Descriptive Statistics

Descriptive statistics

- Used to describe the main trends in the data
- Used to summarise the raw data from research into a more meaningful form.

What does this include?

- measures of central tendency e.g. mean
- Measures of dispersion e.g. range
- Graphical representations of data e.g. bar chart

- When choosing the most appropriate descriptive statistics for their data, the researcher must remember that their main aim is to communicate the major features of a set of data.

- The statistics chosen must be representative of the data.

Measures of central tendency

- MEAN
- Also known as the average; calculated by adding all the values in a data set and dividing by the number of values in the data set.
- Advantage: most sensitive and powerful tool.
- Disadvantage: Because it's so sensitive to each value in a data set, it can be distorted if extreme scores are present. No longer representative.

Median

- The median is the central value of a set of data. It is calculated by putting the values in rank order, and selecting the middle value. If there are two middle values, use their mean.
- Strength: Not distorted by extreme scores at one end of a distribution
- Limitations: less sensitive to the value of each piece of data than the mean; less meaningful with small data sets

Mode

- The mode is the most frequently occurring value or category.
- Strength: It is most useful with data in the form of frequency counts; not influenced by extreme scores in skewed data.
- Limitations: Not sensitive to exact value of each piece of data; not useful when the data set has many modes.

Measures of dispersion

- Allow us to see whether our scores are clustered together closely around the mean, or whether they are widely spread.

Range

- Distance between the largest and smallest scores in the set of data.
- Strength: quick and easy to calculate
- Weakness: Less effective measure of dispersion when used with data where there are extreme values at one end of the set.

Standard deviation

- Gives a measure of how much on average each of the scores in a data set deviates from the mean.
- Strength: uses all the information in a set of data; therefore is the most powerful measure and likely to be representative of the data.
- Limitation: less effective when used with skewed data.