



Sales Data Analysis Report

- **Dataset:** Sample - Superstore
- **Prepared by:** Mustafa Oun
- **Date:** June 26, 2025
- **Tools Used:** Python, Power BI

Project Objectives

- **This project aims to analyze Superstore sales data to uncover actionable insights for marketing and managerial decisions.**
- **We utilized Python for comprehensive exploratory data analysis (EDA) and Power BI for creating interactive, dynamic visualizations.**
- **The goal is to provide a data-driven foundation for strategic business improvements.**



Dataset Overview

The Superstore dataset provides a rich source of sales information, enabling a deep dive into various business dimensions.

Key Columns:

- Order Date
- Sales
- Profit
- Customer Name
- Product Name
- Category
- Region

Number of Rows	9,994
Number of Columns	21
Time Period	2015-2018

Data Cleaning & Preparation

Date Format Conversion

Converted 'Order Date' and 'Ship Date' to datetime for temporal analysis.

Data Type Adjustments

Changed 'Postal Code' to string format to preserve leading zeros.

Redundancy Removal

Removed 'Row ID' as it was non-informative for our objectives.

Data Integrity Check

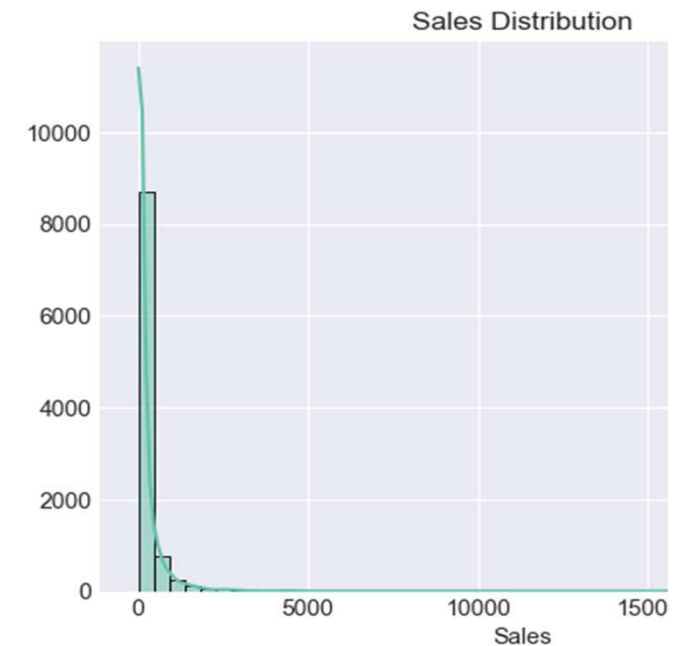
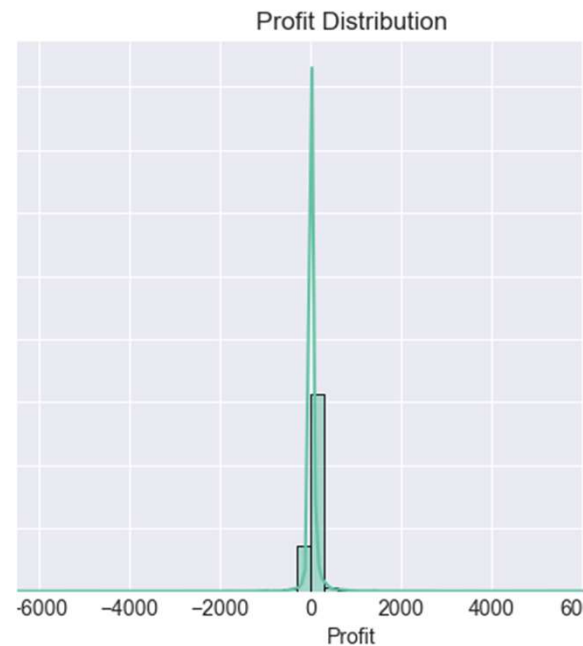
Confirmed the absence of null values and duplicate entries across the dataset.

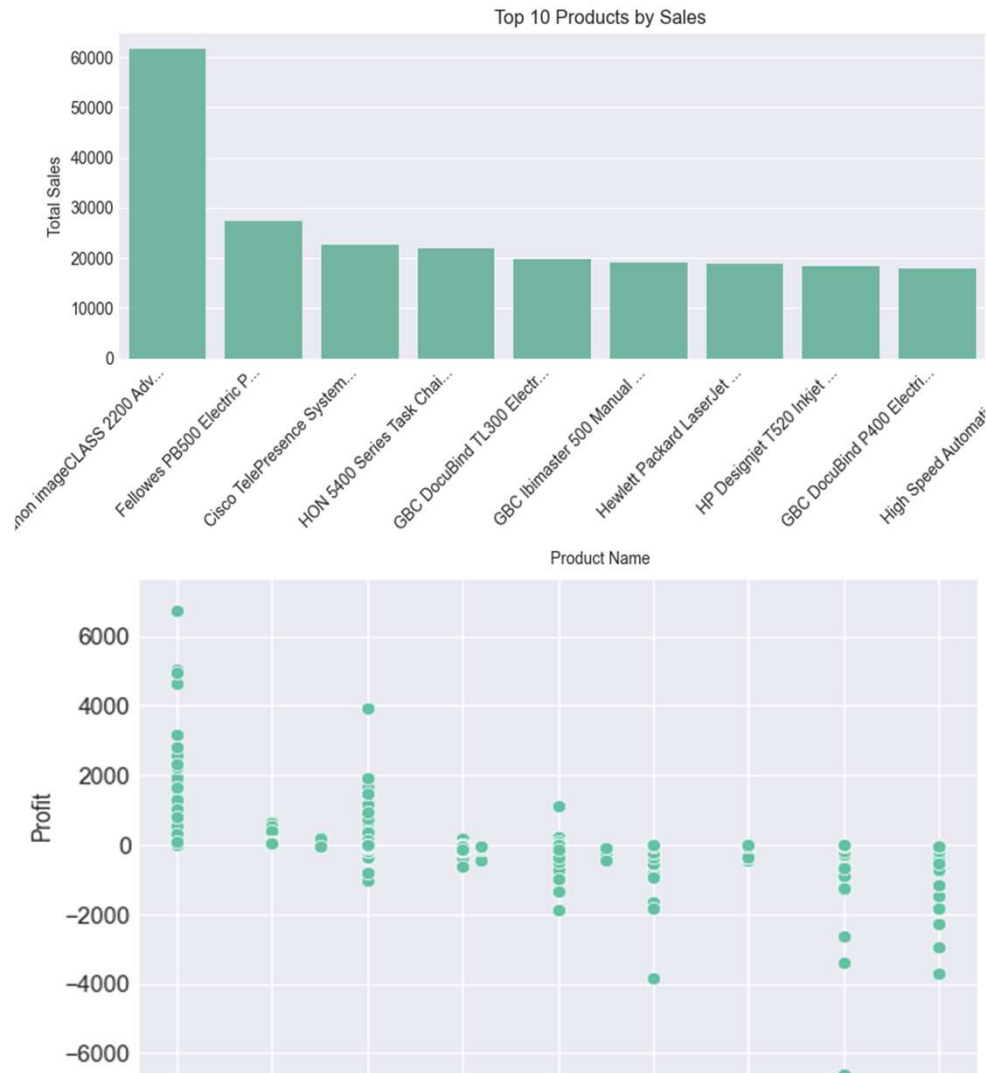
Derived Column Creation

Generated 'Shipping Days', 'Month', 'Year', and 'Month-Year' for enhanced trend analysis.

Exploratory Data Analysis: Profit & Sales

- **Distribution Insights**
- Sales range from \$0.44 to \$22,638.48.
- Profit ranges from -\$6,599.98 to \$8,399.98.
- The distribution is right-skewed, indicating most transactions are low-value.
- Identified 1,881 profit outliers, highlighting transactions with extreme gains or losses.

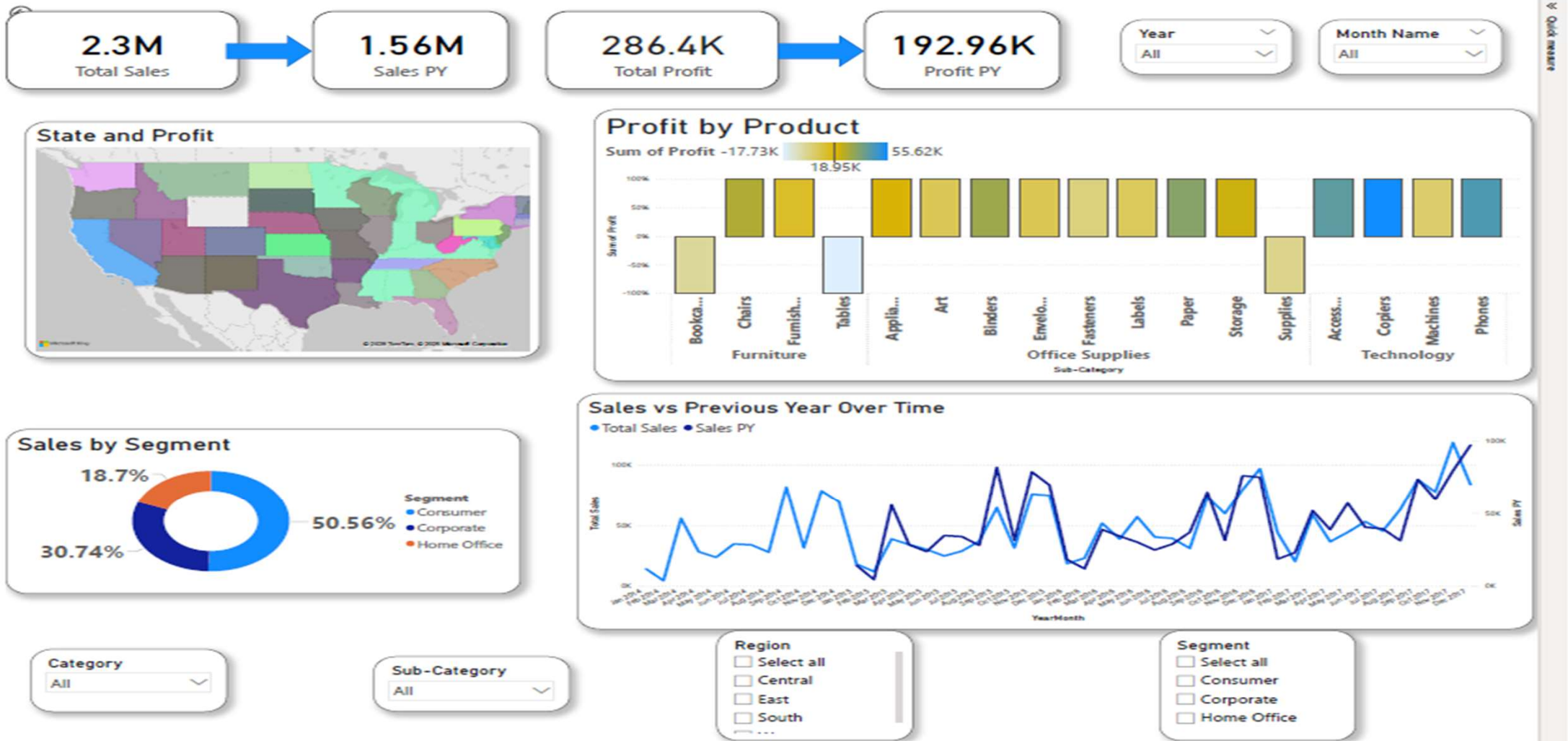




EDA: Discount Impact & Top Products

- The scatterplot clearly shows a **negative correlation between discounts and profitability**, indicating that higher discounts often lead to reduced profits. This suggests a need to re-evaluate discount strategies.
- "Canon imageCLASS 2200 Advanced Copier" stands out as the **highest-selling product**. However, high sales do not always translate to high profits, underscoring the importance of analyzing profitability alongside sales volume for individual products.

Power BI Dashboard



Key Strategic Insights



Discount Optimization

High discounts significantly reduce profitability; strategic review of pricing is crucial.



Category Performance

Technology, especially Copiers, drives the highest profits; Furniture, particularly Tables and Bookcases, incurs losses.



Regional Focus

The West and East regions are the most profitable, indicating strong market presence and customer base.



Customer Value

A small group of key customers contributes disproportionately to overall profits.



Seasonal Opportunities

Sales peak at year-end, presenting clear seasonal marketing opportunities.



Profit vs. Sales

High sales don't always mean high profits; re-evaluate pricing or cost structures for specific products.

Thank You



<https://www.linkedin.com/in/mustafa-oun/>



<https://github.com/mustafaoun>