



PROBLEM STATEMENT

Customer Goods Purchase Data Analysis



PROBLEM STATEMENT

The problem statement involves analyzing a dataset containing transaction details of customer purchases from a retail business, with the goal of deriving meaningful insights through univariate, bivariate, and multivariate analysis.

The dataset consists of various attributes, including Invoice ID, Branch, City, Customer type, Product line, Unit price, Quantity, Tax, Total, Payment method, Gross income, and Customer ratings, among others.

By analyzing these various aspects of customer purchasing behavior, the business can make data-driven decisions to optimize branch operations, marketing strategies, inventory management, and customer satisfaction. The end goal is to enhance profitability, streamline operations, and improve customer experiences.

BUSINESS PROBLEM OVERVIEW

Several business challenges are addressed that are faced by a retail company in analyzing customer purchase data to optimize operations and improve sales performance. The dataset provides detailed transactional data across different branches, product lines, and customer demographics, including attributes such as Branch, Product line, Payment method, Customer ratings, Quantity sold, Unit price, and Gross income.

The overarching business problem is to derive actionable insights from this data that can guide decision-making and enhance business performance.

One primary challenge is understanding the sales performance across different branches. By identifying which branches generate the highest sales and profits, the business can allocate resources effectively, optimize staffing, and tailor marketing efforts to maximize revenue in underperforming locations.

Ultimately, the analysis aims to address these business challenges by providing data-driven insights that can inform strategic decisions to improve profitability, operational efficiency, and customer engagement across all branches.

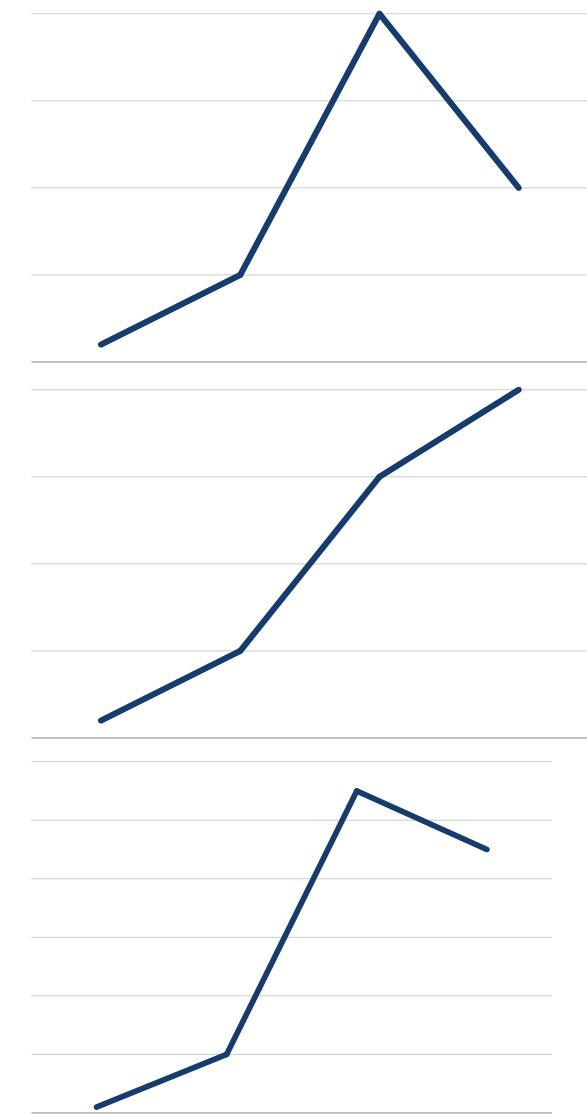


Customer goods purchase refers to the transactional behavior and patterns of customers as they buy goods from a retail store or online platform. The dataset captures several aspects of these purchases, including the type of products purchased, the quantities bought, the payment methods used, and the customer ratings for products.

Essentially, customer goods purchase refers to the decision-making process of consumers as they choose products, make transactions, and provide feedback on their buying experience.

The code analyzes various dimensions of customer purchases, such as product lines (e.g., "Health and Beauty," "Food and Beverages"), to determine which categories generate the highest sales and gross income.

It also evaluates the quantity of products sold across different product lines and branches, helping the business understand which products are in demand and which are not. By identifying trends in customer purchases, the company can optimize inventory, adjust product offerings, and create targeted marketing strategies for the most popular or profitable goods.



UNDERSTANDING & DEFINING DATASET

PROJECT PIPELINE

The project pipeline can be briefly summarized in the following steps:

- Data Understanding: Here, we need to load the data and understand the features present in it. This would help us choose the features that we will need for your final model.
- Exploratory data analytics (EDA): Normally, in this step, we need to perform univariate and bivariate analyses of the data, followed by feature transformations, if necessary. For the current data set, because Gaussian variables are used, we do not need to perform Z-scaling. However, you can check if there is any skewness in the data and try to mitigate it, as it might cause problems during the model-building phase.





THANK YOU