1. What is the difference between Data analysis and Machine learning?

Data analysis is a process of understanding the data, find patterns and try to obtain inferences due to which underlying patterns are observed, whereas Machine learning is to train a system to learn those patterns and try to predict the upcoming patterns.

1. What is Big Data?

Big Data is used to describe the massive volume of both structured and unstructured data that is so large in size. It is difficult to process using traditional techniques. Mobile apps, Emails, Documents are the few examples of Big Data.

1. What are the four main things we should know before studying Data analysis?

* Descriptive Statistics
* Inferential Statistics
* Distribution
* Hypothesis Testing

1. What are the common characteristics used in Descriptive statistics?

Descriptive Statistics are broken down into **measures of central tendency and measures of spread.** Measures of central tendency include the mean, median and mode, while Measures of spread include standard deviation, variance, minimum and maximum variables, kurtosis and skewness.

1. What is Quantitative and Qualitative Data?

Quantitative Data is the type of data whose value is measured in the form of numbers or counts, with a unique numerical value associated with each data set (eg: How many, How much, How often).Quantitative Data can be divided into **Discrete Data and Continuous Data**.

**Discrete data** is a type of data that consists of counting numbers only and as such cannot be measured.

Examples:

* + Number of students in a class
  + Number of days in a year
  + Age of an individual

**Continuous Data** is a data type that takes on numeric values that can be meaningfully broken down into smaller units. It can be placed on a measurement scale (Weight, Length, Time).

Example: Let us consider the cumulative Grade Point (CGPA) of students in a class, measured on a 5 point scale. A student can score any grade between 0 points to 5 points.

Qualitative Data is defined as the data that approximates and characterizes. It can be observed and recorded. This data type is non-numerical in nature (who, what, where).Qualitative can be divided into **Nominal and Ordinal**.

**Nominal Data** is a classification of categorical variables, that do not provide any quantitative value.

Examples:

* Gender
* Colour
* Marital status

**Ordinal Data** is a type of qualitative data where the variables have natural, ordered categories and the distances between the categories are not known.

Examples:

* Customer satisfaction on a variable scale – “satisfied, indifferent, dissatisfied”.
* Grade of students – “A, B, C, D”.