

ASSIGNMENT-1

1. What is difference between data analysis and machine learning?

Ans: Data Analysis is a process of understanding the data, find patterns and try to obtain inferences due to which the underlying patterns are observed. Machine Learning is when you train a system to learn those patterns and try to predict the upcoming pattern.

2. What is big data?

Ans: Big Data is used to describe the massive volume of both structured and unstructured data that is so large it is difficult to process using traditional techniques.

3 What are the four main things we should know before studying data analysis?

Ans: a) Descriptive statistics

b) Inferential statistics

c) Distribution

d) Hypothesis testing

4. Most common characteristics used in descriptive statistics?

Ans: a) Measure of central tendency (includes mean, median, mode)

b) Measure of spread (includes range, quartiles and the interquartile range, variance and standard deviation.)

5. What is quantitative data and qualitative data?

Ans: a) Quantitative Data:-

It is related with quantity. It Deals with numbers, data which can be measured.

It is divided in to two:-

- i. Discrete:-Means integers (ex:-number of students, how much runs in the match etc.)
- ii. Continuous:-Means varying thing (ex:-height, square feet, speed of vehicle, amount of time taken to complete a work, interval and ratio etc.)

b) Qualitative Data:-

It is related with quality. It deals with descriptions. Data that can be observed but cannot be measured. It is also divided in to two:-

- i. Nominal:-Means names (ex:-category, gender, hair colour, marital status etc.)
- ii. Ordinal:-means order or rank (ex:-first, second, grades, economic status etc.)