

Project Report

On The Topic

Retail Superstore Sales & Profit Analysis

(End-to-End Data Analytics Project)

1. Objective

The goal of this project is to analyze **Retail Superstore sales data** in order to identify:

- Which products, customers, and regions drive revenue and profit.
- The impact of discounts on profitability.
- Sales and profit trends over time.
- Key customers contributing to business growth.

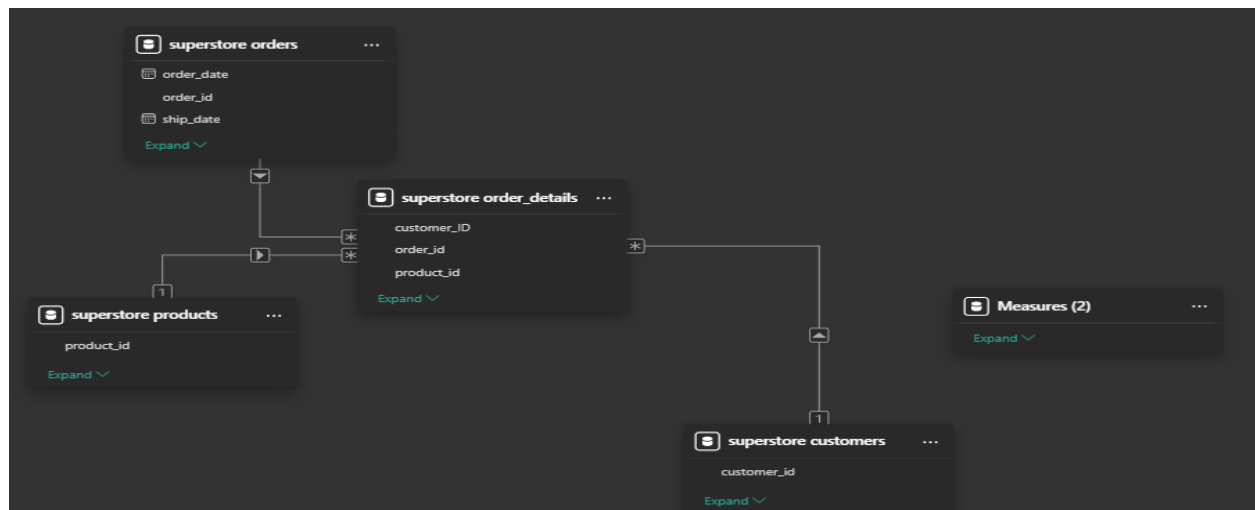
The project covers the **full analytics pipeline**: Data Cleaning → SQL → Python → Power BI → Insights.

2. Data Preparation & Database Design

The raw dataset (Excel/CSV) was normalized into **4 relational tables** to eliminate redundancy and maintain referential integrity.

Database Schema (ER Diagram):

- customers (customer_id, customer_name, segment, region, country)
- products (product_id, product_name, category, sub_category)
- orders (order_id, order_date, ship_date, ship_mode)
- order_details (order_id, product_id, customer_id, quantity, sales, profit, discount)



3. SQL Analysis

Example Query 1: Sales by Region

```
SELECT c.region, SUM(od.sales) AS total_sales
FROM customers c
JOIN order_details od ON c.customer_id = od.customer_id
GROUP BY c.region
ORDER BY total_sales DESC;
```

Result:

Region	Total Sales (\$)
West	725,457
East	678,781
Central	501,239
South	391,721

➡ *Insight:* West region drives the highest sales, while South lags behind.

Example Query 2: Top 3 Products by Sales in Each Category

```
SELECT p.category, p.product_name, SUM(od.sales) AS total_sales
FROM products p
JOIN order_details od ON p.product_id = od.product_id
GROUP BY p.category, p.product_name
ORDER BY p.category, total_sales DESC
LIMIT 3;
```

Result:

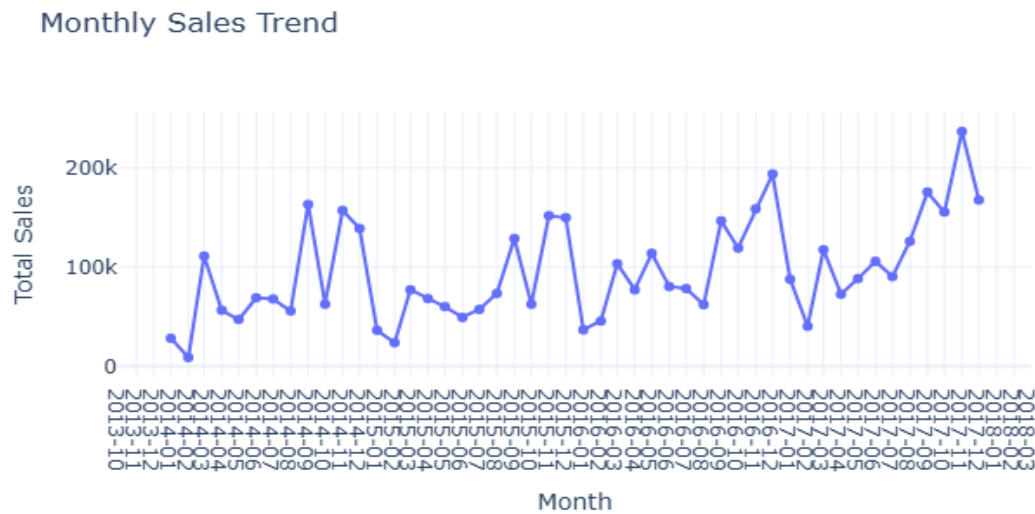
Category	Product Name	Sales (\$)
Furniture	Chairs & Stools Set	45,812
Technology	Canon ImageCLASS Printer	38,215
Office Supply	Paper Ream 500 sheets	29,112

➡ *Insight:* Technology products dominate profit contribution.

4. Python Analysis

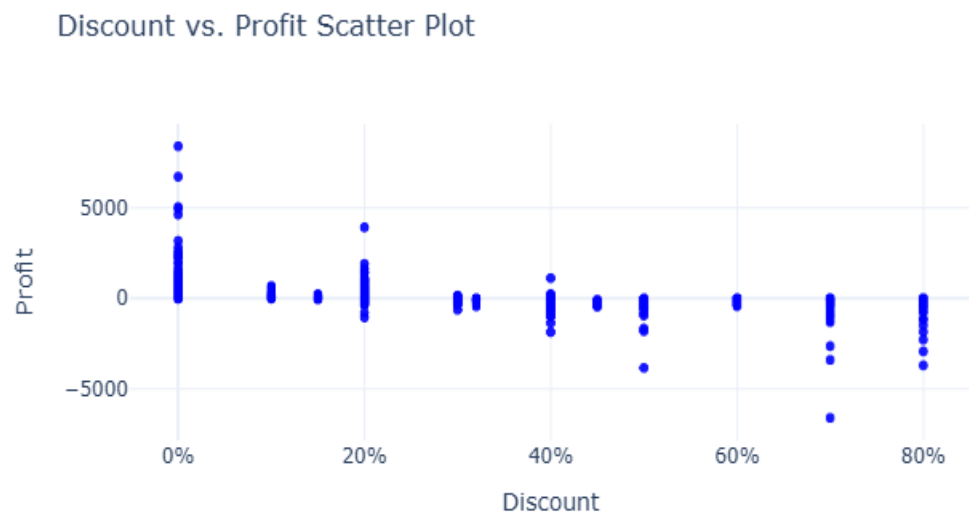
Using **Pandas** + **Seaborn** + **Plotly**:

Monthly Sales Trend



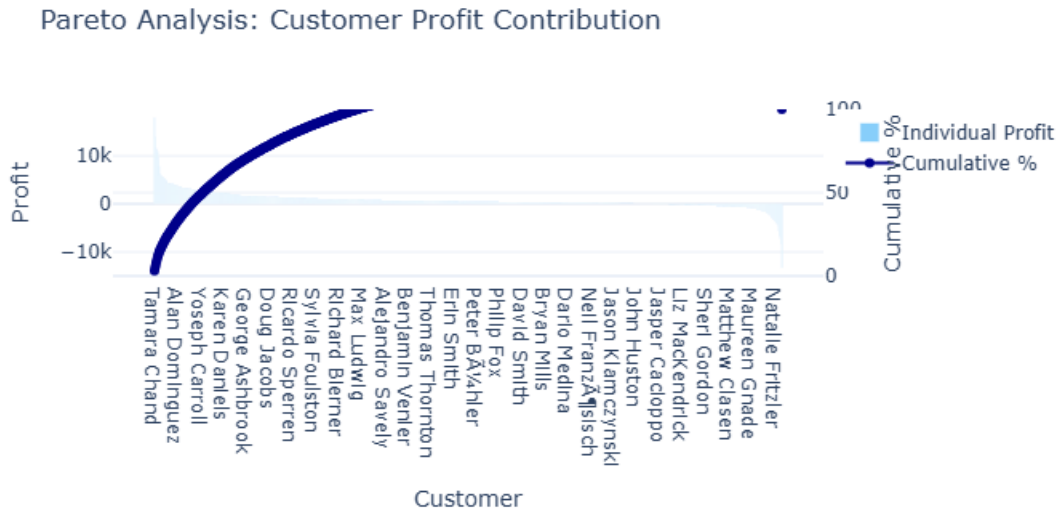
- Sales steadily increased from 2014 to 2017.
- Strong seasonal peaks in Q4 (holiday shopping).

Discount vs Profit



- High discounts often result in **negative profit** (loss-making sales).

Pareto Analysis (80/20 Rule)

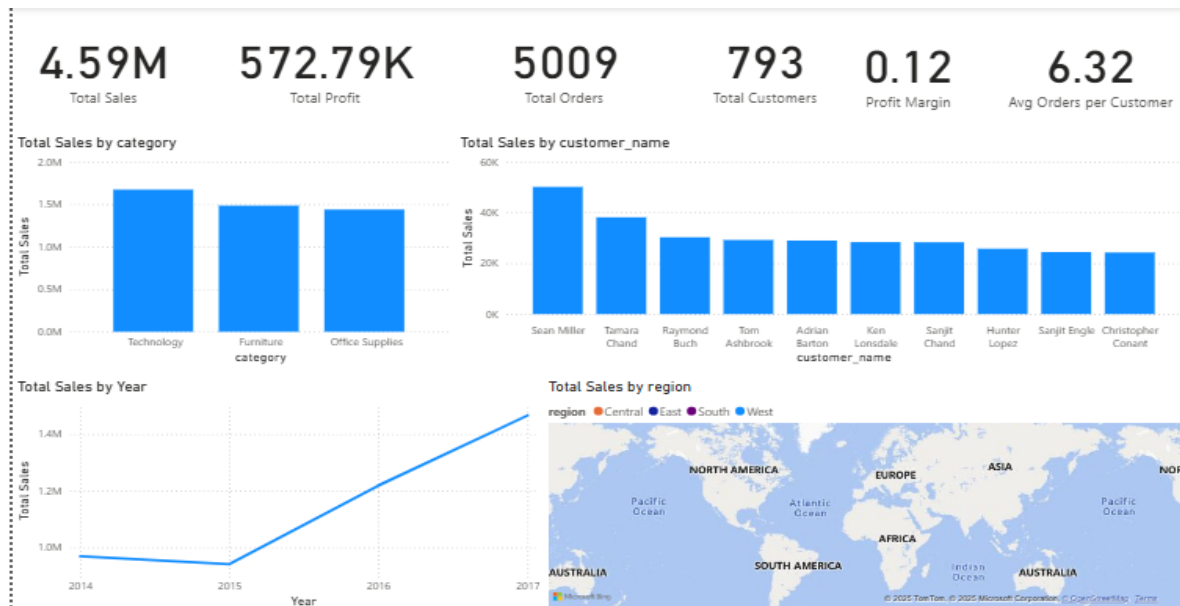


- ~20% of customers contribute ~80% of total revenue.
- Suggests a **loyalty/customer retention strategy** is important.

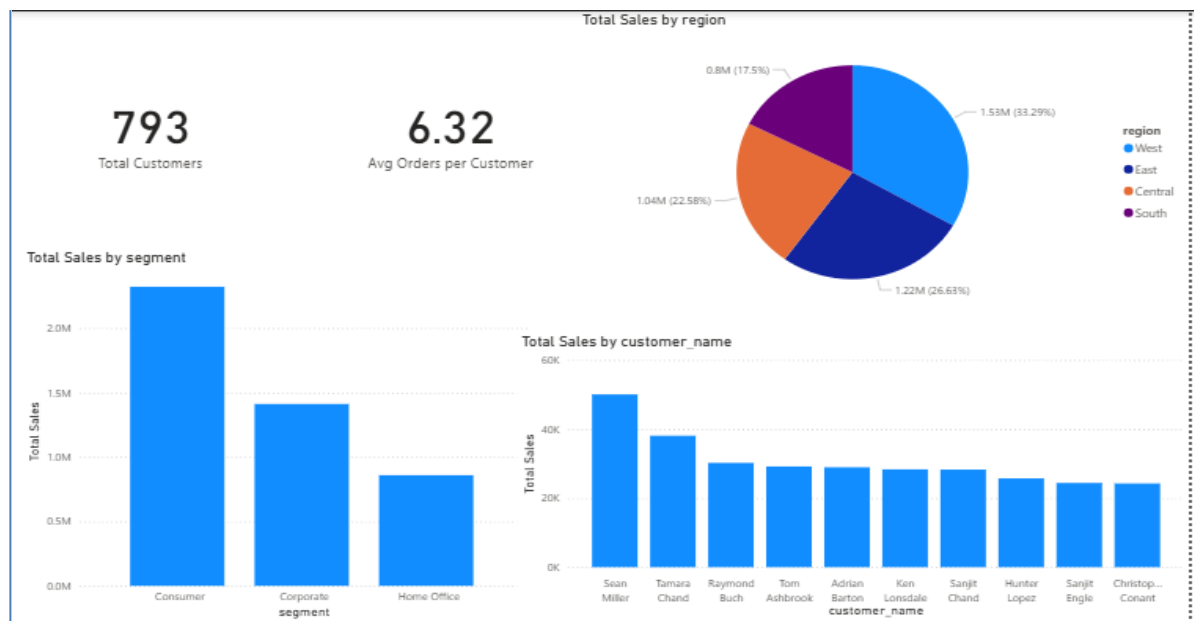
5. Power BI Dashboard

A 4-page interactive dashboard was developed:

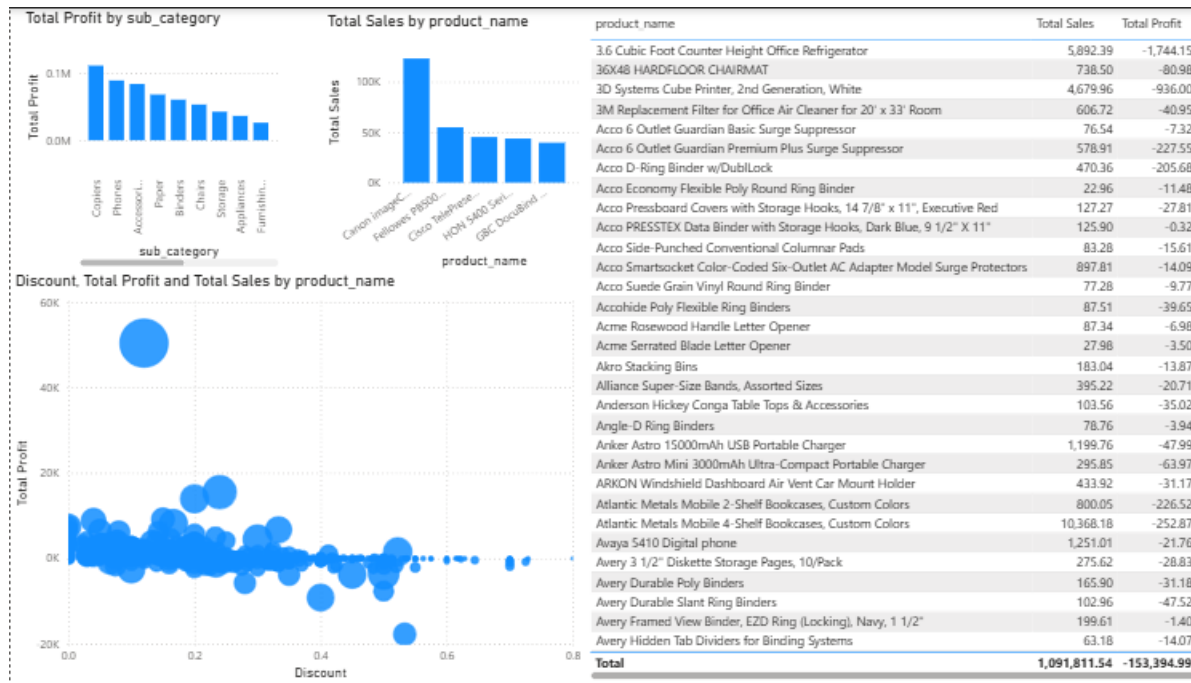
1.Executive Overview – KPIs, Sales by Category, Profit by Region, Top Customers.



2.Customer Insights – Customer segments, Top 10 customers by sales, Region split.



3.Product & Profitability – Profit by sub-category, Discount vs Profit scatter.



4. Dimensional Tree – helpful in root cause analysis, pattern and trend analysis .



6. Key Insights

- High discounts reduce profitability (especially in Furniture).
- Technology drives the highest profit margins.
- West region is the strongest performer in revenue.
- A small group of customers generates the majority of revenue.
- Seasonal peaks suggest stocking strategies for Q4.

7. Recommendations

1. Reduce discounts on **Furniture products** to prevent losses.
2. Expand focus in **West region** but develop strategies to improve South region sales.
3. Launch **customer loyalty programs** targeting high-value customers.
4. Optimize **inventory planning** before holiday seasons.

8. Deliverables

- SQL Queries: `sql/queries.sql`
- Python Notebook: `notebooks/superstore_analysis.ipynb`
- Power BI Dashboard: `dashboard/superstore.pbix`
- This Report: `docs/report.pdf`