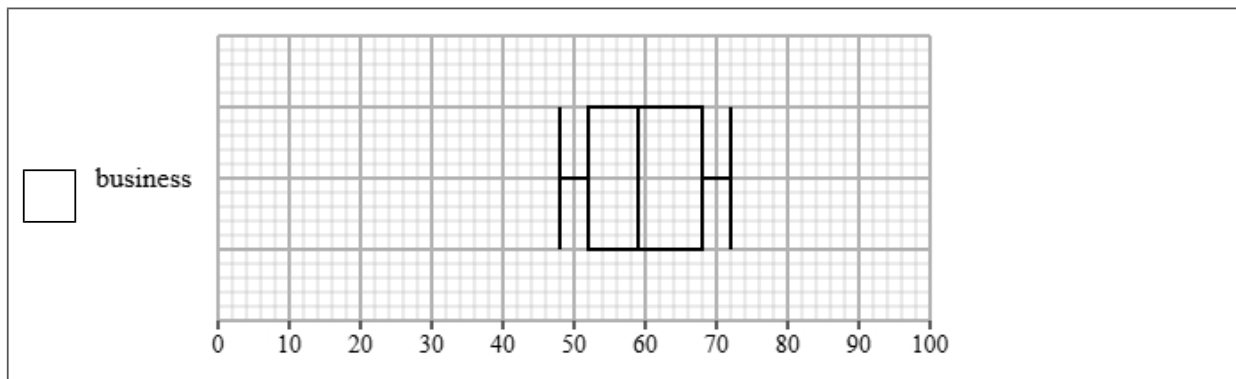
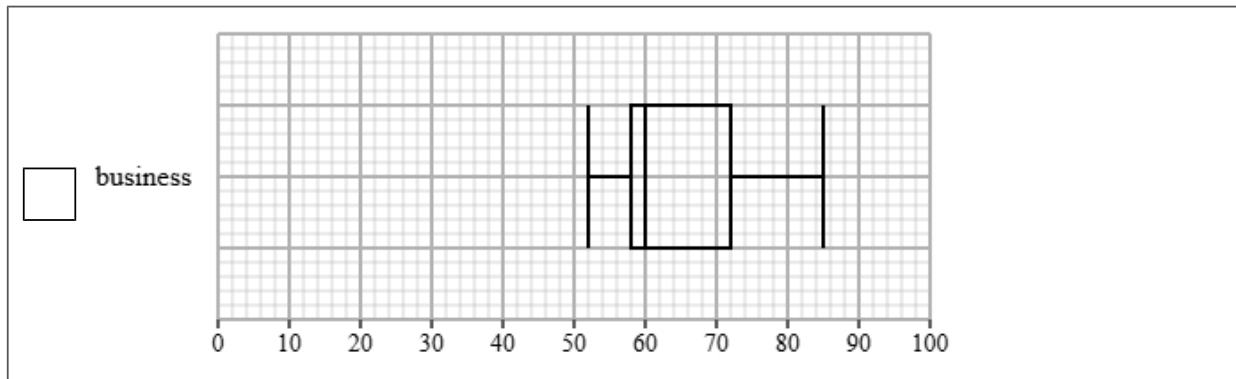
The table gives information about the marks for the business students.

| Least tall | Lower quartile | Median | Upper quartile | Most tall |
|------------|----------------|--------|----------------|-----------|
| 52 | 58 | 60 | 72 | 85 |

- (a) Draw a box plot for the marks for the business students.

(2 marks)

Select the correct answer.



(b) Compare the two distributions of marks.

Give three comparisons and interpret one of these comparisons.

(4 marks)

Select **one** box.

- The median is bigger.
- The median marks for physics students is greater than business students.
- The median marks for physics and business students are equal.
- The median marks for physics students is lower than business students.

Select **one** box.

- The IQR is bigger.
- The IQR for the marks of the physics and business students are equal.
- The IQR for the marks of the physics students is greater than business students.
- The IQR for the marks of the physics students is lower than business students.

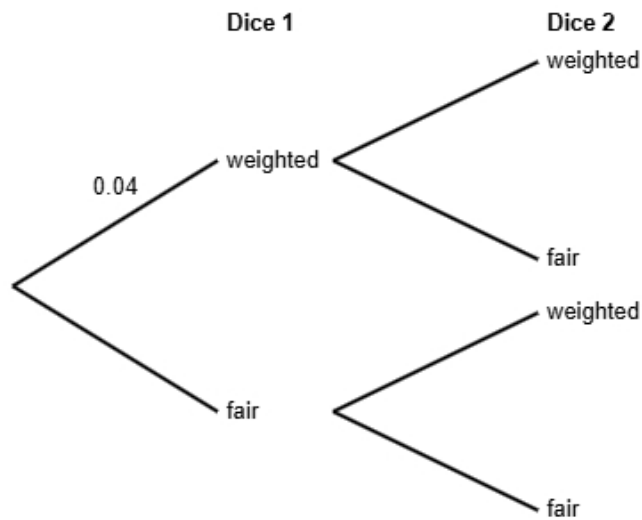
Select **one** box.

- The skews for the marks of the physics and business students are both positive.
- The skews for the marks of the physics and business students are both symmetrical.
- The skew for the marks of the physics students is symmetrical and the skew for the business students is negative.
- The skew for the marks of the physics students is symmetrical and the skew for the business students is positive.

Select **one** box.

- The marks for the physics students are less spread out than the business students.
- The physics students are more skewed than business students.
- The physics students are on average did better on the statistics test than the business students.
- The physics students are on average did worse on the statistics test than the business students.

- 11** A company found that 4% of its dice are slightly weighted and do not roll fairly.
The rest of the dice are perfectly balanced.
Jack picks two dice from a random batch.
He does not know if each die is weighted or fair.



- (a) Complete the probability tree diagram.

(2 marks)

The branches for each stage must add up to 1.
Each test is independent so will have the same probabilities.

(b) Find the probability that both of Jack's dice are fair.

(2 marks)

You will need to find $P(\text{fair})$ AND $P(\text{fair})$.

Remember, AND means \times in probability.

(c) Jack states that the probability that exactly one dice is weighted is less than 8%

Find out whether or not Jack is correct.

(3 marks)

Find the probability of exactly one dice is weighted (there are two outcomes on the tree diagram).

Select **one** box.

- The probability that exactly one dice is weighted is more than 8%, so Jack is not correct.
- The probability that exactly one dice is weighted is less than 8%, so Jack is not correct.
- The probability that exactly one dice is weighted is more than 8%, so Jack is correct.
- The probability that exactly one dice is weighted is less than 8%, so Jack is correct.

12 The table shows information about laptops for sale in London.

| screen size (inches) | number of laptops |
|----------------------|-------------------|
| 13 | 660 |
| 15 | 540 |
| 17 | 240 |
| 19 | 150 |
| 20 or more | 210 |
| Total | 1800 |

A researcher wants to investigate the price of these laptops and takes a stratified sample of 60 laptops according to the screen size (inches).

(a) The researcher says the mode of the screen size (inches) for these laptops is 13.

Explain how the researcher knows this.

(1 mark)

Select **one** box.

- 13 is the middle number.
- 13 inches has the highest frequency.
- 13 is the difference between the largest and smallest number.
- 13 laptops has the highest frequency.

(b) Work out the number of laptops in the sample for each screen size.

| screen size (inches) | number of laptops in the sample |
|----------------------|---------------------------------|
| 13 | |
| 15 | |
| 17 | |
| 19 | |
| 20 or more | |

(3 marks)

Find the divisor for the stratified sample: $\frac{\text{total}}{\text{sample size}}$

Divide each frequency by this number to find the required sample in each group

(c) Describe how the 60 laptops in the sample should be selected.

(3 marks)

Select the **three** correct statements (**three** statements are incorrect).

- Ensure that all 1800 laptops are included in the sample.
- Complete two of the strata.
- Select the first 60 laptops.
- Generate random numbers, remove repeats or numbers out of range.
- Number each of the laptops, and then use the random numbers to select the required amount of laptops.
- Use a sampling frame for each strata.

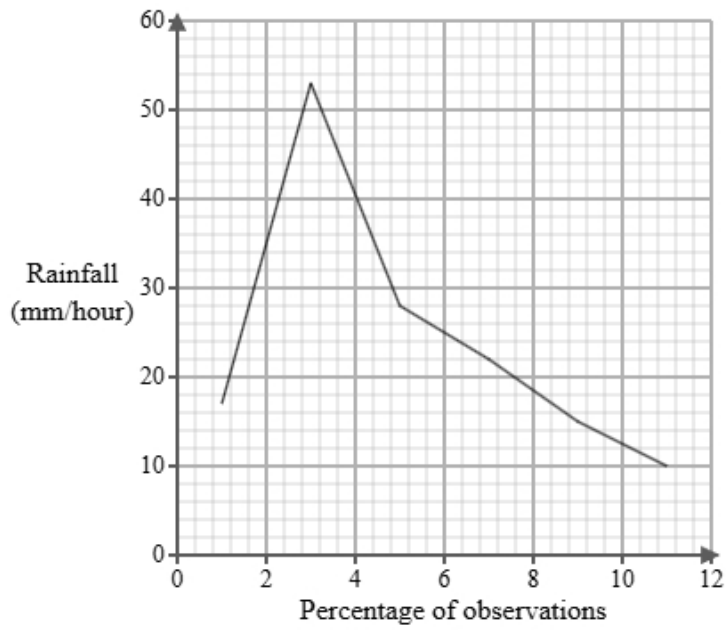
- 13 Priya works for a meteorological office. She has been tasked with investigating rainfall intensity in June. Below is a section of the spreadsheet she used to record her findings.

| Rainfall (mm/hour) | Percentage of observations |
|--------------------|----------------------------|
| $0 < r \leq 2$ | 5 |
| $2 < r \leq 4$ | eight |
| $4 < r \leq 6$ | 7 |
| $6 < r \leq 8$ | 120 |
| $8 < r \leq 10$ | 47 |
| $10 < r \leq 12$ | 13 |
| Total | 100 |

Priya cleans the data to create the table below.

| Rainfall (mm/hour) | Percentage of observations |
|--------------------|----------------------------|
| $0 < r \leq 2$ | 5 |
| $2 < r \leq 4$ | 8 |
| $4 < r \leq 6$ | 7 |
| $6 < r \leq 8$ | 20 |
| $8 < r \leq 10$ | 47 |
| $10 < r \leq 12$ | 13 |
| Total | 100 |

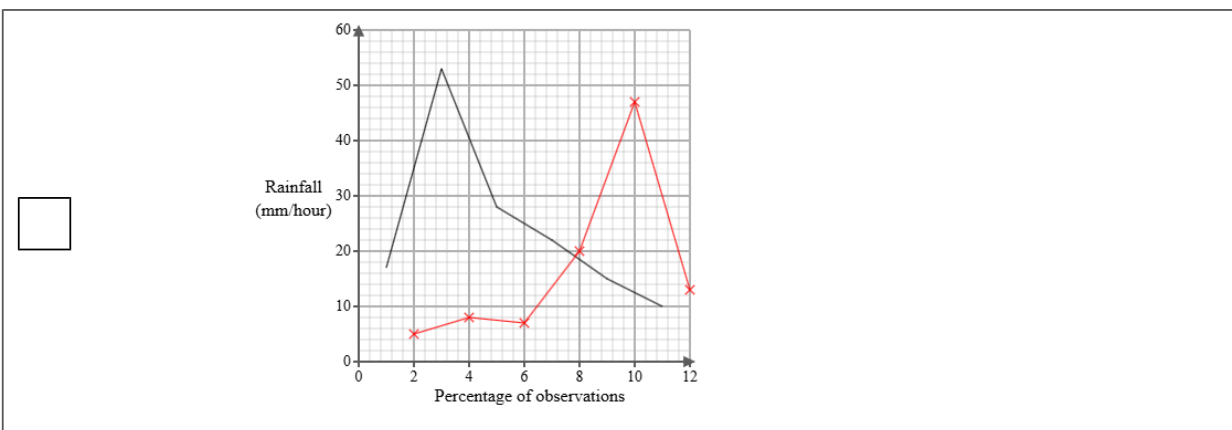
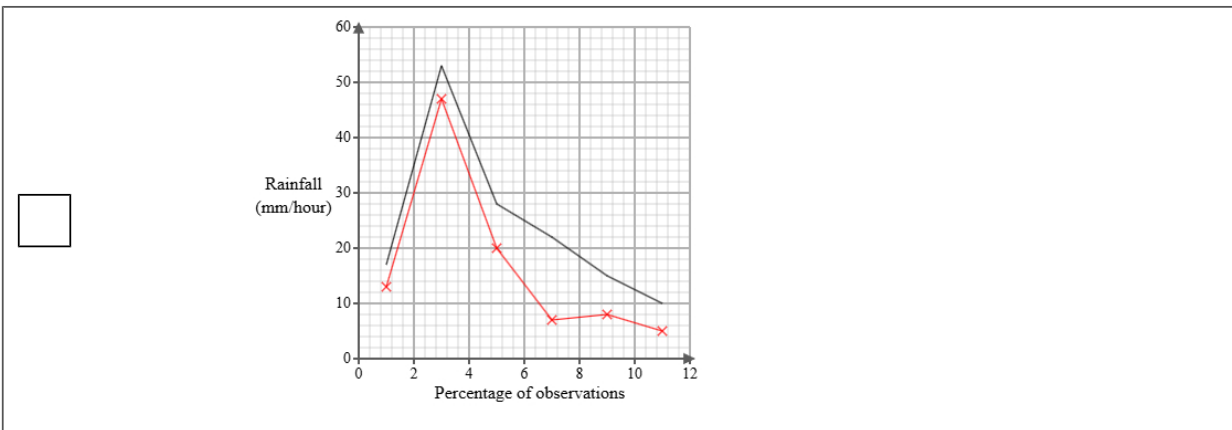
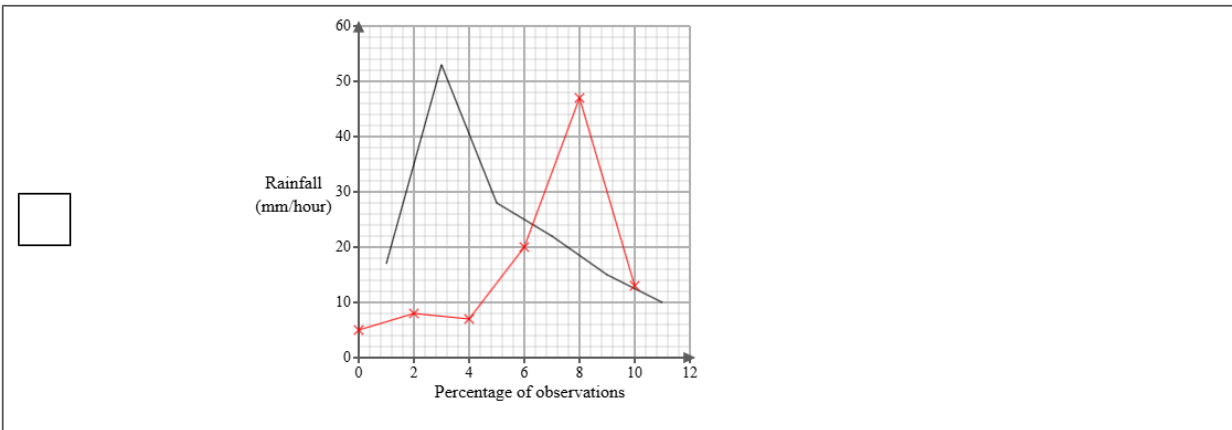
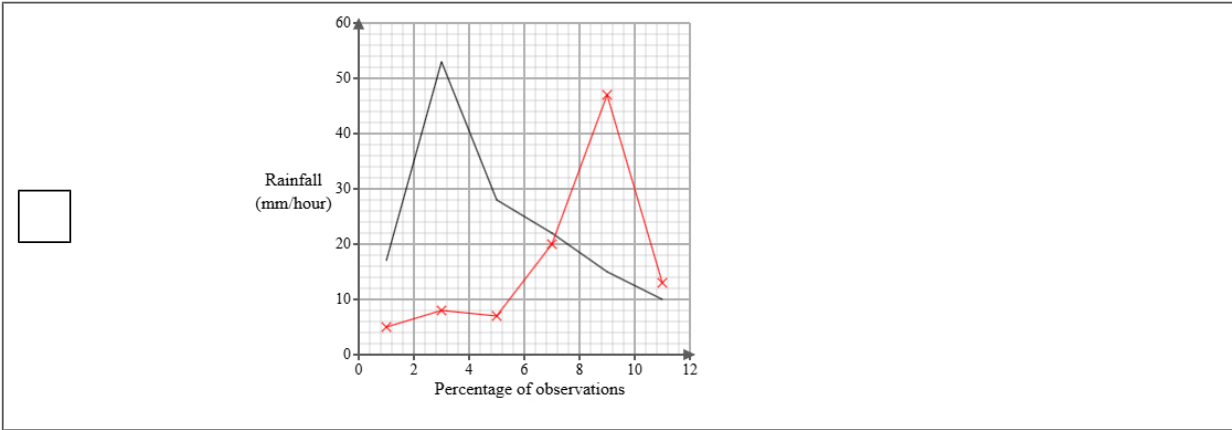
A frequency polygon has been drawn for rainfall intensity in August.



- i) On the same graph, draw the frequency polygon for rainfall intensity in June.
- ii) Using the two frequency polygons, compare the skew of the distributions and explain what your comparison means in context.

(4 marks)

Select the correct answer.



Select the **two** correct statements (**four** statements are incorrect).

- The distribution of rainfall intensity in August is negatively skewed whereas the distribution of rainfall intensity in June is positively skewed.
- This means that in August the rainfall is mainly at the upper end of the distribution and in June the rainfall is mainly at the lower end of the distribution.
- This means that in August the rainfall is mainly at the lower end of the distribution and in June the rainfall is mainly at the upper end of the distribution.
- The distribution of rainfall intensity in August is symmetrical whereas the distribution of rainfall intensity in June is positively skewed.
- This means that in August the rainfall is equally spread out on either side of the median and in June the rainfall is mainly at the upper end of the distribution.
- The distribution of rainfall intensity in August is positively skewed whereas the distribution of rainfall intensity in June is negatively skewed.

- 14 The table shows information about the retail price index (RPI) and NHS Prescription Charges (£) in England for Jan 1995, Jan 2005 and Jan 2015.

| | Jan 1995 | Jan 2005 | Jan 2015 |
|------------------------------|----------|----------|----------|
| retail price index | 100 | 121 | 173 |
| NHS Prescription Charges (£) | 5.25 | 6.5 | 8.05 |

Describe how the increase in NHS Prescription Charges (£) compares with the RPI over the ten years to Jan 2005 and over the twenty years to Jan 2015.

(5 marks)

Select the **four** correct statements (**four** statements are incorrect).

$\frac{6.5}{121} \times 100 = 5$ (nearest integer)

$\frac{8.05}{173} \times 100 = 5$ (nearest integer)

Between Jan 1995 and Jan 2015 the change in price was less than the RPI.

Between Jan 1995 and Jan 2015 the change in price was more than the RPI.

Between Jan 1995 and Jan 2005 the change in price was less than the RPI.

$\frac{6.5}{5.25} \times 100 = 124$ (nearest integer)

$\frac{8.05}{5.25} \times 100 = 153$ (nearest integer)

Between Jan 1995 and Jan 2005 the change in price was more than the RPI.

15 Aisha is investigating how the age in years, x , affects the resale price (£), y for two types of smartphones, model X and model Y.

She found ten smartphones of each type and recorded their age and resale price and plotted each on scatter diagrams.

She then drew a line of best fit on each diagram and found the gradient and y-intercept of each line.

Here are the results:

| Model | Gradient of line of best fit | y-intercept of line of best fit |
|-------|------------------------------|---------------------------------|
| X | -80 | 900 |
| Y | -65 | 1100 |

Interpret and compare these results in context.

(5 marks)

Select the **five** correct statements (**five** statements are incorrect).

- Model X reduces in resale price by £80 per year.
- Model X has a greater initial resale price.
- Model X changes in resale price by £900 per year.
- Model Y has a greater initial resale price.
- Model Y changes in resale price by £1100 per year.
- Both smartphones decrease in resale price as the age increase.
- Model X reduces in resale price less than Model Y.
- Model X reduces in resale price more per year than Model Y.
- Model Y reduces in resale price by £65 per year.
- Both smartphones increase in resale price as the age increase.