

Business Problem Statements and Questions

The company is experiencing inconsistent revenue growth and wants to understand what drives customer spending across different segments. Management needs clarity on which customer groups generate the most value, which products perform best, and how factors such as discounts, subscriptions, shipping preferences, and age groups influence revenue. The goal is to uncover spending patterns, identify high-value customer segments, optimize product listings, and strengthen retention strategies to improve profitability and guide future marketing and sales decisions.

Given: a dataset ‘customer_behavior_dataset.csv’

Questions:

1. Total revenue generated by male vs female customers?
2. To count the number of customers used the discount but still spent more than the average purchase amount?
3. Which are the top 5 products with the highest average review rating?
4. Compare the average purchase amounts between standard and express shipping?
5. Do subscribed customers spend more? Compare average spend and total revenue between subscribers and non-subscribers.
6. Which 5 products have the highest percentage of purchases with discounts applied?
7. Segment customers into New, Returning, and Loyal-based on their total number of previous purchases, and show the count of each segment.
8. What are the top 3 most purchased products within each category?

9. Are customers who are repeat buyers (more than 5 previous purchases) also likely to subscribe? A9: Subscription_yes= 958buyers, Subscription_no= 2518buyers

10. What is the revenue contribution of the each age group?