

Name: Rushikesh Kolekar

Roll No: 282020

Batch: B3

Practical 3

Problem Statement:

Visualize the data using R/Python by plotting graphs for assignments 1 and 2. Select a suitable dataset and generate the following visualizations: Scatter plot, Bar plot, Box plot, Pie chart, and Line chart.

Objectives:

1. To introduce and explore basic data visualization techniques in Python using Seaborn and Matplotlib.
2. To demonstrate the use of different plot types, including Scatter plot, Bar plot, Box plot, Pie chart, and Line chart.
3. To analyze a dataset using various visualizations for enhanced insights.

Resources Used:

Software:

- Visual Studio Code
- Anaconda (Jupyter Notebook) **Libraries:**

- Pandas
- Matplotlib
- Seaborn **Theory:**

1) Seaborn:

Seaborn is a high-level Python visualization library built on Matplotlib. It simplifies the creation of informative and visually appealing statistical graphics, integrating seamlessly with Pandas DataFrames.

Key Features:

- Built-in themes and color palettes.
- Direct integration with Pandas for DataFrame compatibility.
- Enhanced default plots, including categorical and regression plots.

2) Matplotlib:

Matplotlib is a widely used Python library for generating static, animated, and interactive visualizations. It provides a flexible and extensive interface for creating various types of plots.

Key Features:

- Supports multiple plot types, including line, scatter, bar, and pie charts.
- Highly customizable for precise control over figure aesthetics.
- Works seamlessly with NumPy and Pandas.

Methodology:

For this assignment, data will be visualized using the following plot types:

1. **Bar Plot:** Used to represent categorical data with rectangular bars, where the length of each bar corresponds to its value.
2. **Scatter Plot:** Displays the relationship between two continuous variables, with each point representing a data value on the x and y axes.
3. **Box Plot:** Illustrates the distribution of data, highlighting the median, quartiles, and outliers, making it useful for understanding data spread.
4. **Pie Chart:** Represents proportions of a whole, where each segment indicates a category's contribution to the total.
5. **Line Chart:** Shows trends over time or across continuous data points, making it ideal for visualizing time series data.

Conclusion:

Data visualization is an essential component of data analysis and communication. By leveraging Python libraries like Seaborn and Matplotlib, we can generate diverse visualizations that simplify complex datasets. Scatter plots, bar plots, box plots, pie charts, and line charts each offer unique insights into data patterns. Effective visualization not only enhances data interpretation but also supports better decision-making and facilitates clear result sharing.