

Statistics

Quick revision notes [Open it online](#)

Name: _____

Date: _____

At a glance

Averages and spread, grouped data, charts, histograms, cumulative frequency and scatter diagrams.

1. Averages and range
2. Grouped data
3. Charts and diagrams
4. Histograms
5. Cumulative frequency, quartiles and IQR
6. Scatter diagrams

1. Averages and range

☐ I revised this

- **Mean** = total $\div$ number of values. **Median** = middle value in order (average the middle two if there is an even count).
- **Mode** = most common value. **Range** = largest – smallest (a measure of spread).
- From a frequency table: mean = $\Sigma fx \div \Sigma f$; median position = $(n + 1) \div 2$.
- Missing value from a mean: total = mean $\times$ number of values.

Formulas

mean = $\Sigma fx \div \Sigma f$	frequency table
median position = $(n + 1) \div 2$	

Worked example

Scores 1, 2 and 3 have frequencies 4, 5 and 1. Find the mean and the median.

1. $\Sigma fx = 4 + 10 + 3 = 17$; $\Sigma f = 10$.
2. Mean = 1.7.
3. Median: position 5.5 — the 5th and 6th values are both 2.

Mean 1.7, median 2

2. Grouped data

☐ I revised this

- You don't know the exact values, so use each class's **midpoint**: the mean is an **estimate**.
- The **modal class** is the class with the highest frequency (for unequal widths, the highest frequency density).
- The class containing the median: find the $(n/2)$ th value by adding up frequencies.

Worked example

Heights 150–160 cm ($f = 4$), 160–170 ($f = 10$), 170–180 ($f = 6$). Estimate the mean.

1. Midpoints 155, 165, 175.
2. $\Sigma fx = 620 + 1650 + 1050 = 3320$; $\Sigma f = 20$.

Estimated mean 166 cm

3. Charts and diagrams

☐ I revised this

- Bar charts (gaps, equal widths), pictograms (with a key), pie charts (angle = frequency $\div$ total $\times 360^\circ$).
- Stem-and-leaf: always give a key ($3 \mid 4 = 34$); leaves in order; read the median by counting.
- Compare two distributions with an average AND a measure of spread, in context.

Worked example

A stem-and-leaf diagram has key $3 \mid 4 = 34$. What does $5 \mid 1$ mean?

1. Stem = tens, leaf = units.

51

4. Histograms

☐ I revised this

- With unequal class widths the vertical axis is **frequency density** = frequency $\div$ class width.
- The **area** of each bar is the frequency.
- To read a frequency: frequency density $\times$ class width. For part of a class, take that fraction of the bar.

Formulas

frequency density = frequency $\div$ class width

frequency = density $\times$ width

Worked example

A bar for $10 < x \leq 25$ has frequency density 2. A bar for $25 < x \leq 30$ has frequency 15. Find the first frequency and the second density.

1. Width 15: $2 \times 15 = 30$.
2. Width 5: $15 \div 5 = 3$.

Frequency 30; density 3

5. Cumulative frequency, quartiles and IQR

☐ I revised this

- Cumulative frequency is a running total. Plot each total at the **upper end** of its class and join with a smooth curve.
- Read the median at $n/2$, the lower quartile at $n/4$, the upper quartile at $3n/4$ (on the cumulative frequency axis).
- Interquartile range = UQ – LQ: the spread of the middle half, not affected by extremes.
- 'How many more than x ?' — read the cumulative frequency at x and subtract from the total.

How to do it

1. Make a running total column.
2. Plot (upper class boundary, running total).
3. Join with a smooth curve.
4. Read across from $n/4$, $n/2$, $3n/4$ and down to the x-axis.

Worked example

80 students: the graph gives LQ = 31 marks and UQ = 46 marks. Find the IQR. 52 students scored 40 or less: how many scored more than 40?

1. IQR = $46 - 31 = 15$.
2. $80 - 52 = 28$.

IQR 15 marks; 28 students

6. Scatter diagrams

☐ I revised this

- Correlation: positive (both increase), negative (one up, one down), or none.
- Line of best fit: one ruled line following the trend, with points roughly balanced either side; it need not pass through (0, 0).
- Use it to estimate within the data (interpolation). Estimates outside the data (extrapolation) are unreliable.
- An outlier is a point that doesn't fit the pattern.

Worked example

As temperature rises, sales of hot drinks fall. Describe the correlation.

1. One goes up while the other goes down.

Negative correlation

Common mistakes

Don't write	Write instead
✗ Median of 3, 4, 5, 7, 9, 12 is 5	✓ Even count: average the middle two, 6
✗ Giving the frequency as the modal class	✓ Give the class, e.g. 10–20
✗ Plotting cumulative frequency at midpoints	✓ Plot at the upper end of each class
✗ Histogram bar height = frequency with unequal widths	✓ Height = frequency density
✗ Dividing Σfx by the number of classes	✓ Divide by Σf , the total frequency

In the exam

- Show Σfx and Σf in a table — they earn method marks.
- When comparing data, compare an average AND a spread, and say what it means in context.
- Read graphs carefully — check the scale on each axis first.

My notes

<hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/> <hr/>
