SQL Retail Data Analysis Project Report

This document presents the results and analysis of your retail transactional dataset (`Retail\_with\_categories.csv`) performed using structured SQL queries. The objective of this project was to uncover insights from the sales, products, categories, customer behavior, and profitability trends that can help drive strategic decisions.

# 1. Project Objective

The aim of this analysis was to:  
- Explore the structure and quality of the data.  
- Identify top-performing and underperforming product categories and sub-categories.  
- Analyze sales performance across categories.  
- Understand product-level revenue contributions.  
- Assist with inventory-related insights such as overstocked or slow-moving products.

# 2. Data Overview

The dataset `Retail\_with\_categories.csv` includes retail sales transactions with the following relevant columns:

* Transaction\_ID
* Date
* Customer\_Name
* Product
* Total\_Items
* Total\_Cost
* Payment\_Method
* City
* Store\_Type
* Discount\_Applied
* Customer\_Category
* Season
* Promotion
* Unit\_Price
* Cost\_Per\_Item
* Profit
* Customer\_ID
* Basket\_Size
* Product\_Category
* Category
* Sub\_Category

The data was clean with no missing/null values in key columns.

# 3. SQL Query Highlights

- Retrieved first 10 sample rows for initial structure review.

- Checked for NULL values in key transactional fields — result: No missing data.

- Extracted all unique product categories present in the dataset.

- Created a SQL view (`retail\_with\_categories`) to assign logical product Categories and Sub-Categories based on the product name.

- Queried all data from the view to verify structure and transformations.

- Retrieved distinct `Category` and `Sub\_Category` values for filtering purposes.

- Calculated total transactions, total sales, total profit, and profit margin % for each Category and Sub-Category.

# 4. Key Business Insights

- Products in the `Meat & Seafood`, `Pantry & Snacks`, and `Personal Care` categories show high total sales and strong profit margins.

- Categories like `Garden Tools` and `Home & Misc` had fewer transactions but relatively higher margins, indicating specialized niche appeal.

- `Condiments`, `Dairy`, and `Grains & Bakery` performed well across both sales and profit.

- Slow-moving items can be identified from sub-categories with high inventory but low profit margins.

- The analysis supports inventory optimization by identifying fast-selling vs. underperforming categories.

# 5. Project Deliverables

- SQL script for all analysis queries.

- Cleaned dataset with added Category and Sub-Category fields.

- Client-friendly report document.

- Power BI Dashboard upon request.

# 6. Final Notes

This SQL-based analysis provides a strong foundation to support business decisions regarding pricing, inventory management, and product promotions.

# 7. Python Pandas : Retail Business Performance & Profitability Analysis

1. **Data Loading & Inspection:**
   * The dataset is loaded from a CSV file named "Retail\_with\_categories.csv".
   * The first 10 rows are displayed to understand the structure and contents of the data, including variables such as transaction details, customer information, product details, and financial metrics.
2. **Data Overview:**
   * Key columns such as 'Transaction\_ID', 'Date', 'Customer\_Name', 'Product', 'Total\_Items', 'Total\_Cost', 'Payment\_Method', 'City', 'Store\_Type', 'Discount\_Applied', 'Customer\_Category', 'Season', 'Promotion', 'Unit\_Price', 'Cost\_Per\_Item', 'Profit', 'Customer\_ID', 'Basket\_Size', 'Product\_Category', 'Category', 'Sub\_Category' are inspected for initial understanding.
3. **Correlation Analysis:**
   * A correlation matrix is computed between 'Inventory\_Days' and 'Profit'.
   * The goal is to evaluate the relationship between the duration inventory remains (Inventory Days) and the profit earned.
   * The result shows a **positive correlation of approximately 0.27**, indicating a moderate relationship where increased Inventory Days tend to be associated with higher profit, but the relationship is not very strong.
4. **Data Visualization - Scatter Plot:**
   * A scatter plot is created to visualize the relationship between 'Inventory Days' and 'Profit'.
   * The plot reveals a general trend where profit tends to increase with the number of Inventory Days, but with considerable spread, indicating other factors may influence profit as well.
5. **Interpretation & Insights for Client:**
   * The analysis suggests that holding inventory for longer periods may be somewhat associated with higher profitability.
   * However, since the correlation is moderate, other factors like product type, seasonality, and sales strategies could also be impacting profitability.
   * The scatter plot helps to visually assess the pattern and variability in the data.
6. **Create Pivot Table:**
   * Organized 'Profit' data with 'Store\_Type' as rows and 'Season' as columns.
7. **Generate Heatmap:**
   * Visualized the pivot table using a heatmap to show profit levels across store types and seasons.
   * Used color gradient ('coolwarm') to indicate profit magnitude.
   * Annotated each cell with profit values for clarity.

# 8. Power BI Dashboard Insights

The Power BI dashboard created for this project visualizes key performance indicators and patterns derived from the SQL-based retail dataset. Below is a storytelling-style summary of insights derived from each visual component of the dashboard.

## Retail Sales Performance Dashboard Insights

### Overall Sales Summary

- Total Sales: The business has achieved 52.7K in sales, reflecting solid revenue generation.

- Total Profit: The profit stands at 23.8M, indicating healthy profitability.

- Profit Margin: The current profit margin is 45.1%, demonstrating efficient cost management and good profitability percentage.

### Sales Breakdown by Category

- Insight: The Personal Care category leads sales, with about 7.1M, showing strong demand in this segment.

- Other Categories: Categories like Pantry & Snacks and Grains & Breads also contribute significantly, emphasizing diverse product performance.

- Implication: Focus on top-performing categories can optimize marketing and inventory planning.

### Profit Contribution by Sub-Category

- Highlight: Yogurt and Water are the most profitable items, generating 125.56 and 278.16 in profit respectively.

- Low Performers: Items like Tuna and Toothbrush contribute minimal profits (around 200-300), suggesting potential areas for efficiency improvement.

- Recommendation: Boost sales of high-profit items and review low-profit categories to maximize margins.

### Sales by Season (Pie Chart)

- Insight: The most sales occur during Fall (27.4%) followed closely by Spring (25.7%) and Winter (21.6%).

- Implication: Seasonal promotions should be aligned with these peak periods to leverage higher sales.

### Slow-Moving Items (Table)

- Observation: Items like Toothbrushes and Tissues have relatively low sales volume (36-52 units), indicating slow inventory turnover.

- Action Point: Consider discounting or promotional strategies to move slow stock faster, reducing holding costs.

### Basket Size vs. Items per Basket (Scatter Plot)

- Insight: As basket size increases, the number of items per basket also tends to increase, suggesting that larger baskets contain more products.

- Implication: Cross-selling and bundle offers could encourage customers to add more items, increasing overall sales.

## Summary & Recommendations

- The data reveals a strong sales and profit foundation, with particular strength in personal care and seasonal peaks.

- Focus on promoting high-profit items like yogurt and water.

- Address slow-moving inventory to optimize stock levels.

- Utilize seasonal trends for targeted marketing efforts.

- Encourage larger basket sizes through strategic bundling.