Marketing Campaign Customer Segmentation

• • •

December 16, 2024
Humphrey Hui
MIT ADSP Capstone Final Presentation

Agenda

Executive Summary

What was done

- Addressed null values and removed duplicates
- Categorical variables refined by merging similar categories
- Feature Engineering: Created additional columns e.g. Age, Total Purchases, Average Spend
- Exploratory Analysis: Univariate and Bivariate analysis, data scaling
- Five clustering methods tested, best one determined through silhouette score and cluster profiles

Outcome

- Identified correlations e.g. purchases negatively correlated with having children
- Found that **K-Means with 3 clusters** provided the optimal clusters
- Determined that income was the biggest driver in behaviour
- Characterised 3 clusters:
  - Higher-Income Spenders
  - Deal-Oriented Average Shoppers
  - Budget-Conscious Families

Next Steps

- Recommendations
  - White-Label Product Line focusing on family items
  - Refer-a-Friend program
- Identified action items for different teams
  - Marketing
  - Product
  - Sales
  - Data
  - E-Commerce

Problem Definition

Context

In today's competitive market, understanding customers is essential for strategic growth. Ignoring differences between groups leads to generic experiences that fail to engage customers, resulting in missed opportunities, lower satisfaction, and reduced revenue, especially for a company with a broad product portfolio.
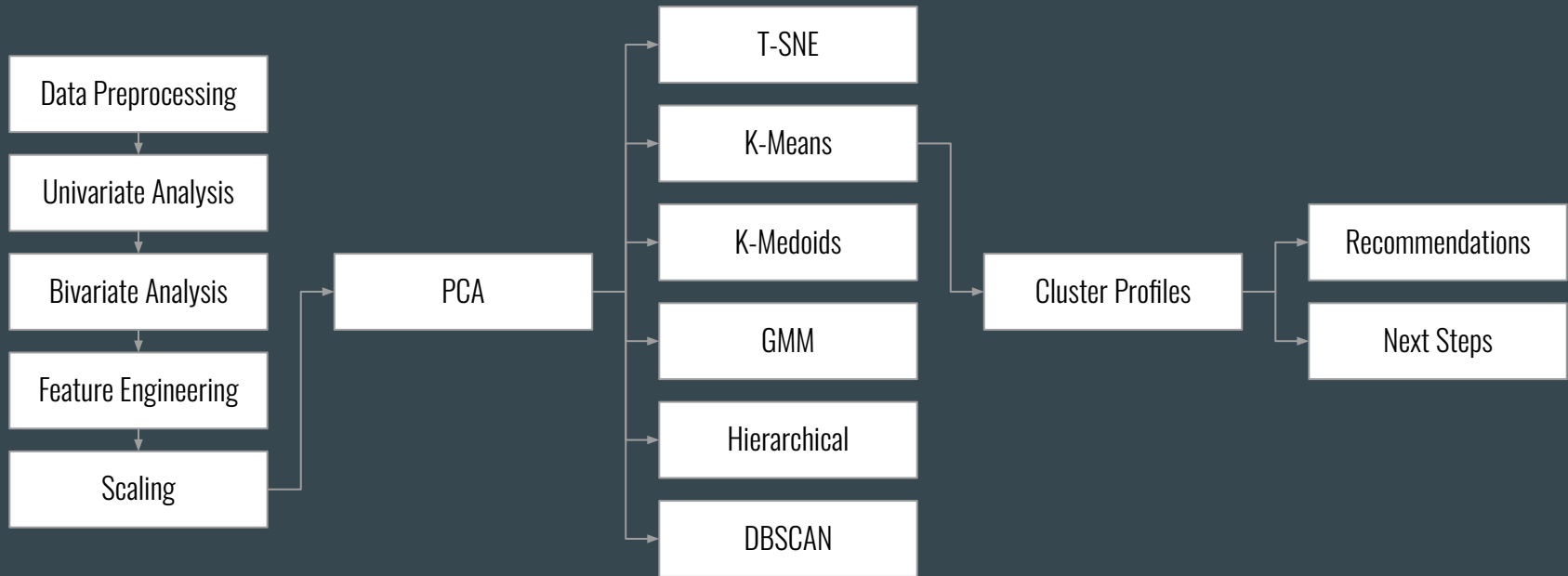
Objective

Our goal is to segment the customer base using behavioral and demographic data to create a framework for personalized engagement. This segmentation will help us allocate resources effectively to maximize impact. For instance, if we identify a group that responds well to discounts, we can design targeted promotions to boost conversions. For segments less motivated by price, we can create value-based offers, protecting margins.

Key questions

- How do our customer demographics influence purchasing behaviour?
- What clustering techniques deliver the most actionable customer segments? What are their profiles?
- What are the next steps for stakeholders to improve their approach to these segments?

# Solution Approach

```
Data Preprocessing
        ↓
Univariate Analysis
        ↓
Bivariate Analysis  ──→  PCA  ──→  ┌─ T-SNE
        ↓                          ├─ K-Means ──→ Cluster Profiles ──→ ┌─ Recommendations
Feature Engineering                ├─ K-Medoids                        └─ Next Steps
        ↓                          ├─ GMM
     Scaling                       ├─ Hierarchical
                                   └─ DBSCAN
```

Data Cleaning, Feature Engineering, and EDA

Data Cleaning + Feature Engineering

- Null values in income: replaced with median
- Removed ID column
- Unnecessary categories in Marital Status and Education: combined with larger categories
- Created new features:
  - Age, Childrenhome, Household size, Total Spend (MntTotal), and Total Purchases

Exploratory Data Analysis

Univariate Analysis

- Many variables were right skewed
- 15% have responded to campaigns, only 1% have made complaints
- Graduation (bachelors) and being married are the most common education and marital statuses
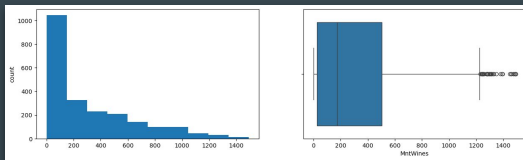
Bivariate Analysis

- Channels were negatively correlated with each other
- Income positively correlated with purchasing across categories
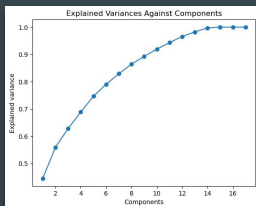- Purchasing negatively correlated with kids
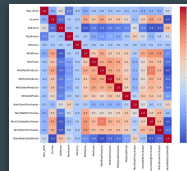
Clustering Methods

PCA
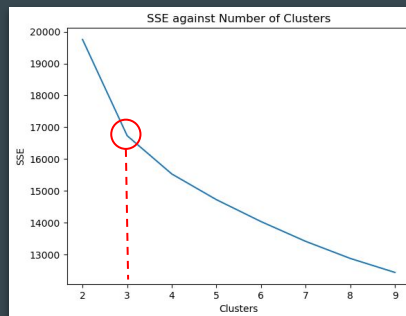
PCA

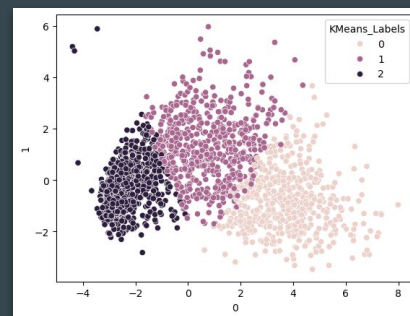- 10 components chosen (91.93% explained variance)

Chosen Method

K-Means

- Elbow identified at 3 clusters
- Highest silhouette score obtained: 0.29
- No cluster overlap in PCA visualization
- Clear cluster characteristics found through group summary statistics

SSE Plot

PCA Visualization

Group Summary Statistics

|	Group 0 Mean	Group 1 Mean	Group 2 Mean	Group 0 Median	Group 1 Median	Group 2 Median
Year_Birth	1967.880392	1966.115385	1971.069215	1969.000000	1966.000000	1972.0
Income	76665.962745	57517.082168	36345.072314	76807.500000	57958.000000	35482.0
Kidhome	0.033333	0.272727	0.766529	0.000000	0.000000	1.0
Teenhome	0.219608	0.812937	0.485537	0.000000	1.000000	0.0
Recency	49.954902	48.005245	48.919421	52.500000	48.000000	49.0

Group 0

25%

Higher-Income Spenders

Profile
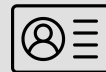
- Less kids
- Makes the most web purchases
- Makes more catalog purchases
- Higher response to campaign offers

Targeting Strategy

- Maintain the current experience
- Create a refer-a-friend program: higher income individuals tend to know other higher income individuals
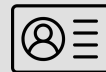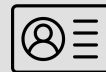
Group 2

50%

Budget-Conscious Families

Profile

- More kids
- Slightly younger
- Less web purchases and but more web visits
- Highest proportion (50%)

Targeting Strategy

- White-label store product brand focused on family goods
- Use online store to generate interest
- Target for campaigns, remove barriers of entry

Recommendations and Action Items

White-Label Product Line ★

Launch affordable products, focusing on family and kids' items.

Action Items
- **Product Team:** Identify a white label producer to partner with.
- **Marketing Team:** Create digital campaigns to attract budget-conscious families.
- **E-commerce Team:** Improve visibility of white-label products on the website.
- **Data Team:** Monitor adoption rates and analyze campaign success.

Refer-a-Friend Program

Leverage their campaign responsiveness by introducing a referral program.

Action Items
- **Marketing Team:** Develop referral campaigns targeting high-income individuals.
- **Product Team:** Ensure premium products are highlighted.
- **Sales Team:** Track and optimize the referral program.
- **Data Team:** Monitor engagement metrics to refine targeting.

Potential Benefits and Costs

Benefits

Increased Revenue

Customer Loyalty

Optimized Marketing

New Customers

Costs

Implementation Costs

Existing Infrastructure

Training Needs

Marketing Budget

Risk and Mitigation Strategies

Inaccurate Segment Identification

- Risk: Misidentifying segments could lead to ineffective campaigns
- Mitigation: Regularly update segments based on fresh data.

Too Much Personalization

- Risk: Customers may find excessive targeting invasive.
- Mitigation: Set limits on communication frequency and provide opt-out options.

Low Adoption of White-Label Products

- Risk: Budget-conscious customers may not adopt the new products
- Mitigation: Test products with small groups and collect feedback

Questions