Assignment 1

1. What is the difference between Data Analysis and Machine Learning?

- Analytics relies on existing information to find patterns that ultimately shape decisions. Whereas machine learning leverages existing data that provides the base for the machine to learn for itself.
- Analytics reveals patterns through the process of classification and analysis while ML uses the algorithms to do the same as analytics but in addition, learns from the collected data.
- Data analytics ultimately aims to find patterns whereas ML aims to learn from data and make estimates and predictions.

2. What is a big data?

- Big data is a term that describes large, hard-to-manage volumes of data – both structured and unstructured – that inundate businesses on a day-to-day basis.
- It is not just the type or amount of data that's important, it's what organizations do with the data that matters.
- Big data can be analysed for insights that improve decisions and give confidence for making strategic business moves.

3. What are the four main things to know before data analysis?

The four main things to know before data analysis are,

- Descriptive Analysis
- Diagnostic Analysis
- Predictive Analysis
- Prescriptive Analysis

DESCRIPTIVE ANALYSIS

- Answer the question of "what happened"
- It will juggle with the raw data from multiple data sources to gain better insights.

DIAGNOSTIC ANALYSIS

- Historical data can be measured against another other data to answer the question.
- Diagnostic data analytics help answer why something occurred. Like the other categories, it too is broken down into two more specific categories: discover and alerts and query and drill downs.

PREDICTIVE ANALYSIS

- It tells what is likely to happens.
- It uses the findings of descriptive and diagnostic analytics to detect the tendency, clusters and expectations and to predict the future.

PRESCRIPTIVE ANALYSIS

- Prescriptive analytics is where AI and big data combine to help predict outcomes and identify what actions to take.
- It prescribes what action to take and to eliminate the future problem.

4.What are the four main characteristics in descriptive Statistics?

There are four major types of descriptive statistics,

Measures of Frequency:

- Count, Percent, Frequency
- Shows how often something occurs
- Use this when you want to show how often a response is given

Measures of Central tendency:

- Mean, Median, and Mode
- Locates the distribution by various points
- Use this when you want to show how an average or most commonly indicated response

Measures of Dispersion or Variation:

- Range, Variance, Standard Deviation
- Identifies the spread of scores by stating intervals
- Range = High/Low points
- Variance or Standard Deviation = difference between observed score and mean

Measures of Position:

- Percentile Ranks, Quartile Ranks
- Describes how scores fall in relation to one another. Relies on standardized scores
- Use this when you need to compare scores to a normalized score

5. What is Quantitative and Qualitative Analysis?

There are two types of data.

- Qualitative data
- Quantitative data

QUALITATIVE DATA: It is non-numerical data.

Eg: the texture of the skin, the color of the eyes, etc.

QUANTITATIVE DATA: Quantitative data is given in numbers. Data in the form of questions such as “how much”, “how many”, gives the quantitative data.

Eg: length, weight, speed, etc.