Statistics Assignment 1

# What exactly is the difference between descriptive and inferential statistics?

Descriptive statistics provides information describing attributes of a dataset. This consists of organizing or summarizing of data. (eg. Measures of central tendency and measures of dispersion).

In inferential statistics, we make conclusions or inferences from the data by using some experiments (eg. Testing of hypothesis). Here we derive conclusions or inferences about the population by exploring the sample.

# I'm not sure what is the difference between a sample and a population?

Population is the whole or complete dataset of elements. Whereas sample is a subset from the population. A sample is mostly used to infer about the population as it is practical and cost-effective to do so. Therefore care must be taken to ensure the quality of the sample so that the generalizations derived accurately represent the population.

# What distinguishes descriptive statistics from other types of statistics?

Descriptive statistics provides information describing attributes of a dataset. This consists of organizing or summarizing of data. (eg. Measures of central tendency and measures of dispersion). Using descriptive statistics in our analysis is the basic step and unavoidable. This is the foundation that we use to understand a dataset. All other inferential or predictive analysis build up over descriptive statistics and comes subsequent to descriptive statistics. Even to effectively explore a dataset, we need to first look at its descriptive statistics.

# What is the difference between quantitative and qualitative data?

Quantitative data can be quantified, counted, or measured. It always has a numerical value. Quantitative can be either discrete or continuous.

Qualitative or categorical data cannot be quantified or measured. It could be either text, number, etc.

# What is the definition of a percentile?

A percentile is a value below which a certain percentage of observations lie.

Percentile rank of x = (# of values below x / sample size) \* 100