1. What is the difference between descriptive statistics and inferential statistics?

Descriptive statistic is the branch of statistics that is concerned with describing the population under study whereas inferential statistics focuses on drawing conclusions about the population on the basis of sample analysis and observation.

Descriptive statistics helps to organize, analyze and present data in a meaningful way whereas inferential statistics compares, test and predicts data.

Descriptive statistics is used to describe a situation. Inferential statistics is used to explain the chances of occurrence of an event.

There is a diagrammatic or tabular representation of final result in descriptive statistics whereas the final result is displayed in the form of probability.

Descriptive statistics explains the data, which is already known, to summarise sample. Conversely, inferential statistics attempts to reach the conclusion to learn about the population; that extends beyond the data available.

1. What is the difference between population and sample in inferential statistics?

 A **population**simply describes *all* of a particular group of individuals that are being studied whereas a sample is the subset of a population.

Population refers to all of the individuals that the study wants to describe. In a study where a sample of college students describe their eating habits, the population of interest may be all college students. Usually, the sample is some of the individuals who satisfy the certain criteria, while the population is all such individuals.

 The sample is only those individuals that give data. So, it is the people who actually responded to a poll, the people who participated in the study, or the objects which we actually measured. If no data is gathered from an individual, that individual is not in the sample.

1. Most common characteristics used in descriptive statistics

**Descriptive statistics** summarize and organize characteristics of a data set. A data set is a collection of responses or observations from a sample or entire population.

There are 3 essential characteristics of descriptive statistics – scales of measurement , measures of central tendency and measures of variability.

Scales of measurement consist of nominal ,ordinal ,interval and ratio scales. Measures of central tendency includes mean, median and mode.

Measures of variability consists of range , variance and standard deviation.

1. How to calculate range and interquartile range?

The range often is defined as the difference between the largest and smallest numbers of the data. Another measure of variability is the interquartile range. The interquartile range is the range of values between the first and third quartile. Essentially, it is the range of the middle 50% of the data and is determined by computing the value of Q3 - Q1. The interquartile range is especially useful in situations where data users are more interested in values toward the middle and less interested in extremes.

1. How is the statistical significance of an insight assessed?

Statistical significance of an insight can be accessed using Hypothesis Testing. Statistical significance is often calculated with statistical hypothesis testing, which tests the validity of a hypothesis by figuring out the probability that your results have happened by chance.

Here, a “hypothesis” is an assumption or belief about the relationship between your datasets. The result of a hypothesis test allows us to see whether this assumption holds under scrutiny or not.

A standard hypothesis test relies on two hypotheses.

* **Null hypothesis:** The default assumption of a statistical test that you’re attempting to disprove (e.g., an increase in cost won’t affect the number of purchases).
* **Alternative hypothesis:** An alternate theory that contradicts your null hypothesis (e.g., an increase in cost will reduce the number of purchases). This is the hypothesis you hope to prove.

The testing part of hypothesis tests allows us to determine which theory, the null or alternative, is better supported by data.