

Exercise walk-through

6.1

Sampling, averages and spread (S1–S4)

AQA · GCSE · Mathematics

1. Unit 3.6 at a glance

Unit 3.6 at a glance

Rows S1-S6 in three tiers. Additional foundation: inferring population properties from a sample with the limitations of sampling (S1) and tables plus line graphs for time series (S2). Higher only: histograms with equal and UNEQUAL class intervals and cumulative frequency graphs (S3), with box plots, quartiles and the inter-quartile range added to S4.

Must-know: $\text{mean} = \frac{\sum fx}{\sum f}$ (midpoints for grouped); $\text{range} = \text{max} - \text{min}$; $\text{IQR} = Q3 - Q1$; $\text{pie angle} = \frac{f}{n} \times 360$; $\text{density} = \frac{\text{frequency}}{\text{class width}}$; correlation direction and strength, never causation.

2. Sampling and populations

Sampling and populations

A sample is a subset; a census covers the whole population (S1). Random selection reduces selection bias but never removes sampling variability or non-response: 24 walkers in a random sample of 80 is a proportion 0.3, suggesting about 180 of 600 - an estimate, not a count.

A convenience sample from a sports club overrepresents active students; a voluntary poll misses non-respondents however many answers arrive. Apply statistics to describe a population (S5) while stating the limits: primary vs secondary data, discrete vs continuous data (siblings counted, travel time measured).

5. Averages, spread and box plots

Averages, spread and box plots

Mean = $\frac{\sum fx}{\sum f}$ with exact products: the 2H Q13 anchor $108 + 136.8 + 56 = 300.8$ over 40 gives 7.52 - approximated products lose the mark. Median by cumulative count, mode or modal class; the value 24 in 2, 4, 4, 6, 24 lifts the mean to 8 but leaves the median at 4 (S4).

Higher box plots: minimum, Q1, median, Q3, maximum with $IQR = Q3 - Q1$ - the 2H Q16 anchor plots 2.4, 4, 6, 7, 9. The 1H Q16 anchor compares two clubs: median for the average, IQR for consistency - BOTH statements, in context.

Exercises 2.1·2.3

2.1 Find the mean of 2, 4, 4, 6, 9. [1]

2.2 Values 1, 2 and 3 have frequencies 2, 5 and 3. Find the mean and the mode. [2]

2.3 In a random sample of 80 students, 24 walk to school. Estimate the number of walkers in a school of 600. [2]

Exercises 2.1·2.3

2.1 Find the mean of 2, 4, 4, 6, 9. [1]

2.1 $(2 + 4 + 4 + 6 + 9) \div 5 = 25 \div 5 = 5.$

2.2 Values 1, 2 and 3 have frequencies 2, 5 and 3. Find the mean and the mode. [2]

2.2 $\sum fx = 2 + 10 + 9 = 21$ over 10 values: mean = 2.1; mode = 2.

2.3 In a random sample of 80 students, 24 walk to school. Estimate the number of walkers in a school of 600. [2]

2.3 $\frac{24}{80} = 0.3$; $600 \times 0.3 = 180$ walkers.

Exercises 2.4·2.6

2.4 For the data 2, 4, 4, 6, 24, find the mean and the median, and explain why they differ. [2]

2.5 Classify each as discrete or continuous: number of siblings; travel time to school. [2]

2.6 Classes $0 \leq x < 10$, $10 \leq x < 20$, $20 \leq x < 30$ have frequencies 2, 5, 3. Estimate the mean and write down the modal class. [3]

Exercises 2.4·2.6

2.4 For the data 2, 4, 4, 6, 24, find the mean and the median, and explain why they differ. [2]

2.4 Mean = $40 \div 5 = 8$; median = 4. The single large value 24 pulls the mean up while the median resists it.

2.5 Classify each as discrete or continuous: number of siblings; travel time to school. [2]

2.5 Siblings: discrete (counted); travel time: continuous (measured, even if recorded to whole minutes).

2.6 Classes $0 \leq x < 10$, $10 \leq x < 20$, $20 \leq x < 30$ have frequencies 2, 5, 3. Estimate the mean and write down the modal class. [3]

2.6 Midpoints 5, 15, 25: $\sum fm = 10 + 75 + 75 = 160$, mean $\approx 160 \div 10 = 16$; modal class $10 \leq x < 20$.

Exercises 3.1-3.3

3.1 (*June 2024 2H Q13, adapted.*) A table of lengths shows 15 values of 7.2, 18 values of 7.6 and 7 values of 8. Work out the mean. [3]

3.2 (*June 2024 2H Q16, adapted.*) Draw a box plot for: shortest value 2.4 cm, lower quartile 4 cm, median 6 cm, upper quartile 7 cm, longest value 9 cm. [3]

3.3 (*June 2024 1H Q16, adapted.*) The swimming club's median age is lower than the cycling club's, and its inter-quartile range is higher. Write two sentences comparing the clubs' ages. [2]

Exercises 3.1·3.3

3.1 (*June 2024 2H Q13, adapted.*) A table of lengths shows 15 values of 7.2, 18 values of 7.6 and 7 values of 8. Work out the mean. [3]

3.1 $\sum fx = 15 \times 7.2 + 18 \times 7.6 + 7 \times 8 = 108 + 136.8 + 56 = 300.8$; mean = $300.8 \div 40 = 7.52$.

3.2 (*June 2024 2H Q16, adapted.*) Draw a box plot for: shortest value 2.4 cm, lower quartile 4 cm, median 6 cm, upper quartile 7 cm, longest value 9 cm. [3]

3.2 A rectangle from 4 to 7 with a median line at 6, whiskers to 2.4 and 9 – five plotted points.

3.3 (*June 2024 1H Q16, adapted.*) The swimming club's median age is lower than the cycling club's, and its inter-quartile range is higher. Write two sentences comparing the clubs' ages. [2]

3.3 The swimming club's average age is lower (its median is lower); its ages are less consistent (its inter-quartile range is higher).

Exercises 3.4·3.5

3.4 (*Original, Higher.*) A cumulative-frequency graph of 80 observations reads $Q_1 = 12$ and $Q_3 = 21$. State where these readings were taken on the cumulative axis, and work out the inter-quartile range. [3]

3.5 (*Original.*) An optional online poll finds that 70% of 400 respondents cycle to school. Explain why this does not establish that 70% of the school cycles. [2]

Exercises 3.4·3.5

3.4 (*Original, Higher.*) A cumulative-frequency graph of 80 observations reads $Q_1 = 12$ and $Q_3 = 21$. State where these readings were taken on the cumulative axis, and work out the inter-quartile range. [3]

3.4 Quartiles read at cumulative frequency $80 \div 4 = 20$ and $3 \times 80 \div 4 = 60$; $IQR = 21 - 12 = 9$.

3.5 (*Original.*) An optional online poll finds that 70% of 400 respondents cycle to school. Explain why this does not establish that 70% of the school cycles. [2]

3.5 The poll is voluntary, so keen cyclists respond more – selection bias that more responses cannot remove; the sample is not random.

Local review guidance: exact $\sum fx$ products before dividing (approximated products forfeit the mark); comparisons state an average AND a spread in context; sampling limits named as selection bias or non-response, not “not enough people”. This is authored classroom guidance.