

# Business Problem Statement

A retail company wants to better understand how its customers shop in order to improve sales and customer retention. The company has data about customer purchases, product categories, discounts, subscriptions, reviews, and basic demographics.

Management is interested in answering questions such as:

- Which product categories generate the most revenue?
- Do subscribed customers spend more than non-subscribed customers?
- How do discounts affect customer spending?
- Which customer groups are the most valuable?

You are asked to analyze the customer behavior dataset and provide insights that can help the company improve its marketing strategy and customer engagement.

## Deliverables

### 1. Data Preparation (Python)

Clean and preprocess the raw dataset using Python (Pandas) by handling missing values, correcting data types, and preparing the data for analysis.

### 2. Data Analysis (SQL)

- Write SQL queries to analyze revenue, customer segments, subscriptions, and purchase behavior
- Use basic aggregations (SUM, AVG, COUNT) and GROUP BY to extract insights
- Apply simple subqueries where needed to compare customer segments and performance metrics

### 3. Visualization (Power BI)

Build an interactive dashboard with KPIs, filters, and segment-level insights to support business decision-making.

### 4. Report and Insights

Summarize key findings and translate them into clear, actionable insights for marketing and product teams.

### 5. GitHub Repository

Store all Python scripts, SQL queries, and dashboard files in a structured and well-documented repository.

### 6. Business Recommendations

Provide data-driven recommendations on pricing, discount strategies, subscription models, and customer engagement.