Measures of central tendency and spread

Lesson 2.3.4 | AS research methodology

Core ideas

Measures of central tendency describe a typical score in a set of data.

The mean is calculated by adding all scores and dividing by the number of scores.

The median is the middle score when scores are placed in order.

The mode is the most frequent score in a data set.

The range shows spread by subtracting the lowest score from the highest score.

Outliers can distort the mean and range, so the median may sometimes be a better summary.

Key terms

Central tendency: A summary of the typical or middle value in a data set.

Mean: The total of all scores divided by the number of scores.

Median: The middle score when values are ordered.

Mode: The most common score in a data set.

Range: The highest score minus the lowest score.

Measures of central tendency and spread

Lesson 2.3.4 | AS research methodology

Continued

Outlier: An unusual score that is far from most other scores.

Data set: A collection of scores or values.

Spread: How far apart or varied the scores are.

Standard deviation: A measure of the dispersion of scores around their mean; larger values indicate greater variability.

Further detail and evidence

Standard deviation describes how dispersed scores are around the mean: a smaller standard deviation means scores are more closely clustered; a larger one means greater variability. Compare it with the mean rather than treating it as a measure of the average.

Two groups can have the same mean but different standard deviations. If both mean recall scores are 10, SD 1 indicates more consistent scores than SD 4; it does not mean the first group has higher average recall.

Mean summarises all scores; median is the middle ordered score; mode is most frequent. Range is the highest minus lowest score. Standard deviation describes dispersion around the mean: smaller values indicate tighter clustering. Know how these measures are found and interpret them.

Measures of central tendency and spread

Lesson 2.3.4 | AS research methodology

Applying the ideas

Two conditions can both have a mean recall score of 10, while SD 1 versus SD 4 indicates different consistency, not a difference in average recall. A very high outlier can raise the mean and range while having less effect on the median. Compare groups meaningfully, keeping units clear.

Worked example: Calculate the mean and range for the scores 4, 6, 6, 8 and 11. The mean is 7 because the total is 35 and there are 5 scores, so 35 divided by 5 equals 7. The range is 7 because the highest score is 11 and the lowest score is 4, so 11 minus 4 equals 7. These two measures summarise a typical score and the spread of scores.

Worked example: Assess why a researcher might report the median instead of the mean. A researcher might report the median when a data set contains an outlier. For example, if most reaction times are close together but one participant is extremely slow, the mean could be pulled upward and become less representative. The median is less affected by the extreme score, so it may better show the typical performance.

Evaluation and connections

A mean without spread can conceal important variation. The mode suits the most common category; a mean is not meaningful for arbitrary category labels. The 2027 syllabus does not require candidates to perform calculations or interpret unspecified inferential tests. Optional arithmetic practice must not be presented as an extra exam requirement.