**DMPM Lab Assignment -1**

**Data Exploration & Visualization**

1. Read the file **"pva97nk.csv"** that is supplied to you.

2. Identify the variables in the file "pva97nk.csv" and determine whether any variable has any missing values.

3. Impute some of the variables that have missing values using their corresponding mean values. Verify whether your task has been correctly done.

4. Compute the Kurtosis and Skewness of the variables and interpret the results obtained.

5. Determine the "summary" information for the numerical variables.

6. Identify the "distributions" of the numerical variables and plot the distributions.

7. Transform the numeric variables into their natural log values and scale [0 - 1] values.

8. Check whether the numeric variables follow normality conditions.

9. Find the correlation matrix for all the variables in the dataset and plot the graph of the correlation matrix.

10. From the given dataset partition the data into 70-15-15 divisions so to construct the training, validation and test datasets.

11. **Any additional ways of Data Exploration & Visualization will be highly appreciated.**