

STUDY TASK ON MARKET SEGMENTATION

-Simran Kaur

Introduction to Market Segmentation:

Market segmentation is a marketing term that refers to aggregating prospective buyers into groups or segments with common needs and who responds similarly to a marketing action. Market segmentation enables companies to target different categories of consumers who perceive the full value of certain products and services differently from one another. Or in other words Market segmentation is a process that consists of sectioning the target market into smaller groups that share similar characteristics, such as age, income, personality traits, behaviour, interests, needs or location.

Companies can generally use three criteria to identify different market segments:

1. Homogeneity or common needs within a segment.
2. Distinction or being unique from other groups.
3. Reaction or a similar response to the market.

Steps to Implement Market Segmentation:

Step 1: Deciding (not) to Segment

At this step of market segmentation, you should focus on discovering how big the market is, where your brand fits and if your products have the capacity to solve what it promises. Market segmentation is a popular marketing strategy that allows businesses to identify and target specific customer groups with customized marketing tactics. However, not all businesses may find segmentation necessary or appropriate for their needs. Here are some reasons why a business may choose not to segment its market:

1. The product or service has a broad appeal: Some products or services may appeal to a wide target audience with similar needs and preferences. For example, basic household products like soap or toothpaste may have a broad appeal and may not require segmentation.
2. The cost of segmentation outweighs the benefits: Conducting market research and segmentation can be time-consuming and expensive. If the cost of segmentation exceeds the potential benefits, a business may choose not to segment its market.
3. The market is too small or niche: In some cases, the market may be too small or niche to justify segmentation. For instance, if a business operates in a small town with a limited customer base, it may not make sense to segment the market further.
4. The business has limited resources: Smaller businesses may not have the resources to conduct market research and segmentation. In such cases, it may be more cost-effective to use a broad marketing approach that appeals to a wide range of customers.

While segmentation can be a valuable strategy in many cases, businesses should carefully evaluate their goals and resources before deciding to segment or not.

Step 2: Specifying the Ideal Target Segment

Step 2 of market segmentation involves developing two sets of segment evaluation criteria: knock-out criteria and attractiveness criteria.

Knock-out criteria:

Knock-out criteria are the essential, non-negotiable features that a potential target segment must meet in order for the organization to consider targeting them. These criteria can include factors such as the size and growth potential of the segment, the organization's ability to effectively reach and communicate with the segment, and the profitability and potential return on investment of targeting the segment.

Attractiveness criteria:

After identifying the knock-out criteria, the organization can move on to the second set of evaluation criteria, known as attractiveness criteria. Attractiveness criteria are used to evaluate the relative attractiveness of the remaining market segments that comply with the knock-out criteria.

Developing these two sets of segment evaluation criteria enables the organization to effectively analyse and evaluate potential target segments, and ultimately select one or more segments to target in the following steps of market segmentation. It also guides many of the following steps, most critically Step 3 (data collection) and Step 8 (selecting one or more target segments).

Step 3: Collecting Data

Empirical data forms the basis of both commonsense and data-driven market segmentation. Empirical data is used to identify or create market segments and – later in the process – describe these segments in detail. In commonsense segmentation, the segmentation variable is typically one single characteristic of the consumers in the sample, such as gender. Other personal characteristics are used as descriptor variables to describe the segments in detail. Data-driven market segmentation is based on multiple segmentation variables, which serve as the starting point for identifying naturally existing or artificially created market segments. Good empirical data is critical for developing a valid segmentation solution, and data quality determines the quality of the extracted market segments and the descriptions of the resulting segments. Empirical data for segmentation studies can come from a range of sources, and the source that delivers data most closely reflecting actual consumer behaviour is preferable.

Segmentation Criteria:

Before extracting segments or collecting data, an organization must choose a segmentation criterion, which can be a specific construct or nature of information used for market segmentation. The most common criteria are geographic, socio-demographic, psychographic, and behavioural. It is recommended to use the simplest possible approach that works for the product or service at the least possible cost.

1. Geographic Segmentation

Geographic segmentation is one of the most commonly used market segmentation criteria. This approach involves dividing the market based on the consumer's location of residence, which serves as the only criterion to form market segments. The key advantage of geographic segmentation is that each consumer can easily be assigned to a geographic unit, making it easy to target communication messages and select communication channels. However, the key disadvantage is that living in the same country or area does not necessarily mean that people share other characteristics relevant to marketers, such as benefits they seek when purchasing a product.

2. Socio-Demographic Segmentation

Socio-demographic segmentation criteria include age, gender, income, and education. Socio-demographic segments can be very useful in some industries, such as luxury goods, cosmetics, and retirement villages. While the socio-demographic criterion may offer an explanation for specific product preferences in some instances, it is not always the cause for product preferences, thus not providing sufficient market insight for optimal segmentation decisions.

3. Psychographic Segmentation

Psychographic segmentation involves grouping people based on psychological criteria, such as their beliefs, interests, preferences, aspirations, or benefits sought when purchasing a product. Benefit segmentation, which is based on people's activities, opinions, and interests, is one of the most popular kinds of psychographic segmentation. The psychographic approach has the advantage that it is generally more reflective of the underlying reasons for differences in consumer behaviour. However, psychographic criteria are more complex than geographic or socio-demographic criteria because it is difficult to find a single characteristic of a person that will provide insight into the psychographic dimension of interest.

4. Behavioural Segmentation

Behavioural segmentation divides the market based on consumer behaviour, including their product usage, brand loyalty, and purchase decision-making. This approach is based on the premise that the best predictor of future behaviour is past behaviour. Behavioural segmentation can be useful in developing marketing strategies for consumers who have already demonstrated a preference for a specific product or brand. It can also be helpful in identifying consumers who are likely to make a purchase in the near future.

Overall, marketers can use one or more of these market segmentation criteria to develop a comprehensive understanding of their target market and create effective marketing strategies.

Step 5: Extracting Segments

Data-driven market segmentation analysis is exploratory in nature, as consumer data sets are typically unstructured and consumer preferences are spread across the entire plot, making it difficult to identify clear groups of consumers. The results of market segmentation analysis depend on the extraction algorithm chosen and the underlying data. Therefore, it is important to explore market segmentation solutions derived from different clustering methods and understand how different algorithms impose structure on the extracted segments. There is no single best algorithm for all data sets, and investigating and comparing alternative segmentation solutions is critical to arriving at a good final solution. Data characteristics and expected or desired segment characteristics allow a pre-selection of suitable algorithms to be included in the comparison.

5.1 Distance-Based Methods

5.1.1 Distance Measures:

Distance measures are an essential part of cluster analysis and market segmentation. A data matrix is used to represent the data, where each row corresponds to an observation and each column represents a variable. Mathematically, this can be represented as an $n \times p$ matrix, where n is the number of observations and p is the number of variables. The vector corresponding to the i^{th} row of matrix X is denoted as x_i , where X is the set of all observations.

There are numerous ways to measure the distance between two vectors. A distance is a function $d(\cdot, \cdot)$ with two arguments: the two vectors between which the distance is being calculated. A distance measure must be symmetric, which means that $d(x, y) = d(y, x)$. It must also fulfill the criterion that the distance of a vector to itself is 0, and only to itself. Additionally, most distance measures fulfill the triangle inequality, which says that if one goes from x to z with an intermediate stop in y , the combined distance is at least as long as going from x to z directly.

The most common distance measures used in market segmentation analysis are the Euclidean distance, Manhattan distance, and asymmetric binary distance.

5.1.2 Hierarchical Methods

Hierarchical clustering methods are the most intuitive way of grouping data because they mimic how a human would approach the task of dividing a set of n observations

(consumers) into k groups (segments). Hierarchical methods are further classified as agglomerative and divisive.

Divisive hierarchical clustering methods start with the complete data set X and splits it into two market segments in a first step. Then, each of the segments is again split into two segments. This process continues until each consumer has their own market segment. Agglomerative hierarchical clustering approaches the task from the other end. The starting point is each consumer representing their own market segment (n singleton clusters). Step-by-step, the two market segments closest to one another are merged until the complete data set forms one large market segment

5.1.3 Partitioning Methods

5.1.3.1 k-Means and k-Centroid Clustering:

The k-means clustering method is the most popular and involves dividing consumers into subsets such that those in the same subset are as similar to one another as possible, while those in different subsets are as dissimilar as possible. The representative of a market segment is referred to as the centroid. The k-means algorithm finds centroids that consist of column-wise mean values across all members of the market segment, which represents the average response pattern across all segmentation variables for all members of the segment. The algorithm is iterative and involves specifying the desired number of segments, randomly selecting k observations from the data set as an initial set of cluster centroids, assigning each observation to the closest cluster centroid to form a partition of the data, recomputing the cluster centroids by holding cluster membership fixed, and minimising the distance from each consumer to the corresponding cluster centroid, and repeating until convergence or a pre-specified maximum number of iterations is reached.

5.2 Model-Based Methods:

Distance-based methods use similarities or distances to group consumers with similar characteristics into market segments, whereas model-based methods use the assumption that each market segment has a certain size, and if a consumer belongs to market segment A , they will have characteristics specific to members of market segment A . The goal of model-based methods is to find the values for segment sizes and segment-specific characteristics that best reflect the data. Model-based methods use a finite mixture model, where the number of market segments is finite, and the overall model is a mixture of segment-specific models. The segment membership of a consumer is determined by the multinomial distribution with segment sizes, and each segment has segment-specific characteristics captured by the vector θ , containing one value for each segment-specific characteristic. The goal is to estimate the parameters of the finite mixture model to determine the segment sizes and segment-specific characteristics, and then assign consumers in the empirical data set to segments based on the probability of each

consumer being a member of each segment. Model-based methods offer an alternative extraction technique to distance-based methods and should be used in conjunction with other extraction methods to determine the most suitable approach for the data at hand.

5.3 Algorithms with Integrated Variable Selection

Some segmentation algorithms assume that all segmentation variables contribute equally to the solution, but this may not always be the case due to redundancy or noise. Pre-processing methods can identify and remove these variables, but it may not be possible with binary data. In such cases, algorithms with integrated variable selection can be used to simultaneously extract segments and identify suitable segmentation variables. Examples include biclustering, the variable selection procedure for clustering binary data (VSBD), and factor-cluster analysis.

5.3.1 Biclustering Algorithms:

Biclustering simultaneously clusters both consumers and variables by extracting market segments containing consumers who all have a value of 1 for a group of variables. Biclustering algorithms exist for any kind of data, including metric and binary, but they have gained significant popularity due to their ability to handle large genetic and proteomic datasets. The biclustering algorithm starts by rearranging rows and columns of the data matrix to create a rectangle with identical entries of 1s. Next, it assigns the observations falling into this rectangle to one bicluster, and removes the rows containing the assigned consumers to repeat the procedure until no more biclusters of sufficient size can be located. Different algorithms search for different patterns in biclusters.

5.3.2 Variable Selection Procedure for Clustering Binary Data (VSBD):

Brusco proposed a variable selection procedure for clustering binary data sets. His VSBD method is based on the k-means algorithm as clustering method, and assumes that not all variables available are relevant to obtain a good clustering solution. In particular, the method assumes the presence of masking variables. They need to be identified and removed from the set of segmentation variables. Removing irrelevant variables helps to identify the correct segment structure, and eases interpretation

5.4 Data Structure Analysis:

The process of extracting market segments is exploratory, so traditional validation methods are not possible. Instead, validation typically involves assessing the reliability or stability of solutions across repeated calculations. This approach is referred to as stability-based data structure analysis, which provides valuable insights into the properties of the data and helps guide subsequent methodological decisions. If natural,

distinct, and well-separated market segments exist in the data, they can be easily revealed. If not, analysts need to explore a large number of alternative solutions to identify the most useful segments for the organization. Data structure analysis can also help to choose a suitable number of segments to extract if there is structure in the data.

Four different approaches to data structure analysis are: cluster indices, gorge plots, global stability analysis, and segment level stability analysis

Case study on McDonalds's:

[FeyNN-labs-internship/Task2.ipynb at main · SimranKachle/FeyNN-labs-internship \(github.com\)](#)