Task 3 – Data Visualization

Introduction

Data visualization is the graphical representation of information and data.

In this task, visualizations are created using Python libraries such as Matplotlib and Seaborn to understand patterns and trends from a sample dataset.

Objectives

1. Load a dataset.
2. Clean and preprocess the dataset.
3. Create different visualizations including bar charts, line graphs, and histograms.
4. Interpret results.

Tools Used

1. Python
2. Pandas
3. Matplotlib
4. Seaborn

Dataset

A sample dataset containing Sales, Month, Profit, and Category was used for visualization.

Code Used

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

data = {
"Month": ["Jan","Feb","Mar","Apr","May","Jun"],
"Sales": [12000,15000,18000,17000,22000,25000],
"Profit": [3000,4000,4500,4200,5000,6000]
```

```
}

df = pd.DataFrame(data)

plt.figure(figsize=(8,5))
plt.plot(df["Month"], df["Sales"])
plt.title("Monthly Sales Trend")
plt.savefig("sales_trend.png")

plt.figure(figsize=(8,5))
sns.barplot(x=df["Month"], y=df["Profit"])
plt.title("Profit by Month")
plt.savefig("profit_barplot.png")
```

Output

Two visualizations were generated:

1. Line chart showing monthly sales trend.

2. Bar chart showing monthly profit.

Conclusion

Data visualization helps in identifying key business insights.

The charts clearly show a growth in sales and profit over six months.