Analyzing sales data

A Mahesh Reddy

[Unified Mentor]

1. Introduction

The goal of this project is to analyze sales data and uncover insights that can help optimize sales strategies and improve decision-making. Sales data provides valuable information about revenue trends, customer behavior, and product performance, which are crucial for business growth. This report presents an analysis of sales data using Python and visualization tools.

2. Dataset Description

The dataset contains sales records, including attributes such as:

- **Order ID**: Unique identifier for each sale.
- **Product**: The name of the product sold.
- **Quantity Ordered**: The number of units sold.
- **Price Each**: Price per unit of the product.
- **Order Date**: Date and time of the sale.
- **Purchase Address**: Address where the order was delivered.
- **Total Sales**: Derived from multiplying the quantity ordered by the price per unit.

The data spans multiple months and covers various geographic regions.

3. Data Preprocessing

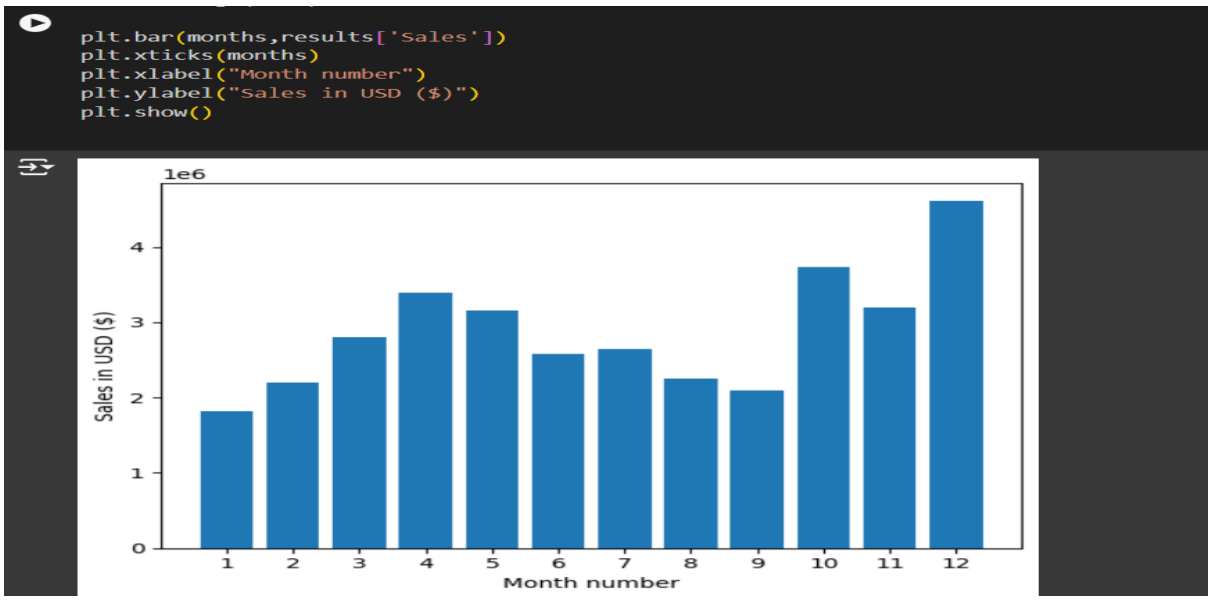
Before analyzing the data, several preprocessing steps were performed:

- **Handling Missing Values**: Rows with missing or incomplete data were identified and removed.
- **Data Type Conversion**: Ensured that columns such as dates and prices were of the correct data type for analysis.
- **Deriving New Features**:
    - Extracted month, day, and hour from the Order Date column to analyze trends over time.
    - Created a Total Sales column by multiplying Quantity Ordered by Price Each.

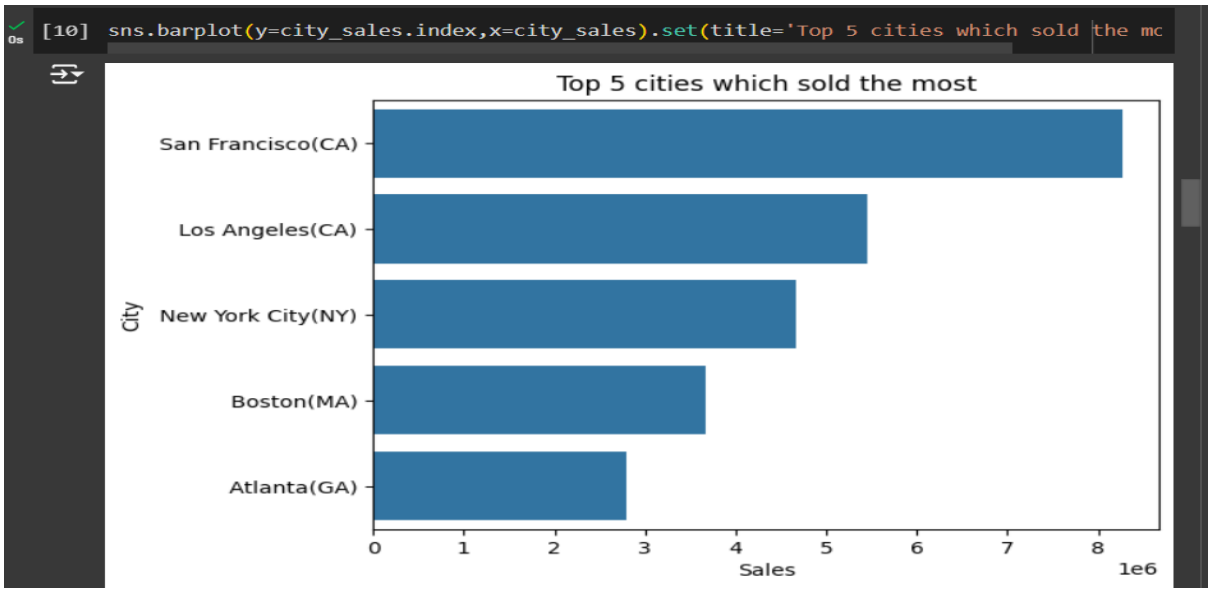
4. Exploratory Data Analysis (EDA)

In this step, the following insights were derived from the sales data:

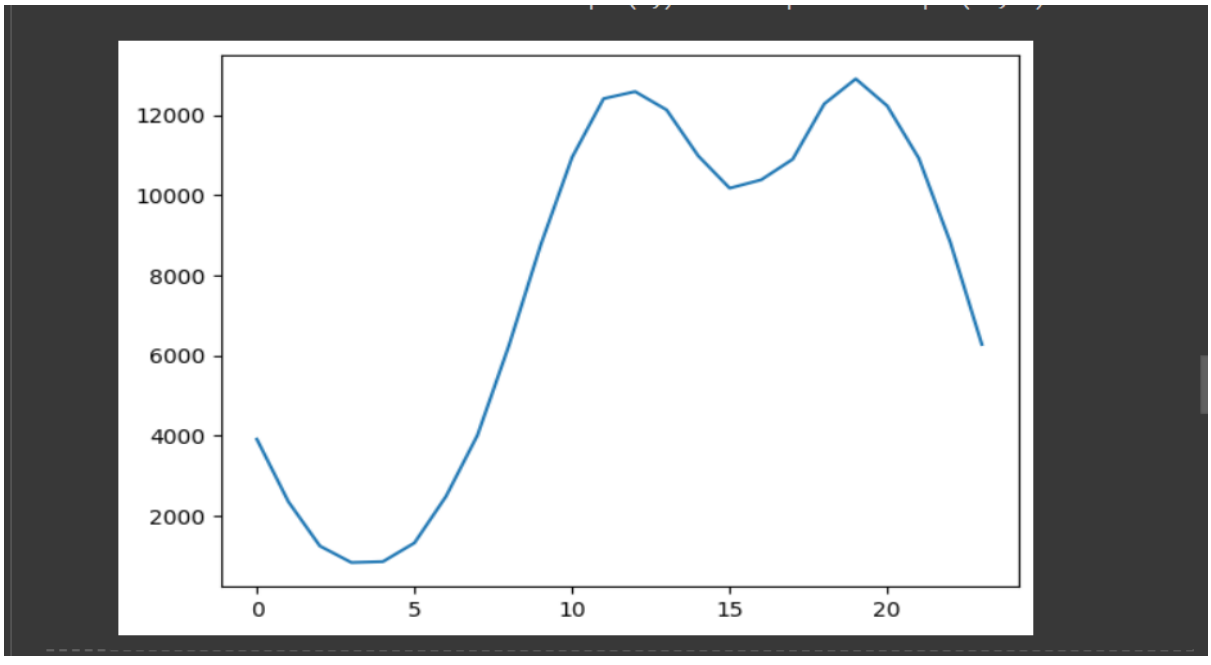
- **Monthly Sales Trend**: Sales were aggregated by month to identify peak sales periods. The analysis showed that sales were highest during the holiday season (November and December), likely due to increased consumer spending during Black Friday and Christmas.



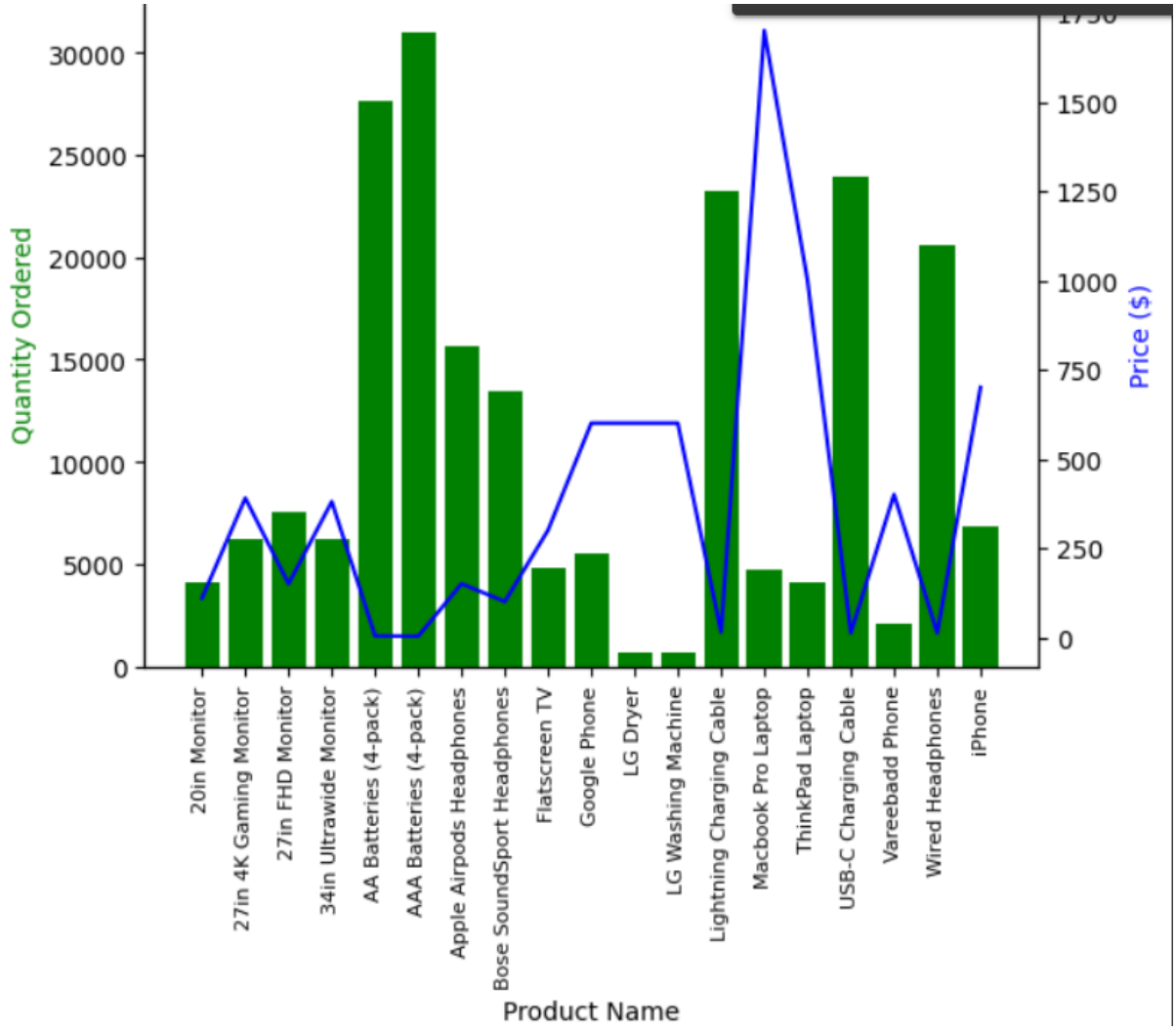
- **Best-Selling Products**: By aggregating the sales data by product, the most popular items were identified. Products with the highest quantity ordered and revenue generated were highlighted. Certain products, such as electronics and home appliances, were among the best-sellers.



- **City-Wise Sales**: Using the purchase address, sales were grouped by city. This allowed for geographic analysis, revealing which cities contributed the most to revenue. Major cities like New York and Los Angeles showed the highest sales figures.



- **Time of Day for Sales**: Analyzing the hour of purchase revealed peak times for sales. Most sales occurred between 12 PM and 8 PM, with a noticeable drop-off during early morning hours.



5. Insights and Recommendations

From the analysis, several actionable insights were discovered:

- **Seasonal Promotions**: Since sales peak during the holiday season, businesses should focus marketing efforts and discounts in November and December to maximize revenue.
- **Product Inventory Management**: Best-selling products should be kept in stock, especially during peak months, to avoid losing sales due to inventory shortages.

- **Geographical Focus**: Target marketing campaigns in cities that generate the highest revenue, and investigate ways to increase sales in underperforming areas.
- **Optimal Sales Hours**: Since most sales happen in the afternoon and evening, scheduling advertising and promotional campaigns during these hours can increase visibility and conversion rates.

6. Conclusion

This project provided a detailed analysis of sales data to uncover key trends in product performance, geographic distribution, and customer purchasing behavior. By leveraging these insights, businesses can make informed decisions to boost sales and optimize their marketing strategies.