7.14 Data Handling: Descriptive Statistics, Graphs and Levels of Measurement: 30-minute Pack

AQA 7182 · about 30 minutes

Total Marks

Name: _____ Date: ____ / ____ / _____

Answer ALL questions.

1 Calculate the mean, median and mode for the following interval-level reaction time data (milliseconds) recorded from eight participants in a lab task: 210, 245, 198, 232, 245, 220, 205, 238. Show your working.

(Total for Question 1 is 4 marks)

2 Explain one strength and one limitation of using the mean and standard deviation together to summarise interval-level psychological data from a laboratory task.

(Total for Question 2 is 4 marks)

3 A researcher wants to present the distribution of scores from a memory test (scores 0 to 10) for a sample of 12 participants. Which graph is most appropriate to show the frequency distribution of these interval-level whole-number scores, and state two labelling conventions that must be followed for that graph.

(Total for Question 3 is 3 marks)

© Revision Library | revisionlibrary.co.uk

4 Calculate the mode and median for these ordinal satisfaction ranks collected from seven participants: 4, 2, 3, 4, 5, 2, 4. Explain briefly why the mean would be less appropriate here.

(Total for Question 4 is 6 marks)

5 Explain two strengths and two limitations of using the median and interquartile range (IQR) to summarise skewed interval-level reaction time data from a small study. Your answer should make clear why these measures might be preferred to the mean and standard deviation for skewed data.

(Total for Question 5 is 8 marks)

© Revision Library | revisionlibrary.co.uk

Mark scheme · 7.14 Data Handling: Descriptive Statistics, Graphs and Levels of Measurement: 30-minute Pack

Question 1

correctly calculates the mean: sum all values = 210+245+198+232+245+220+205+238 = 1793 oe	**M1**
mean = 1793 / 8 = 224.125 ms, allow 224.1 ms or 224.13 ms cao	**A1**
orders data to find median: ordered list 198,205,210,220,232,238,245,245; median is average of 4th and 5th values oe	**M1**
median = (220 + 232) / 2 = 226 ms cao; mode = 245 ms (most frequent) cao	**A1**

Answer: Mean = 224.125 ms (224.1 ms). Median = 226 ms. Mode = 245 ms.

Question 2

identifies strength: the mean uses all data points and is sensitive to small differences in scores oe	**M1**
explains consequence: combining mean with SD gives a precise picture of central tendency and spread, useful for comparing groups and calculating further statistics oe	**A1**
identifies limitation: both mean and SD are affected by outliers and non-normal distributions oe	**M1**
explains consequence: in skewed data or with extreme scores the mean and SD can misrepresent typical performance, so median and IQR might be preferable oe	**A1**

Answer: Strength: mean and SD use all data and together give a precise summary of central tendency and spread, aiding group comparisons. Limitation: both are sensitive to outliers and skew, so they can misrepresent typical scores in non-normal distributions.

Question 3

graph: histogram or bar chart for frequency distribution of whole-number interval scores (histogram preferred for continuous data grouped in bins) oe	**B1**
labelling convention 1: x-axis labelled with the variable name and units, e.g. 'Memory test score (0 to 10)' oe	**B1**
labelling convention 2: y-axis labelled with 'Frequency' or 'Number of participants', and bars touching for a histogram or separated for a bar chart depending on data type oe	**B1**

Answer: Histogram (preferred) or bar chart to show frequency distribution. Label the x-axis with the variable and units, e.g. 'Memory test score (0 to 10)', and label the y-axis 'Frequency' or 'Number of participants'; for a histogram the bars should touch.

Question 4

orders data: 2,2,3,4,4,4,5 oe	**M1**
median is middle value, 4 cao	**A1**
mode is most frequent value, 4 (appears three times) cao	**M1**
identifies that the data are ordinal ranks so mean assumes equal intervals which may not hold oe	**M1**
explains consequence: mean could misrepresent central tendency for ordinal data and be affected by extremes, so median/mode are preferred oe	**A1**
final statement: median = 4, mode = 4, mean less appropriate because interval equality is not guaranteed for ordinal data cao	**A1**

Answer: Ordered data 2,2,3,4,4,4,5. Median = 4. Mode = 4. The mean is less appropriate because these are ordinal ranks and the mean assumes equal intervals between values, which may not be true.

Question 5

identifies a strength: median and IQR are robust to outliers and skewed distributions oe	**M1**
explains consequence: for skewed reaction time data a few slow responses will not distort the median and IQR, so these measures give a better indication of a typical participant and typical spread than mean and SD oe	**A1**
identifies a second strength or expansion: median IQR are appropriate for small samples and ordinal-like distributions within interval data oe	**M1**
explains consequence: they allow meaningful comparisons between groups when assumptions of normality are not met, and they are simple to compute and interpret for practitioners oe	**A1**
identifies a limitation: median and IQR discard some information about individual scores, using only ranks and central 50% oe	**M1**
explains consequence: this can reduce statistical power and conceal details about tails of the distribution, which may be important for understanding variability in reaction times oe	**A1**
identifies a further limitation: some statistical tests and parametric models require means and SDs or assume normality oe	**M1**
explains consequence: using median and IQR may limit options for further parametric analyses unless data are transformed or non-parametric tests are used, so researchers must plan analysis accordingly oe	**A1**

Answer: Strength: median and IQR are robust to outliers and skew, so for skewed reaction time data a few very slow responses will not distort the central value or the measure of spread, giving a better picture of typical performance than mean and SD. They are also simple to compute and suitable for small samples. Limitation: median and IQR ignore some information about individual scores and the tails of the distribution, which can reduce power and hide important variability. Also many parametric analyses require means and SDs, so using median/IQR may restrict further analysis unless non-parametric methods are chosen.

© Revision Library | revisionlibrary.co.uk