Superstore Sales Analysis & Dashboard

By: Luke Bailey

Introduction & About the Dataset

The purpose of this project is to practice, develop, and showcase essential data analysis skills using a real world dataset. I'll be working with the publicly available Superstore Sales dataset from Kaggle to perform data cleaning in **Microsoft Excel** and exploratory data analysis with **Power BI** visualizations. The project will conclude with the creation of an interactive business dashboard that highlights KPIs in **Microsoft Power-BI**.

The chosen dataset for this project is the Superstore Dataset from Kaggle, which represents sales data from a large store. Below are some brief points on it.

- 9994 Entries
- 21 Attributes

Data Cleaning (Excel)

While this dataset comes fairly clean, it's good practice to go through the process. Below is my data cleaning procedure for this project in order to ensure accurate insights. Each of these steps were performed using Excel tools and functions.

1. Remove Duplicate Entries: Ensure all rows are unique to avoid overlap.
2. Format Data Types: Ensure that dates and numeric values are all correctly formatted and validated
3. Handle Null Values: Row is dropped if critical data is missing, otherwise it's filled with a default value.

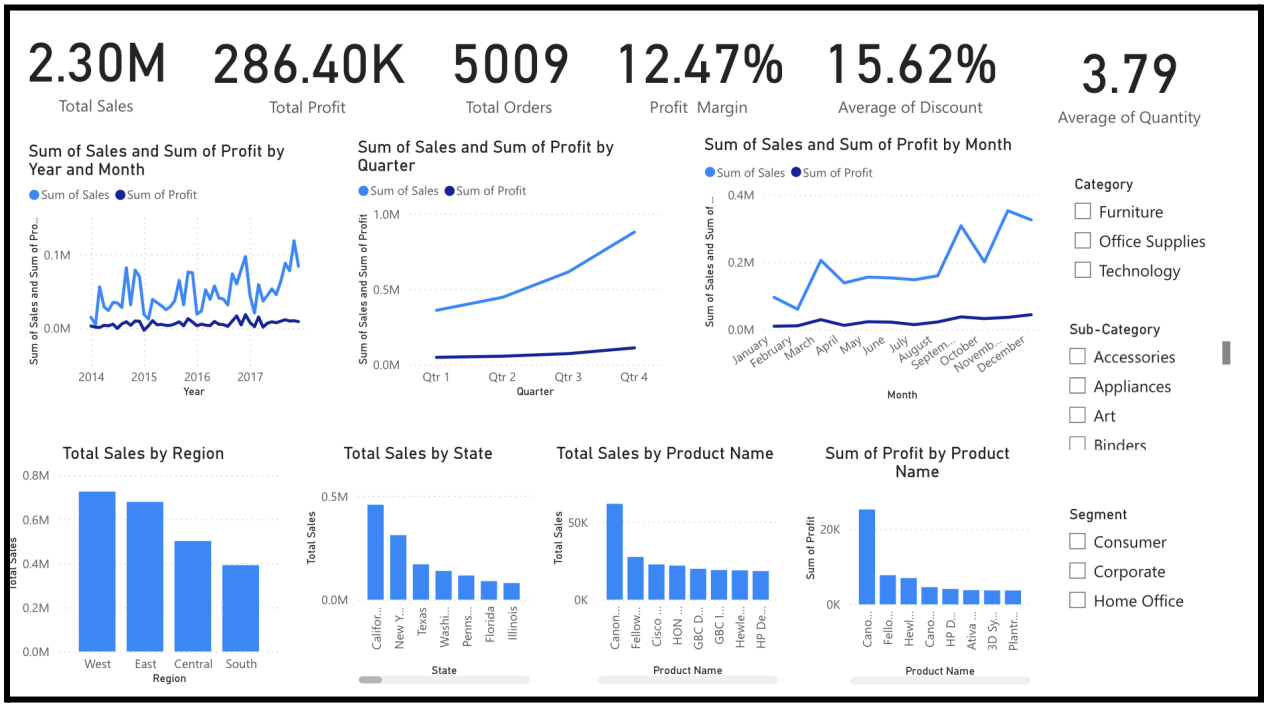
Exploratory Data Analysis

After going through the cleaning process in Excel, I imported my dataset into Power BI to begin the EDA process. It was my intention to draw insights through interactive visualizations, so I began with a list of questions I sought to answer and identified their corresponding visualization. Below you can find a table highlighting this, as well as a screenshot of my EDA dashboard.

Question	Visualization Type
What is the total sales count?	Card
What is the total profit count?	Card
How many orders were placed?	Card
What is the profit margin?	Card
What is the average discount per order?	Card
What is the average quantity per order	Card
How do sales and profit vary over time?	Line Graph
What seasonal patterns are there in sales?	Line Graph
What quarter patterns are there in sales?	Line Graph
Which regions generate the (most/least) sales?	Bar Chart
Which states generate the (most/least) sales?	Bar Chart
Which product categories generate the (most/least) (sales/profit)?	Slicer
Which sub categories generate the (most/least) (sales/profit)?	Slicer
Which products have (high/low) profit?	Bar Chart
Which products have (high/low) sales?	Bar Chart
Which customer segments generate the (most/least) sales?	Slicer

After setting up my initial EDA dashboard, I began the examination process. I took note of any relevant trends or insights in order to inform my KPI dashboard design and recommendation report. Below is a screenshot of the dashboard I used for my own personal insights.

Temporary Dashboard for Personal Use



KPI Design + Calculations

Based on the insights I made in the EDA section, I finalized a list of key performance indicators that I intended to highlight which can be seen in the table below.

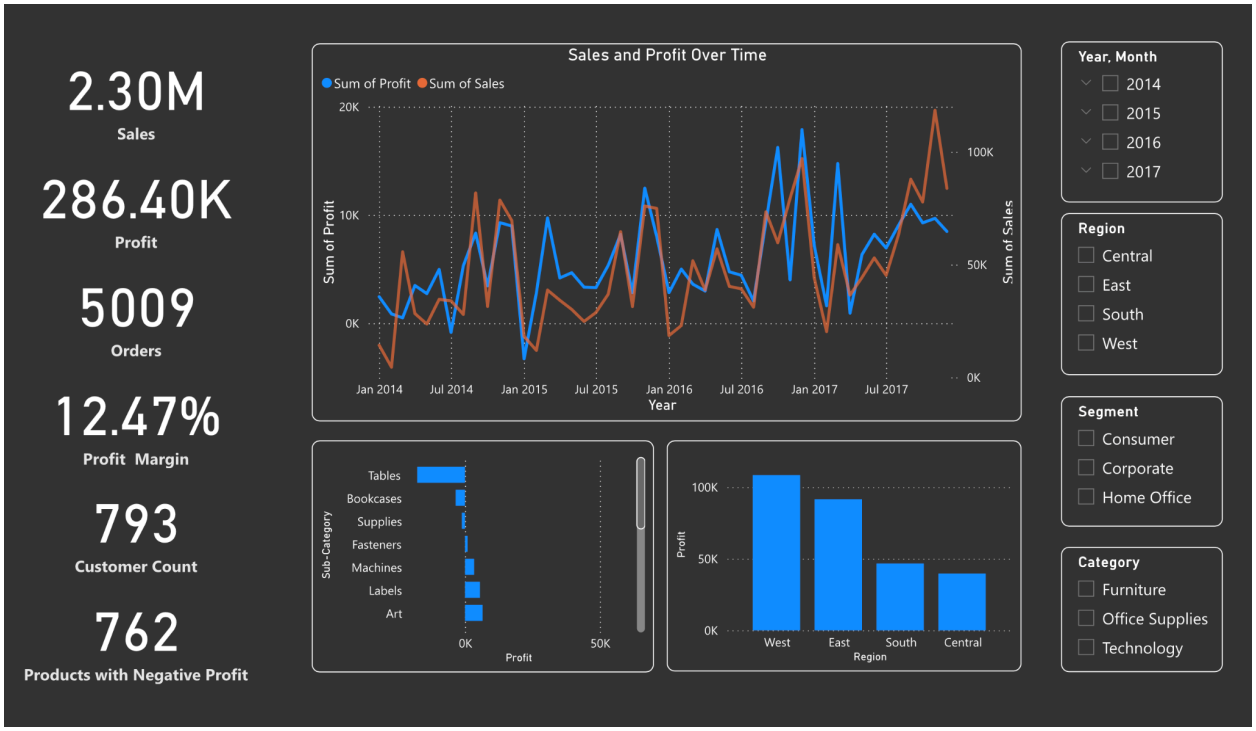
KPI	Formula/ Representation
Sales	SUM(Sales)
Profit	SUM(Profit)
Orders	DISTINCTCOUNT(Order ID)
Profit Margin	DIVIDE(SUM(Profit),SUM(Sales))
Customer Count	DISTINCTCOUNT(Customer ID)
Negative Profit Products	COUNT(Products) where Profit < 0
Sales & Profit VS Time	Line Graph
Profit VS Sub-Category	Bar Chart
Profit VS Region	Bar Chart

Additionally, I incorporated slicers to improve the interactivity aspect of my dashboard. These slicers act like filters to the audience and perform their own analysis on the data. A slicer was added for the following features:

- Date
- Region
- Segment
- Category

Power-BI Dashboard Overview

I used my selected KPIs to inform my final Power-BI dashboard. The intention of this dashboard is to highlight the insights that I gained during the EDA process. To convey a story with the data to the audience, I used design skills to produce an aesthetically pleasing, user-friendly, interactive dashboard. The finished dashboard can be seen below.



Business Insights and Recommendations

Based on my data exploration and analysis, I produced a list of key insights and recommendations which can be seen below.

Relevant Insights

- Sales continuously increased from 2014-2017, however profit margin declined overtime, indicating excessive discounting or rising costs.
- 762 out of 1862 products are generating negative profit, meaning that nearly 41% of the product catalog is underperforming.
- Certain sub-categories like tables and bookcases are among the least profitable, especially when large discounts are applied.
- There is a large spike in sales between August and November, which we can likely attribute to back to school and holiday shopping.
- The western region is consistently outperforming the other regions in various categories. Meanwhile, the southern region is consistently lagging.

Recommendations

- Review product pricing and cost structure for underperforming categories. Especially Tables, Bookcases, and Supplies. Consider a discontinuation of underperforming products after a cost benefit analysis.
- Reduce discounting for underperforming products. Additionally, reevaluate promotional strategies for low margin items.
- Leverage seasonal trends by adjusting stock and preparing targeted marketing campaigns for Q3 and Q4 due to the high demand during back to school and holiday months.
- Focus sales and marketing efforts in the higher performing regions, and explore localized strategies to improve performance in the south.
- Make more profitable products, like technology, a priority by implementing sales techniques like bundles or upselling.