Market Segmentation Analysis

Amborish Sen

July 09, 2025

Step 1: Deciding (not) to Segment

Key Implications of Market Segmentation

- Market segmentation requires **long-term organizational commitment** - it's "a marriage, not a date"
- Organizations must be willing and able to make **substantial changes and investments**
- Costs include research, surveys, focus groups, multiple product packages, and varied communication messages
- Expected increase in sales must justify the implementation costs
- May require developing new products, modifying existing ones, changing pricing/distribution, and restructuring internal organization

Implementation Barriers to Consider

Senior Management Barriers:

- Lack of leadership commitment and involvement
- Insufficient resource allocation for analysis and long-term implementation
- Absence of pro-active championing from executives

Organizational Culture Barriers:

- Lack of market/consumer orientation
- Resistance to change and new ideas
- Poor communication between organizational units
- Short-term thinking and unwillingness to adapt
- Office politics and lack of information sharing

Capability Barriers:

- Lack of training on market segmentation fundamentals
- Absence of qualified marketing experts
- Missing data management and analysis expertise
- Insufficient financial resources
- Inability to make required structural changes

Process-Related Barriers:

- Unclear objectives for segmentation exercise
- Lack of structured planning processes
- Poor responsibility allocation
- Time pressure hindering optimal outcomes

Step 2: Specifying the Ideal Target Segment

Purpose and Importance

- This step requires **significant user input** and cannot be limited to just briefing or final marketing mix development
- The organization must determine **two sets of segment evaluation criteria** that will guide data collection and target segment selection
- This conceptual contribution is critical for producing useful market segmentation results

Two Types of Segment Evaluation Criteria

Knock-Out Criteria (Essential, Non-Negotiable)

These are **binary criteria** that automatically eliminate segments that don't comply:

- **Homogeneous**: Members within the segment must be similar to each other
- **Distinct**: Members must be clearly different from other segments
- **Large enough**: Must contain sufficient consumers to justify customized marketing mix costs
- **Matching strengths**: Organization must have capability to satisfy segment needs
- **Identifiable**: Must be possible to spot segment members in the marketplace
- **Reachable**: Must be able to contact segment members to deliver customized marketing mix

Attractiveness Criteria (Relative Assessment)

These are **graduated criteria** used to evaluate relative attractiveness of remaining segments:

- Not binary - segments can be more or less attractive on each criterion
- Organizations select from a wide range of possible criteria
- Examples include: growth rate, profitability, competitive advantage, market potential, barriers to entry, technological factors, etc.

Implementing a Structured Process

Team-Based Approach:

- Core segmentation team (2-3 people) proposes initial criteria
- Advisory committee (representing all organizational units) discusses and modifies
- Multiple perspectives ensure comprehensive evaluation and organizational buy-in

Segment Evaluation Plot Method:

- Popular structured approach using two axes: segment attractiveness vs. organizational competitiveness
- Cannot be completed in Step 2 (no segments exist yet), but criteria selection enables proper data collection
- Provides framework for systematic segment evaluation in Step 8

Weighting Process:

- Select no more than **6 attractiveness criteria**
- Each team member distributes **100 points** across criteria based on relative importance
- Negotiate until agreement is reached on weightings
- Seek advisory committee approval for final criteria and weights

Step 3: Collecting Data

Based on the document, here's a comprehensive summary of the data collection step in market segmentation analysis:

3.1 Segmentation Variables - Foundation of Data Collection

Core Principles

- **Empirical data** serves as the foundation for both common sense and data-driven market segmentation
- **Data quality** directly determines the success of segmentation analysis
- **Multiple data sources** should be explored to find the most behaviourally accurate information

Two Approaches to Variable Usage

- **Commonsense segmentation**: Uses **one specific characteristic** as the segmentation variable
- **Data-driven segmentation**: Uses **multiple variables simultaneously** for more sophisticated segment identification

Primary Data Sources

- **Survey studies** (most common)
- **Observational data** (scanner data, loyalty program data)
- **Experimental studies** (controlled research environments)

3.2 Segmentation Criteria - Types of Data to Collect

Key Terminology

- **Segmentation variable**: Single measured value (e.g., survey item, expenditure category)
- **Segmentation criterion**: Nature of information used (e.g., benefits sought, lifestyle factors)

Four Main Segmentation Criteria

a. Geographic Segmentation

- **Basis**: Consumer's residential location
- **Characteristics**: Earliest and simplest segmentation approach
- **Application**: Single location-based variable splits market into regional segments

b. Sociodemographic Segmentation

- **Common variables**: Age, gender, income, education
- **Industry value**: Particularly valuable in certain industries
- **Insight potential**: Can provide specific product preference insights
- **Advantage**: Easy to collect and understand

c. Psychographic Segmentation

- **Basis**: Psychological factors (beliefs, interests, preferences, aspirations, desired benefits)
- **Complexity**: More intricate than geographic or sociodemographic approaches
- **Requirements**: Often needs **multiple variables** to capture diverse psychological dimensions
- **Depth**: Provides deeper consumer insights

d. Behavioural Segmentation

- **Variables**: Prior product experience, purchase frequency, spending amounts
- **Key advantage**: Uses **actual behavior** rather than stated or intended behavior
- **Reliability**: Most accurate basis for segment extraction
- **Examples**: Transaction history, usage patterns, brand loyalty

3.3 Data from Survey Studies

Why Surveys Dominate

- **Cost-effective** and **easy to collect**
- **Accessible** to any organization regardless of size
- **Flexible** in terms of information that can be gathered

Critical Survey Design Elements

a. Choice of Variables

- **Focus**: Create focused questionnaires with essential and unique questions
- **Avoid redundancy**: Eliminate duplicate or overlapping questions
- **Research enhancement**: Supplement with exploratory or qualitative research insights
- **Question quality**: Ensure each question serves a specific analytical purpose

b. Response Options Design

- **Preferred formats**: Use **metric or binary response options** when meaningful
- **Analytical compatibility**: Helps avoid complications in data-driven segmentation analysis
- **Distance measures**: Appropriate scaling prevents analytical problems
- **Consistency**: Maintain consistent response formats across related questions

c. Response Styles and Bias Management

- **Response bias definition**: Systematic tendency to respond based on factors other than item content
- **Impact on segmentation**: Algorithms cannot distinguish genuine beliefs from style-influenced responses
- **Common biases**: Social desirability bias, acquiescence bias, extreme response styles
- **Mitigation strategies**: Careful question design and response format selection

d. Sample Size Requirements

- **Minimum standard**: At least **100 times the number of segmentation variables**
- **Optimal data characteristics**:
    - Contains all necessary and no unnecessary items
    - Free of highly correlated items
    - High-quality responses
    - Binary or metric formats
    - Free of response style contamination
    - Suitable sample based on segmentation study goals
    - Adequate sample size for statistical reliability

3.4 Data from Internal Sources

Growing Importance

- Organizations increasingly possess **extensive internal data**
- **Behavioural accuracy**: Represents actual consumer behavior
- **Cost efficiency**: Data already exists within the organization

Key Advantages Over Survey Data

- **Eliminates memory issues**: No reliance on imperfect consumer recall
- **Reduces bias**: Avoids response biases like social desirability
- **Behavioural truth**: Reflects what consumers actually do versus what they say
- **Longitudinal capability**: Can track behavior over time

Types of Internal Data

- **Transaction records**: Purchase history, spending patterns
- **Customer interactions**: Service calls, website behavior, app usage
- **Loyalty program data**: Detailed purchase and preference information
- **CRM data**: Customer relationship and interaction history

3.5 Data from Experimental Studies

Sources of Experimental Data

- **Field experiments**: Real-world controlled studies
- **Laboratory experiments**: Controlled laboratory environments
- **Choice experiments**: Consumers choose among product alternatives
- **Conjoint analyses**: Evaluation of products with various attribute combinations

Experimental Design Features

- **Well-crafted stimuli**: Featuring specific levels of product attributes
- **Controlled conditions**: Eliminate external variables
- **Preference measurement**: Consumers express preferences among product combinations
- **Attribute testing**: Systematic testing of various attribute level combinations

Step 5: Extracting Segments

5.1 Nature of Market Segmentation

- Market segmentation is **exploratory** and heavily depends on the **structure of data** and the **algorithm used**.
- Consumer data is often **unstructured**, and clear segment boundaries usually do not exist visually.

5.2 Impact of Algorithms

- **Segmentation results depend on the algorithm** (e.g., k-means vs single linkage).
- Algorithms impose different **structural assumptions** on the segments.
- Example: Spiral data was handled well by **single linkage** but not by **k-means**.

5.3 Algorithm Selection Factors

- Must consider:
  - **Data characteristics**: size, variable types (nominal, ordinal, metric).
  - **Desired segment features**: similarity within, differences between.
- No one-size-fits-all algorithm; each has **strengths and weaknesses**.

5.4 Types of Methods

Distance-Based Methods

- Based on **similarity/dissimilarity** between data points.
- Examples of distance measures:
  - **Euclidean** (straight-line distance)
  - **Manhattan** (grid-based city block distance)
  - **Asymmetric Binary** (only shared 1s matter)

Model-Based Methods

- Use **probabilistic models** to define segments.
- Not detailed in this excerpt, but referenced for completeness.

5.5 Hierarchical Clustering

- Builds nested groupings:
  - **Agglomerative**: starts with individuals, merges into clusters.
  - **Divisive**: starts with one group, splits recursively.
- Linkage methods:
  - **Single**: closest points between groups.
  - **Complete**: furthest points.
  - **Average**: mean distances.
- **Dendrograms**: visual tree structure showing cluster merges.

5.6 Partitioning Methods

k-Means Clustering

- Most common; assigns points to the **closest centroid**.
- Requires specifying the number of clusters (**k**).
- Sensitive to:
  - Initial centroids
  - Choice of distance measure (Euclidean, Manhattan, etc.)

k-Centroid Clustering

- Generalization of k-means using other distances.

5.7 Advanced Partitioning Techniques

- **Improved k-means**: better initialization to avoid local minima.
- **Hard Competitive Learning**: updates centroids based on one sample at a time.
- **Neural Gas**: updates multiple centroids to varying degrees.
- **Topology Representing Networks (TRN)**: visual neighbourhood graphs.
- **Self-Organizing Maps (SOM)**:
  - Positions centroids on a grid.
  - Visual interpretation of segment similarity.
- **Auto-Encoding Neural Networks**:
  - Encode data via a hidden layer.
  - Used for **non-linear dimensionality reduction** and **clustering**.

5.8 Visualization and Interpretation

- Tools like **bar charts** and **scree plots** help evaluate clusters.
- **Scree plot**: shows within-cluster distance vs. number of segments to find the "elbow point".
- **Cluster means plots**: compare segment behavior across variables.

5.9 Code implementation

Google Colab: - Market Segmentation Analysis Step-5