# Assignment 3 - Data Visualization using Python

Name: Dhananjay Siddhanath Jagtap

Roll No: 282057

Batch: B3

Assignment: 3

## Statement

Q. In this assignment, we have to visualize the data using R/Python by plotting the graphs for Assignment No. 1 and 2. Consider a suitable dataset. Use the following types of graphs:  
- Scatter Plot  
- Bar Plot  
- Box Plot  
- Pie Chart  
- Line Chart

## Objective

1. To learn how to visualize data using Python's visualization libraries such as Matplotlib and Seaborn.  
2. To understand different types of plots and when to use them for data interpretation.  
3. To extract insights and trends from data visually, helping to make informed decisions based on analysis.

## Resources Used

1. Software Used: Jupyter Notebook  
2. Libraries Used:  
 - matplotlib  
 - seaborn  
 - pandas  
 - numpy

## Introduction to Data Visualization

Data visualization is the graphical representation of information and data. Python offers several libraries like matplotlib and seaborn that make it easier to visualize and understand data. Visualization helps in identifying patterns, trends, and outliers in data. Commonly used visualizations include Scatter Plot, Bar Plot, Box Plot, Pie Chart, and Line Chart.

## Some Basic Functions Used

1. sns.scatterplot() – For scatter plots.  
2. sns.barplot() or sns.countplot() – For bar charts.  
3. sns.boxplot() – For displaying distribution and outliers.  
4. plt.pie() – For pie charts.  
5. plt.plot() – For line charts.  
6. df.describe() – To view summary statistics of data.  
7. df['column'].value\_counts() – To get count of unique values.

## Methodology

1. Data Loading: Load a suitable dataset (e.g., Iris or Titanic) using pandas.  
2. Data Preprocessing: Handle missing values if any. Select relevant features for visualization. Understand data types and structure.  
3. Plotting Graphs: Create and customize each required graph using Python libraries. Add titles, labels, legends, and color coding for clarity.  
4. Analysis: Interpret each visualization to understand data insights, such as trends, distributions, or relationships.

## Advantages

1. Makes data easy to understand and interpret.  
2. Enhances decision-making by revealing patterns.  
3. Visually communicates data effectively to non-technical stakeholders.

## Disadvantages

1. Misleading visualizations can cause wrong interpretations.  
2. Some complex relationships in data may not be easily captured in simple charts.

## Conclusion

This assignment provided hands-on experience with data visualization using Python. We explored different plot types and understood their application in analyzing real-world datasets. These visual tools are essential in data analysis workflows, helping turn raw data into meaningful insights. The practice gained from this assignment lays a strong foundation for more advanced visual analytics tasks.