Mini Data Analysis Project Report

Project Title: Superstore Sales Data Analysis

Student Name: Ritesh Bisht

Internship : Dyizan

Objective: The goal of this mini project is to analyze retail sales data from a superstore to identify trends, insights, and patterns that can support better decision-making. The project includes data cleaning, visualization, and interpretation of sales and profit metrics.

Dataset:

➛ Source: Sample Superstore dataset (CSV)
➛ Rows: 9994
➛ Columns: 13
➛ Main Columns Used: Order Date, Sales, Profit, Category, Region

Tools & Libraries Used:

➛ Python 3
➛ pandas
➛ matplotlib
➛ seaborn

Steps Performed:

1. Data Loading: Loaded dataset from a working GitHub Gist CSV link using pandas.
2. Data Cleaning & Exploration: Converted 'Order Date' and 'Ship Date' to datetime objects. Checked for null values and data types.Extracted month and year from 'Order Date' for trend analysis.
3. Data Visualization:

➛ Monthly Sales Trend: Line chart to observe fluctuations in sales over time.
➛ Category-wise Sales: Bar plot to see which product categories generate the most revenue.
➛ Correlation Heatmap: Analyzed how Sales, Profit, Discount, and Quantity are related.

Key Insights:

Sales are generally consistent but spike during specific months, indicating seasonal trends.

The "Technology" category had higher sales than "Furniture" and "Office Supplies."

Discounts do not necessarily improve profits. In fact, high discounts negatively affect profit margins.

Conclusion: The analysis provided valuable insights into the sales dynamics of the superstore. It demonstrates how data visualization and simple statistics can help businesses improve inventory, pricing, and marketing decisions.

Learnings:

➛ Understood the process of loading and cleaning real-world datasets.
➛ Improved skills in using pandas for data analysis and seaborn/matplotlib for visualization.
➛ Gained insights into how business data can be turned into actionable decisions using Python.