Assignment 2

1)What is Difference between inferential statistics and descriptive statistics?

Sr.No.	Descriptive Statistics	Inferential Statistics
1.	It gives information about raw data which describes the data in some manner.	It makes inference about population using data drawn from the population.
2.	It helps in organizing, analyzing and to present data in a meaningful manner.	It allows us to compare data, make hypothesis and predictions.
3.	It is used to describe a situation.	It is used to explain the chance of occurrence of an event.
4.	It explain already known data and limited to a sample or population having small size.	It attempts to reach the conclusion about the population.
5.	It can be achieved with the help of charts, graphs, tables etc.	It can be achieved by probability.

2)What is difference between population and sample in inferential statistics?

Population	Sample
The measurable quality is called a parameter	The measurable quality is called as statistics
The population is complete set	The sample is subset of the population
Reports are a true representation of option	Reports have margin of error and confidence interval
It contain all members of a specified group	It is a subset that represent the entire population

3)Most common Characteristics used in descriptive statistics?

a) Central Tendency

Measures of central tendency focus on the average or middle values of data sets, whereas measures of variability focus on the dispersion of data. These two measures use graphs, tables and general discussions to help people understand the meaning of the analyzed data.

Measures of central tendency describe the center position of a distribution for a data set. A person analyzes the frequency of each data point in the distribution and describes it using the mean, median, or mode, which measures the most common patterns of the analyzed data set.

b) Measures of Variability
Measures of variability (or the measures of spread) aid in analyzing how dispersed the distribution is for a set of data. For example, while the measures of central tendency may give a person the average of a data set, it does not describe how the data is distributed within the set.

4**) How to calculate range and interquartile range?**

Step 1: Order the data

In order to calculate the IQR, we need to begin by ordering the values of the data set from the least to the greatest. Likewise, in order to calculate the median, we need to arrange the numbers in ascending order (i.e. from the least to the greatest).

Step 2: Calculate the median

Next, we need to calculate the median. The median is the "center" of the data. If the data set has an odd number of data points, then the mean is the centermost number. On the other hand, if the data set has an even number of values, then we will need to take the arithmetic average of the two centermost values. We will calculate this average by adding the two numbers together and then dividing that number by two.

Step 3: Upper and lower medians

Once we have found the median of the entire set, we can find the medians of the upper and lower portions of the data. If the data set has an odd number of values, we will omit the median or centermost value of the set. Afterwards, we will find the individual medians for the upper and lower portions of the data.

Step 4: Calculate the difference

Last, we need to calculate the difference of the upper and lower medians by subtracting the lower median from the upper median. This value equals the IQR.

5) How is the statistics significance of an insight assessed?

First, state the null hypothesis and alternative hypothesis. Second, calculate the p-value, the probability of obtaining the observed results of a test assuming that the null hypothesis is true. Last, set the level of the significance (alpha) and if the p-value is less than the alpha then rejects the null — in other words, the result is statistically significant.