**Name:** Apurva Jangle

**Roll no:** 282047

**Batch:** B2

**ASSIGNMENT 3**

**Problem Statement:**

**Visualize the data using R/Python by plotting the graphs for assignment no. 1 and 2. Consider**

**suitable data set. Use Scatter plot, Bar plot, Box plot, Pie chart, Line Chart.**

**Objective:**

The main objective is to visualize data using Python by plotting different types of graphs such as **Scatter Plot**, **Bar Plot**, **Box Plot**, **Pie Chart**, and **Line Chart**. By doing so, we aim to analyze patterns, relationships, and distributions in the dataset effectively. The visualization will be applied to datasets relevant to **assignment no. 1 and 2** to provide clear insights and enhance the interpretability of the data.

**S/W Packages and H/W Apparatus Used:**

**Software Used:**

1. Python 3.x
2. Google Colab / Jupyter Notebook

**Libraries and Packages Used:**

* **NumPy**: For numerical computations and array operations.
* **Pandas**: For data manipulation and preprocessing.
* **Matplotlib**: For creating static, interactive, and publication-quality visualizations.
* **Seaborn**: For advanced data visualization with an aesthetic appeal.

**Graph Types and Their Significance:**

**1. Scatter Plot:**

* **Purpose**: Displays relationships between two variables.
* **Use Case**: Analyze trends or patterns in a bivariate dataset.
* **Example**: Plotting age vs. income to see how they correlate.

**2. Bar Plot:**

* **Purpose**: Compares categorical data.
* **Use Case**: Compare sales performance across different regions or products.
* **Example**: Display total sales of different product categories.

**3. Box Plot:**

* **Purpose**: Shows the distribution of a dataset and identifies outliers.
* **Use Case**: Understand variations in exam scores or monthly expenses.
* **Example**: Visualize spending habits among different age groups.

**4. Pie Chart:**

* **Purpose**: Displays proportional data as segments of a circle.
* **Use Case**: Illustrate the percentage contribution of categories to a whole.
* **Example**: Show the percentage share of different income brackets in a population.

**5. Line Chart:**

* **Purpose**: Displays trends over a continuous interval or time.
* **Use Case**: Monitor changes in sales or website traffic over months.
* **Example**: Visualize stock price movements over time.

**Application:**

**Scatter Plot:**

* Used to visualize relationships between two variables like **Age** and **Spending Score** to determine if there's a correlation between the two.

**Bar Plot:**

* Used to compare data like the **number of students in different departments** or **sales of various product categories**.

**Box Plot:**

* Helpful in identifying **outliers** in numerical datasets like monthly income or scores. It provides insights into the range and distribution of data.

**Pie Chart:**

* Used to represent the **proportion of various expense categories** or the **market share of competing brands** in an industry.

**Line Chart:**

* Best suited for analyzing **trends over time**, like website traffic or annual sales figures, to observe seasonal effects or general trends.

**Limitations:**

* Data visualization heavily depends on the quality and structure of the dataset.
* Inappropriate choice of chart type may lead to misinterpretation of the data.
* Large datasets might require advanced visualization tools for better clarity and scalability.

**Working of Visualization:**

1. **Dataset Selection**: Choose suitable datasets for assignments 1 and 2. Ensure the data is clean and formatted correctly.
2. **Library Setup**: Import the required Python libraries for data manipulation and plotting.
3. **Graph Plotting**:
   * Use **Scatter Plot** for analyzing relationships.
   * Use **Bar Plot** for comparing categorical data.
   * Use **Box Plot** for identifying outliers and visualizing distributions.
   * Use **Pie Chart** for proportion analysis.
   * Use **Line Chart** for trend analysis.
4. **Interpret Results**: Analyze each graph and derive meaningful insights.

**Conclusion:**

In conclusion, data visualization is a powerful tool for analyzing datasets and presenting insights in an intuitive manner. By using **Scatter Plots, Bar Plots, Box Plots, Pie Charts, and Line Charts**, we can better understand the relationships, patterns, and distributions in the data from assignment no. 1 and 2. This process enhances decision-making and provides a strong foundation for further analysis or predictive modeling.