Marketing Campaign Analysis

Rule-Based Customer Segmentation and Business Insights

---

1. Introduction

In today's competitive retail environment, understanding customer behaviour is essential for running successful marketing campaigns. Retail companies often collect large volumes of customer data, but without proper analysis, this data cannot be converted into actionable insights.

This project focuses on analysing customer demographics, spending behaviour, channel usage, and marketing campaign responses to identify valuable customer segments. The aim is to support data-driven decision-making by providing clear insights into customer behaviour and campaign effectiveness.

---

2. Business Problem Statement

The retail company conducted multiple marketing campaigns and recorded detailed customer-level data. Management wants a consolidated analytics solution to answer the following questions:

- Who are the most valuable customers?

- Which customer segments respond best to marketing campaigns?

- How do spending patterns differ across customer groups?

- Which marketing channels perform best for high-value customers?

- Which customer profiles should be targeted in future campaigns?

---

3. Dataset Description

The dataset consists of customer-level information, where each row represents a unique customer. The key categories of variables include:

- **Demographics**: Age, Income, Education, Marital Status, Country

- **Customer Relationship**: Customer tenure, recency

- **Spending Behaviour**: Spending on wine, meat, fruits, fish, sweets, and gold products

- **Channel Usage**: Web purchases, store purchases, catalogue purchases, deal purchases, website visits

- **Campaign Data**: Acceptance of multiple marketing campaigns and overall campaign response

A separate data dictionary was used to understand the meaning of each variable.

---

4. Data Cleaning and Feature Engineering

To prepare the data for analysis, the following steps were performed:

- Removal of records with missing values in critical fields, such as income

- Conversion of date fields into appropriate date formats

- Handling of unrealistic values, such as invalid ages

- Creation of derived features:

  - **Age** (calculated from year of birth)

  - **Children** (sum of kids and teenagers at home)

  - **Total Spend** (sum of spending across all product categories)

  - **Total Purchases** (sum of purchases across all channels)

  - **Customer Tenure** (time since customer enrollment)

These derived features played a crucial role in segmentation and analysis.

---

5. Exploratory Data Analysis (EDA)

Exploratory Data Analysis was conducted to understand customer behaviour and uncover patterns. Key analyses included:

- Distribution of age and income across customers

- Relationship between income and total spending

- Campaign response rates across different income and age groups

- Channel usage patterns for high-value customers

- Identification of customer segments with high engagement but low spending

The insights from EDA helped guide the segmentation strategy and dashboard design.

---

6. Rule-Based Customer Segmentation

Customers were segmented using clear and interpretable business rules rather than complex machine learning models. The segmentation criteria are shown below:

Segment Name	Rule Definition
High Income	Income > ₹75,000
Young Customer	Age < 30
Campaign Responder	Response = 1
High Web Engagement	Website visits > 5 per month
Family Customer	Children > 0
High Spender	Top 10% customers by total spending

This rule-based approach ensures transparency and makes the results easy to explain to non-technical stakeholders.

---

7. SQL Data Modelling and Analysis

A structured SQL schema was created to store cleaned customer data. SQL queries were written to:

- Calculate campaign response rates
  Analyse average spending across segments.

- Evaluate channel usage for high-value customers.

- Generate segment-level summaries for dashboard use.

SQL was used to validate analytical results and support dashboard metrics.

---

8. Dashboard Overview

An interactive dashboard was developed using Streamlit to allow stakeholders to explore the data easily. The dashboard provides:

- Segment-based filtering

- Key Performance Indicators (KPIs)

- Customer-level data previews

- Robust handling of missing or small segment data

The dashboard enables business users to make informed decisions without requiring technical expertise.

---

9. Key Findings

- High-income and high-spending customers show significantly higher campaign response rates.

- Family customers contribute a large share of overall revenue

- Young customers have lower spending but show strong engagement potential. al

- Web-based channels play a critical role for high-value customer segments.

- Certain segments show high engagement but low campaign response, indicating opportunities for improvement.

---

10. Business Recommendations

Based on the analysis, the following recommendations are proposed:

1. Focus premium campaigns on high-income and high-spending customers

2. Increase investment in web-based marketing channels.ls

3. Design personalised offers for family customers.

4. Develop engagement strategies to convert young customers into higher spenders.s

5. Reduce marketing spend on low-response segments

6. Use segmentation-driven targeting for future campaigns.

---

11. Conclusion

This project demonstrates how customer data can be transformed into meaningful business insights using data cleaning, exploratory analysis, rule-based segmentation, SQL analytics, and interactive dashboards. The approach is scalable, interpretable, and suitable for real-world marketing analytics applications.

---

12. Tools and Technologies Used

- Python (Pandas, NumPy)

- SQL

- Streamlit

- Jupyter Notebook

- Data Visualisation Libraries

- Microsoft Word / Google Docs (for reporting)

---

13. Final Note

This project fulfils all capstone requirements, including data preprocessing, EDA, segmentation, SQL analysis, dashboard development, and business-oriented reporting.