CUSTOMER SEGMENTATION

using Data Science

Table of Contents

INTODUCTION

What is Customer Segmentation in Data Science?

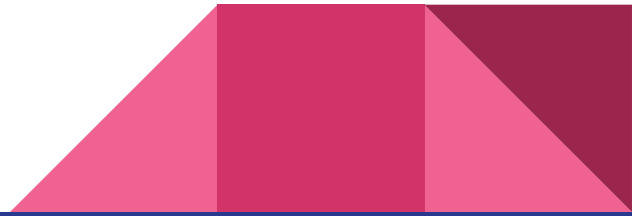
Customer segmentation involves implementing data science methods to divide the customer base into smaller groups based on certain characteristics. It assists marketing managers in better understanding their customers' preferences and presenting them with better-targeted advertisements.

COMMON WAYS TO SEGMENT CUSTOMERS

1. **Demographic information**, such as gender, age, familial and marital status, income, education, and occupation.
2. **Geographical information**, which differs depending on the scope of the company. For localized businesses, this info might pertain to specific towns or counties. For larger companies, it might mean a customer's city, state, or even country of residence.
3. **Psychographics**, such as social class, lifestyle, and personality traits.
4. **Behavioral data**, such as spending and consumption habits, product/service usage, and desired benefits.

STEPS IN CUSTOMER SEGMENTATION

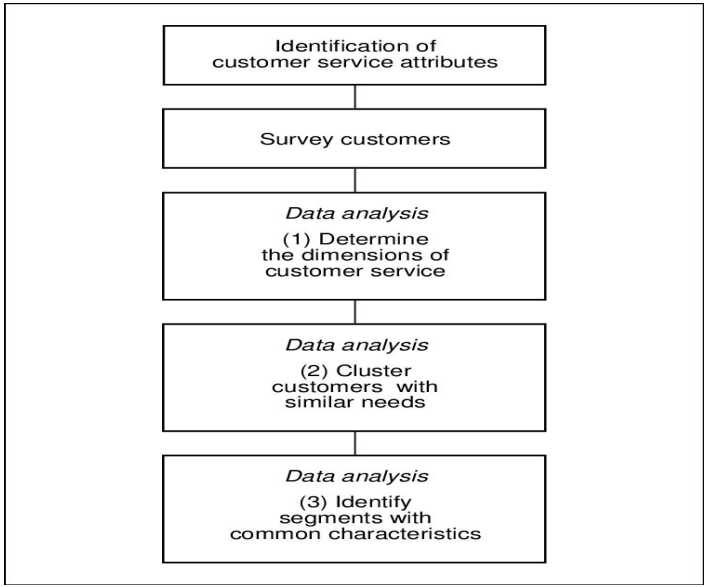
- Step 1: Setting up your customer segmentation project.
- Step 2: Analyzing customer data.
- Step 3: Data collection.
- Step 4: Analysis and prioritization.
- Step 5: Presenting and incorporating feedback.



ENVIRONMENT AND TOOLS

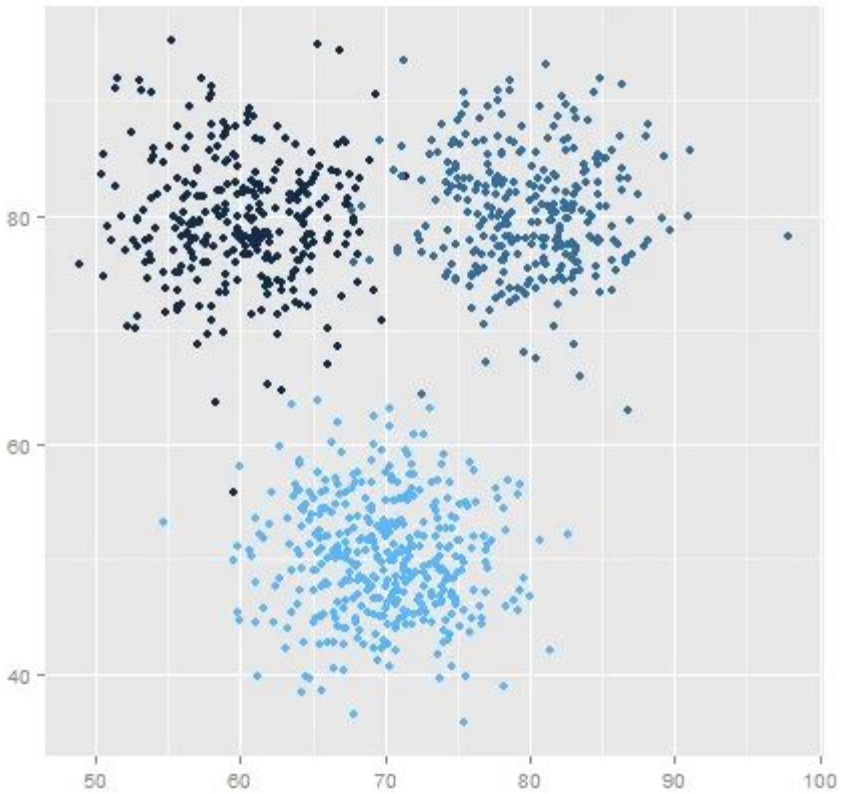
1. scikit-learn
2. seaborn
3. numpy
4. pandas
5. matplotlib

ER DIAGRAM


Identification of customer service attributes
Survey customers
Data analysis
(1) Determine the dimensions of customer service
Data analysis
(2) Cluster customers with similar needs
Data analysis
(3) Identify segments with common characteristics

K MEANS CLUSTERING ALGORITHM

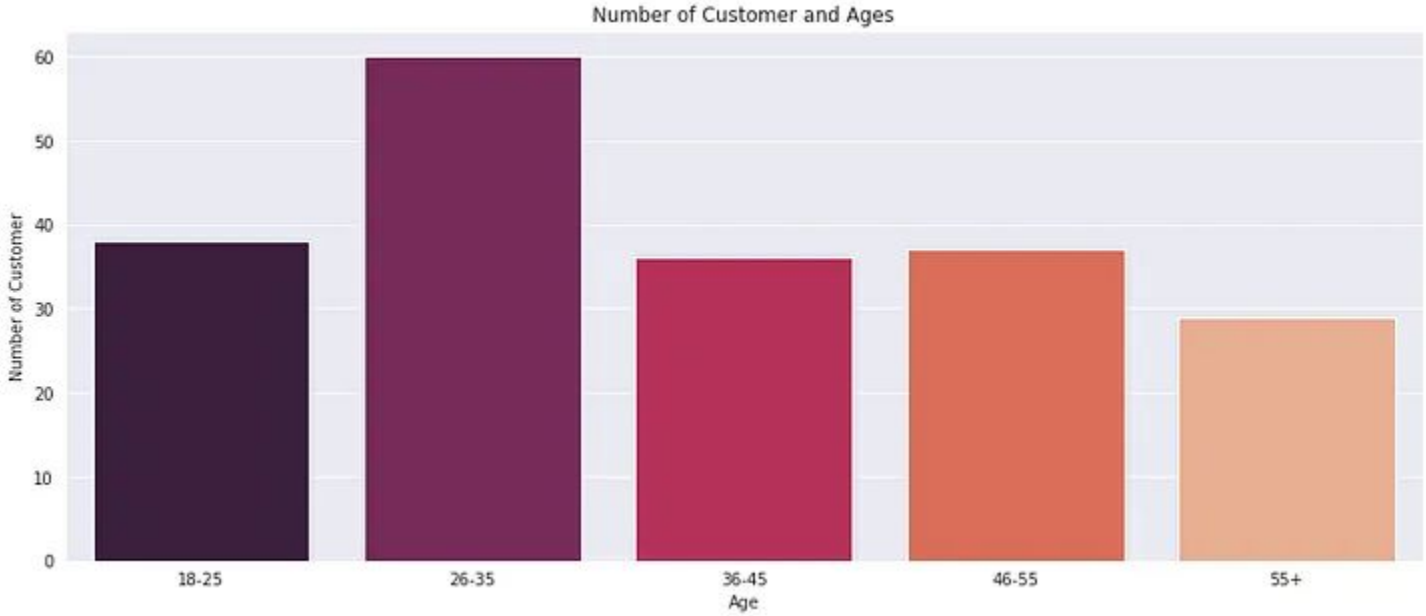
1. Specify number of clusters *K*.
2. Initialize centroids by first shuffling the dataset and then randomly selecting *K* data points for the centroids without replacement.
3. Keep iterating until there is no change to the centroids. i.e assignment of data points to clusters isn't changing.



K Means Clustering where K=3

FOR INSTANCE

A bar plot to check the distribution of number of customers in each age group...


Number of Customer and Ages
Number of Customer
0
10
20
30
40
50
60
18-25
26-35
36-45
46-55
55+
Age

ADVANTAGES

1. Determine appropriate product pricing.
2. Develop customized marketing campaigns.
3. Design an optimal distribution strategy.
4. Choose specific product features for deployment.
5. Prioritize new product development efforts.

CONCLUSION

K means clustering is one of the most popular clustering algorithms and usually the first thing practitioners apply when solving clustering tasks to get an idea of the structure of the dataset. The goal of K means is to group data points into distinct non-overlapping subgroups. One of the major application of K means clustering is segmentation of customers to get a better understanding of them which in turn could be used to increase the revenue of the company.

THANK YOU!