Superstore Sales Analysis – Project Report

Introduction

This project demonstrates a complete data analytics workflow using **PostgreSQL** for data management and **Power BI** for visualization. The goal was to analyze retail sales performance, profitability, and customer dynamics using the popular Superstore dataset.
I chose this project because it closely mirrors **real-world business analysis scenarios**, where analysts must extract insights from transactional data, design data models, and communicate findings through dashboards and reports.

Data Source

- Dataset: **Sample Superstore** (widely used retail dataset).
- Loaded into **PostgreSQL 17** as a relational table (superstore_sales).
- Data cleaning performed during load included:
    - Ensuring date fields were correctly parsed (MDY format).
    - Handling product names with special characters.
    - Converting sales, profit, and discount into numeric formats.

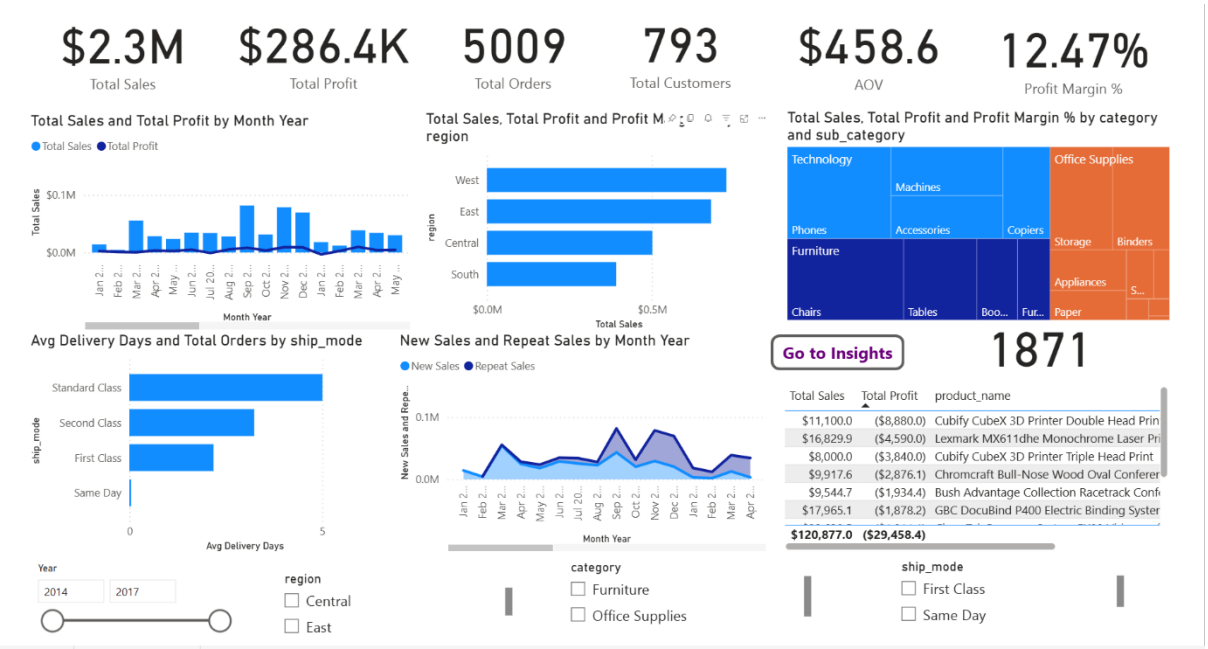
Process

1. **Schema Design**
    - Designed a structured table (superstore_sales) with appropriate data types (dates, numerics, varchar).
    - Created a **Calendar table** to support time-based analysis.
2. **Data Load**
    - Imported CSV into PostgreSQL using the \copy command and Import Wizard.
    - Resolved formatting issues (UTF-8 vs LATIN1 encoding, quoted strings, date parsing).
3. **SQL Preparation**
    - Verified row counts, distinct values, and relationships.
    - Ran exploratory queries for sales by region, category, and profit margins.
4. **Power BI Modeling**
    - Connected PostgreSQL to Power BI.
    - Built relationships (Calendar[Date] → Superstore[Order_Date]).

- Created DAX measures:
    - Total Sales, Total Profit, Total Orders, Total Customers, AOV, Profit Margin %.
- Designed KPIs and visuals across 2 pages.

Key Visuals

- **KPI Cards**: Total Sales ($2.3M), Total Profit ($286.4K), Total Orders (5009), Customers (793), Average Order Value ($458.6), Profit Margin % (12.47%).
- **Sales & Profit Trend (Line + Column)**: Monthly trend of sales vs. profit.
- **Regional Performance (Bar Chart)**: Sales, profit, and margin % by region.
- **Category/Subcategory (Tree Map)**: Contribution of technology, furniture, and office supplies.
- **Shipping Mode Analysis (Bar Chart)**: Average delivery days vs. total orders.
- **Customer Behavior (Line Chart)**: New vs. repeat sales across months.
- **Loss-making Products (Table)**: Highlighted products with negative profitability.



“Figure 1: Executive Dashboard – Superstore Sales & Profitability”

Executive Insights

- **Revenue Concentration:** Sales are driven by the West and East regions, but margin gaps suggest efficiency improvements are needed in the East.
- **Category Dynamics:** Technology delivers the highest revenue, while Office Supplies contributes volume but erodes profitability with thin margins.
- **Shipping Efficiency:** Standard Class dominates, but longer delivery times risk customer satisfaction.
- **Customer Growth:** New customers provide strong initial sales, yet repeat customers sustain long-term growth—highlighting the need for retention strategies.
- **Profitability Risks:** Certain products (e.g., printers & furniture) consistently generate losses, requiring pricing reviews or portfolio rationalization.

"Figure 2: Executive Insights Page – Key Findings & Recommendations"

Conclusion

The analysis shows that **Superstore revenue is growing**, but profitability is uneven across regions, categories, and products. To sustain long-term success, the business must:

- Improve **efficiency in the East region**.
- Rationalize **low-margin products** like certain printers and furniture.
- Focus on **customer retention** to maximize lifetime value.
- Review **shipping methods** to balance cost, speed, and satisfaction.

This project highlights my ability to:

- Design and manage data models in SQL.
- Build KPIs and interactive dashboards in Power BI.
- Translate raw data into actionable **business recommendations**.