

# Business Problem Statement

A leading retail company wants to better understand its customers' shopping behavior in order to improve sales, customer satisfaction, and long-term loyalty. The management team has noticed changes in purchasing patterns across demographics, product categories, and sales channels (online vs. offline). They are particularly interested in uncovering which factors, such as discounts, reviews, seasons, or payment preferences, drive consumer decisions and repeat purchases.

You are tasked with analyzing the company's consumer behavior dataset to answer the following overarching business question:

**“How can the company leverage consumer shopping data to identify trends, improve customer engagement, and optimize marketing and product strategies?”**

## Deliverables

1. **Data Preparation & Modelling (Python):** Clean and transform the raw dataset for analysis.
2. **Data Analysis (SQL):** Organize the data into a structured format, simulate business transactions, and run queries to extract insights on customer segments, loyalty, and purchase drivers.
3. **Visualization & Insights (Power BI):** Build an interactive dashboard that highlights key patterns and trends, enabling stakeholders to make data-driven decisions.
4. **Report and Presentation:** Write a clear project report summarizing your key findings and business recommendations. Prepare a presentation that visually communicates insights and actionable recommendations to stakeholders.
5. **GitHub Repository:** Include all Python scripts, SQL queries, and dashboard files in a well-structured repository.

### Problem

- Q1. What is the total revenue generated by male vs. female customers?
- Q2. Which customers used a discount but still spent more than the average purchase amount?
- Q3. Which are the top 5 products with the highest average review rating?
- Q4. Compare the average Purchase Amounts between Standard and Express Shipping.
- Q5. Do subscribed customers spend more? Compare average spend and total revenue between subscribers and non-subscribers.
- Q6. Which 5 products have the highest percentage of purchases with discounts applied?

- Q7. Segment customers into New, Returning, and Loyal based on their total number of previous purchases, and show the count of each segment.
- Q8. What are the top 3 most purchased products within each category?
- Q9. Are customers who are repeat buyers (more than 5 previous purchases) also likely to subscribe?
- Q10. What is the revenue contribution of each age group?