**Name:- Abhishek Solanke**

**Roll NO :- 281013**

**PRN:- 22310235**

**Assignment-02**

Assignment 2: Data Summary, Visualization & Classification in Python

Title:

Exploratory Data Analysis and Classification using Python

Aim:

To compute summary statistics, visualize distributions, clean and transform data, and apply a classification model on the dataset.

Course Outcome:

CO2 L1 – Understand the importance of data exploration, cleaning, and apply classification models.

CODE- From Visual Studio code.

**Explanation of Steps:**

| **Step** | **Description** |
| --- | --- |
| a) | Calculated mean, std, min, max, range, variance, and percentiles. |
| b) | Plotted histograms to show distribution of features like GRE, TOEFL, etc. |
| c) | Cleaned missing values, standardized features, transformed target for classification. |
| d) | Built a **Decision Tree Classifier** to predict student admission. |

**Conclusion:**

* Summary statistics provided a solid overview of feature distributions.
* Histograms revealed skewness and spread of values.
* Decision Tree effectively classified whether a student would get admitted based on GRE, GPA, and other scores.
* Confusion matrix and classification metrics showed performance (Accuracy, Precision, Recall, F1 Score).