ASSIGNMENT NO. 5

Title: Visualize the data using R/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

OR

b) Perform the data visualization operations using Tableau for the given dataset.

Software/Libraries Used:

- Python
- Matplotlib
- Seaborn

Theory/Methodology:

Data visualization is a critical aspect of data analysis, allowing us to explore relationships, patterns, and distributions within the data. In this practical, we'll utilize Python libraries like Matplotlib and Seaborn to create visualizations such as scatter plots, bar plots, box plots, and histograms.

Advantages:

- Matplotlib and Seaborn offer a wide range of customizable plotting options.
- They provide high-quality, publication-ready visualizations.
- Interactive features can enhance the exploration and presentation of data.

Limitations/Examples:

- Matplotlib's syntax can be verbose for complex plots.

- Seaborn may not offer as much flexibility in customization compared to Matplotlib.

- Both libraries may struggle with very large datasets or complex visualizations.

Working/Algorithm::

1. Scatter Plot: Display the relationship between two numerical variables using points on a 2D plane.

2. Bar Plot: Represent categorical data with rectangular bars, where the length of each bar corresponds to the value of a variable.

3. Box Plot: Visualize the distribution of a numerical variable through quartiles, median, and outliers.

4. Histogram: Display the distribution of a numerical variable by dividing the data into intervals (bins) and counting the number of observations in each bin.

Conclusion:

By utilizing Python libraries like Matplotlib and Seaborn, we can create informative visualizations to explore and communicate insights from the data effectively. Each type of plot offers unique advantages in representing different aspects of the data, allowing for a comprehensive analysis. However, it's essential to choose the appropriate visualization techniques based on the nature of the data and the insights we aim to convey.