**Name:** Janhvi Katakdhond

**Roll no:** 282067

**Batch:** B3

**ASSIGNMENT 3**

**Problem Statement:**

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

**Objective:**

The objective of this assignment is to explore and understand the underlying patterns, trends, and relationships within a selected dataset through data visualization techniques. Using R or Python, various types of plots and graphs will be created to effectively represent the data, enhance interpretability, and provide meaningful insights.

To achieve this, I’ve collected two distinct datasets : anxiety attack dataset and admission dataset.

**S/W Packages and H/W apparatus used:**

Software used:

1. Python 3.x

2. Google Colab

Libraries and packages used: NumPy, Pandas

**Theory:**

**Data visualization** is a fundamental step in the data analysis process, allowing analysts and researchers to represent complex data in a graphical format. Visual tools such as charts, graphs, and plots help simplify large datasets, making it easier to identify patterns, trends, anomalies, and relationships among variables.

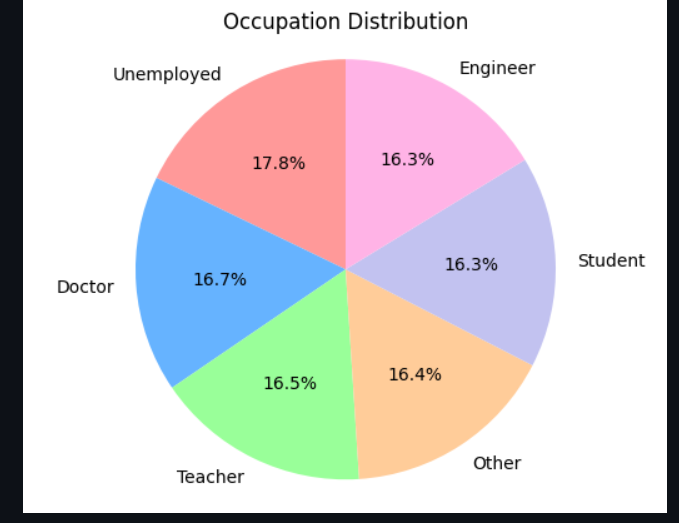
**Importance of Data Visualization**

1. **Simplifies Complex Data**: Large datasets often contain numerous variables that can be hard to interpret in raw tabular form. Visualizations condense this complexity into understandable formats.
2. **Identifies Patterns and Trends**: Graphs can reveal insights such as increasing or decreasing trends, clusters, or outliers that might not be obvious through numerical analysis.
3. **Supports Better Decision-Making**: By presenting data visually, stakeholders can make more informed decisions based on clear evidence.
4. **Enhances Communication**: Visuals are easier to understand for non-technical audiences, making it a powerful communication tool in reports and presentations.

**Common Visualization Techniques Used**

* **Bar Charts**: For categorical comparisons (e.g., gender-wise anxiety).
* **Histograms**: For distribution analysis (e.g., GPA scores).
* **Scatter Plots**: To explore relationships between two numeric variables (e.g., stress level vs. anxiety frequency, or test score vs. admission chance).
* **Box Plots**: To detect outliers and show distributions.
* **Pie Charts**: For showing proportions (e.g., percentage of students admitted).

Diagrams:



*Fig 1. Pie chart of Anxiety Attack Dataset across various occupations*

**Conclusion:**

Data visualization plays a critical role in transforming raw data into meaningful insights. Through the use of various graphs and plots, this assignment provided a visual exploration of the Anxiety Attack and Admission datasets. The visualizations helped uncover patterns such as correlations between stress levels and anxiety, and the impact of academic scores on admission outcomes. By interpreting these patterns, we gain a deeper understanding of the underlying factors affecting mental health and admission decisions. Overall, data visualization not only aids in data analysis but also enhances the clarity, communication, and effectiveness of data-driven conclusions.