**Statistics in Psychology**

Statistics is the practice or science of collecting and analyzing numerical data in large quantities, especially for the purpose of inferring proportions in a whole from those in a representative sample.

**Descriptive Statistics**

Descriptive statistics use numerical data to measure and describe characteristics of groups. This includes measures of central tendency and measures of variation.

Measures of Central Tendency

**Mode:** most frequently occurring score (Mo)

**Mean:** arithmetic average of all the scores (sum divided by number of scores) (x)

**Median:** the midpoint of the scores, half will be above and half will be below. (Md)

These scores can sometimes tell a very different story, especially if there is a **skewed distribution** where scores are not gathered symmetrically around the average. There are lots are strange outliers.

Measures of dispersion

**Range**: the difference between the highest and lowest scores in a distribution (R)

**Standard deviation**: a computed measure of how much scores vary around the mean score (SD)



Large amounts of data tend to be captured in a **normal curve.** This is a symmetrical, bell-shaped curve that describes the distribution of many types of data; most scores fall near the mean (about 68 percent fall within one standard deviation of it) and fewer and fewer to the extremes.

Curves can sometimes have a positive or negatives skew.





**Inferential Statistics**

Inferential statistics is numerical data that allow one to generalize—to infer from sample data the probability of something being true of a population. It is used to make predictions about the rest of the population.

Reliability

1. Representative samples are better than biased ones.
2. Less-variable observations are more reliable than those that are more variable.
3. More cases are better than fewer

Significance

**Statistical significance** is a statistical statement of how likely it is that an obtain result occurred by chance. If the difference between two groups is large and consistent, is reliable.

Statistical significance can be represented by using *p,* which tells use the probability the results occurred by chance. Example:

P < .10

Tells us that the probability of the results occurring by chance is less than 10 in 100 (10%). In other words, 90 out of 100 of the results were due to treatment given (independent variable).