



Customer Segmentation

Overview

- Customer Segmentation is the “practice of dividing a customer base into groups of individuals that are similar in specific ways relevant to marketing.” (Source: Wikipedia)
- There are 5 types of segmentation in general:
 - geographic segmentation
 - demographic segmentation – B2B
 - demographic segmentation – B2C
 - psychographic segmentation
 - behavioral segmentation

Source: <https://www.thebridgecorp.com/customer-segmentation/>

Problem Statement

To segment customers using cluster analysis from a large collection of data.

- Collection of 8 datasets with customer transaction data is provided
- Collected for 2500 household
- Varied demographic factors like age, income, household size etc.
- Product information
- Coupon redemption information
- Transaction data over 2 years
- Campaign information

Approach

- 1) Data cleaning including removing missing values and recoding values into group bins
- 2) Transforming data to remove skewness
- 3) Addressing the issue of multi-collinearity
- 4) Clustering analysis based on recency, frequency and monetary factors.
 - a) Decide clustering variables and techniques
 - b) Define number of clusters
 - c) Perform cluster analysis
 - d) Interpret the cluster features

Modelling Variables

- Recency
 - How recently did the customer purchase the products
- Frequency
 - How frequently did the customer purchase the products
- Monetary
 - How much did the customer spend on each visit

Clustering techniques

- K - Means

- Groups similar data points together to discover underlying patterns

- Latent Class Analysis

- Assess the probability a data point belongs to a particular class

Findings

We derived the following segments of the customers from the cluster analysis

- ◉ **Promising customers** – characterized by:
 - High recency, high frequency and low spenders
 - These customers performed transactions between 5 PM and 6 PM.
 - They shopped for grocery often and bought soft drinks most number of times.
 - Most effective campaign is Type B

Findings

- **Champion customers** – characterized by:

- High recency, high frequency and high spenders
- These customers performed transactions between 6 PM and 7 PM. They are mostly working class.
- They shopped for grocery often followed by drugs and fresh produce
- Top 3 products for this group are soft drinks, fluid milk products and baked bread
- Comprise of young customers
- Most effective campaign is Type C

Findings

- **Loyal customers** – characterized by:

- High recency, average frequency and average spenders
- These customers performed transactions between 3 PM and 6 PM. They are mostly non-working class.
- They shopped for grocery often followed by drugs and kiosk gas
- Top 3 products for this group are coupon/miscellaneous items, soft drinks and bag snacks.
- Most effective campaign is Type B

Findings

- **About to lose customers** – characterized by:

- Low recency, high frequency and high spenders
- These customers performed transactions between 5 PM and 6 PM. They are mostly working class.
- They shopped for grocery often followed by drugs and fresh produce
- Top 3 products for this group are soft drinks, fluid milk products and baked bread
- Most effective campaigns are Type A and Type B

Similar trend has been observed for about to sleep customers.

Recommendations

Customer Segment	Recommendations
Champions	Reward this group. Promote new products
Loyal	Upsell high value products. Engage them by asking for their reviews
Promising	Offer trials on products
About to sleep	Offer discount on renewals, share product information
At risk of losing	Engage through personalized messages. Offer discount on renewals.

Future work: Market Basket Analysis can be performed to get deeper insights into the purchasing patterns.

Source: <https://www.putler.com/rfm-analysis/>