

# **ASSIGNMENT NO: 5**

## **Problem Statement -**

Visualize the data using Python by plotting the graphs for assignment no. 1 and 2. Consider a suitable data set.

a) Use Scatter plot, bar plot, Box plot and Histogram.

## **S/W Packages and Libraries used:**

For the following assignment, the interpreter used was Google Collab and the Primary Library used was

- Seaborn: Seaborn is a statistical data visualization library in Python that provides a high-level interface for creating attractive and informative visualizations. It works well with Pandas DataFrames and offers a variety of plot types and customization options.

## **Theory-**

- Data Selection:
  - Choose a suitable dataset that contains numerical and categorical variables, allowing for the creation of different types of plots such as scatter plots, bar plots, box plots, and histograms.
- Data Preprocessing:

- Load the dataset into Python using Pandas.
- Perform any necessary data cleaning and preprocessing steps, such as handling missing values or encoding categorical variables.
- Visualization:
  - Utilize the Seaborn library for creating visualizations.
  - Create scatter plots to visualize the relationship between two numerical variables

## **Applications:**

- Exploratory Data Analysis (EDA): Visualization techniques provide insights into the underlying structure and patterns within the data, aiding in understanding relationships and identifying trends.
- Data Communication: Visualizations serve as powerful tools for communicating findings and insights to stakeholders or decision-makers clearly and concisely.

## **Limitations:**

- Data Quality: Visualizations are only as good as the quality of the underlying data. Inaccurate or incomplete data may lead to misleading visualizations and erroneous conclusions.

- Interpretation: While visualizations can highlight patterns and relationships in the data, they may not always imply causation. Care must be taken to interpret visualizations within the context of the data and domain knowledge.

## **Conclusion:**

By following the outlined methodology and utilizing the Seaborn library, you can effectively visualize the selected dataset using scatter plots, bar plots, box plots, and histograms. These visualizations will help in exploring relationships, understanding distributions, and identifying patterns within the data, thereby facilitating further analysis and decision-making processes.