Introduction

The project focuses on examining consumer shopping data to derive insights that support data-driven decision-making. Using a combination of Python for data preparation, SQL for structured analysis, and Power BI for visualization, the study will transform raw transactional data into meaningful information. The dataset contains various customer attributes, purchase records, and behavioral indicators. Through systematic cleaning, analysis, and visualization, the project will help uncover the patterns behind customer preferences, spending tendencies, and loyalty behaviors that influence the company's business performance.

Business Problem Statement (sample)

A leading retail company seeks to better understand its customers' shopping behavior to improve overall sales, satisfaction, and brand loyalty. Recent trends have shown shifts in purchasing patterns across different demographics, product categories, and sales channels. The management team wants to identify which factors—such as discounts, product reviews, seasons, and payment preferences—most influence buying decisions and repeat purchases. By analyzing customer data, the company aims to uncover key behavioral insights that can help optimize marketing strategies, strengthen customer engagement, and drive sustained revenue growth.

Objectives

The main objective of this project is to analyze customer shopping behavior to identify trends and actionable insights that can guide business improvements. The project aims to prepare and structure the data for analysis, discover key factors that affect purchase amounts and frequency, and visualize results to help stakeholders understand customer patterns easily. By the end of the project, the findings should enable the company to make informed marketing, pricing, and product placement decisions that align with customer expectations and maximize profitability.

Approach

The project will begin with comprehensive data preparation using Python, where the dataset will be cleaned, standardized, and enhanced with new analytical features such as customer age groups and purchase frequency measures. After cleaning, the data will be imported into a relational SQL database for deeper analysis, allowing for queries that reveal relationships between gender, subscription status, discount usage, and revenue. The results from SQL will then be visualized in Power BI to create an interactive dashboard that highlights key insights like revenue by category, season, and customer type. This integrated approach will ensure that the data is accurate, interpretable, and visually meaningful for decision-making.

Expected Outcome

By the end of the project, the company will have a complete analytical view of customer purchasing patterns and loyalty trends. The final output will include a well-cleaned dataset, analytical insights from SQL queries, and an interactive Power BI dashboard. These outcomes will reveal which customer groups contribute most to revenue, how discounts and reviews influence purchases, and which products and seasons perform best. Ultimately, the insights will guide the company in enhancing marketing strategies, improving customer engagement, and driving business growth through data-backed decisions.