Customer segmentation

What is customer segmentation ?

- Market segmentation involves dividing a large homogenous market of potential customers into clearly identifiable segments.
- Customers are divided based on meeting certain criteria or having similar characteristics that lead to them having the same product needs.
- Segments are made up of customers who will respond similarly to marketing strategies. They share common interests, needs, wants and demands.

What types of customer segments can a company have ?

- Demographic segmentation
- Behavioral segmentation
- Geographic segmentation

What are the benefits of customer segmentation ?

- Greater company focus
- Better serve a customer’s needs and wants
- Market competitiveness
- Targeted communication

Why customer segmentation is important ?

- When marketers use market segmentation it makes planning campaigns easier, as it helps to focus the company on certain customer groups instead of targeting the mass market.

- Segmentation helps marketers to be more efficient in terms of time, money and other resources.

- Market segmentation allows companies to learn about their customers. They gain a better understanding of customer's needs and wants and therefore can tailor campaigns to customer segments most likely to purchase products.

RFM segmentation

What is RFM ?

- RFM stands for Recency, Frequency and Monetary
- It is the easiest form of customer segmentation
- Often used for reactivation campaigns, high valued customer programs, combating churn etc.

RFM metrics

Recency
The freshness of customer activity

e.g. time since last activity

Frequency
The frequency of customer transactions

e.g. the total number of recorded transactions

Monetary
The willingness to spend

e.g. the total transaction value

RFM by hand

- Step 1 : Calculate the RFM metrics for each customer
- Step 2 : Find the distribution for each metric by split values into bins (using quantiles)
- Step 3 : Use these metrics to target your Premium, Gold, Silver and standard customers

Segmentation project

Customer segmentation

- You are a datascientist working as a consultant for Olist, a sales solution for online marketplaces.
- **Olist** wants you to develop a **customer segmentation** that its marketing teams can use on a routine basis for their communication campaigns.
- You have an initial meeting with Olist's Marketing Director, Sophia, to better understand the needs of the project. Here are some notes you took:
  - Olist marketing teams need **an actionable description** of the segmentation that allows them to thoroughly grasp its underlying logic (Notebook format).
  - The database contains information such as order history, products purchased, customer satisfaction comments, and location.
  - The quality of the data is somewhat poor. The idea is to create the best possible features for making optimum use of them.
  - Customer segmentation means understand the different types of users. Unsupervised algorithm is probably the best to use.
  - The segmentation needs to be robust: adding a few new customers to the database and creating a new segmentation using the same model should not result in totally different conclusions.

Step of the project

- Join the different datasets into one dataset.
- Understand, analyze and clean the dataset
- Create variables recency, frequency and amount.
- Choose the variables of interest in your dataset and add them to the previously calculated RFM variables to create your segmentation dataset.
- Choose the optimal number of clusters. Use two different segmentation algorithms, compare them and choose the most interesting one.
- Visualize your data by using PCA and coloring the data according to their class.
- Create the personas of the previously calculated classes. Explain your personas in a powerpoint of maximum 5 slides.

Tips

- Don't use all your variables for your segmentation, use only the variables that you think are the most relevant to your marketing team.
- The goal is to create a segmentation that your marketing team can use, so don't use too many clusters.
- Use only recent data to train your segmentation algorithm to capture the most interesting behaviors (be careful not to cut off a trend, a sliding year of history is often a good choice).
- Beware of categorical variables in a segmentation. Most of the algos are based on distances, the distances between modalities are not meaningful, prefer quantitative data.
- You can use additional variables for determining the behavior of your segments than those used for segmentation.

Deliverables

- Powerpoint (5 slides max)