**ASSIGNMENT 1: DATA ANALYTICS**

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

|  |  |
| --- | --- |
| 1. Data analytics finds patterns and insights from data. 2. It discovers patterns for predicting future outcomes. 3. Human intervention is required. 4. Used in sectors like banking, sales, marketing, etc. | 1. Machine learning uses data to train and test algorithms. 2. It uses complex algorithms to predict the future outcomes. 3. No human intervention is required. 4. Real world applications include speech recognition, image recognition, etc. |

1. What is big data?

Big data is a combination of structured and unstructured data used in predictive modeling and other advanced analytics applications. It is characterized by 5 V’s:

1. Velocity
2. Volume
3. Veracity
4. Value
5. Variability
6. What are the four main things we should know before studying data analysis?
7. Defining the right question.
8. Collecting data from different sources.
9. Categorize data into qualitative and quantitative.
10. Cleaning the data.
11. What is the difference between inferential and descriptive statistics?

|  |  |
| --- | --- |
| 1. Inferential statistics compares, test and predicts future outcomes. 2. Final result is the probability score. 3. Draws conclusion based on data observed in samples. | 1. Descriptive statistics organize, analyze and present data in meaningful way. 2. Final results are in the form of graphs and charts. 3. Summarizes data. |

1. What is the difference between population and sample in inferential statistics?

|  |  |
| --- | --- |
| 1. Population is a complete set of data. 2. Example: Students in a school. 3. The measurable quality is called as parameter. | 1. Sample is a subset of population. 2. Example: Students of age greater than 14. 3. The measurable quality is called as statistics. |