**What is market segmentation?**

Market segmentation is the practise of breaking down a large target market into more narrowly defined subgroups or segments based on shared traits or requirements. The goal of market segmentation is to recognize and better serve the unique tastes, tendencies, and needs of several client groups within a broader market.

**What is market segmentation Analysis?**

The process of examining and assessing the various market segments within a bigger market goal is known as market segmentation. In order to understand each segment's distinctive qualities, actions, hobbies, and calls information and understanding about each segment must be gathered. By using market segmentation analysis, businesses can make sound choices and create marketing plans that are suited to the distinctive wants of each category.

**Steps Involved in market segmentation analysis are :**

**STEP 1 : Deciding (not) to Segment**.

Implications of Committing to Market Segmentation The main repercussion is that a business must commit over the long term to the segmentation plan. It may be necessary to revise all market communications as well as the development of new products, modifications to the designs of already-existing ones, price reductions, changes to the product's distribution channels, and price reductions. These adjustments will therefore probably have an effect on the internal structure of the

company, which could need to change as a result of, say, narrowing the focus to a few different market sectors.

Implementation Barriers The first set of barriers concerns high-ranking employees. Senior leadership's absence from the market segmentation process compromises its Performance by lacking leadership, proactive championing, commitment, and involvement. Another set of obstacles has to do with organisational culture. The following factors have been identified as impeding the successful implementation of the market-oriented approach: a lack of market or consumer orientation, resistance to change and novel ideas, a lack of creative thinking, poor communication and a lack of information and insight sharing across organisational units, short-term thinking, an unwillingness to make changes, and office politics.

**STEP 2 : Specifying the Ideal Target Segment.**

This layer of market segmentation analysis is mainly focused on user input and Participation throughout the process. It highlights the importance of including user input at various phases of the research rather than just at the beginning or end. The organization has to actively engage in the study, specifically in the second phase where segment evaluation criteria are defined.

The initial set of criteria, called knock-out criteria, is non-negotiable characteristics that segments must have for the organization to look into targeting them. These characteristics are important criteria for segment selection. The attractiveness criteria analyze the corresponding appeal of the left portions that satisfy the knock-out criteria. While the literature frequently needs to differentiate between these types of criteria, it does provide a wide range of potential segment assessment criteria at various degrees of detail. These criteria are critical in directing later processes in the analysis, notably data collection and target segment selection. Overall, the user participation in creating and implementing segment assessment criteria, as well as the organization’s input, are crucial to ensuring that market segmentation analysis delivers appropriate knowledge for the organization’s marketing strategy.

**Knock-Out Criteria**

Certain criteria, such as the minimum feasible target segment size, are clear and do not require additional definition, but some, like the knock-out criteria, do. These knock-out criteria must be well understood by senior management, the segmentation team, and the advisory committee. knock-out criteria are used as an initial screening method to decide which market segments should be evaluated further. These criteria guarantee that the segments chosen are significant, quantifiable, accessible, homogenous, distinct, large enough, aligned with organizational strengths, recognizable, and attainable. Clear knowledge and characterization of these criteria are required for the market segmentation process to be successful.

**Attractiveness Criteria**

Market segmentation is the process of examining several factors to evaluate the attractiveness of distinct market segments. These criteria go beyond basic binary ratings and entail measuring the attractiveness of each part in accordance to specified factors. This signifies that segments are neither compliant nor noncompliant with these criteria. Instead, each market sector is given a grade depending on how well it matches a certain criterion. In the last phase of market segmentation analysis, the ratings across all criteria are used to assess if a market segment should be chosen as a

target segment. In summary, selecting target segments entails assessing a wide variety of attractiveness factors and ranking the attractiveness of each segment based on these criteria.

**Step 3: Collecting Data**

The chapter talks about how empirical data can be used to segment markets both intuitively and using data. Market segments are defined and created using empirical data, which also provides in-depth descriptions of these segments.

In a segmentation that makes sense, one consumer trait is employed as the dividing variable. For instance, using gender as the segmentation variable, the sample may be split into segments of men and women. Descriptor variables refer to the additional personal information that is contained in the data, such as age, the quantity of vacations taken, and details on desired benefits while on vacation. To fully describe the segments, these descriptor variables are used. Descriptor variables play a crucial part in gaining this insight, which is necessary for developing effective marketing strategies that are well-aware of the segments. Socio- demographic statistics and information on media usage are examples of descriptor variables that can be used by marketers to interact with their target markets more successfully.

In conclusion, the traditional approaches and data-driven methods to market segmentation employ empirical data as their basic building block. Segmentation factors divide the sample into market segments, and descriptor variables offer more information to the segment characteristics to help marketers create more precise marketing plans.

Rather than using the word segmentation variable, the phrase segmentation criterion is used in this context. The phrase segmentation variable refers to a single measured value, such as a single response in a survey or a single observed spending category. The definition of a segmentation criterion pertains to the type of data that is gathered to segment markets. It may also be related to a single idea, like the advantages desired. Because it involves prior knowledge, choosing which segmentation criterion to utilise cannot be simply delegated to a consultant or data analyst. regarding the market. Geographical, socio-demographic, psychographic, and behavioural segmentation factors are the most popular.

Geographic segmentation is frequently the most suitable method because it is straightforward. For instance, the national tourism organisation of Austria needs to speak Italian, German, Slovenian, Hungarian, and Czech in order to draw visitors from nearby nations. A highly practical justification for dividing tourists from several adjacent countries as separate segments is language disparities across nations. International businesses like Amazon, which sells its Kindle online, offer interesting instances as well. Customers are prompted to select their place of residency before receiving more information that is specific to that nation on a single common web page. IKEA provides a similar selection of goods everywhere it operates, however there are minor variations in offers, prices, and the availability of online shopping. Age, gender, income, and education are among the sociode-mographic segmentation factors education. In some businesses, socio-demographic groups can be quite helpful.

For instance, luxury products are often linked with wealth, while cosmetics are often even when men are the focus, both the female and male portions are handled very differently), gender-specific infant goods, and retirement communities (linked to age), tourism resort goods (linked to having a small whether or not they have children. Geographic or socio-demographic criteria are simpler than psychographic criteria by nature because it is difficult to identify just one aspect of a person. a person who may bring light to the relevant psychographic dimension. As a result, the majority of psychographic segmentation studies make use of many segmentation variables, such as various travel motivations and perceived hazards.

The main benefit of behavioural methods is that segment extraction can be based on the particular behaviour of interest if it is based on actual behaviour rather than declared behaviour or stated anticipated behaviour. Therefore, behavioural segmentation groups individuals based on the similarity that matters the most. A few key aspects that need to be considered when using survey data are :

1. Choice of variable: Considering carefully whether factors to add as segmentation variables for common sense segmentation or for data-driven segmentation, to the effectiveness of the market segmentation solution.

2. Response Options: The scope of the data accessible for subsequent studies depends on the survey response options given to respondents. Not all survey response alternatives are equally acceptable for segmentation analysis because many data analytic approaches are based on distance measures.

3. Response Styles: Responses to surveys exhibit a wide range of response patterns, including respondents' propensity to select extreme response options to choose the neutral option to check every box. Response types have an impact on segmentation outcomes since they frequently used segment extraction techniques are unable to distinguish between data entries expressing the respondent's belief and entries reflecting both the respondent's belief and a manner of replying.

4. Sample size: However, it is relatively simple to identify the number and kind of segments in the data set if the sample size is acceptable. Data from Internal Sources Organisations now have greater access to large volumes of insider information that can be gathered for market segmentation study. The data from grocery store scanners, reservations made through airline loyalty programmes, and online purchases are typical instances. Such data are more reliable because they reflect real customer behaviour rather than consumer assertions about their intentions or behaviour.

Data from Experimental Studies Field or lab experiments may generate experimental data. They may be the outcome of studies on how consumers react to particular adverts, for instance based on the reaction to the advertisement, a segmentation standard.

**Step 5: Extracting Segments**

Consumer data sets are often unstructured and heterogeneous, the process of data-driven market segmentation study is largely exploratory in nature. Preferences tend to be dispersed over the whole plot, as opposed to a two-dimensional plot of customer preferences, which frequently lacks identifiable consumer groupings. Because of the mix of exploratory methodologies and unstructured data, segmentation analysis findings strongly rely on assumptions about the underlying structure of the segments.

The extraction technique and segmentation method used have a substantial impact on the segmentation solution. Many market segmentation approaches are developed from cluster analysis, in which market segments correspond to clusters. Choosing a suitable clustering approach demands matching the data analysis features of the resulting clustering with the researcher's individual requirements.

To have a thorough grasp of market segmentation solutions, it is necessary to investigate the findings of various clustering approaches. This investigation allows researchers to assess how different algorithms impose structure on retrieved segments, assisting in interpretation and decision-making. To summarise, in order to get relevant and actionable insights from a data-driven market segmentation research, both the underlying data and the extraction technique must be carefully considered.

**Distance Measures**

A distance measure has to comply with a few criteria.

One criterion is symmetry, that is:

d(x,y) = d(y,x).

A second criterion is that the distance of a vector to itself and only to itself is 0:

d(x, y) =0 ⇔ x=y.

Euclidean distance is a popular distance measure in market segmentation studies. It denotes the "straight-line" distance between two places in two dimensions. This distance metric takes into account all dimensions of the vectors x and y, which indicate customer attributes or preferences. Manhattan distance, on the other hand, is called after the street grid layout in Manhattan and estimates the distance between two places by taking into account the grid-like pathways that would be used to go between them. It also considers all dimensions of the vectors x and y. In market segmentation analysis, both Euclidean and Manhattan distances are routinely used. The Euclidean distance is the shortest distance between two places in a straight line, whereas Manhattan distance takes into consideration grid-like routes.

These distance measurements are critical for identifying the similarity or dissimilarity of customer preferences or attributes, which aids in market segment identification.

**Hierarchical Methods**

In market segmentation analysis, hierarchical clustering algorithms are frequently regarded as the most natural way to data categorization. These algorithms simulate how a person might split a collection of n observations (consumers) into k groups (segments). Market segmentation analysis, on the other hand, exists between the extremes of specific customer categories and a single, homogenous market.

Divisive hierarchical clustering starts with the entire data set X and divide it into two market groups. This procedure is repeated for each section, with sub-segments added until each consumer gets their own segment. Agglomerative hierarchical clustering, on the other hand, employs the opposite method. It begins with each customer representing their own market segment (n singleton clusters) and gradually combines the two closest segments at each stage until the final segment is formed.

Both algorithms provide a series of nested partitions, with each partition representing a grouping of observations, and each observation belonging to precisely one group. This partitioning sequence might have as little as one group (segment) or as many as n groups (segments). Because the partition with k+1 groups (segments) is produced from the partition with k groups by separating one of the existing groups, the partitions are considered nested.

Overall, hierarchical clustering approaches offer an organised and step-by-step approach to market segmentation study, allowing for the development of nested groups that increasingly separate customers based on their similarities and differences.

**Partitioning Methods**

When dealing with larger data sets in market segmentation analysis, dendrograms become difficult to interpret, and the matrix of pairwise distances may exceed computer memory limitations. In such cases, clustering methods that generate a single partition are more appropriate than hierarchical clustering with nested sequences.

To overcome memory constraints, instead of calculating distances between all pairs of observations at the start of a hierarchical partitioning cluster analysis, a different approach is adopted. Only the distances between each consumer in the data set and the center of the segments are computed. This selective computation reduces the computational burden and allows for more efficient analysis, particularly for data sets containing more than 1000 observations (consumers).

**Clustering with k-Means**

k-Means Clustering is a type of unsupervised machine learning method that divides data into discrete groups based on similarities. The algorithm's core premise is straightforward to grasp and may serve as an excellent introduction to Market Segmentation utilising Clustering techniques.

How does k-Means Clustering function?

1. We specify the hyper-parameter k, which is the number of clusters into which we want our data to be grouped. 2. Next, k centroids, or cluster-means, are chosen at random. 3. Finally, the best centroid positions are identified. The following algorithmic loop does this: a. Assignment procedure: Assign each data point to the nearest centroid (determined as the data point's squared distance from the centroid) b.

As we can realize from our understanding of the algorithm that we need to declare the number of clusters, we want, beforehand for the algorithm to work. But usually in many business cases we do not know beforehand that in how many clusters we should divide our data.

So, a method called Elbow method is used to decide optimum number of clusters. In this method we plot inertia v number of clusters and at which number of clusters the plot’s slope change drastically is taken as optimum number of clusters as shown in Figure 2. Here, the measure of inertia is the sum of squared distance between each point and its cluster center.

Github link - https://github.com/abhishektiwari-ln/Feynn-Labs-Project-2.git