Name: Rujuta Bhanose

Roll No: 281042

Batch: A2

Assignment 3

Problem Statement:

Visualize the data using R/Python by plotting the graphs for assignment no. 1 and 2. Consider a suitable dataset. Use a Scatter plot, Bar plot, Box plot, Pie chart, and Line Chart.

Objectives:

1) To introduce and explore basic visualization techniques in Python using **Seaborn** and **Matplotlib**.
2) To demonstrate how to visualize data using different plot types, including Scatter plot, Bar plot, Box plot, Pie chart, and Line chart.
3) To analyze a suitable dataset using the various plot types for better insights.

Resources used:

1) Software used: Visual Studio Code

2) Libraries used: Pandas, Matplotlib, Seaborn

Theory:

- **Seaborn**: Seaborn is a powerful Python visualization library based on Matplotlib, designed to simplify the process of creating informative and attractive statistical graphics. It provides a high-level interface for drawing various types of plots with simple functions, while also integrating with Pandas DataFrames for seamless data handling.

Key features:

- Built-in themes and color palettes.
- Integration with Pandas for DataFrame compatibility.
- More attractive default plots, including categorical plots and regression plots.
- **Matplotlib**: Matplotlib is one of the most widely used Python libraries for creating static, animated, and interactive visualizations. It provides a low-level interface for creating basic plots and offers extensive customization options.

Key features:

- Extensive support for various plot types such as line, scatter, bar, and pie charts.
- Highly customizable, allowing fine control over figure aesthetics.
- Integration with NumPy and Pandas.

Methodology:

For this assignment, we'll visualize data using the following plot types:

1. **Bar Plot**: Bar plots are used to represent categorical data with rectangular bars, where the length of each bar is proportional to the value it represents.
2. **Scatter Plot**: A scatter plot is used to display the relationship between two continuous variables. Each point represents a data point on the x and y axes.
3. **Box Plot**: Box plots display the distribution of a dataset, highlighting the median, quartiles, and any potential outliers. They are useful for understanding the spread of data.
4. **Pie Chart**: Pie charts are used to represent the proportions of a whole. Each segment of the pie represents a category's contribution to the total.
5. **Line Chart**: Line charts show trends over time or continuous data points. It is useful for visualizing time series data.

Conclusion:

Data visualization is a key aspect of data analysis and communication. Using libraries like **Seaborn** and **Matplotlib** in Python, we can easily create a variety of visualizations to make

sense of complex datasets. The Scatter plot, Bar plot, Box plot, Pie chart, and Line chart each serve a specific purpose and provide unique insights into the data. Proper visualization not only makes it easier to interpret the data but also aids in decision-making and sharing results with others effectively.