Dancing statistics are a beautiful way to visualize statistics so that they are not abstract concepts without an anchor. What makes statistics hard for some of us is because we can’t see what they do behind the scenes. These beautiful videos provided by BPS bring the concepts to life and make the definitions memorable. The definitions below were adapted from What are Your Chances and The Free Dictionary both are great resources!

Correlation and Causation

Correlation is the degree and direction of association of variables and how well one can be predicted from the other. A relationship between variables can be negative (inverse), positive, or curvilinear.

Causation is the relationship of cause to effect. This example about vaccines and autism shows this very nicely

Frequency Distributions

Frequency distribution in statistics, describes the distribution of measurements on a scale for a specific population. Normal distribution is a symmetrical distribution of scores with the majority concentrated around the mean. Probability distribution a mathematical function that assigns to each measurable event in a sample group the probability that the event will occur.
Sampling and Standard Error

Sampling

Sampling is the process of selecting participants to take part in the research on the basis that they can provide detailed information that is relevant to the research question.

- Purposive sampling is selection of participants who have particular knowledge or experience of the area being investigated
- Theoretical sampling is a sampling strategy in which the selection of participants is guided that are emerging from the data analysis.

Standard Error

- Standard Error is the variability in scores that can be expected if measurements are made on random samples of the same size from the same populations. The standard error provides a framework within which a determination of the difference between groups may be made.

Variance

A measure of the spread of a variable about its mean value. In a data set, a single point lies above or below the mean for the entire dataset. A deviation score is the measure of how much each point lies above or below the mean for the entire dataset. It is a way to state how far deviations are from expected results.

Got Questions or Comments, please share!