

Inspect your data

Frequency distribution tables, bar charts, scatter plot

normal distribution means that your data are symmetrically distributed around a center where most values lie, with the values tapering off at the tail ends.

skewed distribution is asymmetric and has more values on one end than the other. The shape of the distribution is important to keep in mind because only some descriptive statistics should be used with skewed distributions.

Outliers are extreme values that differ from most other data points in a dataset. They can have a big impact on your statistical analyses and skew the results of any hypothesis tests.

Calculate measures of central tendency

describe where most of the values in a data set lie.

Mode: the most popular response or value in the data set. To find the mode, order your data set from lowest to highest and find the response that occurs most frequently

Median: the value in the exact middle of the data set when ordered from low to high. To find the median, order each response value from the smallest to the biggest. Then, the median is the number in the middle. If there are two numbers in the middle, find their mean.

Calculate measures of central tendency (cont)

Mean: the sum of all values divided by the number of values. To find the mean, simply add up all response values and divide the sum by the total number of responses. The total number of responses or observations is called N.

Calculate measures of variability

tell you how spread out the values in a data set are. Four main measures of variability are often reported:

Range: the highest value minus the lowest value of the data set. To find the range, simply subtract the lowest value from the highest value.

Interquartile range: the range of the middle half of the data set.

Standard deviation: the average distance between each value in your data set and the mean. The standard deviation (s) is the average amount of variability in your dataset. It tells you, on average, how far each score lies from the mean. The larger the standard deviation, the more variable the data set is.

Variance: the square of the standard deviation. The variance (s²) is the average of squared deviations from the mean. Variance reflects the degree of spread in the data set. The more spread the data, the larger the variance is in relation to the mean.

