Step of Market Segmentation Analysis:

Step 1 - Deciding (not) to Segment

1 Implications of Committing to Market Segmentation

Market segmentation is a significant strategic decision for organizations, impacting various aspects of operations and strategy. Here are the key implications:

- **Long-Term Commitment:** Market segmentation requires a long-term commitment akin to a marriage rather than a casual date. This commitment involves substantial changes such as new product development, pricing adjustments, and tailored communications strategies.
- **Cost Considerations:** Implementing segmentation involves costs like research, surveys, focus groups, and redesigning marketing strategies. Cahill emphasizes that the increase in sales from segmentation must justify these expenses.
- **Organizational Structure:** Croft suggests organizing around market segments rather than products to maximize segmentation benefits, necessitating internal structural adjustments.

2 Implementation Barriers

Several barriers can hinder successful market segmentation implementation:

- **Senior Management Engagement:** Lack of leadership, involvement, and resource allocation by senior management can derail segmentation efforts.
- **Organizational Culture:** Issues like resistance to change, lack of market orientation, and poor communication across departments can impede segmentation initiatives.
- **Training and Expertise:** Insufficient understanding of segmentation principles among senior management and teams tasked with segmentation can hinder effective implementation.
- **Resource Constraints:** Lack of financial resources or the ability to make necessary structural changes can pose significant obstacles.
- **Process-Related Challenges:** These include unclear objectives, poor planning, inadequate process structure, and time pressures.

Step 2 - Specifying the Ideal Target Segment

1 Segment Evaluation Criteria

In Step 2 of market segmentation analysis, the organization defines the criteria for evaluating potential market segments. This step is crucial as it shapes subsequent decisions in the segmentation process:

- **User Involvement:** It is emphasized that user input throughout the process is essential, not just at the beginning or end stages. This involvement ensures that segmentation criteria align closely with organizational goals and market realities.
- **Two Sets of Criteria:**
    1. **Knock-Out Criteria:** These are non-negotiable criteria that segments must meet to be considered viable targets. They include characteristics like homogeneity, distinctiveness, size, match with organizational capabilities, identifiability, and reachability.
    2. **Attractiveness Criteria:** These criteria assess the relative appeal of segments that meet the knock-out criteria. They form a detailed checklist used to evaluate each segment's potential attractiveness.

2 Knock-Out Criteria

Knock-out criteria are fundamental requirements that segments must fulfill to proceed to the attractiveness evaluation stage:

- **Homogeneity:** Segment members should be similar to each other.
- **Distinctiveness:** Segments should be clearly different from each other.
- **Size:** The segment should be large enough to justify targeted marketing efforts.
- **Match:** The organization must have the capability to meet the segment's needs effectively.
- **Identifiability:** Members of the segment should be identifiable in the market.
- **Reachability:** There should be practical ways to reach and engage with segment members.

These criteria are set early in the process and are crucial for focusing the segmentation effort effectively.

3 Attractiveness Criteria

Unlike knock-out criteria, attractiveness criteria are varied and nuanced, offering a checklist for evaluating segments:

- **Non-binary Assessment:** Segments are rated based on their attractiveness across multiple criteria, rather than simply meeting or not meeting criteria.
- **Selection Process:** The segmentation team selects from a wide range of attractiveness criteria those most relevant to their organizational context and strategy.

4.4 Implementing a Structured Process

Structured approaches, such as using a segment evaluation plot, are recommended for systematically assessing and comparing segments:

- **Segment Evaluation Plot:** This tool helps visualize segment attractiveness against organizational competitiveness, facilitating informed decision-making.

- **Criteria Negotiation:** Selection of criteria and their weights involves collaborative negotiation within the segmentation team and validation through advisory committee input.

Step 3 – Gathering Data

1 Segmentation Variables

Empirical data is crucial for market segmentation, identifying and describing market segments through characteristics like gender, age, and benefits sought. Commonsense segmentation uses one variable, such as gender, while data-driven segmentation utilizes multiple variables to identify segments. High-quality data from surveys, observation, and experimental studies ensures accurate segment allocation and effective marketing strategies, with the best data closely representing actual consumer behaviours.

2 Segmentation Criteria

Before extracting segments and collecting data, organizations must choose a segmentation criterion, which is broader than a single measured variable and includes the type of information used, such as benefits sought. This decision requires market knowledge and can't be easily outsourced. Common criteria are geographic, socio-demographic, psychographic, and behavioral factors. Relevant consumer differences include profitability, bargaining power, preferences, barriers to choice, and interaction effects. Although many criteria exist, there are no clear guidelines for selecting the best one. Cahill (2006) recommends using the simplest effective approach to target the product or service.

2.1 Geographic Segmentation

Geographic segmentation, based on consumers' location, is one of the oldest and simplest market segmentation techniques. It's often sufficient for purposes like national tourist organizations segmenting tourists by country of residence to accommodate language differences. Global companies like Amazon and IKEA use it to customize product information, offers, and pricing. The advantages include easy assignment of customers to geographical units and effective targeting of communication messages. However, geographical areas may not guarantee shared characteristics relevant to marketers, as preferences can vary within the same region. This method has seen a revival in international market segmentation studies, challenging researchers to find meaningful variables across regions and account for cultural biases.

2.2 Socio-Demographic Segmentation

Socio-demographic segmentation segments the market based on characteristics like age, gender, income, and education. This method is effective in industries where product preferences closely correlate with these characteristics, such as holiday resorts serving families with small children. The advantages include easy determination of segment membership and direct influence of socio-demographic factors on preferences. However, these factors often do not sufficiently explain consumer behavior and may lack market understanding, as values, tastes, and preferences can be more relevant. Yankelovich and

Meer (2006) argue that values and tastes better influence purchase decisions than socio-demographics.

2.3 Psychographic Segmentation

Psychographic segmentation groups people based on psychological criteria like beliefs, interests, preferences, and aspirations. This method includes benefit segmentation, popularized by Haley (1968), and lifestyle segmentation, based on activities, opinions, and interests. Psychographic criteria are more complex than geographic or socio-demographic ones, often requiring multiple variables. While this method reflects underlying reasons for consumer behavior, it also involves complex segment determination and relies heavily on the reliability and validity of psychographic measures.

2.4 Behavioral Segmentation

Behavioral segmentation focuses on grouping consumers based on behaviors, such as prior experience, purchase frequency, amount spent, and information search behavior. This method is superior to geographic variables for segmentation, as it directly reflects consumer actions. However, behavioral data might not always be available, especially for potential customers who haven't previously purchased the product, limiting the focus to existing customers.

3 Data from Survey Studies

Survey data holds significant potential for market segmentation due to its cost-effectiveness and ease of collection. However, potential biases can impact the quality of segmentation outcomes. Fundamental considerations when using survey data include the choice of variables, response options, styles, and sample size.

3.1 Choice of Variables

Variables should be selected based on their relevance to the segmentation criterion to minimize respondent fatigue and avoid introducing noisy variables that can interfere with segmentation algorithms. Conducting exploratory research beforehand helps identify necessary variables and reduces redundancy in survey questions.

3.2 Nature of Data Collected

Survey data can be categorized into different types:

- **Binary**: Yes or no responses, which are straightforward and effective for segmentation.
- **Metric**: Numeric responses (e.g., age, expenditure) suitable for statistical analysis and segmentation.
- **Ordinal**: Common in questionnaires but less ideal for clustering due to undefined distances between points. Binary or metric options are preferred to ensure accuracy in data-driven segmentation. Visual analogue scales, often used in online surveys, offer precise responses but can be treated as metric data.

3.3 Response Styles

Response styles, such as extreme or midpoint answering, can introduce biases that mislead segmentation algorithms. Minimizing these biases during survey construction is crucial to ensure accurate segmentation results.

3.4 Sample Size

Sample size is critical for robust segmentation. Recommendations vary but generally suggest larger samples to effectively derive segments:

- Formann (1984): At least 2 to 5 times the number of segmentation variables.
- Qiu and Joe (2015): At least 10 times the number of variables times the number of segments.
- Dolnicar et al. (2014, 2016): At least 60 to 100 times the number of segmentation variables, depending on complexity.

4 Data from Internal Sources

Organizations increasingly utilize internal data sources such as scanner data in grocery stores, booking data from airline loyalty programs, and online purchase records for market segmentation analysis. These datasets offer strength in reflecting actual consumer behavior, rather than relying on self-reported behaviors prone to memory errors and response biases like social desirability. They are also readily available without additional data collection efforts if stored in accessible formats.

However, internal data can be biased towards existing customers, potentially overlooking insights into future customer segments with different consumption patterns. This limitation underscores the need for organizations to supplement internal data with external sources to ensure comprehensive market understanding and effective segmentation strategies.

5 Data from Experimental Studies

Experimental studies offer valuable data for market segmentation analysis, derived from controlled field or laboratory experiments. These studies can involve testing consumer responses to advertisements or conducting choice experiments and conjoint analyses.

For instance, experiments may evaluate consumer preferences by presenting them with various product attributes and measuring their responses. This data provides insights into how specific attributes influence consumer choices, which can be utilized as segmentation criteria. Experimental data thus enables organizations to understand consumer behavior more deeply and refine segmentation strategies based on empirical findings.

Step 7: Describing Segments

1 Developing a Complete Picture of Market Segments

 Describing Segments: Uses additional information (descriptor variables) about segment members to develop a comprehensive understanding of each market segment.
 Importance: Essential for developing a customized marketing mix and effective targeted marketing.

2 Using Visualizations to Describe Market Segments

- **Key Advantages of Graphical Statistics:**
  - Simplifies interpretation for analysts and users.
  - Integrates statistical significance to avoid over-interpretation of insignificant differences.
  - Preferred by marketing managers for their intuitiveness and efficiency.

2.1 Nominal and Ordinal Descriptor Variables

- **Cross-Tabulation:** Basis for visualizing and testing differences in nominal or ordinal descriptor variables.
- **Stacked Bar Chart:** Shows segment sizes and gender distribution but makes comparing proportions difficult if segment sizes are unequal.
- **Mosaic Plot:** Visualizes cross-tabulations effectively with bars indicating segment sizes and colors showing statistical significance:
  - Dark red (< -4), Light red (< -2), White (-2 to 2), Light blue (> 2), Dark blue (> 4).
  - Allows easy comparison of proportions across segments.

Key Insights:

- **Gender Distribution:** No significant differences across segments.
- **Income and Segments:** Segment 4 earns more; Segment 3 has fewer low-income members; Segment 6 has fewer high-income members.
- **Moral Obligation:** Segment 3 has low environmental moral obligation; Segment 6 has high moral obligation.

2.2 Metric Descriptor Variables - Key Points

- **Conditional Plots:** Useful for visualizing differences in metric descriptor variables between market segments.
- **Histograms:** Show distributions but can be hard to interpret for differences between segments.
- **Parallel Box-and-Whisker Plot:** Provides clearer insights into distributions and differences across segments.

- **Key Insights:**
  - **Age:** Segment 5 has a lower median age, Segment 6 has a higher median age.
  - **Moral Obligation:** Segment 6 shows high environmental moral obligation; Segment 3 shows low.
- **SLSA Plot:** Visualizes segment stability and includes metric descriptor information with color coding.

3 Testing for Segment Differences in Descriptor Variables

1 Chi-squared Test (χ^2-test):

- Tests association between nominal segment membership and other nominal or ordinal variables.

2 Analysis of Variance (ANOVA):

- Tests for significant differences in means of metric descriptor variables across segments.

3 Post-hoc Tests (Pairwise Comparisons):

- Conducted after ANOVA to identify which segments differ significantly.

4 Tukey's Honest Significant Differences (HSD):

- Visualizes pairwise comparisons from ANOVA, showing confidence intervals for mean differences between segments.

4 Predicting Segments from Descriptor Variables

Regression Models for Prediction:

- Use regression models with segment membership as the categorical dependent variable.
- Independent variables are descriptor variables used to predict segment membership.

Linear Regression Model:

- Assumes a linear relationship between dependent variable (e.g., Age) and independent variables (e.g., segment membership).

Interpreting Coefficients:

- Regression coefficients represent the mean difference in the dependent variable for each segment compared to a reference segment or intercept.

Generalized Linear Models (GLMs):

- Extend regression models to handle a wider range of dependent variable distributions (beyond normal).
- Uses a link function to transform the dependent variable's mean value to an unlimited range suitable for modeling.

GLMs for Classification:

- Can use binomial or multinomial distributions for dependent variables suitable for classification tasks.
- Example distributions include Bernoulli (binary) or multinomial (multiple categories).

Logistic Regression:

- Special case of GLM used when the dependent variable is binary or multinomial, applying the logit function as the link function.

4.1 Binary Logistic Regression

- Binary logistic regression is the use of the logit link function to model the relationship between predictor variables and the probability of an event (in this case, segment membership). The logit function $g(\mu) = \eta = \log(\mu / 1 - \mu)$.transforms the success probability μ into a linear predictor η, which spans the entire real number line.
- This transformation is crucial because it allows the linear regression model to handle binary outcomes effectively by mapping probabilities (which are constrained between 0 and 1) onto a scale where linear relationships can be modeled. This makes interpretation of coefficients straightforward as they represent changes in the log odds of the event (e.g., belonging to a specific segment) with respect to changes in the predictor variables.
- Understanding the logit link function and its application in logistic regression is fundamental for interpreting how changes in independent variables influence the probability of categorical outcomes, such as membership in market segments.

4.2 Multinomial Logistic Regression

multinomial logistic regression is applied to predict categorical outcomes with more than two levels, such as market segments. Using multinom() from the nnet package in R, the model specifies predictors (C6 ~ Age + Oblig2) and estimates coefficients that indicate log odds relative to a baseline category. Model evaluation includes significance tests like likelihood ratio chi-square (Anova()) and model selection (step()), while predict() assesses predictive performance with probabilities and classes. Visualizations with allEffects() illustrate how variables like Age and Oblig2 influence segment membership probabilities across categories.

4.3 Tree-Based Methods

Tree-based methods such as Classification and Regression Trees (CART) are explored for predicting categorical outcomes using independent variables. These methods offer advantages like automatic variable selection, easy interpretation through visualizations, and

handling of interaction effects. However, they can be sensitive to minor data changes, leading to different tree structures.

The tree-building process involves recursive partitioning, where consumers are split into homogeneous groups based on independent variables to predict segment membership. R packages like rpart and partykit implement different algorithms for constructing these trees. For instance, ctree() from partykit fits conditional inference trees, which use statistical tests for variable selection.

Trees are represented with nodes (e.g., root, inner, terminal) that split data based on criteria like environmental behavior or moral obligation scores. Visualizing these trees with plot() provides insights into segment membership predictions for different consumer groups, illustrating how variables influence membership across segments.