Project Report: Retail Business Performance & Profitability Analysis

1. Introduction

The retail industry operates with diverse product categories, customer preferences, and dynamic discounting strategies. Understanding which categories and regions drive profitability is critical for optimizing inventory and ensuring sustainable growth. This project leverages SQL, Python, and Tableau to analyze transactional retail data, identify profit-draining categories, assess the impact of discounts and delivery times, and provide strategic recommendations.

2. Abstract

The dataset used was the *Sample Superstore* dataset, containing 9,994 transactional records. The analysis focused on:

- Profitability by category and sub-category**.**
- Correlation of processing days with profit margins.
- Impact of discounting on profitability.
- Regional performance trends.
- Visualization of sales and profit patterns across time and season.

Findings show that Furniture is a loss driver in many cases, **Office Supplies (esp. West Region) consistently delivers high margins, and discounts above 40% destroy profitability across categories**. **No strong relationship was found between processing days and profitability.**

3. Tools Used

- **SQL (MySQL):** Data import, cleaning, aggregation queries.
- **Python (Pandas, Seaborn, Matplotlib):** Exploratory Data Analysis (EDA), correlation, scatterplots.
- **Tableau:** Interactive dashboard creation with KPIs, filters (Region, Category, Season), and trend charts.

4. Steps Involved

1. **Data Cleaning & Import**
   - CSV cleaned in Excel (dates, nulls, numeric fields).
   - Imported into MySQL using LOAD DATA INFILE.
   - Verified row counts and ensured proper data types.
2. **SQL Analysis**
   - Category/Sub-Category profit margin analysis.
   - Regional profitability breakdown.
   - Discount vs. margin analysis.
3. **Python EDA**
   - Correlation between processing days and profit margins (≈0, no impact).
   - Visualization of discount vs. margin trends.
   - Outlier detection in sales/profit distribution.
4. **Tableau Dashboard**
   - KPI cards for Sales, Profit, Orders, Margin %**.**
   - Seasonal trend chart (line chart with quarters/months).
   - Filters for Region, Product Type, Season**.**

5. Key Insights & Strategic Suggestions

Category level insights

- **Furniture** → Overall margin only 2.49%**,** driven down by Tables (-8.56%) and Bookcases (-3.02%)**.** → Reduce stock, avoid deep discounting.
- **Office Supplies** → Healthy margin 17.04%**;** especially strong in Labels (44%)**,** Paper (43%)**,** Envelopes (42%)**.** → Maintain or expand stock levels.
- **Technology** → Strongest performer (17.40% margin), led by Copiers (37%) and Accessories (25%)**.**

Regional Insights

- **Central Region:** Furniture (-1.75%) is loss-making; focus on Technology (19.77%).
- **East Region:** Furniture nearly breakeven; Office Supplies (19.96%) and Technology (17.91%) are drivers.
- **South Region:** Balanced across categories, Furniture performs relatively better (5.77%).
- **West Region:** Star region; Office Supplies (23.82%) dominates.

Discount Insights

- Margins collapse after >40% discount across all categories.
- Furniture is the most discount-sensitive (negative profits even at moderate discounts).

Office Supplies tolerates moderate discounts better but still loses profitability beyond 50%.

Processing Days

- Correlation with margin = -0.0118 (≈0) → Delivery speed has no direct profitability impact.

6. Conclusion

This project demonstrates the value of data-driven retail strategy.

- Stop deep discounting (>40%) to prevent margin erosion.
- Focus investments in Technology and Office Supplies, especially in the West Region**.**
- Reassess Furniture category, particularly Tables and Bookcases, for stock rationalization.
- Use Tableau dashboards to track KPIs, seasonal sales trends, and regional profitability continuously.

By implementing these strategies, retailers can optimize inventory, reduce losses, and maximize profitability.