1. **NORMAL DISTRIBUTION:**
   It is about the symmetric distribution of the data about the mean, where mean=median=mode. The distribution of data forms a bell curve. It is such that 50% of the data is less than the mean and the other 50% of the data is more than the mean.
2. **Handling Missing Data:**
   Missing data reduces the statistical power of analysis. It reduces the validity of the results, which in turn might produce biased estimates. So the missing data should be handled in a way either by **removing the data** or by various **imputation techniques** available.
   Drop the columns with missing values. Imputation is adding the mean or mode of the values in the data set. Sometimes, it is also about filling the null values with the before or after values in the column.

3. **Imputation** of the missing data with the mean data is acceptable if it is confined to 10% or 20% of the data. If it exceeds the limit the predictions become biased.
4. **Linear Regression:** Simple linear regression is **used to model the relationship between two continuous variables**. Often, the objective is to predict the value of an output variable (or response) based on the value of an input (or predictor) variable.
5. **Branches of statistics:** There are 3 different branches of statistics, which include Data Collection, Descriptive Statistics and Inferential Statisics.