

GCSE / IGCSE Maths – Statistics

GCSE / IGCSE (Higher) · Total marks: 47 · Answers on the last page

Covers averages from lists and tables, cumulative frequency and box plots, histograms and scatter graphs. A calculator may be used. Allow about 60 minutes.

Section A – Averages and spread

1. Here are the numbers of emails received by 8 people in one day: 3, 7, 7, 8, 10, 12, 15, 18.
Find (a) the mean, (b) the median, (c) the mode, (d) the range. [4]
2. The table shows the number of goals scored by a team in 25 matches: 0 goals (5 matches), 1 goal (8 matches), 2 goals (6 matches), 3 goals (4 matches), 4 goals (2 matches).
(a) Write down the mode.
(b) Find the median.
(c) Calculate the mean number of goals per match. [4]
3. The heights, h cm, of 30 plants are grouped: $140 < h \leq 150$ (4 plants), $150 < h \leq 160$ (11), $160 < h \leq 170$ (9), $170 < h \leq 180$ (6).
(a) Write down the modal class.
(b) Calculate an estimate for the mean height.
(c) Explain why your answer to (b) is only an estimate. [4]
4. The mean of five numbers is 9. Four of the numbers are 6, 11, 4 and 13. Find the fifth number. [2]
5. Class A has 12 students with a mean test mark of 64. Class B has 18 students with a mean test mark of 59. Calculate the mean mark of all 30 students. [3]
6. The ages of members of a club are grouped: $10 < a \leq 20$ (8 members), $20 < a \leq 30$ (x members), $30 < a \leq 40$ (5 members).
The estimated mean age is 24 years. Find the value of x . [4]

Section B – Cumulative frequency, histograms, scatter graphs

7. 60 students were timed completing a puzzle. Times t (minutes): $0 < t \leq 10$ (6 students), $10 < t \leq 20$ (14), $20 < t \leq 30$ (22), $30 < t \leq 40$ (12), $40 < t \leq 50$ (6).
(a) Complete the cumulative frequency table.
(b) Describe how you would draw the cumulative frequency graph (which points to plot).
(c) Use linear interpolation (or a cumulative frequency graph) to estimate the median and the interquartile range. [6]
8. Two groups of students sat the same test. Group A: median 38, interquartile range 14. Group B: median 45, interquartile range 9.
Make two comparisons between the marks of the two groups. [2]
9. The table shows the masses, m kg, of 60 parcels: $0 < m \leq 5$ (frequency 10), $5 < m \leq 15$ (30), $15 < m \leq 20$ (8), $20 < m \leq 40$ (12).
(a) Calculate the frequency density of each class.
(b) Use the histogram to estimate the number of parcels with mass between 10 kg and 25 kg. [5]

10. On a histogram, the bar for the class $20 < x \leq 30$ has a frequency density of 1.4.
- (a) Find the frequency of this class.
 - (b) The class $30 < x \leq 50$ has a frequency of 8. Find the height (frequency density) of its bar. [3]
11. A teacher plots a scatter graph of hours revised (x) against test mark (y) for a class. The points show that students who revised longer generally scored higher.
- (a) Describe the type of correlation shown.
 - (b) The line of best fit passes through (2, 40) and (8, 76). Find the equation of the line of best fit.
 - (c) Use your equation to estimate the mark of a student who revised for 5 hours.
 - (d) Explain why the line should not be used to predict the mark of a student who revised for 20 hours. [5]

Section C – Extended

12. 80 students sat an exam. Marks: $0 < m \leq 20$ (4 students), $20 < m \leq 40$ (16), $40 < m \leq 60$ (30), $60 < m \leq 80$ (22), $80 < m \leq 100$ (8).
- (a) Estimate the median mark.
 - (b) The top 25 percent of students are awarded a distinction. Estimate the lowest mark needed for a distinction.
 - (c) A student claims the highest mark was 100. Explain whether the data supports this claim. [5]

Answers

Check your working against these. If an answer does not match, try the question again before looking for the method.

- $80 \div 8 = 10$
 - 9
 - 7
 - 15
- 1
 - 13th value = 1
 - $40 \div 25 = 1.6$
- $150 < h \leq 160$
 - $\frac{145 \times 4 + 155 \times 11 + 165 \times 9 + 175 \times 6}{30} = \frac{4820}{30} = 160.7 \text{ cm}$
 - Midpoints are used because the exact heights are not known
- $45 - 34 = 11$
- $\frac{768 + 1062}{30} = 61$
- $\frac{120 + 25x + 175}{13 + x} = 24 \Rightarrow 295 + 25x = 312 + 24x \Rightarrow x = 17$
- Cumulative frequencies: 6, 20, 42, 54, 60
 - Plot cumulative frequency against the upper class boundary: (10, 6), (20, 20), (30, 42), (40, 54), (50, 60), joined by a smooth curve
 - Median (30th) ≈ 24.5 min; LQ (15th) ≈ 16.4 min; UQ (45th) ≈ 32.5 min; IQR ≈ 16 min (accept readings within ± 1)
- Group B scored higher on average (median 45 compared with 38); Group B marks were more consistent (smaller IQR of 9 compared with 14)
- 2, 3, 1.6, 0.6
 - $\frac{1}{2}(30) + 8 + \frac{5}{20}(12) = 15 + 8 + 3 = 26$
- $1.4 \times 10 = 14$
 - $8 \div 20 = 0.4$
- Positive correlation
 - Gradient = $\frac{36}{6} = 6$, so $y = 6x + 28$
 - 58 marks
 - 20 hours is outside the range of the data (extrapolation) so the relationship may not continue; also marks cannot exceed the maximum
- 40th value: $40 + \frac{20}{30} \times 20 \approx 53$
 - 60th value: $60 + \frac{10}{22} \times 20 \approx 69$
 - Cannot be confirmed: grouped data only shows the highest mark lies between 80 and 100

Need this topic taught, not just practised? Book a trial lesson at 50% off with Mr. Ahmad — online worldwide or in person in Dubai.

WhatsApp +971 55 373 5695 · mathandsciencetutor.com