**Statistics Assignment 1**

**Name = Khushal Tondon**

**Ans no. 1**

Generally , Statistics is divided into two parts one is Descriptive statistics and second is Inferential statistics

**Descriptive statistics**

It consists of organising & summarising data .Here data means the information that been collected from an experiment etc .Basically graphical or pictorial display

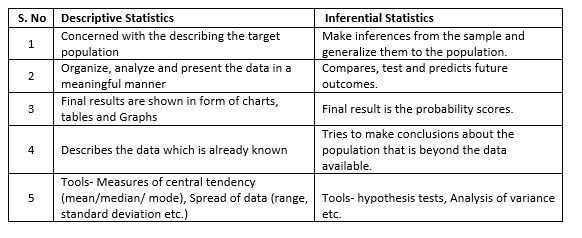
Examples Descriptive statistics describes data (for example, a chart or graph)

like average- mean , mode , median( Measures of central tendency ) , standard deviation , variance ( measure of dispersion )

**Inferential statistics**

Inferential statistics focus on making predictions or generalizations about a larger dataset, based on a sample of those data. Inferential statistics allow you to test a hypothesis or assess whether your data is generalizable to the broader population.

Inferential statistics allows you to make predictions (“inferences”) from that data.



**Ans no. 2**

**Population:**

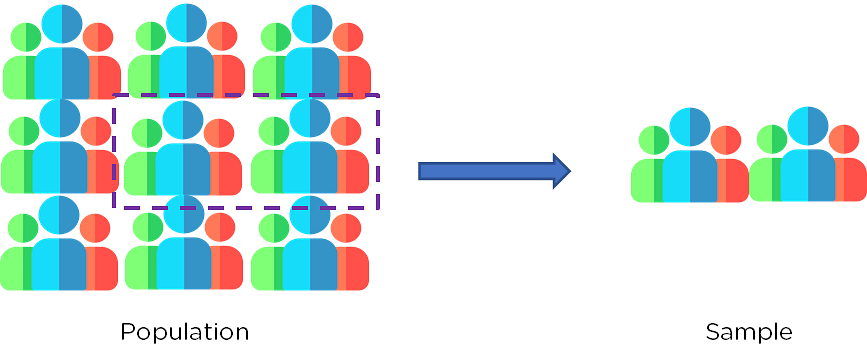
Population is the group that is targeted to collect the data from. Our data is the information collected from the population. Population is always defined first, before starting the data collection process for any statistical study. Population is not necessarily be people rather it could be batch of batteries, measurements of rainfall in an area or a group of people.

We generally denoted the population by N

**Sample :**

It is the part of population which can selected randomly for the study. The sample should be selected such that it represents all the characteristics of the population. The process of selecting the subset from the population is called sampling and the subset selected is called the sample.

Generally denoted by n

infer

And there are many sampling techniques like simple random sampling , stratified sampling , systematic sampling , Convenience sampling

**Ans no .3**

Generally , there are two types of statistics

1. Descriptive statistics
2. Inferential statistics
3. Descriptive Statistics :

Descriptive statistics is the type of statistics that probably springs to most people’s minds when they hear the word “statistics.” In this branch of statistics, the goal is to describe. Numerical measures are used to tell about features of a set of data. There are a number of items that belong in this portion of statistics, such as:

* The average, or measure of the center of a data set, consisting of the mean, median, mode, or midrange
* The spread of a data set, which can be measured with the range or standard deviation
* Overall descriptions of data such as the five number summary
* Measurements such as skewness and kurtosis
* The exploration of relationships and correlation between paired data
* The presentation of statistical results in graphical form

It is simply used for summarizing objects, etc. There are two categories in this as following below.

**(a). Measure of central tendency –**

Measure of central tendency is also known as summary statistics that is used to represents the center point or a particular value of a data set or sample set.

In statistics, there are three common measures of central tendency as shown below:

(i) Mean :

It is measure of average of all value in a sample set.

(ii) Median :

It is measure of central value of a sample set. In these, data set is ordered from lowest to highest value and then finds exact middle.

(iii) Mode :

It is value most frequently arrived in sample set. The value repeated most of time in central set is actually mode.

**(b). Measure of Variability –**

Measure of Variability is also known as measure of dispersion and used to describe variability in a sample or population. In statistics, there are two common measures of variability as shown below:

(i) Variance :

It simply describes how much a random variable defers from expected value and it is also computed as square of deviation.

(iii) Dispersion :

It is measure of dispersion of set of data from its mean.

**Ans no. 4**

Qualitative Data

Qualitative data is a set of information which can not be measured using numbers. It generally consist of words, subjective narratives. Result of an qualitative data analysis can come in form of highlighting key words, extracting information and concepts elaboration. For example, a study on parents perception about the current education system for their kids. The resulted information collected from them might be in narrative form and you need to deduce the analysis that they are satisfied, un-satisfied or need improvement in certain areas and so on .

Sometimes it also known as categorical data . like gender , color etc.

Quantitative Data

Quantitative data is a set of numbers collected from a group of people and involves statistical analysis.For example

If you conduct a satisfaction survey from participants and ask them to rate their experience on a scale of 1 to 5. You can collect the ratings and being numerical in nature, you will use statistical techniques to draw conclusions about participants satisfaction.

No. of bank account , no. of children in family , height , weight , river length etc.

We can further divide this data into discrete data and continuous data .

The main differences between quantitative and qualitative data is

* Quantitative data is countable or measurable, relating to numbers. Qualitative data is descriptive, relating to language.
* Quantitative data tells us how many, how much, or how often (e.g. “20 people signed up to our email newsletter last week”). Qualitative data can help us to understand the “why” or “how” behind certain behaviors, or it can simply describe a certain attribute—for example, “The postbox is red” or “I signed up to the email newsletter because I’m really interested in hearing about local events.”
* Quantitative data is fixed and “universal,” while qualitative data is subjective and dynamic. For example, if something weighs 20 kilograms, that can be considered an objective fact. However, two people may have very different qualitative accounts of how they experience a particular event.
* Quantitative data is gathered by measuring and counting. Qualitative data is collected by interviewing and observing.

**Ans no. 5**

Percentile is a value below which a certain percentage of observation lie .

Suppose I have this dataset

2 , 2 , 3 , 4 , 5 , 5 , 5 , 6 , 7 , 8 , 8 , 8 , 8 , 8 , 9 , 9 , 10 , 11 ,11 , 12

So in this if we want to find percentile rank of 10 so it will be 80 %

( percentile rank of 10 = no. of values below 10 / n \* 100 )

So it means 80 % of the entire distribution is less than 10