In today’s sections, we will cover the below topics.

* What is statistics?
* Different types of statistics

**What is statistics?**

Statistics is the science of Collecting, Organizing, Analyzing the data.

**What is data?**

“Facts or pieces of information”

**Example**: Let us take the ages of students in a classroom as follows

{24, 25, 27, 32, 28….}

By using the above data, we can calculate Mean, Median, mode and standard deviation etc.

**Another example**: We can also that the weights of the students in a classroom

{45, 52, 70, 63, 80….}

**Types of Data:**

**We have 3 types of data.**

1. **Structured data –** Which consists of rows and columns. Ex: Tables, CSV data etc
2. **Unstructured data –** Images, videos, voice, sound, text
3. **Semi – structured data –** XML, JSON etc

* **Discrete variable - Whole numbers. Ex:** number of bank a/cs -1,2, number of children -1,2,3
* **Continuous variable – Decimal Numbers. Ex:** Height, weight, age, rainfall, speed
* **Nominal variable – Variable which is not having Order. Ex.** Male, female
* **Ordinal variable – which is having the order. Ex:** Education (From high school to PhD), ranking – 1, 2, 3 etc.

**Different types of statistics:** There are two types of statistics are there

1. Description statistics [EDA, FE]
2. Inferential statistics

**Description statistics:** It consists of organizing and summarizing the data. This statistics can be widely used in EDA (Exploratory Data Analysis) and FE (Feature Engineering).

**Example**: Histograms, bar charts, pie chart, distribution, candle stick, box plot, scatter plot etc.

**Another type is Inferential Statistics:**

It consists of collecting the sample data and making conclusions about population data using some experiments.

**Example**: Hypothesis testing (Imagination), Confidence Interval, P-value, Z-test, t-test, chi-square test, Annova or F-test.

Let us discuss with another example in detail.

Ex: Lets say there are 20 classrooms in a University and we collected the age of students in one classroom.

Ages = {21, 23, 25, 28, 22, 29, 26, 20 …etc}

**According to the Descriptive statistics**, we can calculate as follow questions.

* What is average age of the students in the classroom
* Relationship between age and weight etc.

**According to the inferential statistics**, we can calculate as follow questions.

* Are the average ages of the students in the classroom less than/greater than /equal to the average age of the students in the University?

Ex: Let’s Class A consists of 1000 students. 50% girls , 50% boys (Hypothesis testing)

Above all explanation in the below simple diagram for better understanding.