Introduction of Statistics and its Types

Statistics simply means numerical data, and is field of math that generally deals with collection of data, tabulation, and interpretation of numerical data. It is actually a form of mathematical analysis that uses different quantitative models to produce a set of experimental data or studies of real life. It is an area of applied mathematics concern with data collection analysis, interpretation, and presentation. Statistics deals with how data can be used to solve complex problems. Some people consider statistics to be a distinct mathematical science rather than a branch of mathematics. Statistics makes work easy and simple and provides a clear and clean picture of work you do on a regular basis.
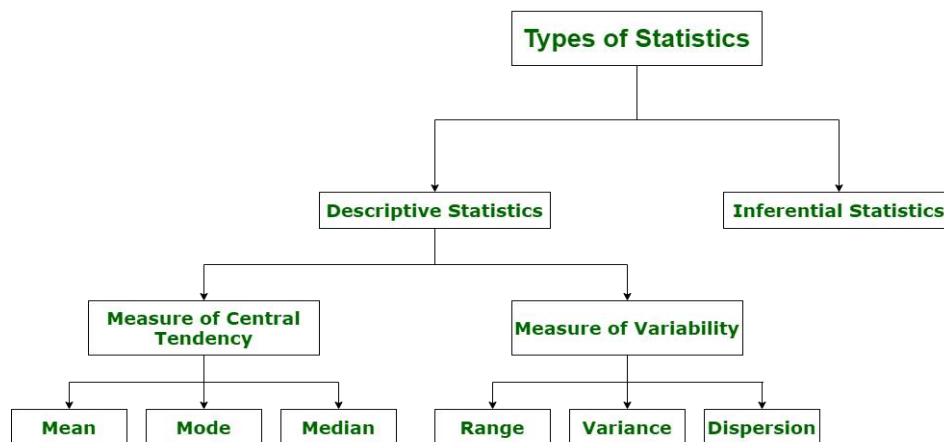
Basic terminology of Statistics:

- **Population –**
  It is actually a collection of set of individuals or objects or events whose properties are to be analysed.
- **Sample –**
  It is the subset of a population.

Types of Statistics:

1. Descriptive Statistics :

Descriptive statistics uses data that provides a description of the population either through numerical calculation or graph or table. It provides a graphical summary of data. 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 centre 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.
    For example,

Cars	Mileage	Cylinder
Swift	21.3	3
Verna	20.8	2
Santro	19	5

$$\text{Mean (m)} = \frac{\text{Sum of all the terms}}{\text{Total no. of terms}}$$

$$m = \frac{21.3 + 20.8 + 19}{3}$$

$$= 20.366$$

- **(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.

  For example,

Cars	Mileage	Cylinder
Swift	21.3	3
Verna	20.8	2
Santro	19	5
i 20	15	4

  Ordering the set from lowest to highest = 15   19   20.8   21.3

  $$\text{Median} = \frac{19 + 20.8}{2}$$

  $$\text{Median} = 23.5$$

- **(iii) Mode :**

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

  For example,

  2 3 4 2 4 6 4 7 7 4 2 4

  Mode = 4

- **(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 three common measures of variability as shown below:

  - **(i) Range :**

    It is given measure of how to spread apart values in sample set or data set.

    Range = Maximum value - Minimum value

  - **(ii) Variance :**

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

    $S^2 = \sum_{i=1}^{n} [(x_i - \bar{x})^2 \div n]$

    In these formula, **n** represent total data points, $\bar{x}$ represent mean of data points and x_i represent individual data points.

  - **(iii) Dispersion :**

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

    $\sigma = \sqrt{(1 \div n)} \sum_{i=1}^{n} (x_i - \mu)^2$

2. Inferential Statistics :

Inferential Statistics makes inference and prediction about population based on a sample of data taken from population. It generalizes a large dataset and applies probabilities to draw a conclusion. It is simply used for explaining meaning of descriptive stats. It is simply used to analyse, interpret result, and draw conclusion. Inferential Statistics is mainly related to and associated with hypothesis testing whose main target is to reject null hypothesis.

Hypothesis testing is a type of inferential procedure that takes help of sample data to evaluate and assess credibility of a hypothesis about a population. Inferential statistics are generally used to

determine how strong relationship is within sample. But it is very difficult to obtain a population list and draw a random sample.

Inferential statistics can be done with help of various steps as given below:

1. Obtain and start with a theory.
2. Generate a research hypothesis.
3. Operationalize or use variables
4. Identify or find out population to which we can apply study material.
5. Generate or form a null hypothesis for these population.
6. Collect and gather a sample of children from population and simply run study.
7. Then, perform all tests of statistical to clarify if obtained characteristics of sample are sufficiently different from what would be expected under null hypothesis so that we can be able to find and reject null hypothesis.

Types of inferential statistics –
Various types of inferential statistics are used widely nowadays and are very easy to interpret. These are given below:
- One sample test of difference/One sample hypothesis test
- Confidence Interval
- Contingency Tables and Chi-Square Statistic
- T-test or Anova
- Pearson Correlation
- Bi-variate Regression
- Multi-variate Regression

Difference between Descriptive and Inferential statistics

Statistics is a branch of mathematics dealing with the collection, analysis, interpretation, and presentation of masses of numerical data. It is basically a collection of quantitative data.
**Types of Statistics:**
- Theoretical Statistics
- Applied Statistics
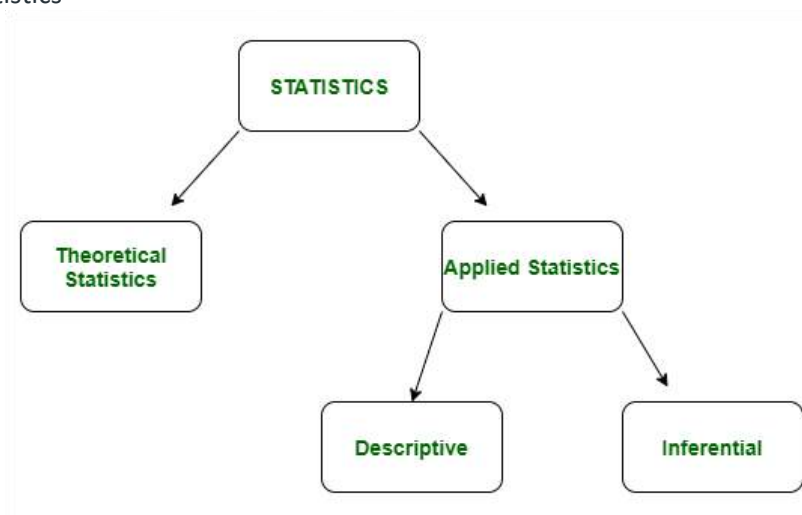
1. Descriptive Statistics: Descriptive statistics is a term given to the analysis of data that helps to describe, show and summarize data in a meaningful way. It is a simple way to describe our data. Descriptive statistics is very important to present our raw data ineffective/meaningful way using numerical calculations or graphs or tables. This type of statistics is applied to already known data.
Types of Descriptive Statistics:
- Measure of Central Tendency
- Measure of Variability
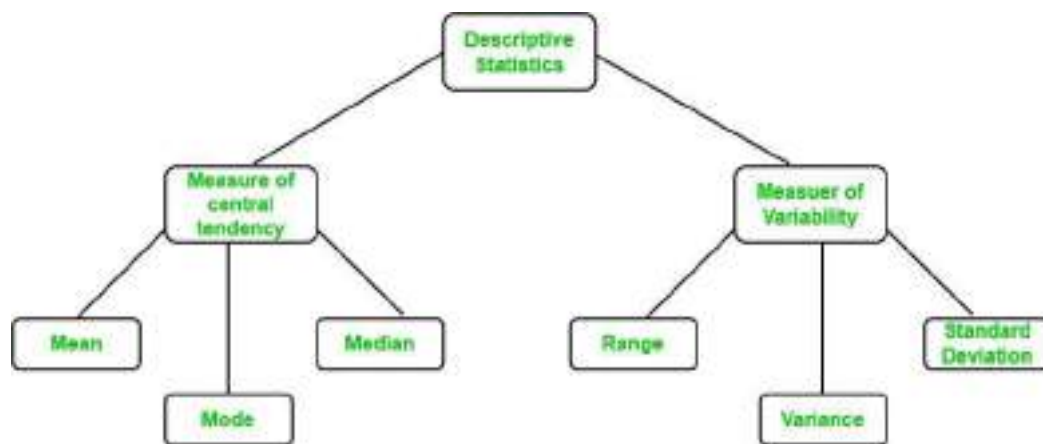
2. Inferential Statistics: In inferential statistics, predictions are made by taking any group of data in which you are interested. It can be defined as a random sample of data taken from a population to describe and make inferences about the population. Any group of data that includes all the data you are interested in is known as population. It basically allows you to make predictions by taking a small sample instead of working on the whole population.
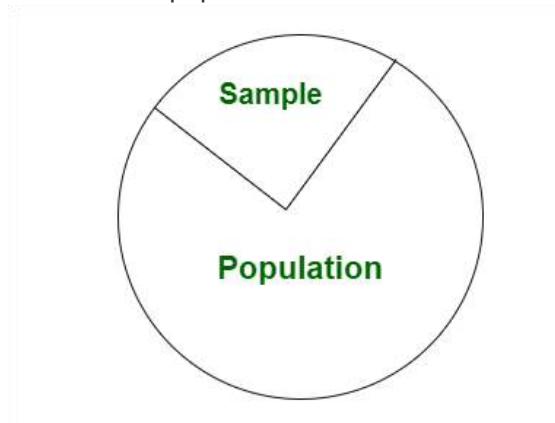
Difference between Descriptive and Inferential statistics:

S.No.	Descriptive Statistics	Inferential Statistics
1.	It gives information about raw data which describes the data in some manner.	It makes inferences about the population using data drawn from the population.
2.	It helps in organizing, analysing, and to present data in a meaningful manner.	It allows us to compare data, and make hypotheses and predictions.
3.	It is used to describe a situation.	It is used to explain the chance of occurrence of an event.
4.	It explains already known data and is limited to a sample or population having a 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.