Sales Data Analysis Report

Name : Anjali Gurjar

University Roll No: 2024011004000036

Branch – CSE AIML

Email id:

anjali.2428cseaiml18@kiet.edu

Introduction

The retail company XYZ Inc. faced challenges in understanding their sales trends, product demand, and revenue patterns. Analyzing this data is crucial to identifying profitable products, forecasting sales trends, and making informed business decisions. This project aims to explore the given sales dataset, extract meaningful insights, and visualize key trends to support data-driven decision-making.

PROBLEM STATEMENT:
Simple Sales Data Visualization – Analyze and plot revenue, product demand, and seasonal sales trends.

Methodology

The analysis was conducted using the following steps:

1. Data cleaning and preparation using Pandas.
2. Visualization of revenue trends and product demand using Matplotlib and Seaborn.
3. Identification of seasonal sales trends through monthly data analysis.

Code

#IMPORTING LIBRARIES

```
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns
```

#READING DATASET

```
df = pd.read_csv('/content/drive/MyDrive/Improved_Sales_Dataset1.csv')
```

#Display the first few rows of the dataset

```
print(df.head())
```

Check the dimensions of the dataset

```
print(df.shape)
```

Get summary statistics of numerical variables

```
print(df.describe())
```

Check the data types of variables

```
print(df.dtypes)
```

#LINEPLOT BETWEEN REVENUE TRENDS OVER TIME

```
plt.figure(figsize=(10, 5))

sns.lineplot(data=df, x='Date', y='Product Revenue', hue='Product Name')

plt.title('Revenue Trends Over Time')

plt.xlabel('Date')

plt.ylabel('Product Revenue')

plt.legend(title='Products')

plt.grid(True)

plt.show()
```

#BARPLOT OF PRODUCT DEMAND VS PRODUCT NAME

```
plt.figure(figsize=(8, 5))

sns.barplot(x='Product Name', y='Product Demand', data=df, estimator=sum)

plt.title('Total Product Demand Distribution')

plt.xlabel('Product Name')

plt.ylabel('Product Demand')

plt.show()
```

#SCATTERPLOT BETWEEN REVENUE VS UNIT SOLD

```
plt.figure(figsize=(8, 5))

sns.scatterplot(data=df, x='Product Demand', y='Product Revenue', hue='Product Name')

plt.title('Revenue vs Units Sold')

plt.xlabel('Product Demand')

plt.ylabel('Product Revenue')
```

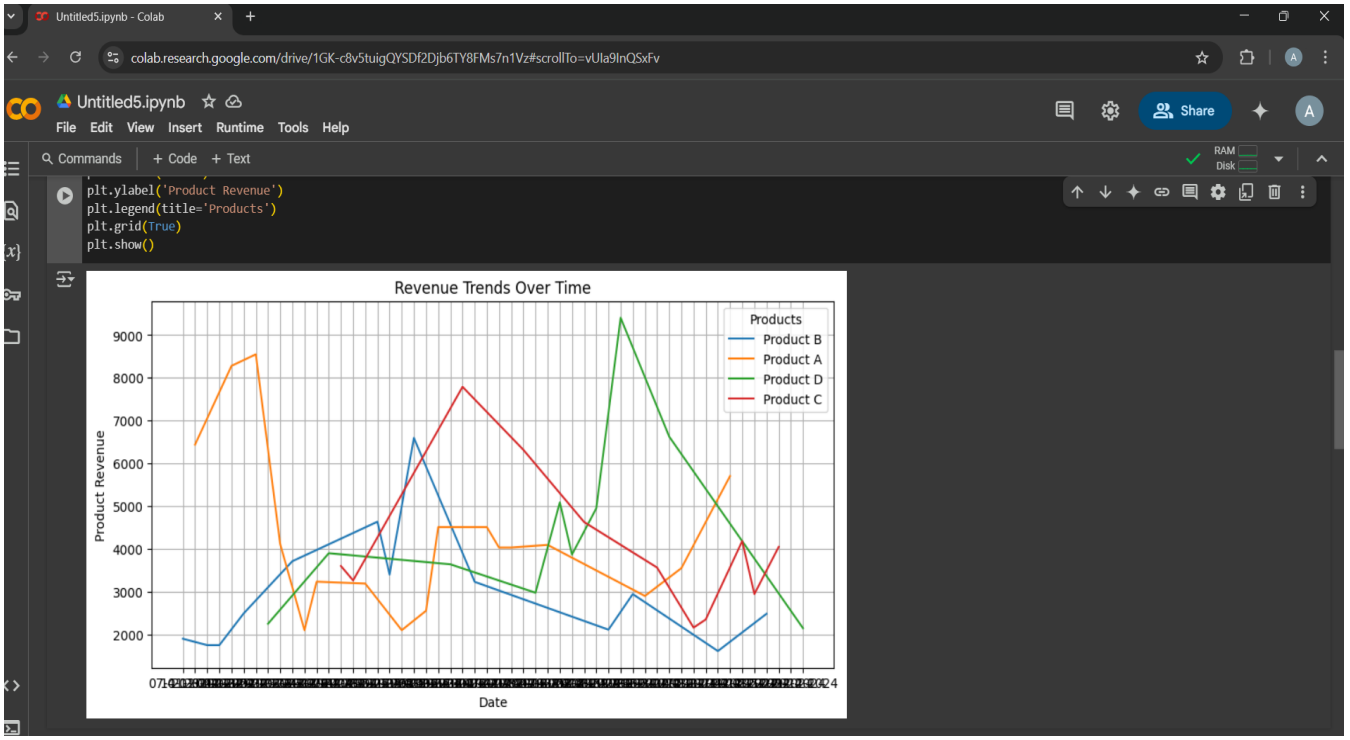
plt.grid(True)

plt.show()

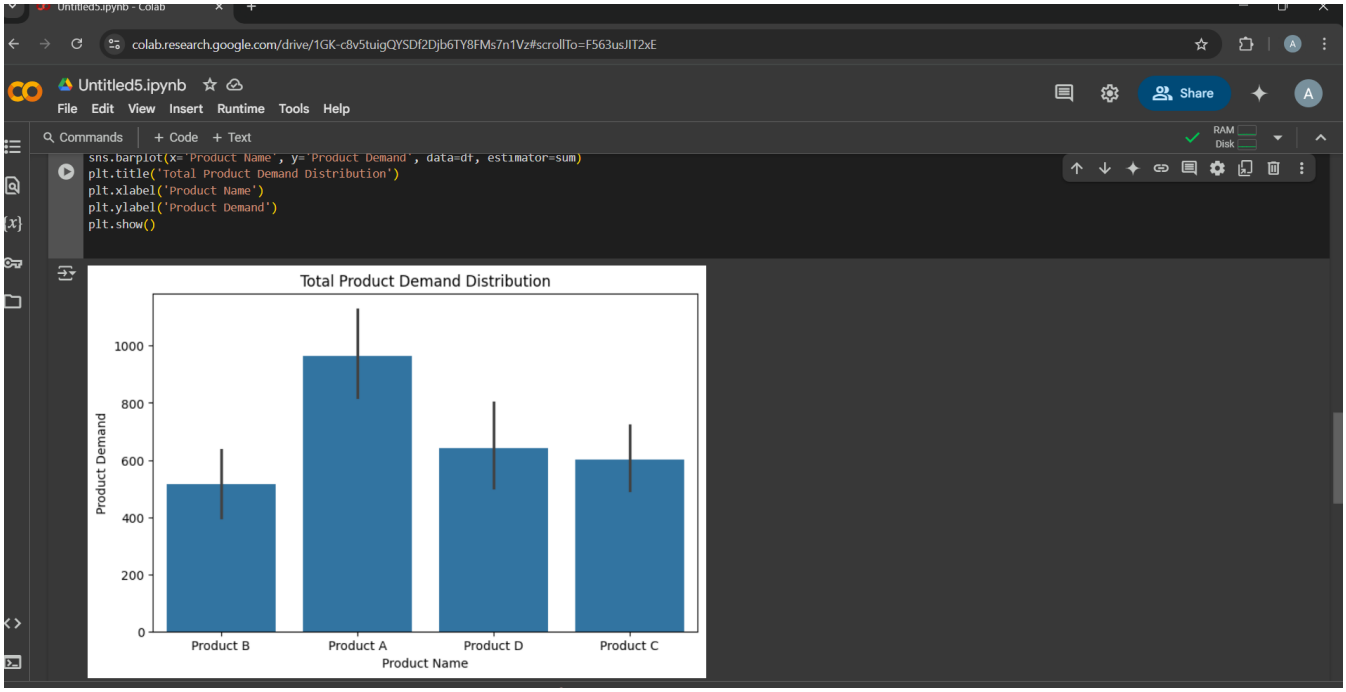
Output/Result

Refer to the attached screenshot for the graphical visualization results:

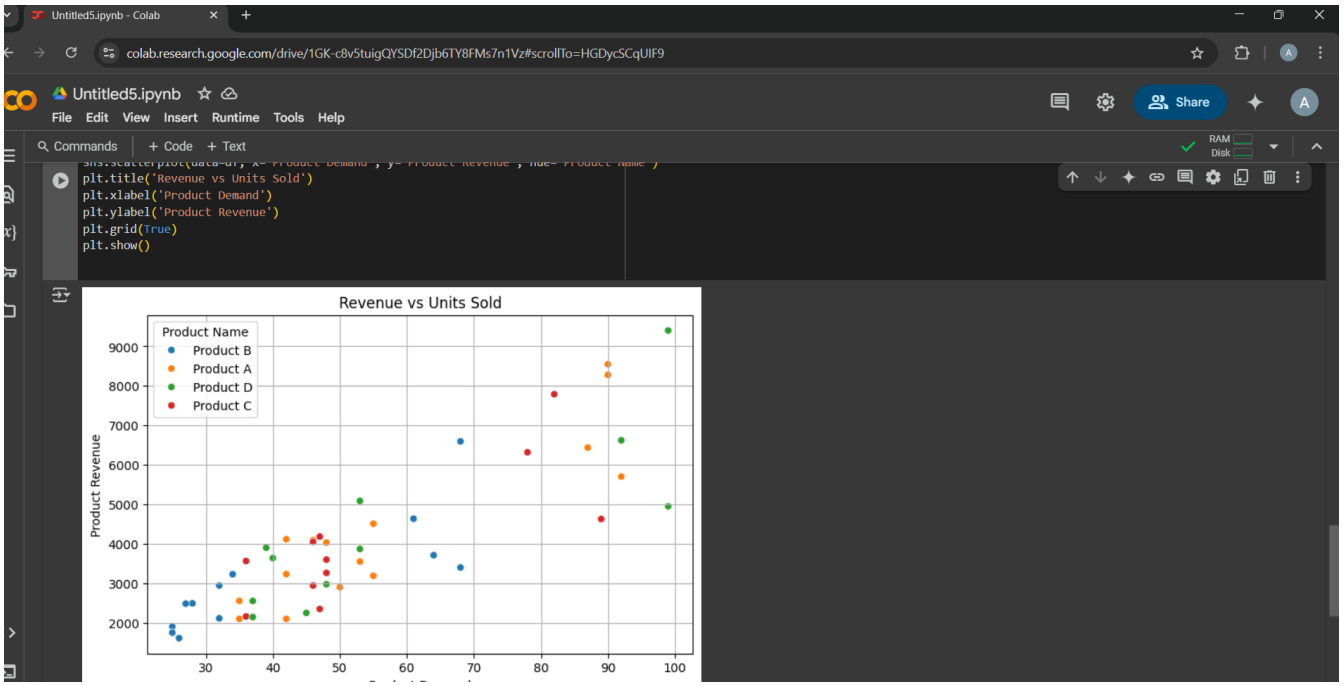
LINE PLOT OF REVENUE TRENDS OVER TIME



BARPLOT OF TOTAL PRODUCT DEMAND DISTRIBUTION



SCATTER PLOT OF REVENUE VS UNITS SOLD



References/Credits

1. Dataset: Improved_Sales_Dataset1.csv
2. Libraries used: Pandas,Numpy Matplotlib, Seaborn
3. Google collab for writing code