**📄 Retail Business Performance & Profitability Analysis  
Internship Phase 2 – Project Report**

**🔹 Introduction**

In the dynamic landscape of retail, understanding which product categories contribute to profitability is crucial for data-driven decision-making. This project aims to analyze retail sales data to uncover profit-draining product categories, improve inventory decisions, and provide actionable recommendations using SQL, Python, and data visualization techniques.

**🔹 Abstract**

We analyzed a simulated retail dataset consisting of transactional records, including sales, quantity, and profit details across different product categories and regions. The primary focus was on calculating profit margins at the sub-category level to identify underperforming product lines. Using SQL for aggregation, Python for visualization, and Tableau for interactive dashboarding, we developed insights that can help optimize retail performance and maximize profitability.

**🔹 Tools & Technologies Used**

* SQL – For data cleaning, transformation, and margin computation
* Python (Pandas, Seaborn, Matplotlib) – For EDA and visual analytics
* Tableau – Optional dashboard for regional/category filters (not shown in this report)

**🔹 Steps Involved**

1. **Data Cleaning & Preparation**
   * Loaded and preprocessed a 1000-row simulated retail dataset with fields: Order Date, Sales, Profit, Category, Sub-Category, etc.
   * Removed missing values and standardized formats.
2. **SQL Analysis**
   * Aggregated total sales and total profit by sub-category.
   * Calculated profit margin using the formula:  
     Profit Margin (%) = (SUM(Profit) / SUM(Sales)) \* 100
3. **Python Visualizations**
   * Generated horizontal bar plots for profit margin by sub-category.
   * Created monthly sales trend analysis.
   * Analyzed discount vs. profit scatter plots to explore profitability drivers.
4. **Insight Derivation**
   * Identified high-revenue but low-margin categories.
   * Highlighted profitable sub-categories suitable for strategic marketing.

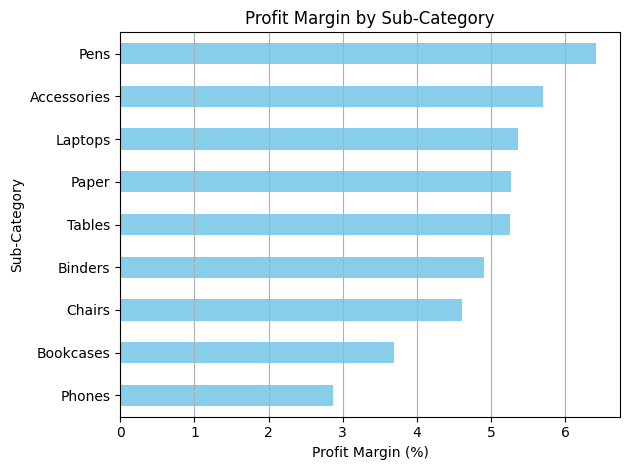
**🔹 SQL Query Used**

*SELECT Category, `Sub-Category`, SUM(Sales) AS Total\_Sales, SUM(Profit) AS Total\_Profit,*

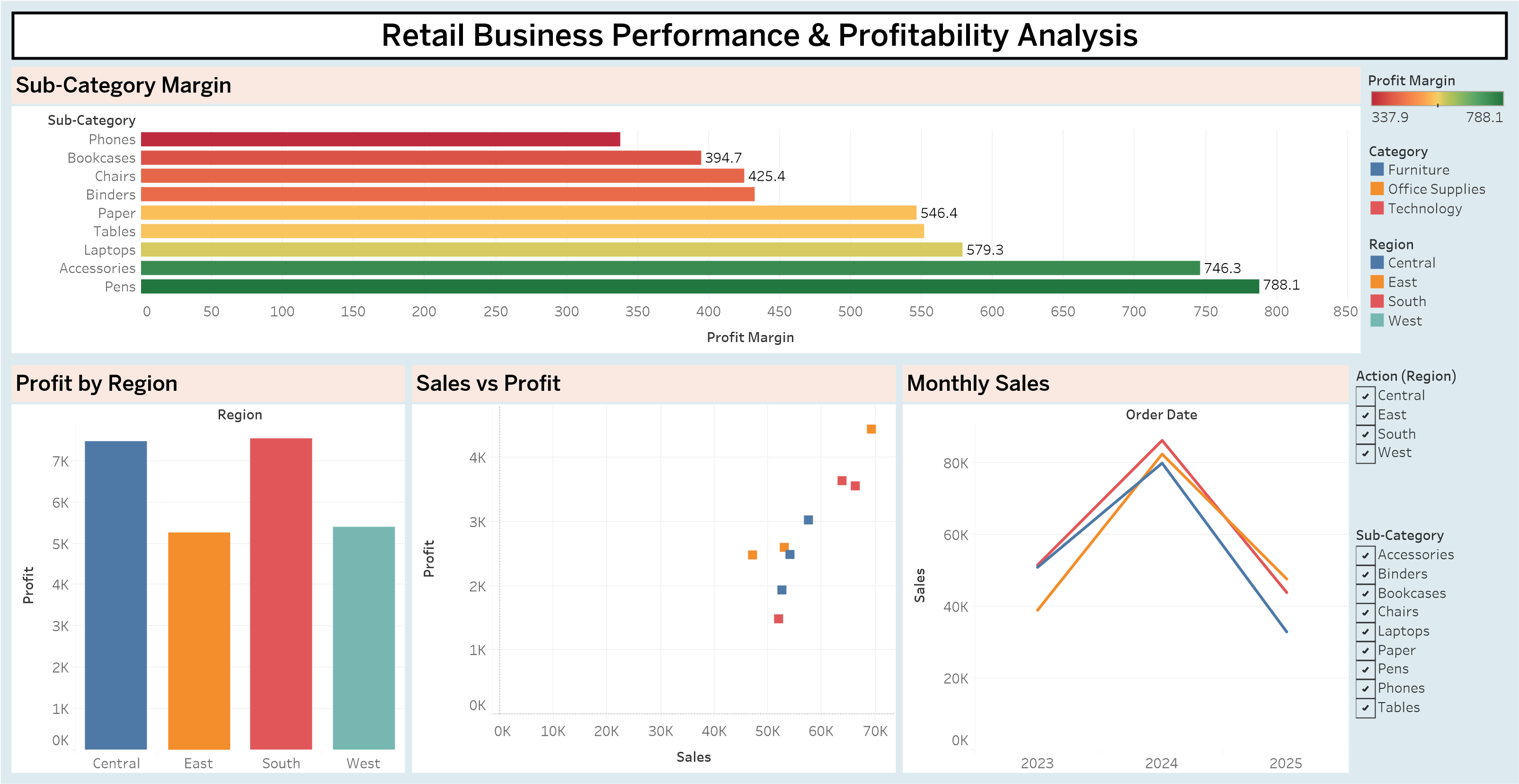
*ROUND(SUM(Profit)/SUM(Sales)\*100, 2) AS Profit\_Margin*

*FROM superstore GROUP BY Category, `Sub-Category` ORDER BY Profit\_Margin;*

**🔹 Key Visual Insight**



**🔹 Dashboard View**



**🔍 Observation:**

* Phones had the lowest profit margin (2.87%) despite high sales, signaling potential over-discounting or cost inefficiencies.
* Pens and Accessories emerged as high-margin products (6.4%+), indicating ideal candidates for focused marketing or bundling.

**🔹 Conclusion**

This analysis demonstrated how combining SQL and Python can surface actionable business insights from transactional retail data. Identifying low-margin products such as Phones helps redirect pricing or procurement strategies, while capitalizing on high-margin items like Pens or Accessories boosts bottom-line performance. Future work may include integrating inventory days, customer segments, and seasonal behaviour into a dashboard for real-time monitoring.