Case Study: Fundamentals of Market Segmentation

Team Members:

Aiswarya Sudhir – Step 1,2,3

Puja Rajput – Step 5

Vedant Aryan – Step 6,7

Gajanan Santosh Purud – Step 8,9

Jada Sunil – Step 4

Step 1: Deciding (not) to Segment

1.1 Implications of Committing to Market Segmentation

The main point emphasized is the necessity for organizations to fully commit to a segmentation strategy over the long haul. Market segmentation is likened to a committed relationship, rather than a casual encounter. This commitment entails being ready and willing to enact significant changes and investments within the organization. These changes may involve developing new products, adjusting existing ones, revising pricing and distribution channels, and refining all market communications. Consequently, such alterations are likely to affect the internal structure of the organization. Due to the substantial impact of this long-term commitment, the decision to explore the potential of a market segmentation strategy must be made at the highest executive level. Moreover, it should be consistently communicated and reinforced across all levels and units within the organization.

1.2 Implementation Barriers

The initial set of obstacles primarily concerns senior management. A lack of leadership, proactive involvement, and commitment from top-level executives undermines the success of market segmentation efforts. Additionally, insufficient allocation of resources by senior management, both for the initial analysis and the long-term implementation of segmentation strategies, can hinder success.

A second set of barriers pertains to organizational culture. Factors such as a lack of focus on markets or consumers, resistance to change, poor communication, short-term thinking, and office politics can impede effective segmentation implementation. Insufficient training and the absence of a formal marketing function or expertise within the organization can exacerbate these challenges. Objective limitations, such as financial constraints or structural constraints, may also pose barriers.

To address resource limitations, organizations should prioritize opportunities judiciously. Simplifying market segmentation analysis and presenting results graphically can aid comprehension by managers.

Many of these barriers can be identified early in the segmentation process and addressed proactively. However, if obstacles persist, reconsideration of pursuing segmentation as a viable strategy may be necessary. For those proceeding, a determined commitment is essential, tempered with patience and an acknowledgment of the challenges inherent in implementing segmentation conclusions.

Step 2: Specifying the Ideal Target Segment

2.1 Segment Evaluation Criteria

The third stage of market segmentation analysis heavily involves user input, emphasizing the necessity for user engagement throughout the process. In this phase, two sets of segment evaluation criteria are established: knock-out criteria and attractiveness criteria. Knock-out criteria are essential features that segments must possess, with no negotiation on their importance. Attractiveness criteria form a more extensive checklist for the segmentation team, allowing them to assess the relative appeal of potential target segments. Team members must select which attractiveness criteria to use and determine their importance to the organization.

2.2 Knock-Out Criteria

Knock-out criteria are used to determine if market segments resulting from the market segmentation analysis qualify to be assessed using segment attractiveness criteria. The criteria's are:
- **Homogeneity:** Segments must consist of similar members.
- **Distinctiveness:** Segments must be different from each other.
- **Size:** Segments should be large enough to justify customized marketing efforts.
- **Alignment with Organizational Strengths:** Segments should match the organization's capabilities to meet their needs.
- **Identifiability:** Members of segments should be identifiable in the marketplace.
- **Reachability:** There should be a way to connect with segment members to deliver customized marketing.

It's crucial for senior management, the segmentation team, and the advisory committee to understand these criteria.

2.3 Attractiveness Criteria

Attractiveness criteria in market segmentation analysis are not binary; segments are not simply classified as either complying or not complying with these criteria. Instead, each market segment is rated on a scale, indicating its level of attractiveness with respect to specific criteria. The collective attractiveness across all criteria determines whether a market segment is chosen as a target segment in the segmentation analysis process.

2.4 Implementing a Structured Process

The structured approach for evaluating market segments, commonly used to select target markets, involves a segment evaluation plot. This plot displays segment attractiveness on one axis and organizational competitiveness on the other. The values for attractiveness and competitiveness are determined by the segmentation team, as there is no universal set of criteria applicable to all organizations. The factors contributing to both attractiveness and competitiveness need to be negotiated and agreed upon, typically through exploring numerous criteria before consensus is reached. Ideally, a team effort, possibly led by a core group of two to three individuals, proposes an initial solution to be discussed and possibly modified by an advisory committee representing various organizational units.

Including representatives from diverse organizational units in the process is crucial for two reasons: first, different units bring varied perspectives, enriching the deliberations, and second, since segmentation strategy impacts all units, they are all stakeholders in the analysis. Selecting attractiveness criteria early ensures that data collection captures pertinent information. By the end of this step, the segmentation team should have around six attractiveness criteria, each assigned a weight indicating its importance relative to others.

2.5 Checklist for the evaluation process

1. Convene Segmentation Team Meeting:
   - Schedule a meeting with the segmentation team to initiate the evaluation process.
2. Discuss and Agree on Knock-out Criteria:
   - Facilitate a discussion among team members regarding homogeneity, distinctness, size, match, identifiability, and reachability criteria.
   - Ensure agreement on these criteria, which will lead to the automatic elimination of non-compliant market segments.
3. Present Knock-out Criteria to Advisory Committee:
   - Share the agreed knock-out criteria with the advisory committee.
   - Allow for discussion and potential adjustments based on committee feedback.
4. Individually Study Market Segment Attractiveness Criteria:
   - Review available criteria for assessing market segment attractiveness.
5. Discuss and Select Subset of Criteria:
   - Engage in discussions with team members to narrow down the criteria to no more than six.
   - Aim for consensus on the subset of criteria that best align with organizational goals.
6. Allocate Points Across Attractiveness Criteria:
   - Individually distribute 100 points across the agreed-upon attractiveness criteria.
   - Ensure points are allocated based on the relative importance of each criterion.
7. Discuss Weightings and Reach Consensus:
   - Collaborate with team members to determine weightings for each attractiveness criterion.
   - Aim for agreement on weightings that reflect the importance of each criterion.
8. Present Selected Criteria and Weightings to Advisory Committee:
   - Share the finalized segment attractiveness criteria and proposed weightings with the advisory committee.
   - Allow for discussion and potential adjustments based on committee input.

By following this checklist, you can systematically progress through the evaluation process, ensuring alignment and agreement among team members and the advisory committee.

Step 3: Collecting Data

3.1 Segmentation Variables

Empirical data in market segmentation is important, distinguishing between commonsense segmentation and data-driven segmentation. In commonsense segmentation, a single characteristic is typically used to divide consumers into segments, with other personal

characteristics serving as descriptors. These descriptors help in detailing the segments, aiding in the development of effective marketing strategies. On the other hand, data-driven segmentation relies on multiple variables to identify or create market segments, enhancing the segmentation process.

Regardless of the segmentation approach, data quality is crucial for accurately assigning individuals to segments and describing them effectively. This accurate description enables customization of products, pricing strategies, distribution channels, and communication methods. The text emphasizes that good empirical data is essential for successful market segmentation analysis.

Empirical data for segmentation studies can be sourced from surveys, observations such as scanner data, or experimental studies. While surveys are common, they may not always accurately reflect consumer behavior, especially if the behavior is socially desirable. Thus, it's important to explore various data sources to ensure alignment with actual consumer behavior.

3.2 Segmentation Criteria

The organization faces a crucial decision: selecting a segmentation criterion for market segmentation. The term "segmentation criterion" refers to the type of information utilized in this process. Four primary segmentation criteria are commonly employed: geographic, sociodemographic, psychographic, and behavioral. The choice of criterion depends on various factors including organizational objectives, product/service nature, available data, and desired segmentation precision. Therefore, organizations must carefully evaluate and choose the most appropriate criterion that aligns with their marketing goals and target market characteristics.

3.2.1 Geographic Segmentation

Geographic information is often regarded as the original criterion for market segmentation. Typically, this approach involves using the consumer's place of residence as the sole basis for forming market segments. While simplistic, geographic segmentation is frequently deemed appropriate. Its primary advantage lies in the ease of assigning each consumer to a geographic unit, facilitating targeted communication and channel selection to reach specific geographic segments.

However, a key disadvantage is that residing in the same geographical area doesn't necessarily imply shared characteristics relevant to marketers, such as purchasing preferences or desired product benefits. Despite this limitation, geographic segmentation has seen a resurgence in international market segmentation studies. These studies aim to extract segments across geographic boundaries. Nevertheless, this approach presents challenges as segmentation variables must be meaningful across diverse geographic regions. Additionally, biases may arise if surveys are conducted with respondents from different cultural backgrounds.

3.2.2 Socio-Demographic Segmentation

Socio-demographic segmentation criteria, including age, gender, income, and education, are commonly utilized in market segmentation. While they can be valuable in certain industries, similar to geographic segmentation, socio-demographic criteria offer the advantage of easily determining segment membership for each consumer. However, socio-demographic characteristics may not always directly explain product preferences. While they might provide insights in some cases, they may not offer sufficient depth for optimal segmentation decisions. Critics argue that values, tastes, and preferences are more influential in driving consumer

buying decisions compared to socio-demographic factors, suggesting a need for a more nuanced approach to segmentation.

3.2.3 Psychographic Segmentation

Psychographic segmentation involves grouping individuals based on psychological criteria like beliefs, interests, preferences, aspirations, or purchasing motivations. Unlike geographic or sociodemographic criteria, psychographic characteristics are inherently complex and often require multiple variables for segmentation. While psychographic segmentation offers insights into the underlying reasons for consumer behavior differences, determining segment memberships is more challenging due to increased complexity. Additionally, the effectiveness of psychographic segmentation relies heavily on the reliability and validity of the empirical measures used to capture these psychological dimensions.

3.2.4 Behavioural Segmentation

Behavioral segmentation involves extracting segments based on similarities in behavior or reported behavior, such as prior product experience, purchase frequency, amount spent, and information search habits. Unlike other approaches, behavioral segmentation uses actual behavior as the basis for segment extraction, which can be advantageous as it focuses on the most relevant similarity. It also eliminates the need for developing measures for psychological constructs. However, obtaining behavioral data may be challenging, particularly when including potential customers who haven't previously purchased the product. This approach is particularly useful for studying existing customers but may require additional effort to include potential customers in the analysis.

3.3 Data from Survey Studies

Market segmentation analyses often rely on survey data due to its affordability and accessibility. However, survey data, unlike data derived from observing actual behavior, is susceptible to various biases. These biases can compromise the quality of segmentation analysis outcomes. Several key considerations must be addressed when using survey data.

3.3.1 Choice of Variables

In data-driven segmentation, it's essential to include all relevant variables related to the segmentation criterion while avoiding unnecessary ones. Including unnecessary variables can increase the complexity of the segmentation problem without adding valuable information, making it more challenging for data analytic techniques to extract optimal market segments. These unnecessary variables, termed noisy or masking variables, divert attention away from critical information needed for accurate segmentation. They often result from poorly developed survey questions or careless selection of segmentation variables. To mitigate this issue, it's crucial to conduct exploratory or qualitative research to identify necessary variables and develop a comprehensive questionnaire. This two-stage process ensures that all critical variables are included, enhancing the quality of segmentation analysis.

3.3.2 Response Options

Survey response options determine the type of data available for analysis, impacting subsequent segmentation analysis. Nominal variables, offering unordered categories like occupation choices, can be transformed into binary data for segmentation. Metrics, such as age or nights spent at a hotel, allow for statistical procedures, including distance measurement, making them ideal for segmentation. While ordinal scales are common, binary or metric options are preferable for segmentation to avoid complications with distance measures. Binary options, especially when formulated level-free, often outperform ordinal scales in various contexts. Therefore, providing binary or metric response options to respondents enhances the effectiveness of segmentation analysis.

3.3.3 Response Styles

Survey data is susceptible to various biases, including response bias and response styles. Response bias refers to a systematic tendency to respond to survey items based on factors other than the item content. If this bias persists consistently over time and regardless of the survey questions asked, it represents a response style. Response styles manifest in various forms, such as using extreme or midpoint answer options, or agreeing with all statements. These response styles can impact segmentation results as algorithms may not differentiate between genuine beliefs and response styles. Minimizing the risk of capturing response styles is crucial in market segmentation data collection. In cases where response patterns indicative of a response style emerge, additional analyses are needed to address this issue. Alternatively, respondents affected by such response styles may need to be excluded before targeting specific market segments.

3.3.4 Sample Size

Market segmentation analysis does not come with specific sample size recommendations, unlike many statistical analyses. However, having a sufficient sample size is crucial for determining the correct number and nature of market segments. The adjusted Rand index is used to measure the correctness of segment recovery, with higher values indicating better alignment between segmentation solutions. Several characteristics of survey data, such as sampling error, response biases, and low data quality, can affect segment recovery. Larger sample sizes generally improve the algorithm's ability to identify correct segmentation solutions, but the extent of improvement varies across different market and data characteristics. Some challenges, like correlation between variables, cannot be effectively compensated for by increasing sample size. Therefore, collecting high-quality, unbiased data is essential for accurate market segmentation analysis.

3.4 Data from Internal Sources

Organizations increasingly leverage internal data, such as scanner data from grocery stores, booking data from airline loyalty programs, and online purchase data, for market segmentation analysis. The strength of such data lies in their representation of actual consumer behavior rather than self-reported intentions, which are prone to biases. Additionally, these data are usually readily available and require minimal effort to collect. However, a potential drawback is that internal data may be biased towards existing customers, overlooking potential future customer segments with different consumption patterns.

3.5 Data from Experimental Studies

Experimental data, derived from field or laboratory experiments, offer another valuable source for market segmentation analysis. These experiments can involve testing consumer responses to advertisements or presenting consumers with product stimuli characterized by different attribute levels. Conjoint analyses and choice experiments yield insights into consumer

preferences regarding specific product attributes and attribute levels. This information can be utilized as segmentation criteria, providing valuable insights into consumer behavior and preferences.

3.6 Checklist

1. Identify Promising Segmentation Variables:
   - Convene a market segmentation team meeting.
   - Discuss potential consumer characteristics that could serve as promising segmentation variables.
   - These variables will be utilized to extract groups of consumers from the data.

2. Determine Descriptor Variables:
   - Discuss additional consumer characteristics required for a comprehensive understanding of market segments.
   - These variables will be used to describe the segments in detail.

3. Data Collection Strategy:
   - Determine the most valid methods to collect data for both segmentation and descriptor variables.
   - Carefully design data collection to minimize data contamination from biases and systematic errors.
4. Data Collection Implementation:
   - Implement the designed data collection strategy.
   - Collect data meticulously to ensure accurate capture of segmentation and descriptor variables.

Step 4: Exploring Data

In Step 4 of the market segmentation process, the focus shifts to exploring and understanding the collected data before proceeding to segment formation. This phase involves various techniques and analyses to make the data more manageable, meaningful, and suitable for subsequent segmentation tasks. Let's delve into the details of each sub-section:

4.1 A First Glimpse at the Data:

At the outset, it's crucial to get an initial overview of the collected data. This involves examining the data's structure, format, and content. Preliminary observations help identify potential issues, such as missing values, outliers, or inconsistencies, which may require further investigation and cleaning.

4.2 Data Cleaning:

Data cleaning is a critical step to ensure the accuracy and reliability of the data. This process involves identifying and addressing errors, inconsistencies, and missing values. Data cleaning may include imputing missing values, correcting inaccuracies, and standardizing formats. Clean data is essential for accurate analysis and reliable segmentation outcomes.

4.3 Descriptive Analysis:

Descriptive analysis involves summarizing and visualizing the data to gain insights into its distribution and characteristics. This step helps identify patterns, trends, and relationships among variables. Techniques such as histograms, bar charts, and scatter plots can provide a clearer understanding of the data's nature.

4.4 Pre-Processing:

Pre-processing prepares the data for more advanced analysis techniques. This step includes transforming and scaling the data to make it suitable for segmentation methods. Two main aspects of pre-processing are:

4.4.1 Categorical Variables:

Categorical variables represent discrete categories or groups. To use them effectively, they may need to be converted into a suitable format. This could involve techniques like one-hot encoding, where each category becomes a separate binary column, or ordinal encoding, which assigns numerical values based on the order of categories.

4.4.2 Numeric Variables:

Numeric variables may require normalization or standardization to bring them to a common scale. Normalization adjusts the values to a specified range, while standardization scales the values to have a mean of zero and a standard deviation of one. This ensures that variables with different scales do not disproportionately influence segmentation outcomes.

4.5 Principal Components Analysis (PCA):

PCA is a dimensionality reduction technique used to simplify complex data sets by transforming them into a smaller set of uncorrelated variables called principal components. This reduces the number of variables while retaining as much meaningful information as possible. PCA can help reveal underlying patterns and relationships within the data.

4.6 Step 4 Checklist:

- Explore the data to determine if there are any inconsistencies and if there are any systematic contaminations.

- If necessary, clean the data.
- If necessary, pre-process the data.
- Check if the number of segmentation variables is too high given the available sample size. You should have information from a minimum of 100 consumers for each segmentation variable.
- If you have too many segmentation variables, use one of the available approaches to select a subset.

- Check if the segmentation variables are correlated. If they are, choose a subset of uncorrelated segmentation variables.
- Pass on the cleaned and pre-processed data to Step 5 where segments will be extracted from it.

Step 5: Extracting Segment

Extracting segments is a crucial step in market segmentation, which involves dividing consumers into distinct groups based on shared characteristics. Extracting segments is important. It helps to Understanding Consumer Behavior, Customized Marketing Strategies and Product Development etc.

5.1 Grouping Consumers

Market segmentation analysis, driven by data, is inherently exploratory. Consumer datasets often lack clear structure, as individual preferences span a wide range. Unlike neatly defined clusters, consumer preferences scatter across a two-dimensional plot. Extracting market segments from such data relies on both the method chosen and the underlying assumptions about segment structure.

Here are key points to consider:

Exploratory Nature: Market segmentation analysis explores consumer behavior. The unstructured nature of consumer data complicates the process.

Algorithm Dependence: The choice of extraction algorithm significantly impacts the segmentation outcome. Assumptions about segment structure play a crucial role.

Cluster Analysis: Many segmentation methods come from cluster analysis. Each market segment corresponds to a cluster. However, the suitability of a clustering method depends on the research context.

Algorithmic Influence: Algorithms impose structure on segments: The same dataset with two spiraling segments is segmented using different algorithms and segment counts.

5.2 Distance-Based Methods

Distance-based methods can use distance measures that account for this asymmetry, and extract segments characterized by common 1s. Market segmentation aims at grouping consumers into groups with similar needs or behaviors.

A distance is a function d with two arguments: the two vectors x and y between which the distance is being calculated. The result is the distance between them (a nonnegative value).

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 $\Leftrightarrow$ x=0

The asymmetric binary distance does not use all dimensions of the vectors. It only uses dimensions where at least one of the two vectors has a value of 1. It is asymmetric because it treats 0s and 1s differently.

5.2.2 Hierarchical Methods

Hierarchical clustering techniques are intuitive for grouping data because they mirror how a human would naturally divide a set of n observations (consumers) into k groups (segments).

Divisive hierarchical clustering begins with the entire dataset X and initially divides it into two market segments. Subsequently, each of these segments is further split into two smaller segments. This recursive process continues until each individual consumer forms their own distinct market segment.

Agglomerative hierarchical clustering takes a different approach. It begins with each consumer as an individual market segment (forming n singleton clusters). Gradually, the two market segments that are closest to each other are merged step-by-step until the entire dataset coalesces into a single large market segment.

5.2.3 Partitioning Methods

For the analysis of small data sets with up to a few hundred observations used Hierarchical Methods. For larger datasets, dendrograms can become challenging to interpret, and the pairwise distance matrix may not fit into computer memory. When dealing with datasets containing more than 1000 consumers, In such cases, instead of computing distances between all pairs of observations at the outset of hierarchical clustering, only the distances between each consumer and the center of the segments are calculated. This approach helps manage memory constraints and simplifies the process.

5.2.3.1 k-Means and k-Centroid Clustering The most popular method is k-mean a number of algorithms are available in this method. These algorithms use the squared Euclidean distance. The goal is to create segments where consumers are similar to their fellow segment members but dissimilar to members of other segments.

5.2.3.2 "Improved" k-Means Refining the k-means clustering algorithm involves improving its initialization process. Instead of randomly selecting k consumers as starting points, a smarter approach is to choose representative starting values.

5.2.3.3 Hard Competitive Learning It is also known as learning vector quantitation. It is different from standard k-means algorithm. In the Hard competitive learning randomly picks one consumer and moves this consumer's closest segment representative a small step into the direction of the randomly chosen consumer.

5.2.3.4 Neural Gas and Topology Representing Networks The neural gas algorithm is a variation of hard competitive learning by Martinez et al. In this not only segment representative is moved towards the randomly selected consumer Instead, also the location of the second closest segment representative is adjusted towards the randomly selected consumer.

5.2.3.5 Self-Organizing Maps the Self-Organizing Map algorithm provides a powerful way to explore data relationships and reduce dimensionality while considering both local and global context

5.2.3.6 Neural Networks

Auto-encoding neural networks for cluster analysis operate differently from traditional methods. They use a single hidden layer perceptron.

5.2.4 Hybrid Approaches

Hierarchical Cluster Algorithms:

No need to specify the number of market segments in advance, visual representation using dendrograms.

Disadvantages: Requires substantial memory capacity, challenging interpretation with large sample sizes.

Partitioning Cluster Algorithms:

Strengths: Minimal memory requirements, suitable for large data sets.

Disadvantages: Requires specifying the number of market segments in advance, non-nested segmentation solutions.

Combining these approaches aims to leverage their respective benefits while mitigating weaknesses.

5.2.4.1 Two-Step Clustering The two steps clustering consist of run a partitioning procedure

followed by a hierarchical procedure.

5.2.4.2 Bagged Clustering It is also combines hierarchical clustering algorithms and partitioning clustering algorithms, but adds bootstrapping in bagged clustering.

5.3 Model-Based Method

The pioneers of model-based methods in market segmentation analysis are mixture methodologies Model-based methods serve as an additional approach for segment extraction available to data analysts. Since segmenting markets is an exploratory task, it's beneficial to employ a variety of extraction methods to identify the most suitable approach for the specific dataset.

5.3.1 Finite Mixtures of Distributions The simplest case of model-based clustering has no independent variables x, and simply fits a distribution to y. To compare this with distance-based methods, finite mixtures of distributions basically use the same segmentation variables.

There are two types of this method

- Normal Distributions
- Binary Distributions

5.3.2 Finite Mixtures of Regression finite mixtures of distributions provide an alternative perspective on market segmentation, emphasizing the varying relationships between variables across segments. Mixture models can produce both more and less useful solutions compared to

hierarchical or partitioning clustering methods. These models assume a dependent target variable (y) explained by a set of independent variables (x). The functional relationship between dependent and independent variables varies across different market segments.

5.3.3 Extensions and Variations Finite Mixture Models are more intricate than distance-based methods. Their added complexity provides great flexibility, allowing the use of any statistical model to describe market segments. Consequently, finite mixture models can handle a wide range of data characteristics.

5.4 Algorithms with Integrated Variable Selection

Many algorithms assume that each segmentation variable contributes to determining the segmentation solution. However, this is not always the case. there are some Preprocessing Methods for Variable Selection

- Filtering Approach
- Challenges with Binary Data

There are some Algorithms for Simultaneous Variable Selection and Segment Extraction

5.4.1 Biclustering Algorithm Biclustering simultaneously clusters both consumers and variables. Biclustering algorithms exist for any kind of data, including metric and binary.

5.4.2 Variable Selection Procedure for Clustering Binary Data 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.

5.4.3 Variable Reduction: Factor-Cluster Analysis It is involving a two-step process for data-driven market segmentation

- In the first step, segmentation variables undergo factor analysis.
- In the second step, market segments are extracted using the factor scores obtained from the factor analysis

5.5 Data Structure Analysis

Data structure analysis provides valuable insights into the properties of the data. These insights guide subsequent methodological decisions. Most importantly, stability-based data structure analysis provides an indication of whether natural, distinct, and well-separated market segments exist in the data or not.

We discuss four different approaches to data structure analysis:

5.5.1 Cluster Indices It is providing insight into particular aspects of the market segmentation solution. Which kind of insight, depends on the nature of the cluster index used. Generally, two groups of cluster indices are distinguished

5.5.1.1 Internal Cluster Indices Internal cluster indices use a single segmentation solution as a starting point. Solutions could result from hierarchical, partitioning or model-based clustering methods.

5.5.1.2 External Cluster Indices It is assessing a market segmentation solution by incorporating additional external information. Unlike internal indices, which rely solely on the information within a single segmentation solution, external indices consider external factors to evaluate the quality of the segmentation.

5.5.2 Gorge Plot It is a simple method to assess how well segments are separated, is to look at the distances of each consumer to all segment representatives. If natural, well-separated market segments are present in the data, the gorge plot to contain many very low and many very high values. This is why this plot is referred to as gorge plot.

5.5.3 Global Stability Analysis To assess the global stability of any given segmentation solution, several new data sets are generated using resampling methods, and a number of segmentation solutions are extracted.

5.5.4 Segment Level Stability Analysis Relying on global stability analysis could lead to selecting a segmentation solution with suitable global stability, but without a single highly stable segment. It is recommendable, therefore, to assess not only global stability of alternative market segmentation solutions, but also segment level stability of market segments.

5.5.4.1 Segment Level Stability Within Solutions (SLSW) The criterion of segment level stability within solutions (SLSW) is similar to the concept of global stability (see Sect. 7.5.3). The difference is that stability is computed at segment level, allowing the detection of one highly stable segment.

5.5.4.2 Segment Level Stability Across Solutions (SLSA) The second criterion of stability at segment level proposed by Dolnicar and Leisce(2017) is referred to as segment level stability across solutions (SLSA). The purpose of this criterion is to determine the re-occurrence of a market segment across market segmentation solutions containing different numbers of segments.

Step 6: Profiling Segments

Identifying Key Characteristics of Market Segments

- The profiling step aims to understand the resulting market segments from the extraction process, particularly in data-driven segmentation. Unlike commonsense segmentation, where segment profiles are predefined (e.g., age groups), data-driven segmentation requires identifying defining characteristics of segments based on segmentation variables.
- Profiling involves characterizing market segments individually and comparing them to other segments. It helps identify unique characteristics that differentiate segments, ensuring accurate interpretation and strategic decision-making.

- In data-driven segmentation, interpreting segmentation results can be challenging for managers. Studies indicate that a significant percentage of marketing managers struggle to understand data-driven segmentation solutions, often perceiving them as black boxes.
- Traditional methods of presenting segmentation results, such as lengthy reports or spreadsheets, are often deemed insufficient or unclear by marketing managers. Graphical statistics approaches offer a more intuitive and less tedious way of profiling segments, reducing the likelihood of misinterpretation.

Traditional Approaches to Profiling Market Segments

- Data-driven segmentation solutions are typically presented either as oversimplified summaries or as large tables detailing exact percentages for each segmentation variable by segment. However, such tables are challenging to interpret and may not provide a quick overview of key insights.
- Comparing segment percentages with each other or the total values in the table is necessary to identify defining characteristics of segments.
- In scenarios with multiple segmentation solutions, users would have to compare a significant number of pairs of numbers to understand segment characteristics. This task can be overwhelmingly tedious, even for experienced users.
- Providing information on the statistical significance of differences between segments for each segmentation variable is not statistically correct, as segment membership is derived directly from the variables and segments are designed to be maximally different, precluding the use of standard statistical tests.

Segment Profiling with Visualizations

- Traditional tabular representations of market segmentation solutions, whether highly simplified or complex, often neglect the use of graphics, despite the integral role of data visualization in statistical data analysis.
- Graphics play a crucial role in exploratory statistical analysis, such as cluster analysis, by providing insights into complex relationships between variables. Moreover, in the era of big data, visualization offers a simple means of monitoring developments over time.
- Recommendations from scholars advocate for the use of visualization techniques to enhance the interpretability of market segmentation analysis results.
- Haley (1985) highlighted the superior insightfulness of graphical representations compared to tabular forms, while Cornelius et al. (2010) suggested that simpler, two-dimensional graphical formats are preferable for interpreting market structure analysis results.
- Various visualization techniques for cluster analysis and mixture models have been reviewed by Leisch (2008).
- Visualizations play a crucial role in inspecting segments in detail for each segmentation solution, facilitating the interpretation of segment profiles and aiding in the assessment of the usefulness of a given market segmentation solution.

- With the plethora of alternative solutions generated during the segmentation process, selecting the most suitable solution becomes a critical decision, where visualizations of solutions serve as valuable aids for both data analysts and users.

Identifying Defining Characteristics of Market Segments

- A segment profile plot serves as an effective tool to understand the defining characteristics of each market segment by visually representing how each segment differs from the overall sample across all segmentation variables.
- In constructing segment profile plots, the order of segmentation variables in figures and tables need not necessarily follow their order of appearance in the dataset. Instead, variables can be rearranged based on meaningful order or similarity of answer patterns to enhance visualizations.
- Hierarchical clustering of the variables in the dataset can be employed to determine the order of segmentation variables for visualization. This clustering helps identify relationships between variables and aids in organizing them in a meaningful order for analysis.
- Marker variables, which significantly deviate from the overall mean, play a crucial role in segment profile plots. These variables are highlighted in color to distinguish them from others, making it easier to interpret segment characteristics.
- The segment profile plot consists of multiple panels, each representing one segment. These plots display the cluster centers (centroids) for each segment, providing a comparative view of segment characteristics against the overall sample.
- Comparing segment profile plots to traditional tabular representations, reveals that segment profile plots offer a more intuitive and efficient way to interpret segmentation results. They provide a quick overview of segment characteristics, aiding in faster decision-making processes.
- An eye-tracking study comparing interpretation efficiency between traditional tabular formats and graphical statistics, like segment profile plots, showed that participants required less cognitive effort and spent less time extracting information from graphical representations. This suggests that well-designed visualizations significantly ease the interpretation of segmentation results, particularly for managers making strategic decisions.

Assessing Segment Separation

- Segment separation plots provide a visual representation of the overlap between segments across all relevant dimensions of the data space. These plots offer a quick overview of the data situation and segmentation solution.
- While segment separation plots are straightforward when the number of segmentation variables is low, they become more complex as the number of variables increases. Nonetheless, they remain valuable for assessing segment separation even in complex situations.

- Examples of segment separation plots are provided, illustrating two different data sets: one with three distinct, well-separated clusters, and another with an elliptic data structure. These plots consist of scatter plots of observations colored by segment membership, cluster hulls indicating the shape and spread of segments, and neighborhood graphs indicating similarity between segments.
- When dealing with high-dimensional data, such as the 20-dimensional travel motives data, projection techniques like principal components analysis (PCA) are used to reduce dimensions and create segment separation plots. These plots help visualize the separation between segments in lower-dimensional space.
- Segment separation plots assist in identifying distinct market segments and understanding their characteristics. They allow for the assessment of segment overlap and help in making informed decisions based on segmentation results.
- Interpretation of segment separation plots requires careful analysis, considering the specific characteristics of each segment and understanding the context of the data. While a single projection may show overlap between segments, it does not imply overlap in all projections. Each projection offers unique insights into segment separation.

Step 6 Checklist

- After selecting segments from Step 5, the next step involves visualizing segment profiles to understand what distinguishes each segment from the others. This visualization helps in gaining insights into the unique characteristics of each segment.
- Knock-out criteria are then applied to assess whether any of the selected segments should be eliminated from further consideration. These criteria act as filters to identify segments that do not meet specific requirements or standards predefined by the analyst or organization.
- Segments that fail to comply with the knock-out criteria are eliminated from consideration, leaving only the remaining segments to proceed to Step 7 for further description and analysis.
- Visualization of segment profiles aids in identifying key differences between segments, while knock-out criteria serve as a means to ensure that only relevant and viable segments are carried forward in the segmentation process.
- This iterative process of visualizing segment profiles and applying knock-out criteria helps refine the segmentation solution and ensures that only the most meaningful and actionable segments are considered for further analysis and decision-making.

Step 7: Describing Segments

Developing a Complete Picture of Market Segments

• Step 7 focuses on describing market segments to develop a comprehensive understanding of their characteristics.
• While Step 6 involved profiling segments by analyzing differences in segmentation variables, Step 7 delves deeper by examining additional information about segment members, known as descriptor variables.

• Profiling segments is akin to going on dates to get to know a potential spouse, while describing segments is akin to crossing segments with various other variables to understand them better.
• Market segment descriptions encompass various factors such as age, gender, past behavior, preferences, media usage, and expenditure patterns during vacations, among others.
• Detailed segment descriptions are crucial for customizing marketing strategies tailored to each segment's preferences and characteristics.
• Descriptions of market segments are essential for developing a customized marketing mix to effectively target each segment.
• Differences between market segments concerning descriptor variables can be studied using descriptive statistics, visualizations, or inferential statistics. Visualizations are preferred for their user-friendliness, although traditional statistical testing and tabular presentations are also common in the marketing literature.

Using Visualisations to Describe Market Segments

• Visualizations are crucial for describing differences in descriptor variables among market segments.
• Two basic approaches are discussed: one suitable for nominal and ordinal variables (e.g., gender, education level, country of origin), and another for metric variables (e.g., age, number of nights at tourist destinations, money spent on accommodation).
• Utilizing graphical statistics for describing market segments offers two main advantages: simplifying interpretation for both data analysts and users, and integrating information on the statistical significance of differences, thus preventing over -interpretation of insignificant differences.
• According to Cornelius et al. (2010), graphical representations effectively convey the essence of marketing research results and are preferred by marketing managers for their intuitiveness.

Nominal and Ordinal Descriptor Variables

• Statistical Testing for Descriptor Variables:
o Nominal or ordinal variables: χ^2-test suitable.
o Metric variables: ANOVA or Kruskal-Wallis test used.
• χ^2-test:
o Tests association between nominal segment membership and another nominal or ordinal variable.
o Example: Testing gender distribution differences across segments.
o Small p-value (<0.05) indicates significant differences in variable distribution.
• ANOVA:
o Tests differences in means across segments for metric variables.
o Compares variance between segment means with variance within segments.
o Small p-value (<0.05) suggests significant mean differences among segments.
• Pairwise Comparisons:

o Further insights into specific segment differences.
o Tukey's honest significant differences test adjusts for multiple testing.
o Provides point estimates of differences with confidence intervals.
o Non-crossing confidence intervals indicate significant differences.
• Importance:
o Identifies segments with significant differences in descriptor variables.
o Guides targeted marketing strategies and interventions.

Metric Descriptor Variables

• Metric Descriptor Variables: Crucial for understanding market segments.
• Conditional Plots: Useful for describing differences in metric descriptor variables among segments.
• Data Division: Conditional plots divide data into sections, presenting results for different segments.
• R Package Lattice: Provides functions for creating conditional plots, such as histograms and box-and-whisker plots.
• Visualization: Histograms can compare age distribution or moral obligation scores across segments.
• Challenges: Interpreting differences based solely on histograms can be difficult.
• Box-and-Whisker Plots: Offer clearer visualization of metric variable distribution for each segment.
• Comparison: Facilitate easier comparison and identification of outliers.
• Statistical Inference: Elements like 95% confidence intervals for medians can be incorporated to assess significance of segment differences.
• Modified SLSA Plot: Traces metric variable value across multiple segmentation solutions.
• Insights: Visualizations provide insights into segment differences, aiding in development of targeted marketing strategies based on demographic, socioeconomic, and attitudinal attributes.

Testing for Segment Differences in Descriptor Variables

• Testing for Segment Differences in Descriptor Variables:
o Utilizes statistical tests to formally assess differences across market segments.
• Test Methods:
o χ^2-test for nominal or ordinal variables.
o ANOVA or Kruskal-Wallis test for metric variables.
• χ^2-test:
o Tests association between nominal segment membership and another nominal or ordinal variable.
o Example: Testing gender distribution differences across segments.
o Small p-value (<0.05) indicates significant distribution differences.
• ANOVA:
o Tests differences in means across segments for metric variables.
o Compares variance between segment means with variance within segments.
o Small p-value (<0.05) suggests significant mean differences among segments.

o Pairwise Comparisons:
o Provide insights into specific segment differences in means.
o Tukey's honest significant differences test adjusts for multiple testing.
• Tukey's Test:
o Provides point estimates of differences in means with confidence intervals.
o Significant differences indicated by confidence intervals not crossing zero.
• Importance:
o Statistical tests aid in identifying segments with significant differences in descriptor variables.
o Provide insights for targeted marketing strategies and interventions.

Predicting Segments from Descriptor Variables

• Predicting Market Segment Membership:
o Uses regression models with segment membership as dependent variable and descriptors as independent variables.
• Regression Models:
o Applied for classification using statistical and machine learning methods.
• Approach Contrast:
o Prediction models consider all descriptors simultaneously, unlike individual testing methods.
• Basis of Models:
o Regression analysis predicts dependent variable using independent variables.
• Linear Regression:
o Assumes linear relationship between variables, with normal distribution for dependent variable.
• R lm() Function:
o Fits linear regression, specifying variables.
o Correct handling of intercept and categorical variables crucial for interpretation.
• Generalized Linear Models (GLMs):
o Extend linear regression for wider dependent variable distributions.
o Effective for categorical dependent variables.
• GLMs and Link Function:
o Utilize link function for linear function modeling.
• Special Cases:
o Logistic regression for binary/multinomial dependent variables using logit function.
• Predictive Modeling:
o Identifies critical descriptors for segment membership and assesses model performance.

Binary Logistic Regression

• Binary logistic regression models binary data, predicting event likelihood based on predictor variables.
• Logistic regression assumes success probability follows a Bernoulli distribution, using the logit link function.
• In R, the glm() function fits generalized linear models, specifying binomial family and "logit"

link.
• Example: Predicting segment membership based on age and moral obligation score.
• Model interpretation involves understanding regression coefficients' impact on log odds of success.
• Intercept represents log odds of success when all predictors are 0.
• Predicted probabilities obtained using inverse logit function.
• Model assessment involves examining deviance residuals, coefficient significance, and goodness-of-fit measures like AIC.
• Visualizations like effect plots aid in interpreting predictor variable effects.
• Model selection techniques such as stepwise regression help avoid overfitting.
• Boxplots compare predictive performance of different models.
• Model performance evaluated by comparing predicted probabilities for target group against others.

Multinomial Logistic Regression

• Multinomial logistic regression is used when the dependent variable has more than two categories, and it follows a multinomial distribution with the logistic function as the link function.
• In R, the **multinom()** function from the **nnet** package is used to fit multinomial logistic regression models. The model is specified similarly to binary logistic regression, using a formula and a data frame.
• The coefficients in multinomial logistic regression represent the change in log odds for each category of the dependent variable compared to a baseline category (often the first category).
• Model assessment involves examining deviance residuals, coefficients' significance, and goodness-of-fit measures such as AIC (Akaike Information Criterion).
• Function **Anova()** can be used to test if dropping any of the independent variables significantly reduces model fit.
• Model selection techniques, such as stepwise regression, can be applied to choose the best - fitting model.
• Model performance can be evaluated by comparing predicted segment memberships to observed ones. Visualization tools like mosaic plots and boxplots of predicted probabilities can aid in assessing model performance.
• Effect plots can be used to interpret the effects of predictor variables on predicted probabilities for each segment.
• In multinomial logistic regression, each category of the dependent variable is modeled separately against a reference category, allowing for the interpretation of the effects of predictor variables on the likelihood of belonging to each category.

Tree-Based Methods
• Classification and regression trees (CARTs) predict binary or categorical dependent variables based on independent variables.
• Advantages of CARTs include variable selection, ease of interpretation through visualizations, and straightforward incorporation of interaction effects.
• CARTs are suitable for handling a large number of independent variables.

• However, CARTs can produce unstable results due to small changes in data leading to different trees.
• Tree-building process involves recursively splitting dataset into homogeneous groups based on independent variable values.
• Various tree construction algorithms differ in splitting approach, variable and split point selection criteria, stopping criteria, and final prediction at terminal nodes.
• R offers packages like rpart and partykit for implementing tree-building algorithms.
• The ctree() function from partykit fits conditional inference trees, performing unbiased variable selection.
• Fitted tree model output describes tree structure, including nodes, splits, and terminal nodes.
• Visualizing the tree using plot() aids in clearer interpretation of tree structure.
• Parameters adjustment in tree-building algorithm using control argument influences tree construction.
• Classification trees handle binary or categorical dependent variables, with each categorical variable level representing a different segment or category.
• Interpretation involves examining splits and terminal nodes to understand important variables for prediction and their interactions.

Step 7 Checklist

• Utilize market segmentation solutions identified in Step 6 based on appealing profiles.
• Select descriptor variables to enhance segmentation analysis.
• Implement visualization methods to grasp variations among market segments regarding descriptor variables.
• Use appropriate plots like mosaic plots for categorical/ordinal variables and box -andwhisker plots for metric variables.
• Evaluate the statistical significance of descriptor variables.
• Apply corrections for multiple testing if separate statistical tests were utilized to avoid overestimation of significance.
• Present each market segment to team members to assess comprehension.
• Inquire about the necessity for additional insights to improve understanding of specific segments.

STEP 8: Selecting the Target Segment(s)

This step frames an organized methodology for a division group meeting pointed toward choosing potential objective business sectors in view of recently dissected market fragments. The interaction includes a progression of basic assessments and estimations to guarantee the chose sections are suitable and line up with the association's capacities and objectives. Here is a breakdown of each undertaking:

Convene a Segmentation Team Meeting:

Gather key team members involved in market segmentation and strategy. Determine Potential Target Markets: Assess Organizational Competitiveness for Each Segment: Evaluate and agree on the organization's relative competitiveness in each segment (e.g., resources, expertise, brand strength).

Preliminary Selection of Target Segments:

 Based on the evaluations and plot, select the most viable segments as potential targets.
✓ Choosing your ideal market segment is a crucial step, not just picking the biggest or most profitable one. Here's what matters.
✓ Knock-Out Criteria: Before diving in, ensure each segment meets essential criteria like size, internal similarity, distinctiveness from others, identifiability, reachability, and alignment with your company's strengths.
Prioritize Smartly: Recognize scores to prioritize segments that offer both high attractiveness and strong competitiveness. Don't be afraid to eliminate segments with high profit potential but low attractiveness or competitiveness – they might not be a good long-term fit.

The Targeting Decision

Selecting the Target Segments is a critical decision-making point. This step involves selecting one or more target market segments from the options identified earlier in the process.
In Step 8, the team reviews knock-out criteria established in Step 2 to ensure that all considered segments meet these criteria. These criteria include factors like size, homogeneity, distinctiveness, identifiability, reachability, and alignment with organizational capabilities. Once knock-out criteria are confirmed, the team evaluates the attractiveness of remaining segments and the organization's competitiveness for each.

Market Segment Evaluation

Market Segment Evaluation This section discusses the use of decision matrices in the market segmentation process, particularly in Step 8, which involves selecting target market segments. The matrices help visualize relative segment attractiveness and organizational competitiveness.

 In the suggested example, a generic segment evaluation plot is presented, emphasizing two axes: "How attractive is the segment to us?" and "How attractive are we to the segment?" To determine segment attractiveness, the team assigns values for each attractiveness criterion based on profiles and descriptions resulting from previous steps. The weighted values for each criterion are then summed up to represent a segment's overall attractiveness. The same process is applied to determine relative organizational competitiveness. Then the segments with lower attractiveness may be eliminated, despite high profit potential.

STEP 9: Customising the Market Mix

Implications for Marketing Mix Decisions

• The widely accepted modern interpretation is the 4Ps model, focusing on Product, Price, Promotion, and Place as the key components of an effective marketing strategy.
• Market segmentation is integral to strategic marketing and is closely linked with positioning and competition.
• The segmentation-targeting-positioning (STP) approach, emphasizes a sequential process: segmentation involves profiling and describing segments, followed by targeting the selection of a specific segment, and positioning focuses on distinct product perception aligned with segment needs, differentiating it from competitors.

Product: -

✓ One of the key decisions an organisation needs to make when developing the product dimension of the marketing mix, is to specify the product in view of customer needs.
✓ Other marketing mix decisions that fall under the product dimension are: naming the product, packaging it, offering or not offering warranties, and after sales support services. In terms of the product targeted at this market segment, possible product measures may include developing a new product.

Price

• Price is the amount of money customers are willing to pay for a product or service.
• Pricing decisions involve determining the right balance between setting a price that covers production and distribution costs while remaining attractive and competitive in the market.

Place

• Place refers to the distribution channels through which a product or service is made available to customers.
• It involves decisions related to the selection of distribution channels, logistics, inventory management, and ensuring that the product is available at the right place and time for customers to purchase.

Promotion

• Promotion encompasses all the activities that a company undertakes to communicate and promote its products or services to the target audience.

• This includes advertising, sales promotions, public relations, personal selling, and other promotional activities. The goal is to create awareness, generate interest, and persuade customers to make a purchase.

Fast Food Case Study Github Links:

Aiswarya Sudhir - https://github.com/Aiswaryasudhir/Fast-Food-Market-Segmentation-Analysis

Puja Rajput - https://github.com/poojarajput2/Market-Segmentation-with-Clustering

Vedant Aryan - https://github.com/vedyan/FeynnLabs_ML_Intern

Gajanan Santosh Purud - https://github.com/GAJANAN07/Market-Segmentation-project-2-study-task-Mc-Donalds-case-study-feynn-labs-

Jada Sunil - https://github.com/sunilyadav2713/casestudy