

Please check the examination details below before entering your candidate information

|                      |                      |                      |                      |                      |                      |                      |                      |                      |  |
|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|--|
| Candidate surname    |                      |                      |                      |                      | Other names          |                      |                      |                      |  |
| Centre Number        |                      |                      |                      |                      | Candidate Number     |                      |                      |                      |  |
| <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> | <input type="text"/> |  |

**Pearson Edexcel Level 1/Level 2 GCSE (9–1)**

**Friday 13 June 2025**

Afternoon (Time: 1 hour 30 minutes)

**Paper reference** **1ST0/2H**

**Statistics**

**PAPER 2**

**Higher Tier**

**You must have:**  
Ruler graduated in centimetres and millimetres, protractor,  
pair of compasses, pen, HB pencil, eraser, scientific calculator.

Total Marks

## Instructions

- Use **black** ink or ball-point pen.
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions.
- Answer the questions in the spaces provided  
– *there may be more space than you need.*
- Scientific calculators may be used.
- You must **show all your working out** with **your answer clearly identified** at the **end of your solution**.



## Information

- The total mark for this paper is 80.
- The marks for **each** question are shown in brackets  
– *use this as a guide as to how much time to spend on each question.*

## Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.

Turn over ►

P76387A

©2025 Pearson Education Ltd.  
Y:1/1/1/1/1/1/



  
**Pearson**

## Higher Tier Formulae

You must not write on this page.

Anything you write on this page will gain NO credit.

$$\text{Skew} = \frac{3(\text{mean} - \text{median})}{\text{standard deviation}}$$

$$\text{Standard deviation} = \sqrt{\frac{1}{n} \sum (x - \bar{x})^2}$$

*An alternative formula for standard deviation is*

$$\text{standard deviation} = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

Spearman's rank correlation coefficient

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

$$\text{Rates of change (e.g. Crude birth rate} = \frac{\text{number of births} \times 1000}{\text{total population}})$$

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



Answer ALL questions.

Write your answers in the spaces provided.

You must write down all the stages in your working.

1 Elijah has a coin.

Elijah says,

“I predict that the probability of flipping a head on this coin is 50%”

(a) Name the type of probability that Elijah has given.

(1)

Peter flips Elijah’s coin 300 times.  
Here are his results.

| Result | Frequency |
|--------|-----------|
| Heads  | 138       |
| Tails  | 162       |

Peter says,

“The probability of flipping a head on this coin is 46%”

(b) Comment on the differences between these two probabilities.

(2)

(Total for Question 1 is 3 marks)



- 2 Elise wants to investigate the relationship between the size of mammals and the size of the fruits and seeds that these mammals eat. This is her plan.

**Hypothesis**

- Can body mass and skull shape predict the type of seed and fruit eaten by different mammal species?

**Data collection**

- Skull measurements will be collected from museums for two adult males and two adult females for each mammal species.
- The data will be collected for 100 mammals. Any missing data will be sourced from the internet.
- The size of the seeds and fruits that the 100 mammals eat will be recorded using information collected from existing databases.

Discuss the appropriateness of Elise's hypothesis and her data collection strategy.



**DO NOT WRITE IN THIS AREA**

**(Total for Question 2 is 5 marks)**



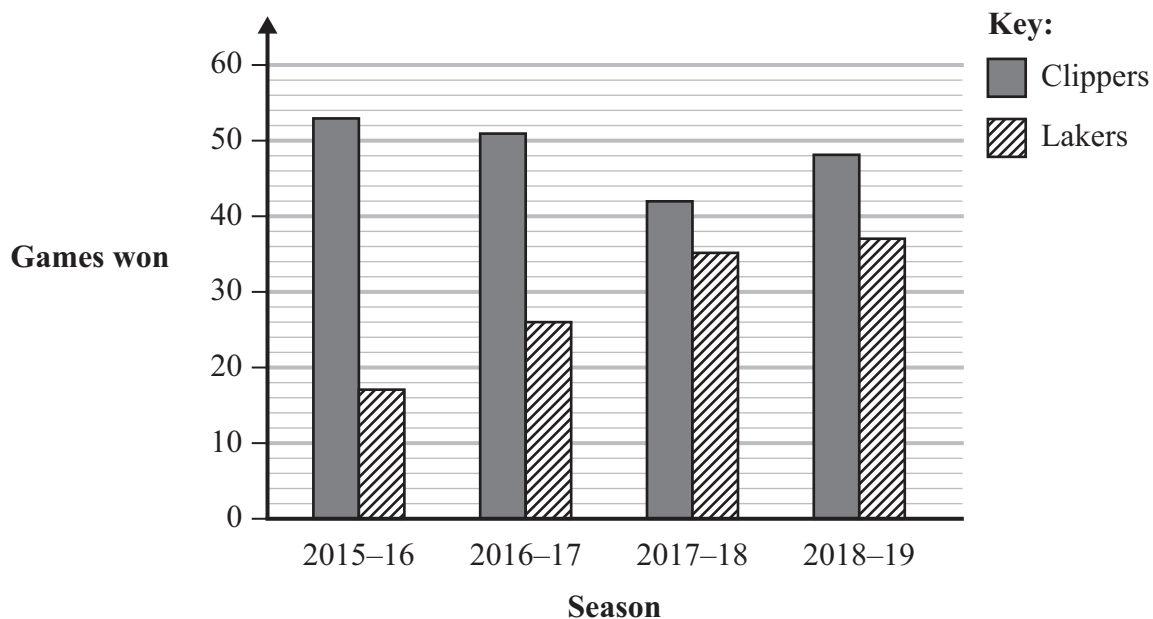
- 3 Divyesh is investigating the number of games won by two Los Angeles basketball teams, the Clippers and the Lakers. He collects some data from the NBA website.

(a) Complete the sentence with an appropriate statistical word.

Divyesh has collected data from the internet and so this is ..... data.

(1)

The information in the bar chart is used to compare the numbers of games won by the Clippers and the Lakers in four basketball seasons from 2015 to 2019



(Source: [www.nba.com](http://www.nba.com))

Using the information in the bar chart to compare the number of games won by each team over the four years,

(b) write down one comparison between the two teams,

(1)

(c) describe the trend in the number of games won by the Lakers over the four seasons.

(1)



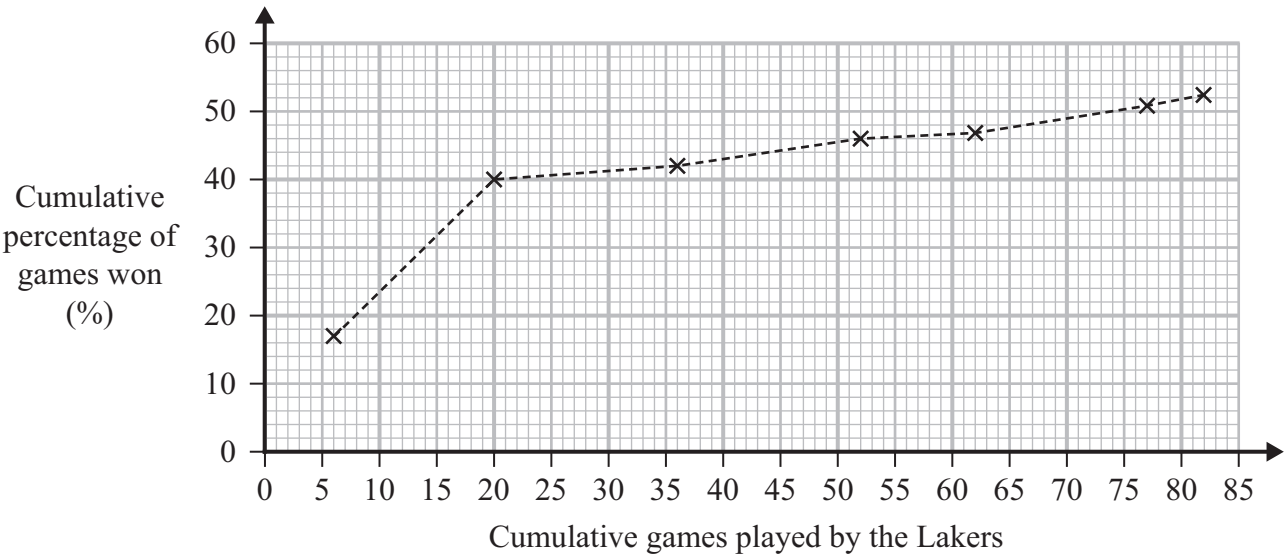
The table contains information about the number of games won by the Clippers in the 2022–23 season.

|                                        | Oct 22 | Nov 22 | Dec 22 | Jan 23 | Feb 23 | Mar 23 | April 23 |
|----------------------------------------|--------|--------|--------|--------|--------|--------|----------|
| Cumulative games played                | 7      | 23     | 38     | 54     | 64     | 78     | 82       |
| Cumulative games won                   | 3      | 13     | 21     | 29     | 33     | 41     | 44       |
| Cumulative percentage of games won (%) | 43     | .....  | 55     | 54     | 52     | .....  | 54       |

- (d) Complete the table.  
Give your answers correct to the nearest whole number.

(2)

The line graph shows the cumulative games played and the cumulative percentage of games won by the Lakers in the same season.



Each team played 82 games in the 2022–23 season.

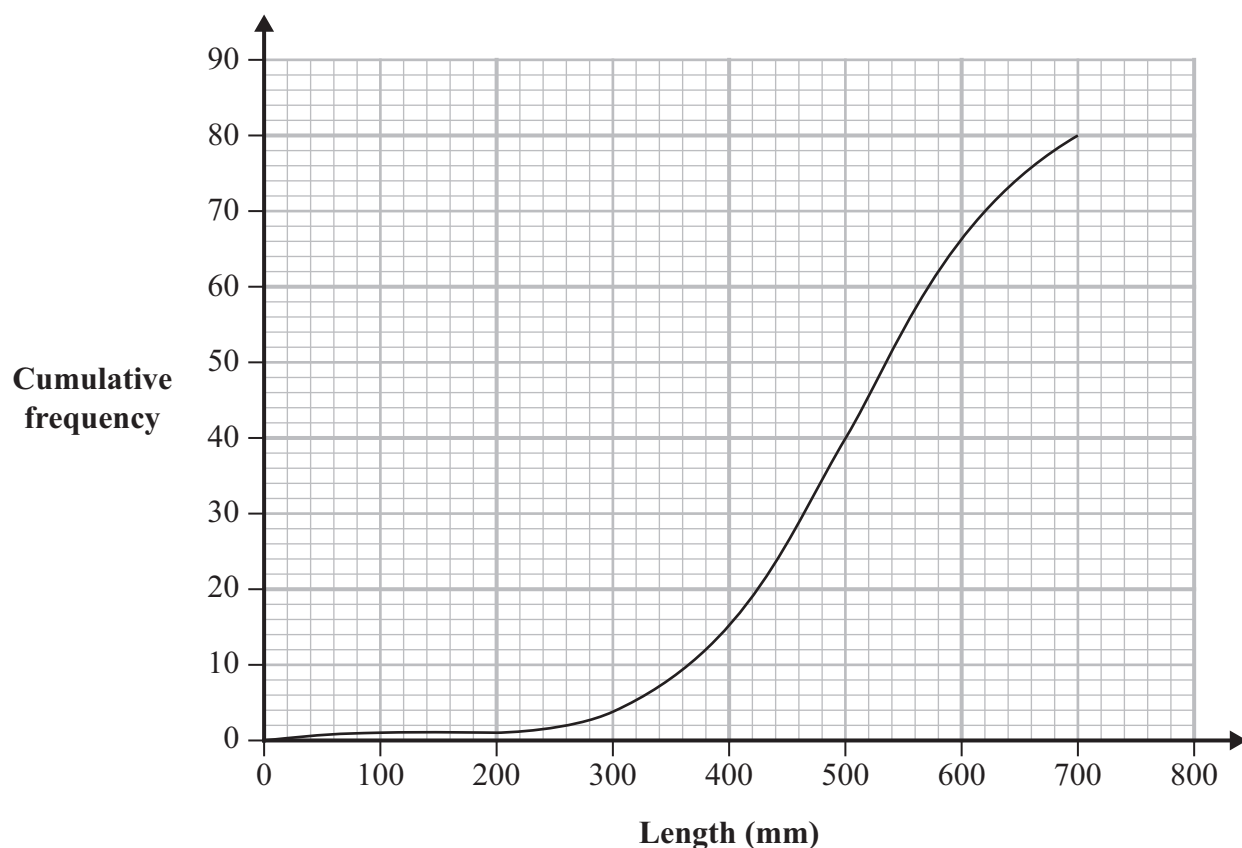
- (e) Which team won the most games by the end of the 2022–23 season?  
You should justify your answer.

(1)

(Total for Question 3 is 6 marks)



- 4 The cumulative frequency diagram gives information about the lengths, in mm, of a sample of 80 pike, caught in the Yukon River in Alaska. The pike were put back into the river.



(Source: [www.dataone.org](http://www.dataone.org))

- (a) Estimate how many of these pike had a length of 440 mm or more.

(2)

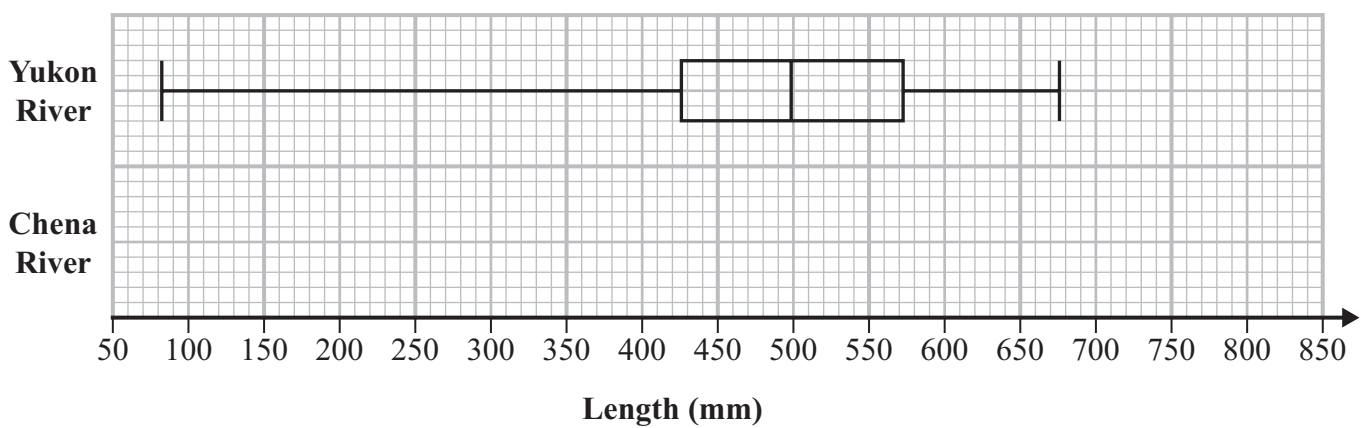
These data were collected using opportunity sampling.

- (b) State an advantage of using opportunity sampling.

(1)



The box plot below shows information about the lengths of the pike sampled from the Yukon River.



Here are the summary statistics, in mm, for a sample of pike caught in the Chena River in Alaska.

| Median | Minimum | Maximum | Q1  | Q3  |
|--------|---------|---------|-----|-----|
| 476    | 210     | 825     | 390 | 556 |

(c) On the grid above, draw a box plot for the lengths of the pike caught in the Chena River.

(2)

(d) Give **three** comparisons between the distribution of the lengths of pike from the Yukon River and the distribution of the lengths of pike from the Chena River. Interpret **one** of these comparisons in context.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(4)

(Total for Question 4 is 9 marks)

DO NOT WRITE IN THIS AREA

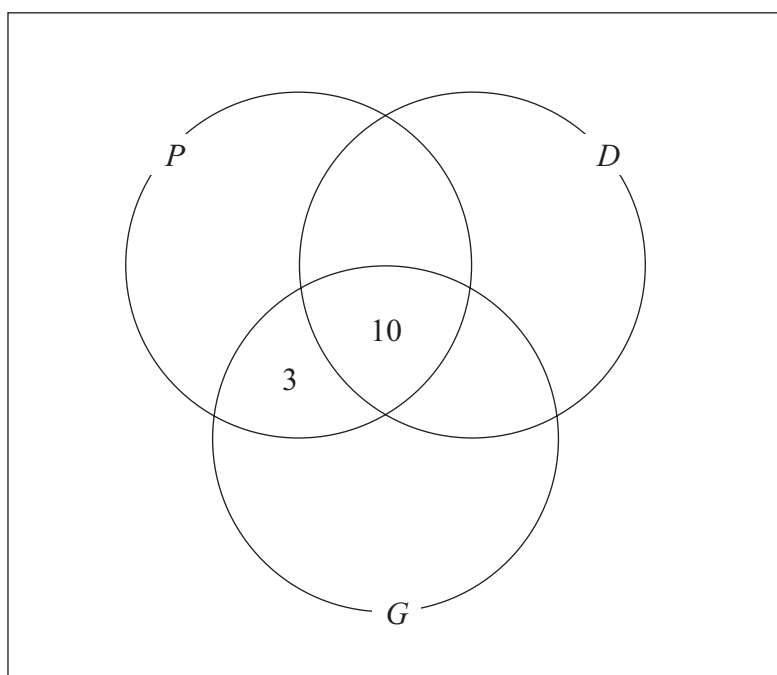
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

- 5 120 musicians were asked what instruments they play.  
The most popular instruments were piano ( $P$ ), drums ( $D$ ) and guitar ( $G$ ).

10 musicians play all three instruments  
13 musicians play both piano and guitar  
17 musicians play both piano and drums  
23 musicians play both guitar and drums  
50 musicians play the piano  
40 musicians play the drums  
35 musicians play the guitar

- (a) Complete the Venn diagram using this information.



(3)

One musician is to be chosen at random.

- (b) Find  $P(G)$

(1)

- (c) Find  $P(G|D)$

(2)



Toga claims that playing the guitar and playing the drums are independent events.

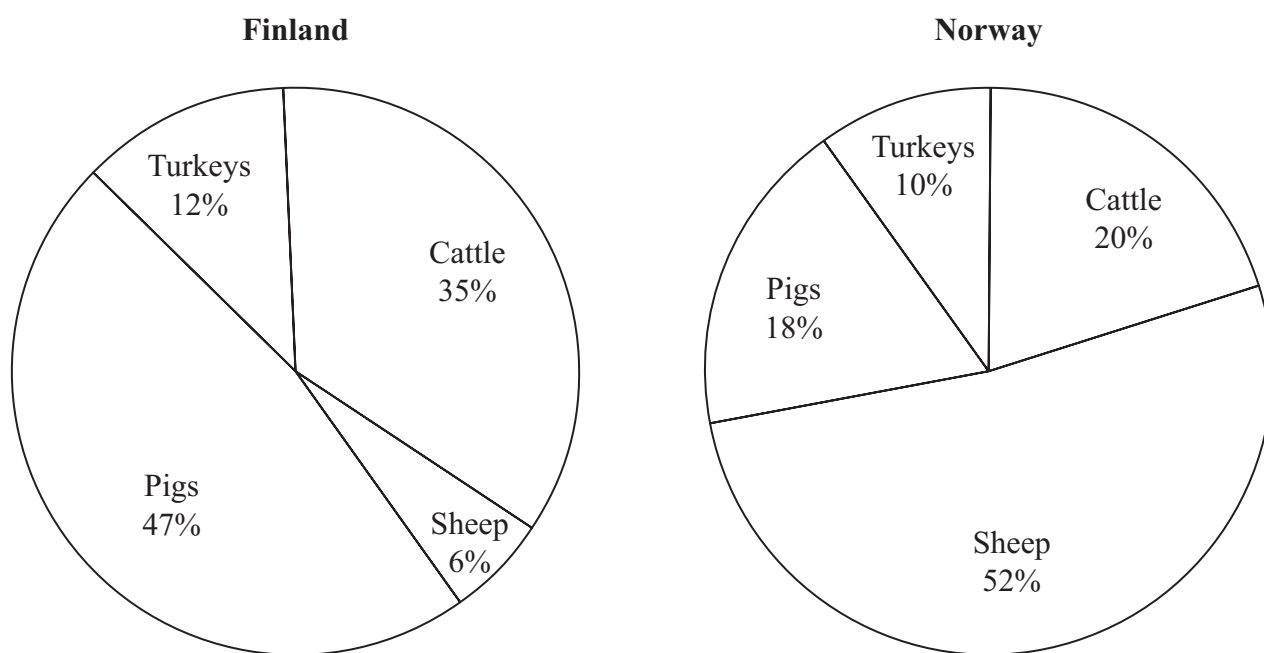
- (d) Do you agree with Toga?  
Give a reason for your answer.

(2)

(Total for Question 5 is 8 marks)



- 6 The pie charts show information about the proportion of livestock in 2021 for Finland and for Norway.



(Source: [www.fao.org](http://www.fao.org))

Marcus concludes that there are more turkeys in Finland than there are in Norway.

- (a) Explain why Marcus might be wrong.

(1)

The pie chart for Finland represents 2 341 780 livestock.

- (b) Work out the number of cattle represented in the pie chart for Finland.

(2)

In 2021, the total number of cattle, sheep, turkeys and pigs in Denmark was 14 888 487

Marcus plans to draw comparative pie charts for Finland and Denmark.

The pie chart for Finland will have a radius of 5 cm.

- (c) Calculate the radius he should use for the pie chart for Denmark.  
Give your answer correct to one decimal place.

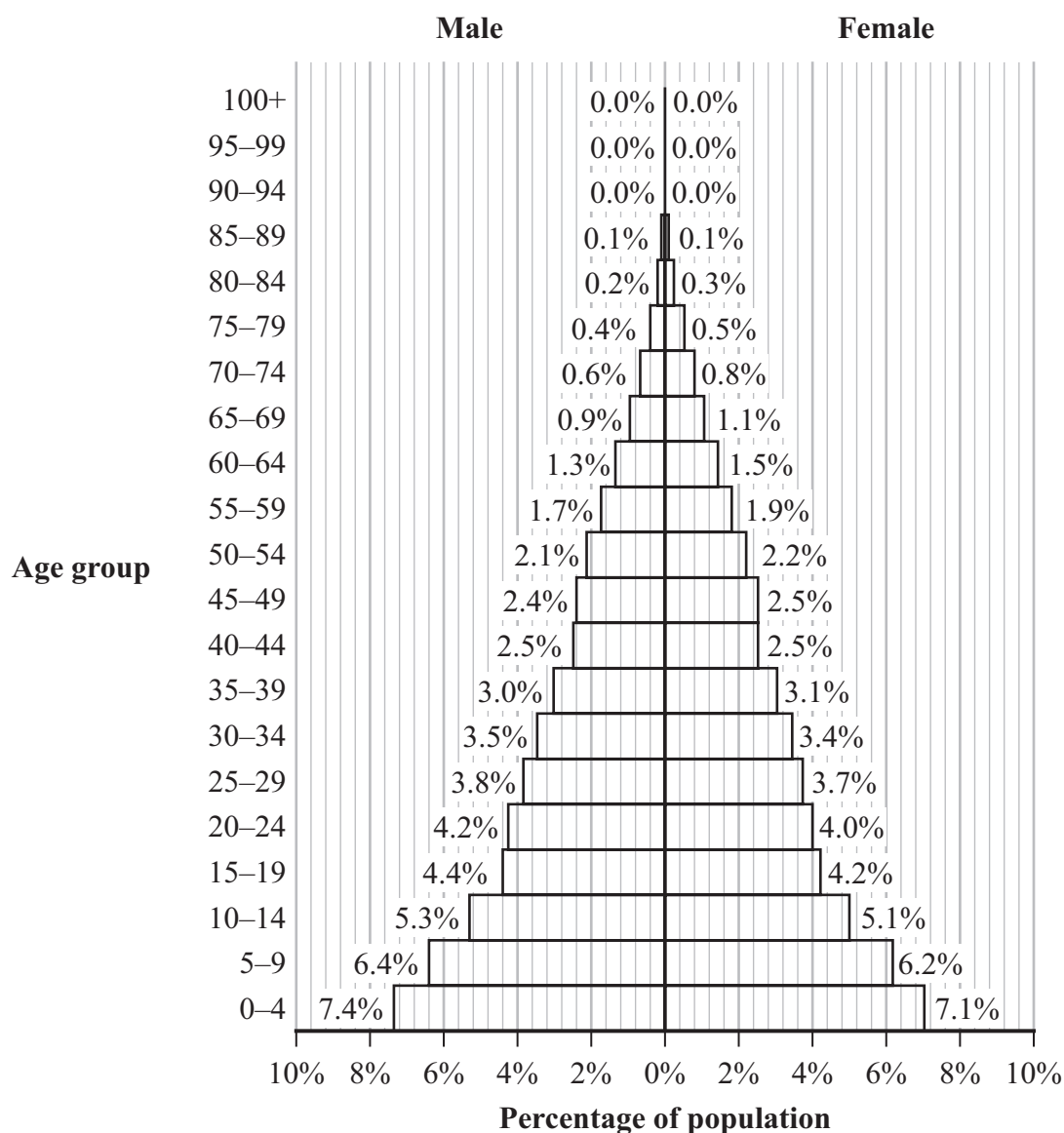
..... cm

(2)

(Total for Question 6 is 5 marks)

- 7 The population pyramid shows the age distribution of the world population in the year 1960

Each percentage is given correct to one decimal place.



(Source: [www.populationpyramid.net](http://www.populationpyramid.net))

- (a) Work out the percentage of males aged 65 or over in 1960

..... %  
(2)



Kiran is carrying out research into how the population of Finland is changing. She uses the data for Finland in 2023

| Finland 2023     |                                  |
|------------------|----------------------------------|
| Age ( $y$ years) | Percentage of the population (%) |
| $0 \leq y < 15$  | 14.91                            |
| $15 \leq y < 30$ | 17.13                            |
| $30 \leq y < 45$ | 19.52                            |
| $45 \leq y < 60$ | 18.39                            |
| $60 \leq y < 75$ | 18.88                            |
| $75 \leq y$      | 11.17                            |

(b) Use linear interpolation to work out an estimate of the median age in Finland in 2023

(2)

In 1960, the median age in Finland was 27 years correct to the nearest year. Kiran claims that Finland’s population is getting older.

(c) Explain whether or not Kiran’s claim is correct. You should justify your answer.

(1)

On a website, Kiran reads that the median age in Finland in 2023 was 42.8 years.  
(Source: *www.worldometers.info*)

(d) Suggest a reason why the median of 42.8 from the website and a median found using linear interpolation would be different.

(1)

(Total for Question 7 is 6 marks)



- 8 In world athletics, an athlete's personal best results in different track and field events are recorded.

Zara wants to compare a sample of athletes' longest recorded long jump against their height. She is going to present this information on a scatter diagram.

- (a) Explain why it is appropriate to use a scatter diagram.

(1)

Zara plotted the athletes' heights on the  $x$ -axis.

- (b) Give a reason why plotting the heights on the  $x$ -axis is appropriate.

(1)

Zara selects 12 athletes ranked as some of the world's best long jumpers. She calculates the Pearson Product Moment Correlation Coefficient (PMCC) for their personal best long jump distances against their heights.

The PMCC value is 0.02

(Source: [www.worldathletics.org](http://www.worldathletics.org))

- (c) Describe the correlation shown by the PMCC.

(1)

Using data for the same 12 athletes, Zara compares their personal best long jump distances ( $m$  metres) to their fastest times in a 100 metre race ( $t$  seconds).

She calculates the PMCC to be  $-0.5$

- (d) Describe and interpret the relationship between the athletes' personal best long jump distances and their personal best 100 metre race times in context.

(2)



The data for the 12 athletes is shown in the table below.

| World rank<br>April 2023 | Long jump athlete | Personal best<br>long jump distance<br>( $m$ metres) | Personal best 100 m<br>race time ( $t$ seconds) |
|--------------------------|-------------------|------------------------------------------------------|-------------------------------------------------|
| 1                        | M Tentoglou       | 8.60                                                 | 10.70                                           |
| 3                        | M Dendy           | 8.42                                                 | 10.31                                           |
| 4                        | S Ehammer         | 8.45                                                 | 10.46                                           |
| 5                        | T Montler         | 8.38                                                 | 10.49                                           |
| 8                        | E Cáceres         | 8.37                                                 | 10.37                                           |
| 10                       | T Gayle           | 8.69                                                 | 10.13                                           |
| 11                       | E Lasa            | 8.28                                                 | 11.03                                           |
| 15                       | H Frayne          | 8.34                                                 | 10.96                                           |
| 16                       | Y Hashioka        | 8.36                                                 | 10.53                                           |
| 17                       | R Juška           | 8.31                                                 | 11.07                                           |
| 26                       | C Mitrevski       | 8.21                                                 | 10.27                                           |
| 30                       | H Santos          | 8.19                                                 | 11.77                                           |

(Source: [www.worldathletics.org](http://www.worldathletics.org))

The linear regression line for these data is  $t = 25 - 1.7m$

In April 2023, M Massó was ranked number 2 in the world for the long jump.  
His personal best long jump distance was 8.39 m.

Zara uses this regression line to predict that M Massó could run the 100 m race  
in 10.7 seconds.

(e) Comment on the reliability of this prediction in context.

(2)

(Total for Question 8 is 7 marks)



P 7 6 3 8 7 A 0 1 7 2 8

- 9 A group of homeowners are investigating the level of lead in their tap water.

Five homeowners each collected a 100 ml sample of their water at the same time each morning for 10 consecutive days.

All samples were sent to the same laboratory where the level of lead in each sample was measured.

- (a) Given these control measures, comment on the reliability of their results.

(1)

In previous tests in this area, the level of lead in the water was found to be normally distributed.

The mean level of lead was  $7 \mu\text{g/l}$  and the upper action limit was  $9.4 \mu\text{g/l}$

- (b) Find the standard deviation used to set the upper action limit.

(2)

“There are strict regulations concerning the amount of lead allowed in drinking water. The current drinking water standard is no more than  $10 \mu\text{g/l}$ ”

(Source: [www.stwater.co.uk](http://www.stwater.co.uk))

One homeowner claims that the lead content in their tap water often exceeds  $10 \mu\text{g/l}$

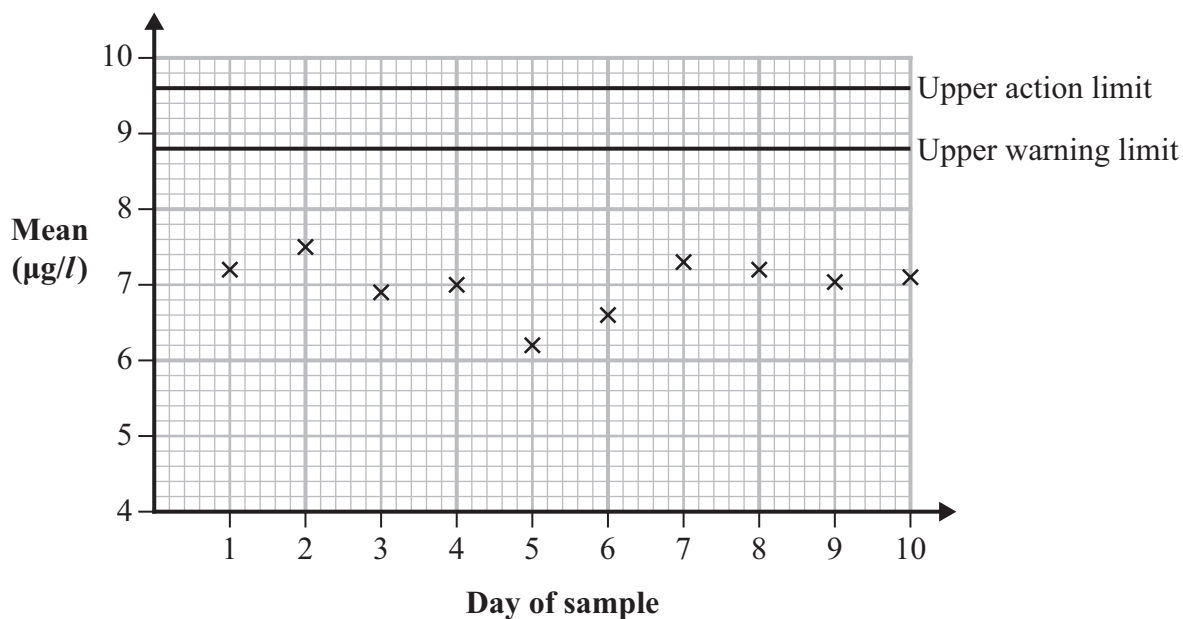
- (c) Explain why this homeowner is likely to be wrong.

(1)

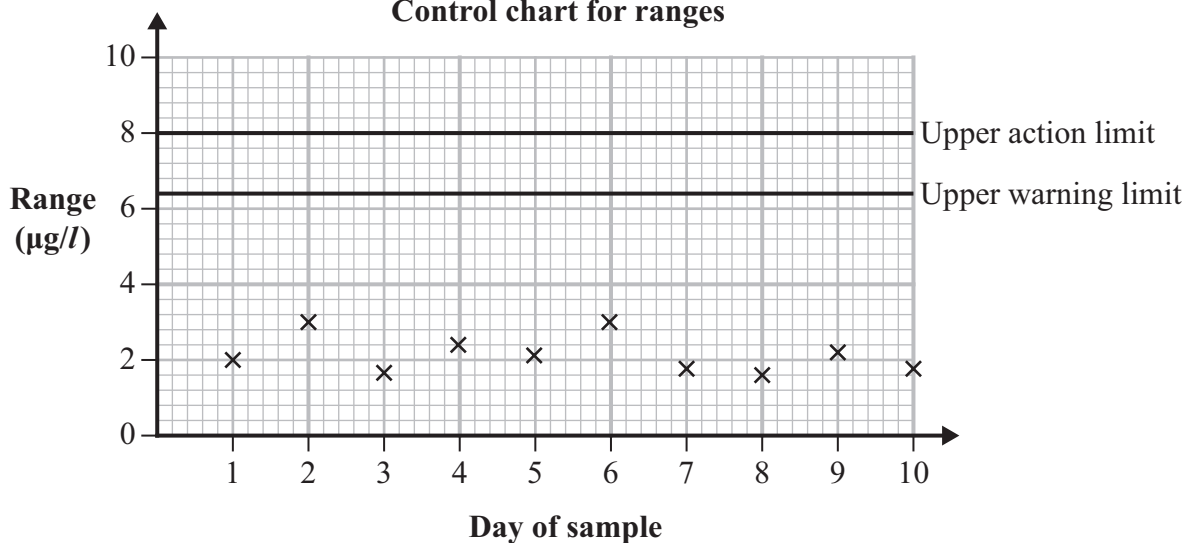


Control charts for the sample means and the sample ranges for the five homeowners over the ten days are shown below.

**Control chart for means**



**Control chart for ranges**



(d) Decide what actions, if any, should be taken.

Justify your answer by referring to both the sample means and the sample ranges.

(3)

(Total for Question 9 is 7 marks)

- 10 The table gives information about the median house price in a town called Bulkington, in England, between the years 2015 and 2020

The table also gives, correct to one decimal place, some of the chain base index numbers for these prices.

|                         | 2015    | 2016    | 2017    | 2018    | 2019    | 2020    |
|-------------------------|---------|---------|---------|---------|---------|---------|
| Median house price (£)  | 170 000 | 188 800 | 201 500 | 212 500 | 224 000 | 222 000 |
| Chain base index number |         | 111.1   | 106.7   | 105.5   | 105.4   |         |

(Source: [www.ons.gov.uk](http://www.ons.gov.uk))

- (a) Calculate the geometric mean of the chain base index numbers from 2016 to 2020 and interpret your result.  
You must show your working.

(5)



DO NOT WRITE IN THIS AREA

Here are some summary statistics for house prices in 2020 for all of England and Wales.

| Mean     | Standard deviation | Median   |
|----------|--------------------|----------|
| £235 884 | £119 492           | £213 250 |

(b) Calculate the skew of this distribution and interpret your result.

(2)

(Total for Question 10 is 7 marks)



- 11 A road safety officer records the speed of each vehicle travelling faster than the speed limit of 30 mph as they pass a site.

Here are the results.

| Speed of vehicles ( $v$ mph) | Frequency |
|------------------------------|-----------|
| $30 < v \leq 35$             | 25        |
| $35 < v \leq 40$             | 18        |
| $40 < v \leq 45$             | $p$       |
| $45 < v \leq 50$             | 14        |
| $50 < v \leq 55$             | 3         |

An estimate for the mean of these data was calculated to be 40.5 mph.

- (a) Find the missing value  $p$  in the table.

(3)

The government recommends that a speed camera should be installed if at least 20% of vehicles travel faster than the speed limit at a site.

(Source: [www.ageas.co.uk](http://www.ageas.co.uk))

- (b) Explain why the data collected by the road safety officer cannot be used as evidence that a speed camera should be installed at this site.

(1)

(Total for Question 11 is 4 marks)



- 12 The table shows some of the percentile eye lengths (mm) in children between the ages of 4 and 8

| Eye length (mm) for boys (B) and girls (G) |            |       |       |       |       |       |       |       |       |       |
|--------------------------------------------|------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Age<br>(years)                             | Percentile |       |       |       |       |       |       |       |       |       |
|                                            | 5th        |       | 25th  |       | 50th  |       | 75th  |       | 95th  |       |
|                                            | B          | G     | B     | G     | B     | G     | B     | G     | B     | G     |
| 4                                          | 21.41      | 20.87 | 21.99 | 21.41 | 22.39 | 21.78 | 22.78 | 22.14 | 23.33 | 22.66 |
| 5                                          | 21.64      | 21.11 | 22.26 | 21.69 | 22.69 | 22.10 | 23.12 | 22.50 | 23.74 | 23.08 |
| 6                                          | 21.86      | 21.33 | 22.51 | 21.96 | 22.97 | 22.41 | 23.45 | 22.85 | 24.15 | 23.50 |
| 7                                          | 22.07      | 21.53 | 22.75 | 22.22 | 23.25 | 22.70 | 23.76 | 23.19 | 24.54 | 23.90 |
| 8                                          | 22.26      | 21.73 | 22.98 | 22.46 | 23.51 | 22.98 | 24.07 | 23.51 | 24.92 | 24.28 |

(Source: <https://bjo.bmj.com>)

- (a) Compare the 5th to the 75th interpercentile range for 6 year old boys with the 5th to the 75th interpercentile range for 6 year old girls.  
You must show your working.

(3)

- (b) Give an interpretation of your comparison in part (a).

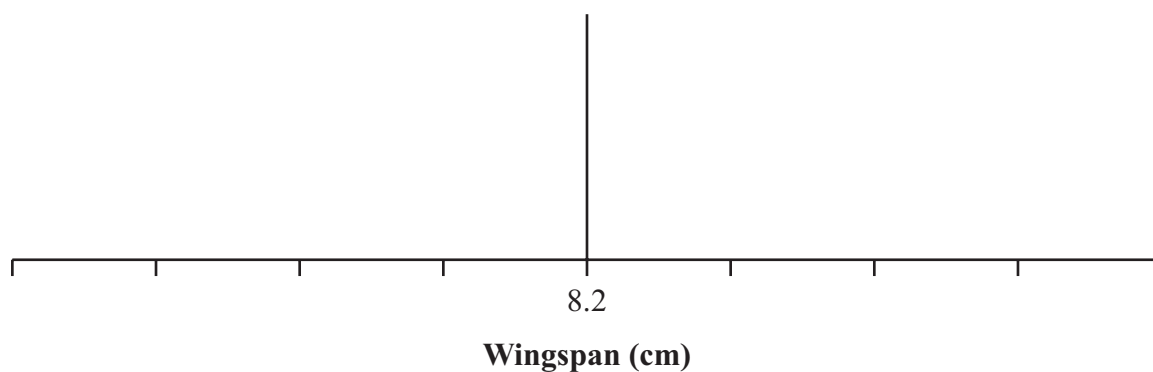
(1)

(Total for Question 12 is 4 marks)

- 13 The wingspans of a species of butterfly are normally distributed.  
The mean length of the wingspans for this species of butterfly is 8.2 cm.  
Only 2.5% of this species of butterfly have a wingspan longer than 8.6 cm.
- (a) Show that the standard deviation of this species of butterfly is 0.2 cm.

(1)

- (b) On the axis below, sketch a diagram for this normal distribution.



(2)

Sarah claims to have caught a butterfly of this species with a wingspan of 7.5 cm.

- (c) Comment on the likelihood of her claim being true.  
Give a reason for your answer.

(1)

The lifespan of this butterfly species is known to be 15 days or less.

Liam wants to know if the lifespan of this butterfly species can be modelled by a normal distribution.

He has collected data for 34 random butterflies of this species.

First, Liam tabulates this data.

He produces this table to show the length of time each butterfly lived.

| Length of time lived ( $d$ days) | Frequency |
|----------------------------------|-----------|
| $0 < d \leq 3$                   | 4         |
| $3 < d \leq 6$                   | 9         |
| $6 < d \leq 9$                   | 12        |
| $9 < d \leq 12$                  | 7         |
| $12 < d \leq 15$                 | 2         |



Liam then produces this plan to check if the lifespan of this butterfly species can be modelled by a normal distribution.

Using the data collected I will draw a histogram.

I will then compare the mean and the median.

I will also use the mean and the standard deviation to work out the amount of data within 1 standard deviation of the mean and the amount of data within 2 standard deviations of the mean.

(d) Comment on whether or not Liam's tabulation and plan are appropriate.

(5)

(Total for Question 13 is 9 marks)

**TOTAL FOR PAPER IS 80 MARKS**



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

BLANK PAGE



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

**BLANK PAGE**



DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

BLANK PAGE

