**Summarized notes**

*Submitted by: Srivathsa SH*

*On: 26-10- 2024*

**Step 1: Deciding (not) to Segment**

Market segmentation is the process marketing managers use to identify and select target markets for their products by dividing a broad market into smaller, more specific groups, or segments, based on shared characteristics. Smith (1956) was the first to suggest using segmentation as a marketing strategy, defining it as breaking down a diverse market into smaller, similar groups. This helps businesses understand their customers better.

***1.1 Implications of Committing to Market Segmentation***

While market segmentation is a valuable strategy for organizations, it’s not universally the best choice. Companies must consider the long-term implications before committing significant time and resources, as effective segmentation. It requires substantial organizational changes and investments in research, surveys, and tailored products and marketing messages; these costs must be justified by expected increases in sales (Cahill, 2006). Organizations may need to develop new products, adjust pricing, modify distribution channels, and enhance communication strategies, which could require changes in their internal structure to focus on different market needs. This decision should be made at the highest levels of the organization and communicated clearly throughout the company to ensure everyone is clear with the terms.

***1.2 Implementation Barriers***

1. **Senior Management Barriers**: Lack of leadership and commitment from senior management can undermine segmentation efforts.
2. **Need for Recognition**: Senior leaders must recognize the need for segmentation, understand the process, and show active interest (McDonald and Dunbar, 1995).
3. **Resource Allocation**: Insufficient resources allocated for both initial analysis and long-term implementation can hinder success.
4. **Organizational Culture Barriers**: A lack of market or consumer orientation within the organization can prevent effective segmentation.
5. **Resistance to Change**: Resistance to change, poor communication, and office politics can obstruct new ideas and creative thinking (Dibb and Simkin, 2008).
6. **Lack of Training**: If senior management and the segmentation team lack understanding of market segmentation principles, the effort is likely to fail.
7. **Absence of a Formal Marketing Function**: Organizations without a qualified marketing expert or formal marketing structure may struggle with implementation.
8. **Objective Restrictions**: Limited financial resources can restrict the ability to pursue segmentation effectively.
9. **Proactive Barrier Identification**: Many barriers can be identified early in the segmentation study and addressed proactively; if barriers cannot be removed, organizations should consider abandoning the segmentation effort.

***1.3 Step 1 Checklist***

This first checklist includes not only tasks, but also a series of questions which, if not answered in the affirmative, serve as knock-out criteria.

|  |  |
| --- | --- |
| Is the organization’s culture market-oriented? | If yes, proceed; if no, reconsider. |
| Is the organization genuinely willing to change? | If yes, proceed; if no, reconsider. |
| Does the organization take a long-term perspective? | If yes, proceed; if no, reconsider. |
| Is the organization open to new ideas? | If yes, proceed; if no, reconsider. |
| Is communication across organizational units effective? | If yes, proceed; if no, reconsider. |
| Is the organization in a position to make significant structural changes? | If yes, proceed; if no, reconsider. |
| Does the organization have sufficient financial resources to support a segmentation strategy? | If yes, proceed; if no, reconsider. |
| Is there visible commitment to market segmentation from senior management? |  |
| Is there active involvement of senior management in the market segmentation analysis? |  |
| Is the market segmentation concept fully understood? | If not, conduct training. |
| Are the implications of pursuing a market segmentation strategy fully understood? | If not, conduct training. |

**Step 2: Specifying the Ideal Target Segment**

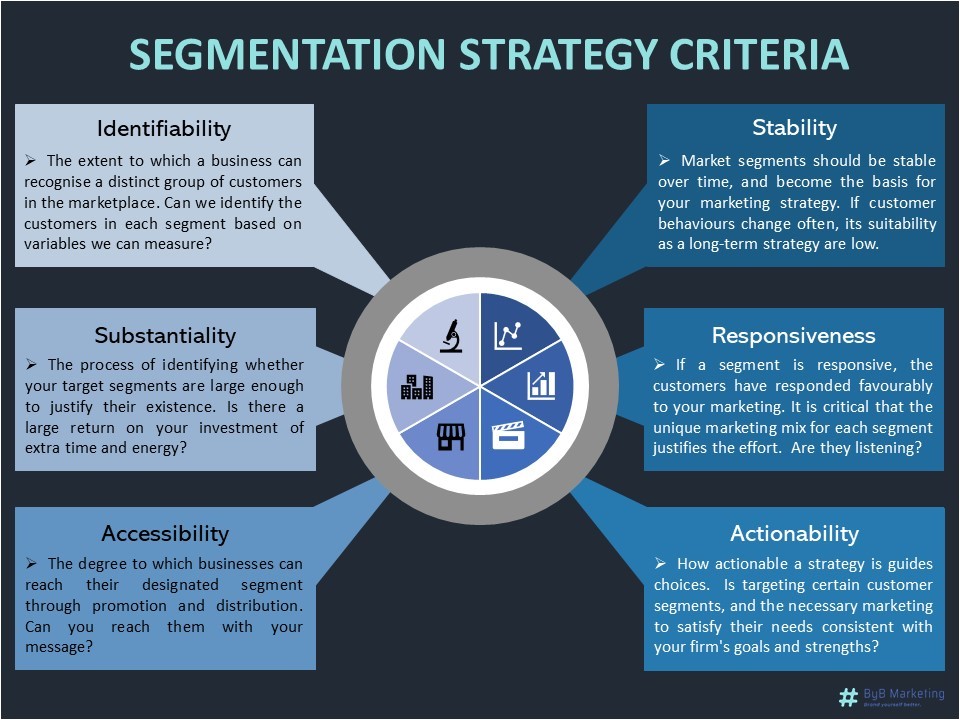
***2.1 Segment Evaluation Criteria***

The third layer of market segmentation analysis relies heavily on user input throughout the process, rather than limiting it to initial briefings or final marketing mix development. User involvement is crucial in most stages, especially after an organization decides to explore a segmentation strategy.

In Step 2 of the analysis, the organization must significantly contribute by establishing two sets of criteria for evaluating segments:

1. **Knock-Out Criteria**: These are essential, non-negotiable attributes that a segment must possess to be considered for targeting. If a segment does not meet these criteria, it is automatically excluded.
2. **Attractiveness Criteria**: Used to assess the appeal of segments that meet the knock-out criteria, these criteria help evaluate and rank the remaining segments for potential targeting.

While literature often discusses a variety of possible evaluation criteria, it does not always clearly differentiate between knock-out and attractiveness criteria. A selection of these criteria is detailed in Table,

Members of the segmentation team need to select which of these criteria they want to use to determine how attractive potential target segments are. The segmentation team also needs to assess the relative importance of each attractiveness criterion to the organisation

***2.2 Knock-Out Criteria***

Knock-out criteria are essential for assessing whether market segments identified in the analysis qualify for further evaluation using attractiveness criteria, the key knock-out criteria include:

* **Homogeneity**: Segment members should be similar.
* **Distinctiveness**: The segment must differ from others.
* **Size**: The segment should be large enough to justify a customized marketing approach.
* **Alignment with Strengths**: The organization must be capable of meeting segment needs.
* **Identifiability**: Segment members should be easily recognizable in the marketplace.
* **Accessibility**: There should be a way to effectively reach segment members.

These criteria must be understood by senior management, the segmentation team, and the advisory committee.

***2.3 Attractiveness Criteria***

In market segmentation analysis, attractiveness criteria are used to evaluate potential segments beyond binary compliance. Each segment is rated on a scale for various criteria rather than simply being categorized as attractive or unattractive. This detailed evaluation helps to provide a clearer view of how appealing each market segment is. In the end, the overall attractiveness ratings for all criteria guide the decision on which segments will be chosen for targeting in the marketing strategy.

***2.4 Implementing a Structured Process***

The literature emphasizes the importance of following a structured approach when assessing market segments. One popular method is to use a segment evaluation plot, which 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.

To establish these criteria, various potential factors related to segment attractiveness and organizational competitiveness must be examined, and a consensus reached on the most critical ones. It is advisable to limit the number of key factors to no more than six. This task should ideally be done by a small core team that drafts initial proposals and presents them to a broader advisory committee with representatives from different organizational units. This inclusion is important as various units offer unique perspectives and will be impacted by the segmentation strategy.

While the segment evaluation plot cannot be completed in the early stages of segmentation analysis due to the absence of available segments, selecting attractiveness criteria early is beneficial. It ensures that relevant information is captured during data collection, facilitating target segment selection later in the process.

By the end of the criteria selection step, the team should have a list of around six attractiveness criteria, each assigned a weight to indicate its relative importance. Team members typically allocate a total of 100 points among the criteria, followed by negotiations to reach conclusion. Approval from the advisory committee is recommended to ensure that multiple viewpoints are considered in defining the segment attractiveness criteria.

***2.5 Step 2 Checklist***

| **Task Number** | **Task Description** |
| --- | --- |
| 1 | Convene a segmentation team meeting. |
| 2 | Agree on knock-out criteria for eliminating segments. |
| 3 | Present knock-out criteria to the advisory committee. |
| 4 | Research criteria for assessing segment attractiveness. |
| 5 | Discuss and agree on a maximum of six attractiveness criteria. |
| 6 | Distribute 100 points across the selected criteria. |
| 7 | Finalize weightings for the criteria collaboratively. |
| 8 | Present selected criteria and weights to the advisory committee. |

**Step 3: Collecting Data**

***3.1 Segmentation Variable***

Segmentation variables are key characteristics from empirical data that help create and describe market segments. Commonsense segmentation uses a single variable, such as gender, while data-driven segmentation utilizes multiple variables, like benefits sought, to identify meaningful segments. The quality of this empirical data is vital for accurately assigning individuals to segments and developing effective marketing strategies. Data can be sourced from surveys, observational studies, and experiments, but should reflect true consumer behaviour. Overall, high-quality empirical data is essential for successful market segmentation analysis in both commonsense and data-driven approaches.

***3.2 Segmentation Criteria***

Before extracting market segments, organizations must choose an appropriate segmentation criterion, which encompasses more than just a single variable and requires market knowledge. Common criteria include geographic, socio-demographic, psychographic, and behavioural factors. Bock and Uncles (2002) highlight the importance of considering profitability and consumer preferences. While there are many potential criteria, the simplest approach is often best.

| **Segmentation Type** | **Criteria** | **Advantages** | **Disadvantages** | **Examples** |
| --- | --- | --- | --- | --- |
| **Geographic Segmentation** | Location of residence (e.g., country, city) | Easy to assign consumers to segments; facilitates local targeting | May not capture shared preferences among consumers in the same area | National tourism campaigns (e.g., Austria targeting neighbouring countries) |
| **Socio-Demographic Segmentation** | Age, gender, income, education | Clear segment membership; can explain some product preferences | Provides limited insight (Haley, 1985); demographics explain ~5% of behaviour | Luxury goods marketing targeted at high-income individuals |
| **Psychographic Segmentation** | Beliefs, interests, preferences, aspirations | Reflects deeper motivations for consumer behaviour | More complex to determine segment memberships; reliant on valid measures | Travel companies targeting cultural tourists based on interests |
| **Behavioural Segmentation** | Purchase behaviour (e.g., frequency, spending) | Uses actual behaviour for segmentation; highly relevant insights | Data may be unavailable for potential customers who haven’t purchased yet | Analysing actual consumer spending patterns for targeted promotions |

***3.3 Data from Survey Studies***

Data from survey studies are commonly used for market segmentation analysis due to its low cost and ease of collection. However, survey data can be biased and negatively impact the quality of segmentation solutions. Careful selection of variables is crucial in both commonsense and data-driven segmentation to ensure optimal market segmentation solutions.

***3.3.1 Choice of variable***

* Carefully selecting the variables for segmentation is crucial for both commonsense and data-driven segmentation quality.
* In data-driven segmentation, all relevant variables related to the segmentation criterion must be included, while unnecessary ones should be avoided to prevent respondent fatigue and lower quality responses.
* Noise variables, which do not contribute to identifying correct market segments, can make extracting the correct solution challenging. This can be avoided by asking necessary and unique questions, without including redundant ones.
* Redundant questions, common in survey research, can interfere with segment extraction algorithms.
* Developing a good questionnaire involves both exploratory and qualitative research to ensure all important variables are included. This process helps in understanding people's beliefs and categorizing insights for questionnaire inclusion.

***3.3.2 Response Options***

* Response options in surveys play a crucial role in determining the scale of data available for analysis. Options that allow respondents to answer in only two ways generate binary data, while those that allow selection from a range of unordered categories correspond to nominal variables.
* Metric data, on the other hand, are generated when respondents indicate a number, such as age. Metric data are ideal for segmentation analysis as they allow for various statistical procedures to be applied.
* Ordinal data, which are commonly used in surveys, arise when respondents are asked to express their agreement using a limited number of ordered answer options. While this format captures fine nuances of responses, it may pose challenges in applying standard distance measures.
* Providing binary or metric response options to respondents is preferred to avoid complications in segmentation analysis
* In some cases, binary response options have been shown to outperform ordinal options, especially when formulated in a level-free manner.

***3.3.3 Response Style***

* Survey data can be influenced by biases, including response bias, where respondents consistently answer in a certain way regardless of the question.
* Various response styles, such as always agreeing or using extreme options, can impact segmentation results.
* For instance, a segment of tourists who say yes to all spending items may seem lucrative, but it could simply be a response style rather than actual spending behaviour. To avoid misinterpretation, it is essential to identify and address response styles during data collection, especially in marketing segmentation
* Additional analysis may be needed to verify genuine market segments and exclude biased responses

***3.3.4 Sample size***

Sample sizes are not typically recommended for market segmentation analysis, but it is crucial that the sample size is adequate in order to determine the correct number and nature of segments within the data. Various studies have explored sample size recommendations for different segmentation variables and segments within the data set.

Formann (1984) suggested that the sample size should be at least 2p (or better five times 2p), where p is the number of segmentation variables, specifically for latent class analysis with binary variables. Qiu and Joe (2015) proposed a recommendation of at least ten times the number of segmentation variables times the number of segments in the data for constructing artificial data sets to study clustering algorithms. Dolnicar et al. (2014) conducted simulation studies and found that a sample size of at least 60p is ideal for correctly identifying segments, with no major improvements beyond a sample size of 70p in more complex scenarios.

Dolnicar et al. (2016) expanded this research to account for market characteristics that impact segmentation algorithms, such as the number and size of market segments, segment overlap, and survey data quality issues like response biases and correlation between items. Larger sample sizes consistently improve an algorithm's ability to identify the correct segmentation solution, but the impact varies based on market and data characteristics. For instance, uncorrelated segmentation variables lead to better segment recovery, while high correlation between variables makes the task challenging even with a larger sample size.

The recommendation from Dolnicar et al. (2016) is to ensure the data contains at least 100 respondents for each segmentation variable to achieve accurate segmentation results. It is crucial to collect high-quality unbiased data with the right items, no unnecessary items, no correlated items, high-quality responses, binary or metric data, no response styles, appropriate sample size, and no correlated items for successful market segmentation analysis.

In conclusion, having a sufficient sample size is essential for accurate market segmentation analysis, and various factors such as market characteristics and data quality can impact the effectiveness of segmentation algorithms. Adhering to recommendations for sample size and data quality parameters can help ensure reliable segmentation results.

***3.4 Data from Internal Sources***

Organizations use internal data, such as scanner records from grocery stores, airline booking data, and online purchase histories, for market segmentation. This data shows real consumer behaviour, which is more reliable than people's self-reported habits that can be biased. Since this data is usually collected automatically, it’s easy to access. However, a major drawback is that it often comes from existing customers, which might not accurately represent potential new customers and their preferences.

***3.5 Data from Experimental Studies***

Experimental data comes from controlled studies, either in the field or in the lab. This includes consumer reactions to specific advertisements and results from choice experiments, where consumers choose between different products with various features. This data helps marketers understand which product attributes are most appealing to consumers, allowing them to create more targeted marketing strategies based on consumer preferences.

**Step 5: Extracting Segments**

***5.1 Grouping Consumers***

**1. Nature of Market Segmentation Analysis**

Market segmentation analysis is inherently exploratory. Consumer data is usually unstructured, and their preferences are widely varied, making it difficult to identify clear clusters when plotted graphically. The process relies heavily on the assumptions embedded in the methods used for segmentation, meaning that different methods can produce varying results depending on how they interpret the structure of the data​

**2. Impact of Clustering Methods**

Clustering methods are a common approach to market segmentation, where the goal is to group consumers with similar characteristics or preferences into clusters. The choice of a suitable clustering technique is crucial, as it must align with the data characteristics and the specific needs of the research. For instance, some algorithms may be better at detecting compact, round clusters, while others might identify more complex shapes in the data.

**3. Example of Algorithm Influence: K-Means vs. Single Linkage**

An example highlights how different algorithms can impose structure on the data. A dataset with two spiralling segments was segmented using both k-means clustering and single linkage hierarchical clustering. The k-means method, which seeks compact clusters, failed to recognize the spiral shapes and grouped consumers based solely on proximity in Euclidean space. Conversely, the single linkage method identified the spiral patterns more effectively. This illustrates how the choice of algorithm can significantly influence the outcome of segmentation analysis, depending on the nature of the underlying data.

**4. No One-Size-Fits-All Solution**

The discussion emphasizes that there is no universally best segmentation algorithm for all datasets. Each method has its strengths and limitations, which become particularly significant when the data is not well-structured. In such cases, the segmentation results are heavily influenced by the method’s inherent tendencies. For instance, k-means may perform well with evenly distributed data but might struggle with non-linear structures, while single linkage might excel in detecting more irregular shapes but could be affected by outliers.

**5. Criteria for Algorithm Selection**

To select an appropriate method for segment extraction, various characteristics of the data and desired segments must be considered. These include:

* **Data Set Characteristics**: The size of the data set, the number of segmentation variables, and the scale level (e.g., nominal, ordinal, metric) all influence which algorithm is most suitable. For example, distance-based methods require the selection of a distance measure that fits the scale of the data.
* **Segment Characteristics**: Researchers need to define what similarities should exist among consumers within a segment and how they should differ across segments. This also helps in determining the structure that the algorithm should impose on the extracted segments.

**6. Challenges with Binary Data**

When segmentation variables are binary (e.g., "yes" or "no" responses to specific questions), additional considerations arise. It may be necessary to treat the variables symmetrically or asymmetrically, depending on whether both the presence and absence of features are important. For example, when using vacation activities as segmentation variables, it may be more interesting to identify tourists who engage in a particular activity than those who do not.

**7. Comparison of Segmentation Methods**

The text stresses the importance of exploring different clustering methods when performing market segmentation. No single method is superior in all cases, and it is often necessary to compare multiple approaches to achieve a robust segmentation solution. The interaction between the chosen algorithm and the data characteristics is pivotal, as it can determine how effectively the segments are identified

**8. Data Characteristics as a Guide**

The segmentation process involves evaluating the sufficiency of the number of consumers for the segmentation variables. The data size impacts how finely segments can be extracted. For instance, larger sample sizes allow for more detailed segmentation, while smaller samples may require algorithms that can simultaneously select key variables during the extraction process.

***5.2 Distance-Based Methods***

These methods measure the **similarity or dissimilarity** between observations to extract segments.

* + 1. **Distance Measures (Euclidean, Manhattan, and Binary):**
  + **Euclidean Distance**: Measures the straight-line distance between two points (used by default in many R functions).
  + **Manhattan Distance**: Calculates the sum of absolute differences, useful in data structured like grids.
  + **Binary Distance**: Handles binary data asymmetrically, focusing on the presence (1s) rather than absence (0s) of characteristics. This is relevant for identifying niche behaviours like rare vacation activities.
    1. **Hierarchical Methods:**
  + **Agglomerative Clustering**: Starts with each data point as its own cluster and merges them step-by-step.
  + **Divisive Clustering**: Begins with one large cluster, splitting it until only single data points remain.
  + **Linkage Methods**:
    - **Single Linkage**: Joins clusters based on the closest points.
    - **Complete Linkage**: Uses the farthest points between clusters.
    - **Average Linkage**: Averages distances between points in clusters.

**5.2.3 Partitioning Methods:**

* + **k-Means Clustering**: Divides data into k clusters by minimizing within-cluster variance.
  + **Improved k-Means**: Uses optimized starting points to avoid getting stuck in local optima.
  + **Hard Competitive Learning**: An advanced variation focusing on minimizing distances iteratively, similar to k-means but with different update rules.

**5.2.4 Hybrid Approaches:**

These methods combine hierarchical and partitioning techniques for more nuanced segmentation solutions.

The analysis concludes with a focus on **implementation in R**, explaining how to compute distances and cluster data efficiently. It emphasizes the importance of selecting appropriate distance measures and clustering algorithms based on the data type and segmentation goals.

***5.3 Model-Based Methods***

1. **Model-Based vs. Distance-Based Methods**:
   * Distance-based clustering relies on measures of similarity to group consumers.
   * Model-based segmentation assumes that each segment has a specific size and members share common characteristics. Instead of distance, it uses statistical models to infer segment memberships.
2. **Finite Mixture Models**:
   * These models blend multiple segment-specific models, representing the market as a mixture of distributions.
   * The segment sizes are determined by a multinomial distribution, and segment-specific characteristics are captured via statistical parameters.
3. **Statistical Techniques**:
   * **Maximum Likelihood Estimation (MLE)**: Estimates parameters by maximizing the likelihood of observed data.
   * **Expectation-Maximization (EM) Algorithm**: Handles the challenge of incomplete data by iteratively updating estimates.
   * **Bayesian Methods**: Utilize Markov Chain Monte Carlo (MCMC) techniques to infer segment memberships.
4. **Model Selection**:
   * Different models vary based on segment-specific covariance structures. For example, **spherical, varying-volume models (VII)** assume the data points form clusters with spherical shapes but different volumes.
   * **BIC, AIC, and ICL criteria** help determine the optimal number of segments by balancing model fit and complexity.
5. **Applications and Example**:
   * An example using Australian vacation motives demonstrates the process. Here, model-based methods are compared with traditional clustering approaches, showing similarities in segment profiles.
   * Various covariance models (like EII, VVV, etc.) are available in the **mclust** R package, helping analysts fine-tune segment solutions.
6. **Challenges**:
   * A key challenge lies in selecting the correct number of segments, as over-segmentation can reduce clarity.
   * Uncertainty plots illustrate segment assignments, highlighting cases where observations are not confidently placed in specific segments, which can reveal artificial segmentation.

This section emphasizes that model-based methods offer a complementary perspective to traditional approaches by using probabilistic models, making them more flexible and capable of capturing complex segment structures.

***5.4 Algorithms with Integrated Variable Selection***

This approach helps improve clustering quality by carefully selecting or reducing variables that define the clusters. Below is a brief breakdown of the subtopics:

**5.4.1 Bi-clustering Algorithms**

* + Bi-clustering simultaneously clusters rows and columns of a data matrix.
  + This technique is useful in identifying subgroups of consumers and their preferences across multiple variables, capturing complex patterns.
    1. **Variable Selection Procedure for Clustering Binary Data (VSBD):**
  + Designed specifically for binary data, this algorithm helps in choosing only the most relevant variables.
  + It ensures that only significant features are retained, enhancing the quality and interpretability of clusters.
    1. **Factor-Cluster Analysis for Variable Reduction:**

This method uses principal component or factor analysis to reduce data dimensionality before clustering.

It simplifies the clustering process by replacing original variables with a smaller set of factors that represent the underlying structure of the data.

These integrated approaches improve segmentation by optimizing the number and type of variables used, ensuring meaningful and manageable results.

***5.5 Data Structure Analysis***

Data Structure Analysis, which is crucial for identifying and extracting market segments. This section emphasizes four key methods for analysing data structures:

1. Cluster Indices:

These indices help determine the optimal number of market segments. They are divided into:

- Internal cluster indices: Use a single segmentation solution to assess how compact or well-separated segments are.

- External cluster indices: Compare multiple segmentation solutions to assess similarity and stability.

2. Gorge Plots:

Visual tools that reveal whether natural clusters exist by displaying patterns in the data. A perfect “gorge” shape indicates strong, well-separated clusters.

3. Global Stability Analysis:

This method involves resampling techniques to evaluate how consistent segmentations are across multiple trials. It helps determine the best segmentation solution.

4. Segment Level Stability Analysis:

Instead of assessing the entire segmentation solution, this method focuses on the stability of individual segments, which is helpful for organizations targeting specific segments.

The section emphasizes that understanding data structures avoids poor segmentation decisions. It suggests that few real-world data sets contain natural clusters, so stability analysis is often needed to find reproducible or managerially useful segments.

**Step 6: Profiling Segments**

***6.1 Identifying Key Characteristics of Market Segments***

* **Goal:** Understand the distinct attributes of segments generated through data-driven segmentation techniques.
* **Process:** Profiling helps compare characteristics across segments to clarify differences. For example, if most tourists ski during their winter vacation, that attribute can characterize a segment but may not distinguish it unless contrasted with other activities or behaviour trends.
* **Challenges:** Managers often struggle to interpret segmentation results accurately due to the complexity of data-driven segmentation.

***6.2 Traditional Approaches to Profiling Market Segments***

* **Demographic Profiling:** Uses age, income, gender, or geographic factors to define segments.
* **Behavioural Profiling:** Focuses on product usage, purchase frequency, or benefits sought.
* **Attitudinal Profiling:** Explores consumer values, interests, or lifestyle attributes.
* **Limitations:** Each traditional method has its pitfalls. For instance, demographic criteria may not adequately capture nuances in consumer behaviour.

***6.3 Segment Profiling with Visualisations***

* **Visual Tools:** Charts and graphs can clarify segment characteristics and relationships between variables.
* **Defining Characteristics:** Variables are analysed to ensure they not only describe but also differentiate segments.
  + **Example:** Identifying the distinct behaviour of consumers who combine skiing with spa activities compared to those who only ski.
* **Assessing Separation:** This step ensures that each segment is distinct enough to warrant targeted marketing.

Code: <https://github.com/srivathsa26/mc-donalds-market-segmentation.git>