Lab Assignment 21

Student Name: Chauhan Vandana Ramdayal

Student Id: AF0411629

Topic: Data Visualisation

What is Data Visualization?

Data visualization translates complex data sets into visual formats that are easier for the human brain to comprehend. This can include a variety of visual tools such as:

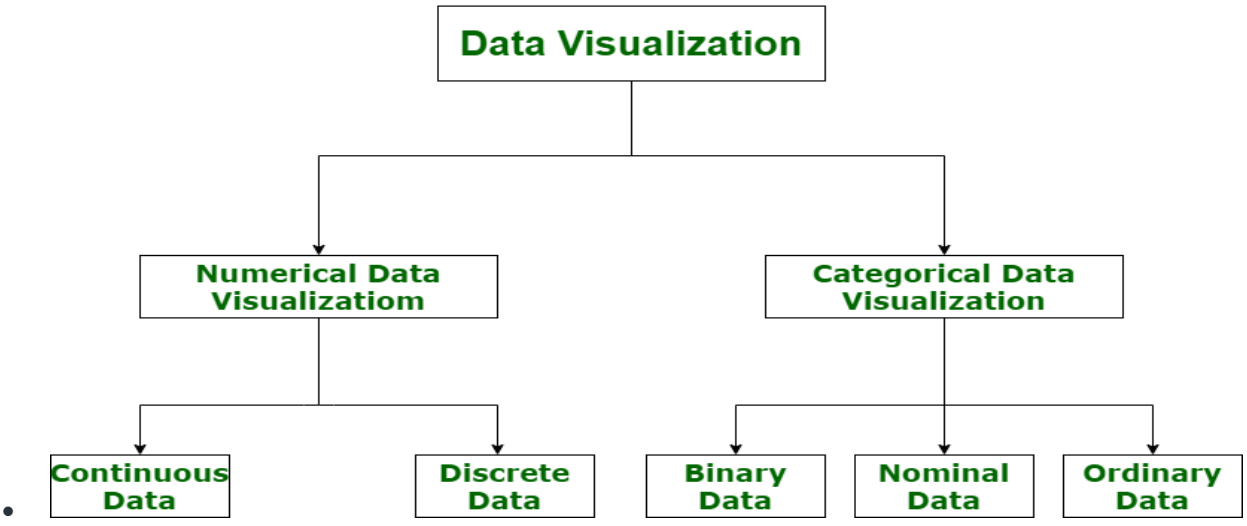
- **Charts**: Bar charts, line charts, pie charts, etc.
- **Graphs**: Scatter plots, histograms, etc.
- **Maps**: Geographic maps, heat maps, etc.
- **Dashboards**: Interactive platforms that combine multiple visualizations.

The primary goal of data visualization is to make data more accessible and easier to interpret, allowing users to identify patterns, trends, and outliers quickly. This is particularly important in the context of big data, where the sheer volume of information can be overwhelming without effective visualization techniques.

Types of Data for Visualization

Performing accurate visualization of data is very critical to market research where both numerical and categorical data can be visualized, which helps increase the impact of insights and also helps in reducing the risk of analysis paralysis. So, data visualization is categorized into the following categories:

- Numerical Data
- Categorical Data
  - Let's understand the visualization of data via a diagram with its all categories.



Question:

1. Create sample line chart using matplotlib

Code:

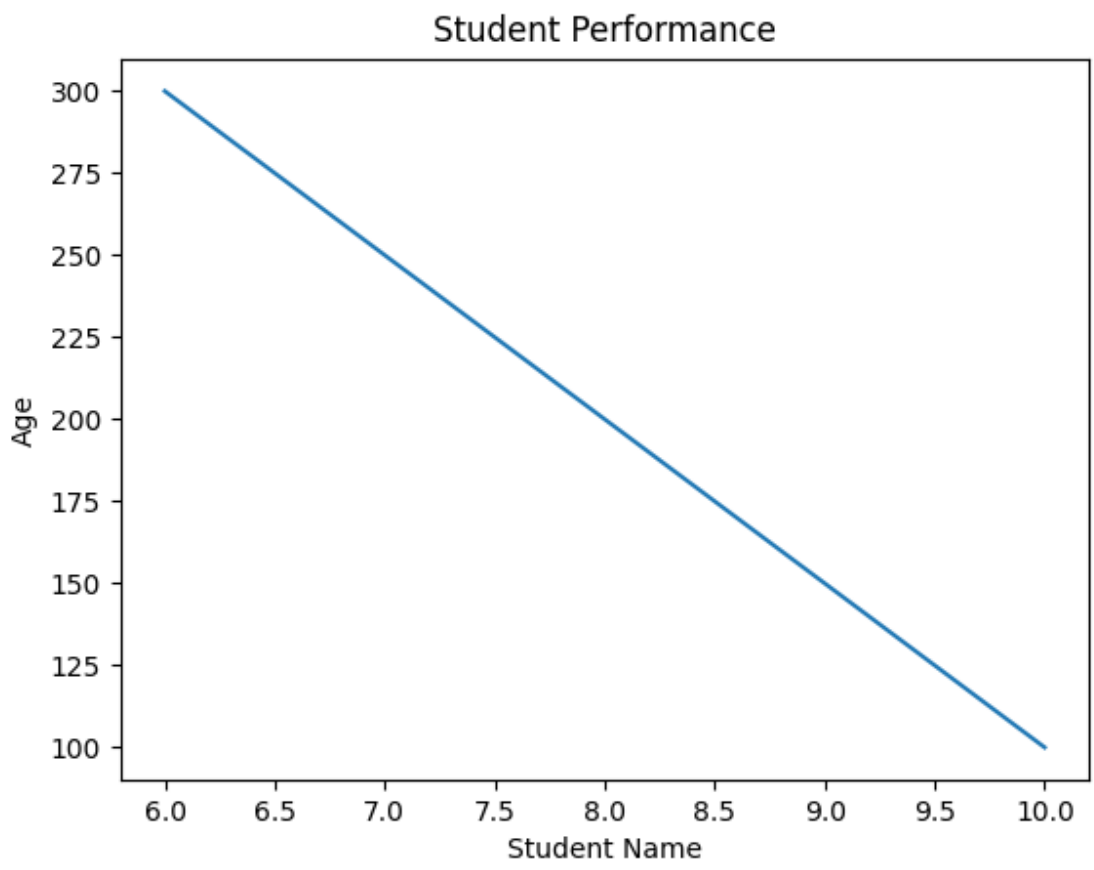
```python
import matplotlib.pyplot as pit
import numpy as np

# X and Y points
xpoints = np.array([10, 6])  # X coordinates
ypoints = np.array([100, 300])  # Y coordinates

# Plot xpoints and ypoints
pit.plot(xpoints, ypoints)  # Pass xpoints and ypoints as separate arguments

# Add labels and title
pit.xlabel("Student Name")  # Label for x-axis
pit.ylabel("Age")  # Label for y-axis
pit.title("Student Performance")  # Title of the plot
pit.show()
```

Output:



2. Create Bar Chart using Sample data of employee salary report with different bar colors, data labels

Code:

```python
import pandas as pd
import matplotlib.pyplot as plt
# Load data from the Excel file
file = pd.read_excel('Data1.xlsx')  # Ensure this file is in the same directory
# Get the employee names and salaries
xpoints = file['Empname']  # Employee names
ypoints = file['Salary']    # Corresponding salaries
print(ypoints)  # Print the salary data for verification
colors = ['#ffca33', '#900C3F', '#db159f', '#38ecdb', '#dd1545', '#8215dd'] # Define colors for the bars
barplot = plt.bar(xpoints, ypoints, color=colors) # Create the bar plot
plt.bar_label(barplot, labels=ypoints, label_type='center') # Add labels on top of each bar
# Set x and y labels
plt.xlabel("Employee Name")
plt.ylabel("Salary")

# Show the plot
plt.title("Employee Salary Report")
plt.show()
```

Output:

```
0    40000
1    60000
2    50000
3    35000
4    40000
5    25000
```