

MARKET SEGMENTATION ANALYSIS

Abhishek RR

STEP 1: Deciding (not) to Segment

- Market segmentation needs an investment of long-term, that involves major alteration in organizations and long sets of resources.
- Segmentation requires research cost, market survey, and specialized marketing material and must be undertaken only when it promises greater profitability
- The strategy can be a change in product or price or distribution or communication and it alters the organizational structure
- The strategic decision on segmentation must be strategic in nature and should be supported at every level of the organization, by the top executives.
- Effective segmentation requires commitment and resource allocation from top management, besides an organizational culture that can accommodate segmentation.
- Experts in marketing as well as analytics along with relevant resources should be present since these can seriously constrain the success of segmentation
- Clear objectives, process-centric approaches and graphical presentation of findings will aid understanding and implementation
- Barrier identification and overcoming must be done early on in the process, if that is not feasible, it is probably time to relook at segmentation.
- Concentration, Perseverance and flexibility will ensure successful implementation of segmentation.

Step 2: Specifying the Ideal Target Segment

- Good market segmentation requires a continuous flow of input from users at all steps in the process and not just as a process of initial briefings and final strategy development.
- The conceptual input at the start by the organization influences many of the most essential parts of the process that consequently influence data gathering and selection of the segments.
- There are two forms of criteria: knock-out criteria-a set of features a segment must have-and attractiveness criteria, to be used in order to evaluate the attractiveness of the segments that pass the knock-out standards.
- The knock-out criteria screen through the segments to qualify for further evaluation on criteria such as homogeneity, uniqueness, large enough, aligned with the company's strengths, identifiable, and accessible.
- Certain knock-out criteria are elaborated further-for instance, in the case of segment size by defining a minimum viable size.
- The attractiveness criteria are graded so that a given segment can be ranked on its attractiveness relative to other segments, rather than accepting-or-rejecting criteria, and which will guide the selection of target segment.
- A methodical approach such as segment evaluation plot makes assessment more effective, presenting the attraction of the segment in contrast with the level of organizational competitiveness.
- A collective approach to deciding critical criteria helps to provide a balanced view as representatives of various organizational units are consulted.
- Even if segments have not yet been defined, making attractiveness criteria early on will streamline data collection and simplify segment selection.
- At the end, the team should have approximately six essential attractiveness criteria that are weighted by importance, normally agreed on by distributing points across the criteria for comprehensive consideration and advisory committee to ensure organizational views are aligned.

Step 3: Collecting Data

- Empirical evidence is the basis of a commonsense and data-driven segmentation.
- Characteristics for consumer division are segmentation variables whereas detailed insights for segment tailoring are descriptor variables
- A single characteristic is applied in commonsense segmentation, whereas the data-driven segmentation uses several characteristics for having a more detailed understanding. Quality data must be used for proper assignment of segments.
- A segmentation criterion is a broad information category, say for example benefits sought, required for any segmentation that demands market knowledge incapable of being outsourced.
- Common criteria for segmentation are often geographic, socio-demographic, psychographic, and behavioral ones, and the easier the approach, the better for practical reasons.
- Geographic segmentation Emphasizes the location of consumers. For it, it is easy to name and communicate segments. It fails to capture so well the preferences, mostly internationally, since cultural nuances should also be regarded.
- Socio-demographic segmentation rests on variables like age, income, education, and household size as grounds for segmentation of the consumer groups. That way, it allows marketers to design their strategy effectively. Psychographic segmentation measures consumers' lifestyles, values, and personalities. This kind of segmentation seeks to study consumer's motivation and preference more profoundly.
- Behavioral segmentation involves grouping consumers according to their behaviors with the product, including usage and purchasing habits. It helps set marketing campaign targets according to consumer behavior.

(until Geographic Segmentation)

Step 4 : Exploring Data

- Exploratory Data Analysis (EDA) is necessary to be performed in a data set after the data is collected. It allows the determination of good algorithms for

market segmentation by describing measurement levels, univariate distributions, and dependency structures among variables.

- Before one analyzes data, it's really important to clean it by checking the accuracy of values recorded and ensuring consistent labels for categories. That is, you need to set plausible value ranges for metric variables such as age to catch errors made in collection time.
- Categorical variables need to be checked for allowed values. A gender variable should only consist of valid choices like "male" and "female." Any errors must be edited during the data cleaning process.
- Whenever categorical variables, like income, are read into R, their levels may be read in alphabetical order. Reality could be different than that. One has the option of re-ordering the levels to make them ordinal variables so proper analysis can be done.
- This process makes it simpler to reproduce the analysis in later sessions or when new data is added. All of the data cleaning and transformation steps have to be done in code rather than through by-hand spreadsheet manipulation.
- Learning about the data is a process where you would use both numeric summarization, such as range, quartiles, and mean, and graphical methods, which include histograms, boxplots, and scatter plots; these helps give an idea about the distribution and what characteristics the data will possess.
- Histograms are helpful for data with numeric variables. They will illustrate the number of times observations fall in specific bins of values. An important decision in the binning process is the width of the bins, as it will affect the visualization and/or interpretation of data.
- For categorical variables, there are two common pre-processing that can be used: merge similar categories and convert them into numeric forms if meaningful. The conversion could make distance-based clustering methods applicable.
- In doing a segmentation analysis one has to make the standardization of numeric variables in order to rectify the effect that variables have with one another.

- Standardization aligns the variables to a common scale by subtracting the mean and dividing it by the standard deviation, hence creating an equal contribution from each variable toward the analysis.
- PCA is a transformation technique that converts multivariate datasets into ordered uncorrelated variables, commonly known as principal components. This is designed to retain the relative positions of observations and allows for dimension reduction, which makes it easier for further simplification of complex relationships to be visualized.
- The major parts of the variances will be held by the first principal components resulting from PCA, allowing the analyst to focus solely on those components to plot and analyze the data. This simplification is particularly useful for understanding high-dimensional data structures.

Step 5: Extracting Segments

- Data-oriented market segmentation analysis is inductive where, most of the time, it is simply built from unstructured consumer datasets and hence, preferences are distributed rather than being sharply defined.
- The kind of segmentation technique adopted has a significant influence on the outcome, and it is important to discuss different clustering techniques since various algorithms differ in their manner of enforcing different structures on segmentation solutions.
- Distances-based measures are of importance in grouping consumers who exhibit similar behaviors such as tourist activity patterns that have an impact on the effectiveness of the segmentation.
- Auto-coding neural networks use one hidden layer perceptron in cluster analysis in which the input nodes are the representation of segmentation variables and the values of the hidden layer are restricted to add up to 1.
- Hybrid techniques use hierarchical and partitioning algorithms. The two techniques have advantages in use; in hierarchical clustering, the number of segments need not be prepecified but it does demand much memory, while

partitioning methods are efficient but require specifying a number of segments beforehand.

- The two-step clustering in IBM SPSS runs a partitioning procedure followed by hierarchical analysis which would help segment mobile phone users although the k-means algorithm in R will extract many clusters so that data size can be reduced down to just the centroids for hierarchical analysis.
- Bagged clustering brings together the potential both of the hierarchical and partitioning methods with bootstrapping to enhance segmentation robustness by creating the bootstrapped datasets that reduce dependence on specific consumer data.
- In fact, model-based methods point to the existence of market segments characterized either by similarity or by difference, and finite mixture models allow for exploration of the segments on the basis of empirical evidence.
- Simple model-based clustering provides a fit for the empirical distribution of consumer characteristics by trying to find a distribution for all the given variables, whereas finite mixture models take into account the size of the segments as well as the size of the probabilities that add up to one.
- For binary data, finite mixtures of binary distributions (latent class models) reflect different probabilities for some activities to determine specific preferences of the consumers over segments.
- Hierarchical clustering methods exhibit intuitive grouping, namely from one segment to one segment per consumer, but for larger datasets, the more efficient partitioning methods are, such as k-means clustering.
K-means clustering reduces intra-cluster distance and maximizes inter-cluster distance; initialization with good starting points has the potential of increasing the quality of the outcome and reducing the chance of local optima.
- All clustering algorithms assume that all variables are of equal importance to the analysis; the use of pre-treatment is also there to detect and filter out redundant or noisy variables.
- Bi clustering algorithms pursue simultaneous clustering of customers and variables, achieving important applications in domains of genetic and proteomic data analysis with irrelevant genes.

- Brusco (2004) reported the variable selection procedure for clustering of binary data, VSBD that enhances the quality of the resulting clustering solution by the identification and removal of the irrelevant masking variables.
- Internal and external cluster indices guide segment extraction in market segmentation analysis where stability is established through corresponding segments obtained at repeated analyses.
- Gorge plots demonstrate the distances between consumers and representatives of segment to estimate the goodness and effectiveness of segmentation
- Segment Level Stability (SLSW) Analysis: It is a procedure developed by Dolnicar and Leisch in 2017, which ensures there are bootstrapping samples as well as repeated calculations therefore identifying segments which are stable.