Market Segmentation – A Case Study

-Siddesh Pardeshi

Date: 10th Oct, 2023

Abstract

Market segmentation is a marketing strategy that involves dividing a larger target market into smaller groups with similar characteristics. It helps businesses customize their offerings and marketing efforts to better meet the diverse needs of these segments, leading to increased competitiveness and success.

Market Segmentation:

Market segmentation is a marketing strategy that divides a diverse target market into smaller, more specific groups based on shared characteristics. This approach helps businesses create more effective marketing campaigns by tailoring their efforts to the unique preferences and needs of each segment. It enhances competitiveness, resource allocation, and customer engagement while acknowledging the complexity of today's consumer landscape.

Benefits of Market Segmentation:

- **Better Customer Understanding**: Deeper insights into customer needs.
- **Targeted Marketing**: Customized strategies for higher engagement.
- **Efficient Resource Allocation**: Maximized ROI on marketing spend.
- **Competitive Advantage**: Tailored offerings for an edge.
- **Product Customization**: More satisfied, loyal customers.
- **Market Expansion**: Discovering new growth opportunities.
- **Improved Product Development**: Products that match market needs.
- **Effective Communication**: Messages that resonate with audiences.
- **Brand Positioning:** Building a strong market presence.
- **Measurable Results**: Data-driven decision-making.
- **Flexibility:** Adaptation to changing market conditions.
- **Risk Management**: Diversification for reduced risk.
- **Customer Retention**: Strategies to retain and satisfy customers.

Segmentation Criteria:

Market segmentation involves dividing a target market into smaller, more homogeneous groups based on specific criteria or variables. The choice of segmentation criteria depends on the objectives of the business and the nature of the market. Here are the four basic marketing strategies:

1. **Demographic Segmentation**: This strategy divides the market based on demographic factors, such as age, gender, income, education, marital status, occupation, and family size. Demographic variables are relatively straightforward to measure and often serve as the foundation for many segmentation efforts. For example, a company may target a specific age group for its products or tailor its marketing campaigns to different income brackets.
2. **Psychographic Segmentation**: Psychographic segmentation focuses on customers' lifestyles, interests, values, attitudes, and behaviours. It delves into the psychological aspects of consumer behaviour. By understanding the psychographics of their target market, businesses can create marketing messages that resonate with customers on a deeper level. For instance, a company may target environmentally conscious consumers by aligning its product with eco-friendly values.
3. **Behavioural Segmentation**: Behavioural segmentation divides the market based on how customers interact with products or services. This can include factors like purchasing history, product usage patterns, brand loyalty, and the decision-making process. For example, a business might target loyal customers who frequently purchase its products or create specialized promotions for first-time buyers.
4. **Geographic Segmentation**: Geographic segmentation is based on the location or geography of customers. This strategy may involve segmenting a market by country, region, city, climate, or even urban/rural distinctions. Companies can tailor their offerings and marketing efforts to address the specific needs and characteristics of customers in different geographic areas. For example, a retail chain may offer different products in regions with varying climate conditions.

Data-Driven Market Segmentation: A Holistic Approach

a) **Data Exploration:**

Data exploration for market segmentation involves the initial investigation and analysis of raw data to understand customer characteristics and behaviours. This process includes collecting, cleaning, and visualizing data to identify patterns, trends, and potential segments. Through data profiling, summary statistics, and visualizations, businesses gain insights into their customer data, paving the way for more effective segmentation strategies and targeted marketing efforts.

b) **Data Cleaning:**

Data cleaning is the essential process of identifying and rectifying errors, inconsistencies, and missing values in datasets. This crucial step ensures that data is accurate and reliable for subsequent analysis. Cleaning involves tasks like removing duplicates, correcting formatting issues, and imputing missing data points. By enhancing data quality, businesses can make more informed decisions, reduce the risk of errors, and improve the effectiveness of market segmentation and other data-driven processes.

Descriptive Analysis:

Descriptive analysis is a statistical technique that involves summarizing and examining data to gain insights into its key characteristics. It provides an overview of data, including measures of central tendency (mean, median, mode), variability (range, variance, standard deviation), and the distribution of data through graphical representations like histograms, box plots, and scatter plots. Descriptive analysis helps in understanding data patterns, trends, and the distribution of values, serving as a crucial step in market segmentation by providing a foundation for decision-making and further analysis.

Data Preprocessing:

a) **Categorical Variables**:

Data preprocessing for categorical variables involves converting categorical data (labels or groups) into a numerical format that machine learning algorithms can process. This can be done using methods like one-hot encoding for nominal data, label encoding for ordinal data, or custom ordinal encoding for ordered categories. Preprocessing enables accurate analysis and segmentation of data.

b) **Numeric Variables:**

1. **Scaling:** Scaling numeric variables aims to adjust their range without changing the underlying structure of the data. This is crucial when working with algorithms that are sensitive to the scale of features, such as distance-based methods. Common scaling methods include Min-Max scaling, which transforms data into a specific range (e.g., 0 to 1), and Standardization, which scales data to have a mean of 0 and a standard deviation of 1. Scaling ensures that all numerical variables have similar influence in analysis and model training.

2. **Normalization**: Normalization, on the other hand, is a technique to transform data into a common scale, often between 0 and 1, by adjusting for the minimum and maximum values. This method is particularly useful when the data distribution is not necessarily Gaussian or when variables have varying ranges. Normalization maintains the relationships between data points while making them more amenable to analysis.

Both scaling and normalization are crucial for data preprocessing as they enhance the quality and compatibility of numeric variables in a dataset. They ensure that the numerical features are processed uniformly, making them suitable for various machine learning algorithms used in market segmentation and other data-driven applications.

Principal Components Analysis:

Principal Component Analysis (PCA) is a dimensionality reduction technique used in data preprocessing. It simplifies complex datasets by identifying and retaining the most critical information while eliminating redundant or less significant features. PCA does this by creating new variables, known as principal components, that capture the maximum variance in the data. These components are uncorrelated, making data easier to analyse and visualize. PCA is valuable in market segmentation by reducing the dimensionality of datasets, simplifying analysis, and enhancing the efficiency of algorithms, ultimately leading to more effective customer segmentation.

K- means Clustering Algorithm:

K means clustering is one of the simplest and popular unsupervised machine learning algorithms. A cluster refers to a collection of data points aggregated together because of certain similarities. The algorithm tries to group similar items in the form of clusters. The number of groups is represented by You'll define a target number k, which refers to the number of centroids you need in the dataset. A centroid is the imaginary or reallocation representing the centre of the cluster. Every data point is allocated to each of the clusters through reducing the in-cluster sum of squares. In other words, the K-means algorithm identifies number of centroids, and then allocates every data point to the nearest cluster, while keeping the centroids as small as possible. The 'means' in the K-means refers to averaging of the data i.e., representing the centroid.
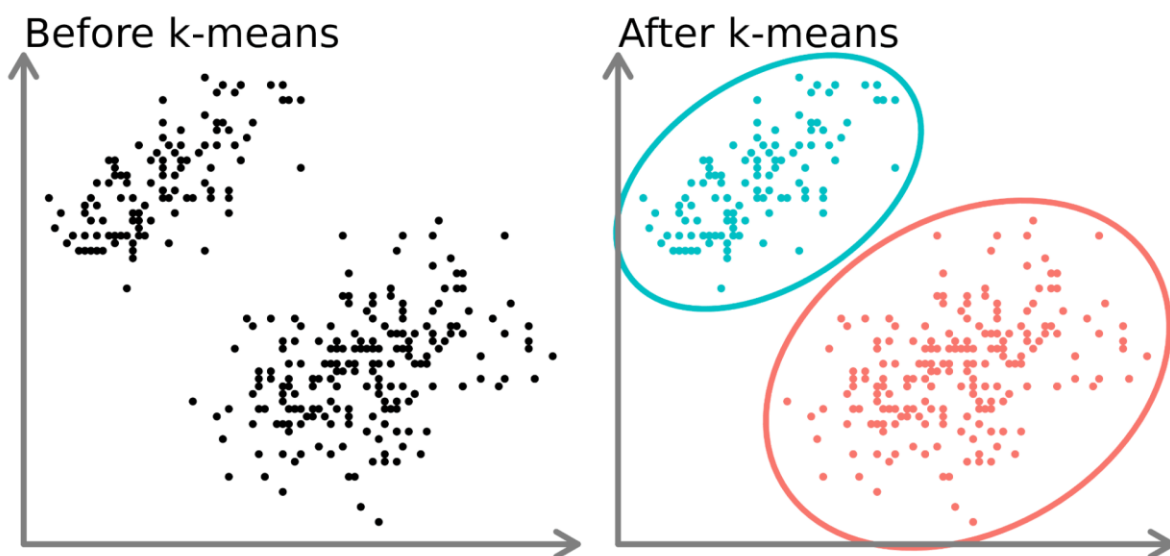
A recipe for k-means:

The recipe for k-means is quite straightforward.

1. Decide how many clusters you want, i.e., choose k.

2. Randomly assign a centroid to each of the k clusters

3. Calculate the distance of all observation to each of the k centroids.

4. Assign observations to the closest centroid.

5. Find the new location of the centroid by taking the mean of all the observations in each cluster.

6. Repeat steps 3-5 until the centroids do not change position.

Elbow Method - Optimal Value:

It is one of the most popular methods to determine the optimal value of K. We use it to choose a K when we observe negligible change in the inertial values between different values of K.

Elbow Method for selection of optimal "K" clusters
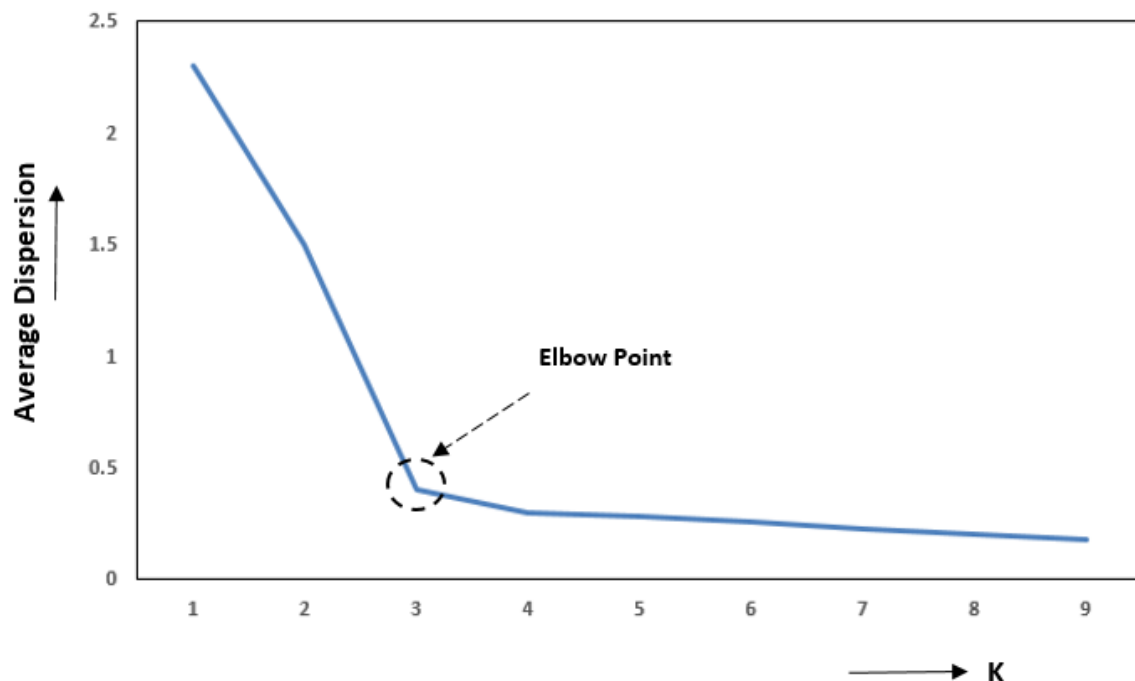
Advantages of K-means clustering:

1. Simplicity and ease of implementation.

2. Fast and efficient for large datasets.

3. Suitable for various data types, including numerical and categorical.

4. Clusters are interpretable and easy to visualize.

5. Usually converges to a solution with random initializations.

6. Can handle clusters of different shapes.

7. Valuable for preprocessing by reducing dimensionality.

8. Widely applied in customer segmentation, image analysis, and more.

Market Segmentation case study on McDonalds Dataset:

GitHub link : https://github.com/SiddeshPardeshi/Feynn-Labs/blob/main/McDonalds_clustering.ipynb

Dataset Link : https://homepage.boku.ac.at/leisch/MSA/datasets/mcdonalds.csv

Conclusion:

In conclusion, market segmentation is a fundamental strategy in modern business that involves dividing a target market into distinct and meaningful segments. This approach allows companies to better understand their customers, create tailored marketing strategies, and allocate resources more efficiently. By focusing on the unique needs and preferences of each segment, businesses can enhance customer satisfaction, improve product development, and gain a competitive edge. Market segmentation is a dynamic and data-driven process that requires ongoing analysis and adaptation to evolving market conditions and consumer behaviours. When executed effectively, market segmentation can be a key driver of success, enabling companies to connect with their audience in a more meaningful and impactful way.

References:

https://www.iteratorshq.com/blog/data-cleaning-in-5-easy-steps/

https://www.datacamp.com/tutorial/k-means-clustering-python

https://domino.ai/blog/getting-started-with-k-means-clustering-in-python

https://www.oreilly.com/library/view/statistics-for-machine/9781788295758/c71ea970-0f3c-4973-8d3a-b09a7a6553c1.xhtml