Market Segmentation Analysis

By

Prajwal Raut

Date: 28th Oct 2023

Abstract

This report provides a concise overview of Market Segmentation (MS), a widely adopted marketing strategy that involves categorizing an audience into subgroups sharing common characteristics. The document not only introduces the fundamental concepts associated with MS but also delves into the theoretical underpinnings and the practical execution of this strategy using the Python programming language.

Step 1: Deciding (not) to Segment

Implications of Committing to Market Segmentation:

- Market segmentation is a strategic approach used in organizations but requires a long-term commitment.
- It necessitates substantial changes, investments, and organizational restructuring.
- Commitment is crucial and must be driven from the highest executive level.
- Market segmentation should be more profitable than marketing without it to justify the costs.

Implementation Barriers:

- Senior management involvement, leadership, and resource allocation are critical.
- Organizational culture, including market orientation, resistance to change, and communication, plays a significant role.
- Lack of training and understanding of market segmentation concepts can hinder success.
- A formal marketing function or qualified marketing experts are essential.
- Objective restrictions such as financial resources and structural changes can be barriers.
- Process-related barriers, including planning, lack of structured processes, and time constraints, may impede progress.

Step 1 Checklist:

- Assess the organization's culture for market orientation.
- Determine willingness to embrace change.
- Verify commitment to long-term perspective.
- Evaluate openness to new ideas.
- Ensure effective communication between organizational units.
- Check the capacity for significant structural changes.
- Assess financial resources availability.
- Secure visible commitment, involvement, and financial support from senior management.
- Provide training to ensure a clear understanding of market segmentation concepts and implications.
- Assemble a 2–3-member segmentation team including experts in marketing, data, and data analysis.
- Establish an advisory committee representing affected organizational units.
- Clarify objectives and develop a structured process for market segmentation analysis.
- Assign responsibilities within the segmentation team.
- Allocate sufficient time for analysis without time pressure

Step 2: Specifying the Ideal Target Segment:

In this step, the organization prepares for market segmentation by defining the criteria that will be used to evaluate potential market segments. The criteria are divided into two categories:

Knock-Out Criteria:

1. Homogeneous: Segment members should be similar to each other.
2. Distinct: Segment members must be distinct from members of other segments.
3. Large enough: The segment should have a sufficient number of consumers.
4. Matching the strengths of the organization: The organization must have the capability to meet the needs of segment members.
5. Identifiable: It must be possible to identify segment members in the marketplace.
6. Reachable: There must be a way to communicate with segment members effectively.

Attractiveness Criteria:

These criteria are not binary and require rating for each market segment:
- Segment Attractiveness and Organizational Competitiveness (values determined by the segmentation team).
- Specific factors that contribute to attractiveness and competitiveness need to be determined by the organization.
- No more than six factors should be used as a basis for calculating these criteria.
- The advisory committee, consisting of representatives from all organizational units, should be involved in the selection and weighting of these criteria.
- The criteria chosen in this step will guide data collection in Step 3 and target segment selection in Step 8.

Step 2 Checklist:

1. Convene a segmentation team meeting.
2. Discuss and agree on the knock-out criteria: homogeneity, distinctness, size, match, identifiability, and reachability. These criteria will be used to automatically eliminate market segments that do not comply (at the latest in Step 8).
3. Present the knock-out criteria to the advisory committee for discussion and potential adjustments.
4. Individually study available criteria for the assessment of market segment attractiveness.
5. Discuss these criteria with other members of the segmentation team and agree on a subset of no more than six criteria.
6. Individually distribute 100 points across the segment attractiveness criteria agreed upon by the segmentation team. Allocate points to reflect the relative importance of each attractiveness criterion.
7. Discuss weightings with other members of the segmentation team and reach an agreement.
8. Present the selected segment attractiveness criteria and the proposed weights to the advisory committee for discussion and potential adjustments.

Step 3: Collecting Data for Market Segmentation

Segmentation Variables

- Empirical data forms the basis for both commonsense and data-driven market segmentation.
- Segmentation variables are the characteristics used in commonsense segmentation to divide the sample into market segments.
- In commonsense segmentation, a single characteristic (e.g., gender) is often used as the segmentation variable.
- Descriptor variables are additional characteristics used to describe the segments in detail (e.g., age, number of vacations, benefits sought).
- Data-driven market segmentation uses multiple segmentation variables to identify or create market segments.

Quality of Empirical Data

- The quality of empirical data is crucial for developing a valid segmentation solution.
- Good data quality is necessary for correctly assigning individuals to segments and accurately describing the segments.
- Data quality is particularly important for data-driven segmentation.
- Empirical data can come from various sources, including surveys, observations (e.g., scanner data), or experimental studies.

Segmentation Criteria

- The organization must choose the segmentation criteria, which relate to the nature of the information used for market segmentation.
- Common segmentation criteria include geographic, socio-demographic, psychographic, and behavioral factors.
- Factors such as profitability, preferences, barriers to choice, and consumer interaction effects can be relevant for segmentation.
- The recommendation is to use the simplest possible segmentation approach based on the specific marketing context.
- Geographic segmentation is based on location and is useful when differences in languages or regions impact product appeal.
- Socio-demographic segmentation includes criteria like age, gender, income, and education and can be valuable in various industries.
- Psychographic segmentation focuses on psychological criteria such as beliefs, interests, preferences, and benefits sought when buying a product.
- Behavioral segmentation directly considers consumer behavior, including purchase frequency, amount spent, and brand choices.

Data from Survey Studies

- Survey data is a common source for segmentation analysis but can be subject to various biases.
- Careful selection of variables is essential, including including necessary items and avoiding unnecessary or noisy variables.
- The choice of response options in surveys impacts the scale of the data and its suitability for analysis.
- Binary or metric data is preferable for segmentation analysis.
- High-quality, unbiased responses are critical for valid segmentation.
- An adequate sample size (at least 100 respondents per segmentation variable) is necessary for segment recovery.

Data from Internal Sources

- Organizations can use internal data sources, such as scanner data, booking data, or online purchase data, for segmentation.
- Internal data often represent actual consumer behavior.
- However, they may be biased by over-representing existing customers, potentially missing information about potential new customers.

Data from Experimental Studies

- Experimental data can be gathered through field or laboratory experiments, such as response to advertisements or choice experiments.
- This data can be used as segmentation criteria in market segmentation analysis.

Step 3 Checklist:

1. Convene a market segmentation team meeting.
2. Discuss potential segmentation variables for consumer characteristics.
3. Identify the additional variables required to describe market segments.

4. Plan data collection methods for segmentation and descriptor variables.
5. Ensure data collection methods maintain validity.
6. Prioritize data quality by minimizing biases and systematic errors.
7. Use clear and unbiased survey questions.
8. Avoid including unnecessary variables in your data.
9. Choose suitable response options for your data collection.
10. Initiate the data collection process.

Step 4: Exploring Data

A First Glimpse at the Data

- After data collection, conduct exploratory data analysis to clean and preprocess the data.
- Use data exploration to identify measurement levels of variables, investigate univariate distributions, and assess dependencies between variables.
- Pre-process data if needed to prepare it for segmentation algorithms.
- Insights from data exploration guide the selection of appropriate segmentation methods.

Descriptive Analysis

- Utilize descriptive analysis to become familiar with the data and avoid misinterpretation.
- In R, use commands like `summary()` to obtain numeric summaries for variables and frequency counts for categorical variables.
- Visualize data using histograms, boxplots, scatter plots for numeric variables, and bar plots for categorical variables.
- Mosaic plots are useful for illustrating associations between multiple categorical variables.
- Histograms are essential for visualizing the distribution of numeric variables.
- It is critical to create bins for histograms (binning) that cover the entire range of observations.
- Use the bin ranges on the x-axis and frequencies on the y-axis to plot histograms.

Pre-Processing

Categorical Variables

- Consider merging levels of categorical variables if there are too many categories.
- Conversion of ordinal data to numeric may be appropriate if distances between scale points are approximately equal.
- Use binary answer options when appropriate, as they are less prone to capturing response styles.

Numeric Variables

- Standardize numeric variables to put them on a common scale.
- This step balances the influence of different variables in distance-based segmentation methods.

- Principal Component Analysis (PCA) can be used to reduce dimensionality, although it should not replace original variables in the segmentation process.
- PCA can help identify highly correlated variables, which can be pruned to reduce dimensionality.

Step 4 Checklist:

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

2. If necessary, clean the data.

3. If necessary, pre-process the data.

4. 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.

5. If you have too many segmentation variables, use one of the available approaches to select a subset.

6. Check if the segmentation variables are correlated. If they are, choose a subset of uncorrelated segmentation variables.

7. Pass on the cleaned and pre-processed data to Step 5 where segments will be extracted from it.

Step 5: Extracting Segments

Grouping Consumers:

  - Market segmentation is exploratory and data-driven, making it inherently flexible.
  - Consumer data is often unstructured and spread across a broad range.
  - The choice of segmentation method and underlying assumptions greatly influence the segmentation results.

Distance-Based Methods:

  - These methods categorize consumers into segments based on their similarities in terms of distances between data points.
  - Hierarchical methods create nested segments, while partitioning methods generate distinct segments.
  - k-means clustering is a widely used partitioning method that minimizes the squared Euclidean distance.

Distance Measures:

  - Euclidean distance is a common metric for measuring the distance between data points.
  - It computes the straight-line distance between two points in a multi-dimensional space.

- Other distance measures, such as Manhattan distance or Mahalanobis distance, may be appropriate depending on the data.

Hierarchical Methods:

  - Hierarchical clustering approaches mimic how a human would group data.
  - The number of segments can vary, from one large segment to one segment per consumer.
  - Hierarchical methods provide a visual representation of segment relationships through dendrograms.

Partitioning Methods

  - These methods are more suitable for larger datasets and generate distinct segments.
  - k-means clustering is the most popular partitioning method.
  - It minimizes the sum of squared distances between data points and their assigned cluster centroids.

Hybrid Approaches:

  - Hybrid methods combine hierarchical and partitioning algorithms to leverage their respective strengths.
  - They provide more memory-efficient solutions while allowing for flexible numbers of segments.
  - The combination of approaches helps mitigate the limitations of individual methods.

Model-Based Methods:

  - Model-based methods, like mixture models, offer an alternative to distance-based approaches.
  - They have gained interest among researchers and marketing practitioners.
  - Mixture models consider latent class membership and the probability distribution of segmentation variables.

Algorithms with Integrated Variable Selection:

  - Some algorithms include variable selection to identify relevant segmentation variables.
  - Variable selection reduces the impact of redundant or noisy variables on segmentation results.
  - Metrics like clusterability can help assess the suitability of variables for segmentation.

Data Structure Analysis:

  - Market segmentation is exploratory and lacks clear optimization criteria for validation.
  - Traditional validation methods are not feasible as they require running multiple strategies simultaneously.
  - Validation in segmentation analysis often relies on the usefulness and practicality of the resulting segments in real-world applications.

Step 5 Checklist:

1. Pre-select extraction methods that suit your data's properties.

2. Apply chosen extraction methods to group consumers.

3. Evaluate global and segment-level stability to identify promising segmentation solutions and segments.

4. Select the most promising segments based on stability assessments.

5. Assess these selected segments against the knock-out criteria defined in Step 2.

6. Forward the remaining promising segments to Step 6 for detailed profiling.

Step 6: Profiling Segments

Identifying Key Characteristics of Market Segments:

- Profiling is essential when using data-driven market segmentation. For commonsense segmentation, the segment profiles are predefined (e.g., age groups).
- In data-driven segmentation, defining characteristics of segments are not known until after data analysis. Profiling aims to identify these defining characteristics.
- Profiling involves characterizing each market segment individually and in comparison, to other segments.
- It's particularly critical when no natural segments exist in the data, and a reproducible or constructive segmentation approach is required.
- Profiling is essential for the correct interpretation of the segments, which, in turn, influences strategic marketing decisions.

Traditional Approaches to Profiling Market Segments:

- Data-driven market segmentation solutions are often presented as tables with exact percentages for each segmentation variable. However, these can be challenging to interpret.
- Segment profile plots are a graphical alternative that show how each market segment differs from the overall sample for each segmentation variable.
- These plots make interpretation more accessible than tables and facilitate quick understanding of defining characteristics.
- Segment profile plots also highlight "marker variables," which deviate significantly from the overall mean, aiding in the identification of key features.
- These marker variables are defined based on a threshold, such as an absolute difference or relative difference from the overall mean.
- Use of graphical statistics is recommended for better interpretation of segmentation results.

Segment Profiling with Visualizations:

- Segment separation plots provide visual insight into the overlap of segments in multi-dimensional spaces.

- These plots combine scatter plots of observations with neighborhood graphs to indicate similarity between segments.
- When working with high-dimensional data, projection techniques like principal component analysis can be used to create segment separation plots.
- Segments can be well-separated or overlapping, and visualization helps assess their distinctiveness.
- The choice of projection can affect the visual interpretation, and multiple projections may be necessary to fully understand segment relationships.
- Segment separation plots can be modified to enhance readability, such as adjusting colors and highlighting specific areas.

Identifying Defining Characteristics of Market Segments:

- Segment profile plots visually display how each market segment differs from the overall sample.
- Highlighted "marker variables" in these plots help identify key features that distinguish each segment.

Assessing Segment Separation:
- Segment separation plots visualize the overlap or distinctiveness of segments in multi-dimensional data.
- Neighborhood graphs show the similarity between segments based on the distance between segment centers.
- The choice of projection affects how segments appear in the plot.

Step 6 Checklist:

1. Use the Selected Segments from Step 5.
2. Visualize Segment Profiles to learn about what makes each segment distinct.
3. Use knock-out criteria to check if any of the segments currently under consideration should already be eliminated because they do not comply with the knock-out criteria.
4. Pass on the remaining segments to Step 7 for describing.

Github Link for the Case study implementation:

https://github.com/PrajwalRaut8/Mcdonalds_Casestudy/blob/main/Case_Study_McDonalds.ipynb