**Name:** Shubhankar D. Jakate

**Roll no:** 282019 **PRN:** 22310371

**Class :** SY CSE AI **Batch:** B1

**Assignment 3: Data Visualization**

**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.

**Task Overview:**   
In this assignment, we visualize the **Heart Disease Dataset** using Python. The key visualization techniques include:

1. **Scatter Plot** – To analyze relationships between two numerical variables.
2. **Bar Plot** – To compare categorical data.
3. **Box Plot** – To visualize data distribution and identify outliers.
4. **Pie Chart** – To represent categorical proportions.
5. **Line Chart** – To show trends in numerical data over time or ordered sequences.

**Objective**

1. Gain hands-on experience with **Matplotlib** and **Seaborn** for data visualization.
2. Understand the importance of visual analysis in extracting meaningful insights.
3. Explore different types of plots and their applications in analyzing health-related datasets.

**Tools and Resources**

* **Software Used**: Google Colab / Jupyter Notebook
* **Libraries Used**: Pandas, Matplotlib, Seaborn

**Key Functions Used**

1. **Scatter Plot**: *Used to examine the correlation between Age and Cholesterol levels.*
2. **Bar Plot**: *Helps compare average blood pressure levels between males and females.*
3. **Box Plot** : *Used to analyze cholesterol level distribution and detect outliers.*
4. **Pie Chart**: *Displays the proportion of patients diagnosed with heart disease.*
5. **Line Chart**: *Shows trends in maximum heart rate with age progression.*

**Methodology**

1. **Loading the Data**
   * Read the **Heart Disease Dataset** using Pandas.
   * Check for missing values and clean data if necessary.
2. **Visualization Process**
   * Choose appropriate graphs to represent relationships and trends.
   * Use Matplotlib and Seaborn to generate various plots.
   * Customize graphs with labels, titles, and colors for better readability.

**Advantages of Data Visualization**

1. **Simplifies Complex Data** – Helps in understanding patterns and trends.
2. **Detects Outliers** – Box plots highlight extreme values.
3. **Enhances Data Interpretation** – Scatter and bar plots reveal relationships between variables.
4. **Improves Decision-Making** – Pie charts provide an overview of proportions in categorical data.

**Challenges:**

* Selecting the right visualization for different data types.
* Handling large datasets efficiently for plotting.

**Conclusion**

This assignment provided a hands-on approach to data visualization using Python. I explored:

* Different types of plots for analyzing heart disease data.
* The importance of graphical representation in data analysis.
* Practical applications of visualization techniques in healthcare datasets.