

Retail Store Data Analysis: Sales, Profitability, and Customer Behavior

1. Project Title:

Customer Classification and Sales Profitability Analysis for Retail Store

2. Project Overview:

This project analyzes customer data and sales transactions to understand key drivers of profitability, identify high-value customer insights, and optimize sales strategies. By analyzing various factors such as **order priority**, **discounts**, **shipping costs**, and **profit margins**, the project will help the retail store make data-driven decisions to increase revenue and customer satisfaction.

3. Objectives:

The primary objectives of this analysis are:

- Identify customer segments that generate the most revenue and profit.
- Analyze the relationship between **order priority** and **profitability**.
- Understand the impact of **discounts** on **sales** and **profit margins**.
- Explore the effect of **shipping costs** and **ship mode** on overall profitability.
- Identify top-selling products and categories.
- Segment customers by geographical location and their purchasing behavior.
- Evaluate sales performance over time based on **order date** and **ship date**.

4. Data Collection:

The dataset contains that was sourced from kaggle.com, it does not contain real patient store data.

The dataset contains the following key fields:

- **Customer Details:** Customer name, customer ID, and location data (country, region, state, city, postal code).
- **Sales Information:** Order ID, order priority, unit price, discount, shipping cost, sales, profit, quantity ordered, and product details (product category, sub-category, container type, and product name).
- **Order Timing:** Order date and ship date.
- **Shipping Details:** Ship mode and shipping cost.
- **Profit Metrics:** Product base margin, profit, and sales.

5. Data Cleaning and Preprocessing:

Steps involved in preparing the data for analysis:

- **Handle Missing Data:**
 - Check for missing or incomplete values in fields like **shipping cost**, **profit**, and **order date**.
 - Impute missing values where necessary or remove records with missing critical fields.
 - Transformed the **Profit** column to two decimals places to make it resemble money format.
- **Data Standardization:**
 - Standardize date formats for **order date** and **ship date**.
 - Ensure **product names** and **categories** are consistently formatted.
- **Feature Engineering:**
 - Calculate the **profit margin** for each transaction by dividing **profit** by **sales**.
- **Remove Duplicates:** Ensure there are no duplicate rows based on **order ID**.

6. Data Analysis:

Begin by conducting a broad analysis to understand the overall distribution of key variables and potential relationships.

- **Sales Trends Over Time:**
 - Plot total sales by **order date** to see sales patterns over months/quarters.

- Create a line chart to track total sales and profit by **order date**.
- **Top Products:**
 - Bar chart of the **top-selling products** by total sales and profit.
 - Analyze the relationship between **product category** and **total sales** to identify high-performing categories.
- **Customer Segments:**
 - Pie chart or bar plot to visualize the distribution of sales and profit across **customer segments** (e.g., **Consumer, Corporate, Home Office**).
 - Compare the total **sales** and **profit** across different **regions** (e.g., **North America, Europe**).
- **Discount Impact:**
 - Scatter plot to explore the relationship between **discount percentage** and **profit**.
 - Compare **profit margins** for products sold with high discounts versus those with low or no discounts.

7. Key Findings:

Based on the analysis, summarize the most important insights and patterns.

- **Customer Segmentation:**
 - The **Corporate** segment generates the highest total revenue and profit, followed by **Home Office** segment.
 - **East** region accounts for the most sales, with the **West** right behind it.
- **Discount and Profit Margins:**
 - Higher **discounts** (above 10%) tend to reduce **profit margins**.
- **Order Priority and Profit:**
 - **High priority orders** (urgent) tend to generate higher sales and profits due to quicker shipping but have slightly higher **shipping costs**.
 - **Low priority orders** generate the lowest profit margins, often because of delayed shipping and lower-order quantities.
- **Shipping Mode and Sales:**
 - **Express Air shipping** results in higher shipping costs but is linked to more expensive products and higher **profit margins**.
 - **Delivery Truck shipping** is more cost-effective.
- **Top-Performing Products:**
 - The most profitable products are in the **technology** and **furniture** categories.

9. Data Visualizations:

Present the insights using visualizations:

- **Sales Trends:** Line graph showing **total sales** over time, highlighting seasonal trends.
- **Top Products:** Bar chart of **top-selling products** by profit.
- **Profit vs. Discount:** Scatter plot showing **profit margins** at varying **discount** levels.
- **Customer Segments:** Pie chart showing the proportion of total **sales** by **customer segment**.
- **Sales by Region:** Map visualization showing **sales by state** or **city**.
- **Order Priority Impact:** Comparing **sales** and **profit** across different **order priorities**.

10. Recommendations:

Based on the findings, provide actionable recommendations for the retail store.

- **Increase Targeting of High-Value Customers:** Focus on **Corporate** customers, who generate the highest revenue and profit.
- **Optimize Discount Strategy:** Avoid large discounts on low-margin products. Offer smaller discounts on higher-margin items to boost profitability.
- **Improve Shipping Efficiency:** Consider revising the **shipping policy** for **low priority orders** to reduce costs. Invest in faster shipping options for **high priority orders** to maintain high sales and customer satisfaction.
- **Expand Successful Products:** Increase the stock and marketing of top-selling products.
- **Regional Expansion:** Explore expanding physical stores or targeted marketing in the east and west region.

11. Conclusion:

Summarize the key insights and their implications for business strategy.

This analysis has revealed critical insights into customer behavior, sales trends, and profitability. By focusing on high-value customer segments, optimizing discounts, and improving shipping efficiency, the retail store can significantly increase profitability. Furthermore, regional expansion in high-performing areas can lead to more sales opportunities, and targeting specific products with higher profit margins will optimize inventory and marketing strategies.