MARKET SEGMENTATION ANALYSIS

TEAM – HIMANSHI

MEMBERS – 1. HIMANSHI GARG (TEAM LEADER)

2. ARYAN KUMAR

3. P SADHASIVAM

4. KUSHAGRA BHATNAGAR

The purpose of marketing is to match the needs and desires of consumers with the offers of suppliers particularly suited to satisfy those needs and desires. A marketing plan consists of two components: a strategic and a tactical marketing plan. Strategic long-term direction of an organisation, but does not provide much detail on short Term marketing action required to move in this long-term direction. The tactical marketing plan contains instructions on what needs to be done to get there.

What is market segmentation?

- Market segmentation is a decision-making tool for the marketing manager in the crucial task of selecting a target market for a given product and designing an appropriate marketing mix
- In simple terms market segmentation means cutting markets into slices
- Consumers belonging to the same market segments are very similar to one another with respect to the consumer characteristics deemed critical by management
- And consumers belonging to different market segments are very different from one another with respect to those consumer characteristics
- Consumer characteristics deemed critical to market segmentation by management are referred to as segmentation criteria
- Concentrating entirely on satisfying the needs of one market segment is known as a concentrated strategy and it is attractive for organisations who are resource-poor, but are facing fierce competition in the market. But it is risky as it is depending on one single market segment entirely
- If all aspects of the marketing mix would have to be customised for each of the target segments such a strategy is known as Differentiated strategy
- When an organisation decides not to use market segmentation, it chooses to pursue an undifferentiated market strategy, where the same product is marketed using the same marketing mix to the entire market

Benefits of Market segmentation

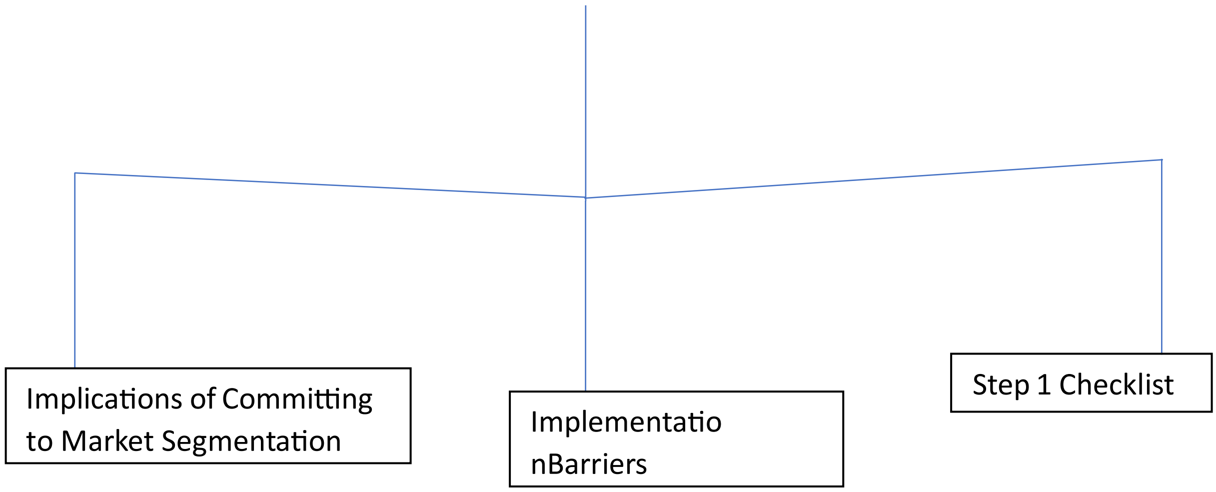
- Makes organisations to reflect on what they are particularly good at compared to competitors, and make an effort to gain insights into what consumers want
- Market segmentation also leads to tangible benefits including a better understanding of differences between consumers, which improves the match of organisational strengths and consumer needs
- Extreme market segmentation is able to offer a customised product or service to very small groups of consumers. This approach is referred to as micro marketing or hyper-segmentation
- At an organisational level, market segmentation can contribute to team building If this is achieved successfully, it can also improve communication and information sharing across organisational units.

Approaches to Market segment analysis – Based on choice of segmentation variables

- A more technical way of systematising segmentation approaches is to use as a basis the nature of consumer characteristics used to extract market segments
- When one single segmentation variable is used, the segmentation approach is referred to as a priori, convenience-group or commonsense market segmentation
- When commonsense segmentation is conducted, the provider of the product usually has a reasonably good idea of the nature of the appropriate segment or segments to target
- Commonsense segmentation results from splitting consumers up into groups using one segmentation variable first. Then, one of the resulting segments is selected and split up further using a second segmentation variable
- The aim of the segmentation analysis is not to identify the key defining characteristic of the segment, but to gain deeper insight into the nature of the segments
- Another approach which exploits multiple segmentation variables, is referred to as a posteriori, cluster based or post hoc segmentation
- These terms indicate that the nature of the resulting market segments is not known until after the data analysis has been conducted also know as Data driven segmentation
- The aim of data-driven segmentation is first to explore different market segments that can be extracted using the segmentation variables chosen, and second to develop a detailed profile and description of the segments selected for targeting. data-driven/data-driven segmentation is the result of combining two sets of segmentation variables

Ten steps in Market Segmentation Analysis

STEP 1 – DECIDING(NOT) TO SEGMENT



Implications of Committing to Market Segmentation

Although market segmentation has developed to be a key marketing strategy applied in many organizations, it is not always the best decision to pursue such a strategy. Before investing time and resources in a market segmentation analysis, it is important to understand the implicationsof pursuing a market segmentation strategy.

Cahill recommends not to segment unless the expected increase in sales is sufficient to justify implementing a segmentation strategy, stating that One of the truisms of segmentation strategyis that using the scheme must be more profitable than marketing without it, net of the expenseof developing and using the scheme itself.

Implementation Barriers

The first group of barriers relates to senior management. Lack of leadership, pro-active championing, commitment, and involvement in the market segmentation process by senior leadership undermines the success of market segmentation.

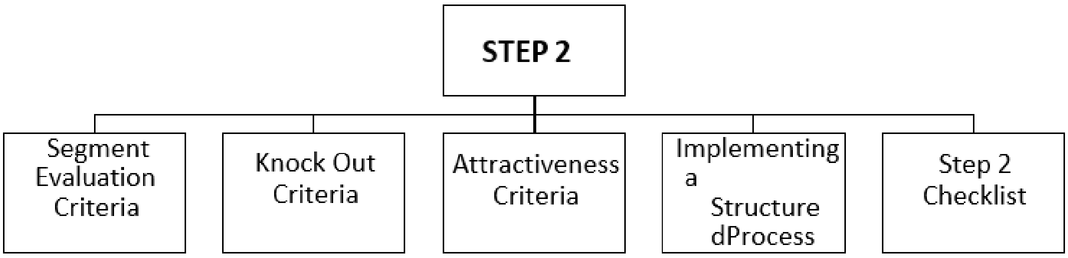
A second group of barriers relates to organizational culture. Lack of market or consumer orientation, resistance to change and new ideas, lack of creative thinking, bad communication,and lack of sharing of information and insights across organizational units, short-term thinking,unwillingness to make changes and office politics have been identified as preventing the successful implementation of market segmentation.

Another potential problem is lack of training. If senior management and the team tasked with segmentation do not understand the very foundations of market segmentation, or if they are unaware of the consequences of pursuing such a strategy, the attempt of introducing market segmentation is likely to fail.

Another obstacle may be objective restrictions faced by the organization, including lack of financial resources, or the inability to make the structural changes required.

If barriers cannot be removed, the option of abandoning the attempt of exploring market segmentation as a potential future strategy should be seriously considered.

STEP 2 – SPECIFYING THE IDEAL TARGET SEGMENT



- Market segmentation depends on user input.
- User inputs cannot be limited to the beginning or end or development of the process.
- Users have to be involved in all stages.
- Orgs. must determine two set of segment evaluation criteria.
- One set is called *knock-out* criteria, these are essential, non-negotiable features that are to be considered.
- Other is called *attractiveness* criteria, these are used to relative attractiveness of remaining market segmentation.
- Different books give different sets of criteria that are given in table 4.1 of book (pg. 32)
- Knock out criteria:
    - The segments must be homogeneous
    - Must be distinct.
    - Must be large enough, must have enough consumers to make it worthwhile to spend money.
    - Must match the strengths of the organisation; the segment members must be satisfied too.
    - Must be reachable.
- Attractiveness criteria:
- Each segment is rated differently.
- Can be more or less attractive with respect to specific criteria.
- Attractiveness across all criterion decides if segment is selected as target or not.
- Structured process is beneficial.
- Segment attractiveness VS Organisational competitiveness plot is used to select target segments.
- These values are determined by segmentation teams.
- While the plot cannot be completed in this stage because there is no segment, The attractiveness criteria should be selected in this step.
- This will help in knowing precisely what it is about market segments that matter to the organisation, ensuring all info is captured whilst collecting data in next step.
- At the end, the segmentation team should have a list of approx. six segment attractiveness criteria.
- Each criteria should have a weight marking its importance.

STEP 3 – COLLECTING THE DATA

Segmentation Variables

- Empirical data forms the basis of both commonsense and data-driven market segmentation. Empirical data is used to identify or create market segments and – later in the process – describe these segments in detail
- If the commonsense segmentation uses gender as the segmentation variable, Market segments are created by simply splitting the sample using this segmentation variable into a segment of women and a segment of men
- The other variables which describe the segment in detail are known as descriptor variables
- When commonsense segments are extracted – even if the nature of the segments is known in advance – data quality is critical to both assigning each person in the sample to the correct market segment, and being able to correctly describe the segments
- Similarly for data-driven market segmentation where data quality determines the quality of the extracted data-driven market segments, and the quality of the descriptions of the resulting segments. Good market segmentation analysis requires good empirical data
- Surveys should not be seen as the default source of data for market segmentation studies. A range of possible sources should be explored, the source that delivers data most closely reflecting actual consumer behaviour is preferable

Segmentation Criteria

- The term segmentation criterion relates to the nature of the information used for market segmentation. It can also relate to one specific construct, such as benefits sought
- The most common segmentation criteria are geographic, socio demographic, psychographic and behavioural

Geographic Segmentation

- Geographic segmentation is using the consumer's location of residence for the criterion to form market segments
- The key advantage of geographic segmentation is that each consumer can easily be assigned to a geographic unit. And so it is easy to target communication messages, and select communication channels (such as local newspapers, local radio and TV stations) to reach the selected geographic segments
- The key disadvantage is that living in the same country or area does not necessarily mean that people share other characteristics relevant to marketers, such as benefits they seek when purchasing a product

Socio Demographic Segmentation

- Socio-demographic segmentation criteria include age, gender, income and education
- Useful in industries like luxury goods (associated with high income), cosmetics (associated with gender; even in times where men are targeted, the female and male segments are treated distinctly differently), baby products (associated with gender), retirement villages (associated with age), tourism resort products (associated with having small children or not)

- Socio-demographic segmentation criteria have the advantage that segment membership can easily be determined for every consumer.

Psychographic Segmentation

- People grouped according to psychological criteria, such as their beliefs, interests, preferences, aspirations, or benefits sought when purchasing a product, such a thing is called as psychographic segmentation
- Psychographic criteria are, by nature, more complex than geographic or socio Demographic criteria because it is difficult to find a single characteristic of a person that will provide insight into the psychographic dimension of interest
- The psychographic approach has the advantage that it is generally more reflective of the underlying reasons for differences in consumer behaviour
- The disadvantage of the psychographic approach is the increased complexity of determining segment memberships for consumers

Behavioural Segmentation

- Another approach to segment extraction is to search directly for similarities in behaviour or reported behaviour. A wide range of possible behaviours can be used for this purpose, including prior experience with the product, frequency of purchase, amount spent on purchasing the product on each occasion (or across multiple purchase occasions), and information search behaviour
- The key advantage of behavioural approaches is that – if based on actual behaviour rather than stated behaviour or stated intended behaviour – the very behaviour of interest is used as the basis of segment extraction
- Behavioural data is not always readily available, especially if the aim is to include in the segmentation analysis potential customers who have not previously purchased the product, rather than limiting oneself to the study of existing customers of the organisation

Data from Survey Studies

Choice of variables

- In data-driven segmentation, all variables relevant to the construct captured by the segmentation criterion needs to be included. At the same time, unnecessary variables must be avoided
- Unnecessary variables included as segmentation variables divert the attention of the segment extraction algorithm away from information critical to the extraction of optimal market segments. Such variables are referred to as noisy variables or masking variables
- Noisy variables do not contribute any information necessary for the identification of the correct market segments, their presence makes it more difficult for the algorithm to extract the correct solution
- Noisy variables can result from not carefully developing survey questions, or from not carefully selecting segmentation variables from among the available survey items
- The problem created by noisy variables can be avoided at the data collection and the variable selection stage of market segmentation analysis

Response Options

- Answer options provided to respondents in surveys determine the scale of the data available for subsequent analyses
- Options allowing respondents to answer in only one of two ways, generate binary or dichotomous data. Such responses can be represented in a data set by 0s and 1s.
- Options allowing respondents to select an answer from a range of unordered categories correspond to nominal variables
- Options allowing respondents to indicate a number, such as age or nights stayed at a hotel, generate metric data
- The most commonly used response option in survey research is a limited number of ordered answer options larger than two, this answer format generates ordinal data, meaning that the options are ordered
- Using binary or metric response options prevents subsequent complications relating to the distance measure in the process of data-driven segmentation analysis

Response Styles

- Response styles manifest in survey answers, including respondents' tendencies to use extreme answer options (STRONGLY AGREE, STRONGLY DISAGREE), to use the midpoint (NEITHER AGREE NOR DISAGREE), and to agree with all statements
- Response styles affect segmentation results because commonly used segment extraction algorithms cannot differentiate between a data entry reflecting the respondent's belief from a data entry reflecting both a respondent's belief and a response style

Sample Size

- Viennese psychologist Formann (1984) recommends that the sample size should be at least 2p (better five times 2p), where is the number of segmentation variables. This P rule of thumb relates to the specific purpose of goodness-of-fit testing in the context of latent class analysis when using binary variables
- According to Qiu and Joe(2015), the sample size should – in the simple case of equal cluster sizes – be at least ten times the number of segmentation variables times the number of segments in the data (10 ・ p ・ k where p represents the number of segmentation variables and k represents the number of segments)
- Increasing the sample size improves the correctness of the extracted segments. But the biggest improvement is achieved by increasing very small samples
- De Craen et al. (2006) show that the presence of unequally sized segments makes it more difficult for an algorithm to extract the correct market segments. Steinley (2003) shows the same for the case of overlapping segments
- Larger sample sizes always improve an algorithm's ability to identify the correct market segmentation solution
- If the variables are not correlated at all, the algorithm has no difficulty extracting the correct segments. And if variables are highly correlated, the task becomes so difficult for the algorithm, that even increasing the sample size dramatically does not help
- The recommendation by Dolnicar et al. (2016) is to ensure the data contains at least 100 respondents for each segmentation variable
- Data used in market segmentation analyses should
  ✓ contain all necessary items;
  ✓ contain no unnecessary items;

✓ contain no correlated items;
✓ contain high-quality responses;
✓ be binary or metric;
✓ be free of response styles;
✓ include responses from a suitable sample given the aim of the segmentation study;
✓ include a sufficient sample size given the number of segmentation variables (100 times the number of segmentation variables)

Data from internal sources

- Advantage of using internal data is that such data are usually automatically generated and – if organisations are capable of storing data in a format that makes them easy to access– no extra effort is required to collect data
- The danger of using internal data is that it may be systematically biased by overrepresenting existing customers

STEP 4 – EXPLORING DATA

A First Glimpse at the Data

- After data collection, exploratory data analysis cleans and – if necessary – preprocesses the data. This exploration stage also offers guidance on the most suitable algorithm for extracting meaningful market segments.
- Data exploration helps to (1) identify the measurement levels of the variables; (2) investigate the univariate distributions of each of the variables; and (3) assess dependency structures between variables.
- Exploring any dataset includes – viewing all the features(columns) ,data types of columns, dimension of dataset, statistical measurements like mean , median , mode , max ,min values, no. of missing values etc.

Data Cleaning

- The first step before commencing data analysis is to clean the data. This includes checking if all values have been recorded correctly, and if consistent labels for the levels of categorical variables have been used.
- Any other values are not permissible, and need to be corrected as part of the data cleaning procedure.
- Data cleaning includes checking for null values and filling it accordingly , removing the extreme values i.e. Outliers , sorting the necessary columns
- Cleaning data using code requires time and discipline, but makes all steps fully documented and reproducible.

Descriptive Analysis

- Descriptive numeric and graphic representations provide insights into the data.
- Histograms visualise the distribution of numeric variables. They show how often observations within a certain value range occur.

- Histograms reveal if the distribution of a variable is unimodal and symmetric or skewed.
- Helpful graphical methods for numeric data are histograms, boxplots and scatter plots. Bar plots of frequency counts are useful for the visualisation of categorical variables. Mosaic plots illustrate the association of multiple categorical variables.
- The boxplot is the most common graphical visualisation of unimodal distributions in statistics. The simplest version of a boxplot compresses a data set into minimum, first quartile, median, third quartile and maximum

Pre-Processing

Categorical Variables

- Two pre-processing procedures are often used for categorical variables. One is merging levels of categorical variables before further analysis, the other one is converting categorical variables to numeric ones.
- Merging levels of categorical variables is useful if the original categories are too differentiated (too many).
- The distance-based clustering methods assume that data are numeric, and measured on comparable scales. Ordinal data can be converted to numeric data if it can be assumed that distances between adjacent scale points on the ordinal scale are approximately equal options.
- Binary answer options are less prone to capturing response styles, and do not require data pre-processing. Pre-processing inevitably alters the data in some way. Binary variables can always be converted to numeric variables, and most statistical procedures work correctly after conversion if there are only two categories.

Principal Component Analysis

- Principal components analysis (PCA) transforms a multivariate data set containing metric variables to a new data set with variables – referred to as principal components – which are uncorrelated and ordered by importance
- The first variable (principal component) contains most of the variability, the second principal component contains the second most variability, and so on
- Principal components analysis works off the covariance or correlation matrix of several numeric variables
- In most cases, the transformation obtained from principal components analysis is used to project high-dimensional data into lower dimensions for plotting purposes
- Principal components analysis is also used for the purpose of reducing the number of segmentation variables before extracting market segments from consumer data
- Principal components analysis to explore data, and identify highly correlated variables
- Insights gained from such an exploratory analysis can be used to remove some of the original – redundant – variables from the segmentation base. This approach also achieves a reduction in dimensionality, but still works with the original variables collected

STEP 5 – EXTRACTING SEGMENTS

Grouping Consumers

- Many segmentation methods used to extract market segments are taken from the field of cluster analysis
- selecting a suitable clustering method requires matching the data analytic features of the resulting clustering with the context-dependent requirements that are desired by the researcher. It is important to explore market segmentation solutions derived from a range of different clustering methods
- k-means cluster analysis aims at finding compact clusters covering a similar range in all dimensions
- A single linkage method constructs snake-shaped clusters. When asked to return too many (8) segments, outliers are defined as micro-segments, but the two main spirals are still correctly identified
- kmeans cluster analysis fails to identify the spirals because it is designed to construct round, equally sized clusters. So the k-means algorithm ignores the spiral structure and, instead, places consumers in the same market segments if they are located close to one another (in Euclidean space), irrespective of the spiral they belong to
- If consumer data is well-structured, and well-separated, distinct market segments exist, tendencies of different algorithms matter less
- Distance-based methods use a particular notion of similarity or distance between observations (consumers), and try to find groups of similar observations (market segments)
- Data characteristics and expected or desired segment characteristics allow a pre-selection of suitable algorithms to be included in the comparison.
- If the target segment is expected to be a niche segment, larger sample sizes are required. Larger samples allow a more fine-grained extraction of segments.
- The structure of segments extracted by the algorithm needs to align with these expected characteristics, we distinguish directly observable characteristics from those that are only indirectly accessible. Benefits sought are an example of a directly observable characteristic
- An example of an indirect characteristic is consumer price sensitivity. If the data contains purchase histories and price information, and market segments are based on similar price sensitivity levels, regression models are needed. This, in turn calls for the use of a model based segment extraction algorithm.

Distance Based Methods

Distance Measures

- Numerous approaches to measuring the distance between two vectors exist; several are used routinely in cluster analysis and market segmentation.

- A distance is a function with two arguments: the two vectors x and y between which the distance is being calculated. The result is the distance between them (a nonnegative value)
- A distance measure has to comply with a few criteria. One criterion is symmetry, that is: d(x, y) = d(y, x).
- A second criterion is that the distance of a vector to itself and only to itself is 0: d(x, y) = 0 ⇔ x = y.
- In addition, most distance measures fulfil the so-called triangle inequality: d(x, z) ≤ d(x, y) + d(y, z). The triangle inequality says that if one goes from x to z with an intermediate stop in y, the combined distance is at least as long as going from x to z directly.
- Let x = (x1, . . . , xp) and y = (y1, . . . , yp) be two p-dimensional vectors. The most common distance measures used in market segmentation analysis are:

  Euclidean distance:

  $$d(x, y) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2}$$

  Manhattan or absolute distance:

  $$distance = \sum_{i=0}^{n-1} |(x[i] - y[i])|$$

  - Asymmetric binary distance: applies only to binary vectors, that is, all xj and yj are either 0 or 1. d(x, y) = _0, x = y = 0 ({j |xj = 1 and yj = 1})/(#{j |xj = 1 or yj = 1})
  - Euclidean Distance: This distance measure calculates the straight-line distance between two data points in a multi-dimensional space. It is commonly used for numerical variables and is calculated as the square root of the sum of squared differences between corresponding attributes
  - Manhattan Distance: Also known as city-block distance, this measure quantifies the distance between two points by summing the absolute differences of their attributes. It is suitable for categorical variables or when the distribution of the data is not Euclidean
  - Both Euclidean and Manhattan distance treat all dimensions of the data equally; they take a sum over all dimensions of squared or absolute differences. If the different dimensions of the data are not on the same, the dimension with the larger numbers will dominate the distance calculation between two observations.

Hierarchical Methods

- Hierarchical clustering methods are the most intuitive way of grouping data because they mimic how a human would approach the task of dividing a set of observations (consumers) into k groups (segments)
- Divisive hierarchical clustering methods start with the complete data set and splits it into two market segments in a first step. Then, each of the segments is again split into two segments. This process continues until each consumer has their own market segment

- Agglomerative hierarchical clustering approaches the task from the other end. The starting point is each consumer representing their own market segment (n sin-gleton clusters). Step-by-step, the two market segments closest to one another are merged until the complete data set forms one large market segment
- Agglomerative clustering is a measure of distance between groups of observations (segments). This measure is determined by specifying (1) a distance measure d(x, y) between observations (consumers) x and y, and (2) linkage method
- The linkage method generalises how, given a distance between pairs of observations, distances between groups of observations are obtained
- Single linkage: distance between the two closest observations of the two sets
- Complete linkage: distance between the two observations of the two sets that are farthest away from each other
- Average linkage: mean distance between observations of the two sets
- A very popular alternative hierarchical clustering method is named after Ward (1963), and based on squared Euclidean distances. Ward clustering joins the two sets of observations (consumers) with the minimal weighted squared Euclidean distance between cluster centers. Cluster centers are the midpoints of each cluster
- The result of hierarchical clustering is typically presented as a dendrogram. A dendrogram is a tree diagram. The root of the tree represents the one-cluster solution where one market segment contains all consumers. The leaves of the tree are the single observations (consumers), and branches in-between correspond to the hierarchy of market segments formed at each step of the procedure. The height of the branches correspond to the distance between the clusters. Higher branches point to more distinct market segments.

Partitioning Methods

k-Means and k-Centroid Clustering

- The most popular partitioning method is k-means clustering. Within this method, R function kmeans() implements the algorithms by Forgy (1965), Hartigan and Wong (1979), Lloyd (1982) and MacQueen (1967). These algorithms use the squared Euclidean distance
- A generalisation to other distance measures, also referred to as k-centroid clustering, is provided in R package flexclust. The representative of a market segment is referred to in many partitioning clustering algorithms as the centroid. For the k-means algorithm based on the squared Euclidean distance, the centroid consists of the column-wise mean values across all members of the market segment

Improved K Means

- The simplest improvement is to initialise k-means using “smart” starting values, rather than randomly drawing consumers from the data set and using them k as starting points

- Using starting points that are not representative of the data space increases the likelihood of the k-means algorithm getting stuck in what is referred to as a local optimum
- One way of avoiding the problem of the algorithm getting stuck in a local optimum is to initialise it using starting points evenly spread across the entire data space
- Steinley and Brusco conclude that the best approach is to randomly draw many starting points, and select the best set. The best starting points are those that best represent the data.

Hard Competitive Learning

- Hard competitive learning, is also known as learning vector quantisation, differs from the standard k-means algorithm in how segments are extracted
- Hard competitive learning randomly picks one consumer and moves this consumer's closest segment representative a small step into the direction of the randomly chosen consumer
- It is also possible that hard competitive learning finds the globally optimal market segmentation solution, while k-means gets stuck in a local optimum (or the other way around)

Neural Gas and Topology Representing Networks

- A variation of hard competitive learning is the neural gas algorithm proposed by Martinetz et al.
- Here the location of the second closest segment representative (centroid) is adjusted towards the randomly selected consumer too.
- A further extension of neural gas clustering are topology representing networks
- In addition, topology representing networks count how often each pair of segment representatives (centroids) is closest and second closest to a randomly drawn consumer This information is used to build a virtual map in which "similar" representatives – those which had their values frequently adjusted at the same time – are placed next to one other. Almost the same information – which is central to the construction of the map in topology representing networks ,such a graph is known as segment neighbourhood graph

Self-Organising Maps

- Another variation of hard competitive learning are self-organising maps (Kohonen 1982, 2001), also referred to as self-organising feature maps or Kohonen maps
- Self-organising maps position segment representatives (centroids) on a regular grid, usually a rectangular or hexagonal grid
- The self-organising map algorithm is similar to hard competitive learning: a single random consumer is selected from the data set, and the closest representative for this

random consumer moves a small step in their direction. In addition, representatives which are direct grid neighbours of the closest representative move in the direction of the selected random consumer

- The advantage of self-organising maps over other clustering algorithms is that the numbering of market segments is not random. Rather, the numbering aligns with the grid along which all segment representatives (centroids) are positioned

Neural Networks

- Auto-encoding neural networks for cluster analysis work mathematically differently than all cluster methods The most popular method from this family of algorithms uses a so-called single hidden layer perceptron
- A single hidden layer perceptron. The network has three layers. The input layer takes the data as input. The output layer gives the response of the network. In the case of clustering this is the same as the input. In-between the input and output layer is the so-called hidden layer. It is named hidden because it has no connections to the outside of the network.
- The values of the three nodes in the hidden layer h1, h2 and h3 are weighted linear combinations of the inputs for a non-linear function fj . Each weight αij in the formula is depicted by an arrow connecting nodes in input layer and hidden layer. The fj are chosen such that $0 \le hj \le 1$, and all hj sum up to one ($h1 + h2 + h3 = 1$).
- Once the network is trained, parameters connecting the hidden layer to the output layer are interpreted in the same way as segment representatives (centroids) resulting from traditional cluster algorithms. The parameters connecting the input layer to the hidden layer can be interpreted in the following way: consider that for one particular consumer $h1 = 1$, and hence $h2 = h3 = 0$. In this case ˆxi = β1i for = 1, . . . , 5. This is true for all consumers where h1 is 1 or close to 1. The network predicts the same value for all consumers with $h1 \approx 1$. All these consumers are members of market segment 1 with representative β1i . All consumers with $h2 \approx 1$, are members of segment 2, and so on.
- Neural network clustering is an example of a so-called fuzzy segmentation with membership values between 0 (not a member of this segment) and 1 (member of only this segment). Membership values between 0 and 1 indicate membership in multiple segments.

Hybrid Approaches

The basic idea behind hybrid segmentation approaches is to first run a partitioning algorithm because it can handle data sets of any size. But the partitioning algorithm used initially does not generate the number of segments sought. Rather, a much larger number of segments is extracted. Then, the original data is discarded and only the centres of the resulting segments and segment sizes are retained, and used as input for the hierarchical cluster analysis. At this point, the data set is small enough for hierarchical algorithms, and the dendrogram can inform the decision how many segments to extract.

MODEL BASED METHODS

Model-based methods, such as finite mixture models, have gained significant interest in the field of marketing research. These methods extract market segments by assuming that each segment has a specific size and that consumers belonging to the same segment share specific characteristics.

Unlike distance-based clustering methods that rely on similarities or distances, model-based methods determine segment membership based on the likelihood of observing certain values given the segment-specific characteristics. The finite mixture model is a combination of segment-specific models, with parameters such as segment sizes and segment-specific characteristics to be estimated.

The estimation of finite mixture models is typically done through maximum likelihood estimation or Bayesian methods. To determine the appropriate number of segments, information criteria such as AIC, BIC, and ICL are commonly used. These criteria assess the goodness-of-fit of the model and penalize for the number of estimated parameters.

The advantage of using finite mixture models is their ability to capture complex segment characteristics and allow for various extensions. While initially appearing complex, these models provide a flexible framework for market segmentation analysis.

Finite mixtures of distributions

The finite mixture model reduces to

$$\sum_{h=1}^{k} \pi_h f(y|\theta_h), \quad \pi_h \geq 0, \quad \sum_{h=1}^{k} \pi_h = 1.$$

The formulae are the same , the only difference is that there is no x. The statistical distribution function f () depends on the measurement level or scale of the segmentation variables y.

Normal distribution

The multivariate normal distribution is suitable for capturing covariance between variables.

Examples are given, such as physical measurements on humans (e.g., height, arm length, leg length) and business data (e.g., prices in markets with many players), where the variables follow approximate multivariate normal distributions. Examples of its application in biology and business are provided. The segment-specific parameters, including mean vectors and covariance matrices, are described, and the number of parameters to estimate is determined. The "mclust" package in R is recommended for fitting the models, with model selection based on the BIC. The use of spherical covariance matrices to reduce parameter estimation is explained. The BIC values are plotted to determine the optimal number of segments. Overall, the excerpt provides insights into fitting mixture models and the importance of covariance matrices in market segmentation.

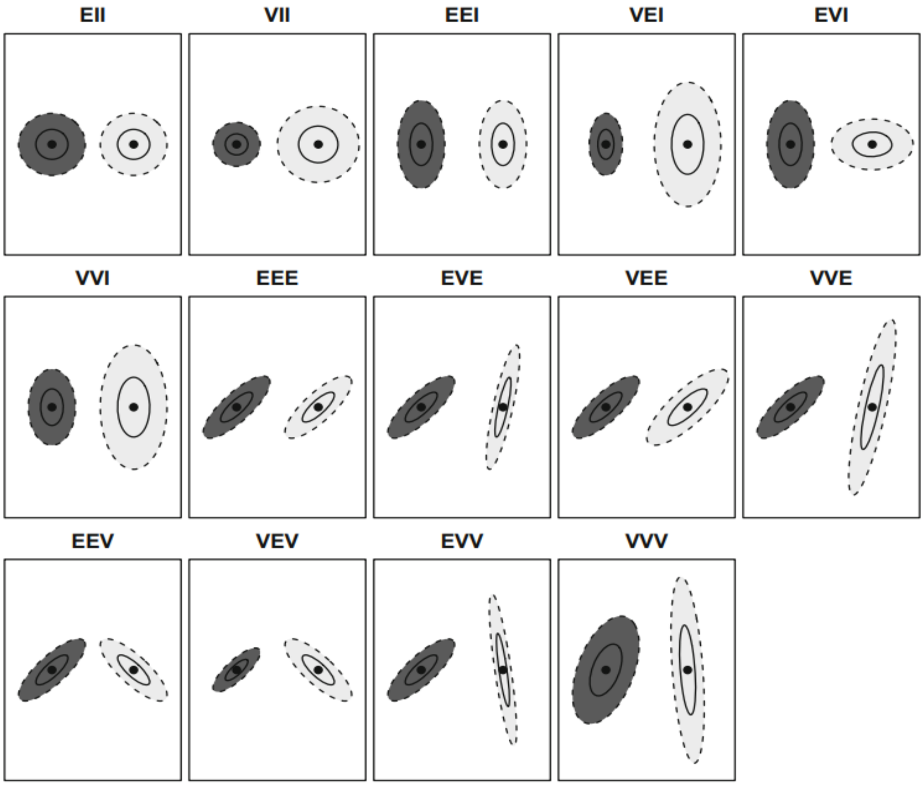


Fig. 7.25 Visualisation of the 14 covariance models available in package mclust

Binary distributions

Binary distributions, also known as latent class models or latent class analysis, for analyzing binary data in market segmentation. The example focuses on winter activities of Austrian tourists, specifically alpine skiing and sight-seeing. The observed frequency patterns indicate an association between the two variables. To better represent the data, a mixture model of binary distributions is fitted using the flexmix package in R. The EM algorithm is employed with multiple random restarts to find the best model. The results suggest a mixture model with two segments, revealing distinct preferences for the activities. The parameters of the segment-specific models, representing the probabilities of observing a certain activity, are obtained. The expected frequencies from the fitted model match the observed frequencies, indicating a successful explanation of the association between the variables through segmentation.

Finite mixtures of regressions Finite mixtures provides an example of a data set consisting of consumers' willingness to pay for a theme park based on the number of rides available. The data set is generated using two linear regression models representing different market segments.

The text explains the process of fitting a finite mixture of regression models using the flexmix package in R. It mentions that the EM algorithm is used for estimation, and the number of segments is specified as two (k = 2). The results of the estimation, including the cluster sizes and the estimated parameters, are presented.

Algorithms with integrated variable selection

Segmentation algorithms often assume that all segmentation variables are relevant, but preprocessing methods can identify redundant or noisy variables. For binary variables, biclustering and the variable selection procedure for clustering binary data (VSBD) simultaneously extract segments and select suitable variables. Factor-cluster analysis compresses variables into factors before segment extraction. These methods enhance meaningful segment identification and reduce the impact of irrelevant variables.

Biclustering Algorithms:

Biclustering algorithms are used for segment extraction in the presence of binary variables. These algorithms aim to simultaneously identify meaningful segments and select suitable segmentation variables. They are effective in handling binary data but may require additional computational resources due to their complexity. Biclustering methods, such as Spectral Biclustering and Plaid Models, provide flexibility in identifying subgroups with specific patterns

Variable Selection Procedure for Clustering Binary Data (VSBD):

VSBD is specifically designed to address the challenge of variable selection in binary data segmentation. It aims to identify the most informative segmentation variables while extracting segments. VSBD evaluates the relevance and contribution of each binary variable to the clustering solution. By focusing on informative variables, it enhances the accuracy and interpretability of the resulting segments.

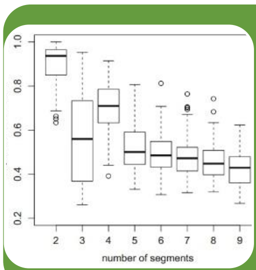
Variable Reduction: Factor Cluster Analysis:

Factor Cluster Analysis is a two-step approach that combines variable reduction and segment extraction. It compresses the original segmentation variables into factors through techniques like factor analysis or principal component analysis. These factors capture the underlying structure and relationships among the variables. Subsequently, clustering algorithms are applied to the reduced factor space to extract segments. This approach simplifies the analysis by reducing dimensionality and enhancing interpretability.

Biclustering algorithms are suitable for binary data segmentation, while VSBD focuses specifically on variable selection in binary data. Factor Cluster Analysis incorporates variable reduction to simplify the segmentation process. The choice of the appropriate method depends on the specific characteristics of the data, research objectives, and desired outcomes in terms of interpretability and efficiency.

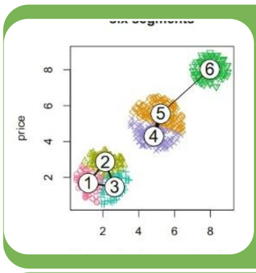
Data Structure Analysis

Stability-based data structure analysis, such as modifying data or algorithms, is commonly used to evaluate the robustness of segmentation results. This analysis provides insights into the existence of distinct market segments and helps guide methodological decisions. Approaches like cluster indices, gorge plots, global stability analysis, and segment-level stability analysis are utilized for data structure analysis.



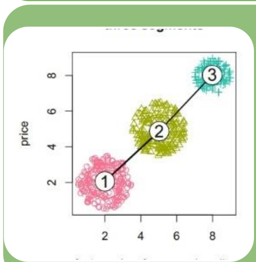
Cluster Indices:

Cluster indices are metrics used to assess the quality and validity of clustering solutions. These indices quantify the compactness and separation of clusters, providing a measure of how well-defined and distinct the segments are. Examples of cluster indices include the Silhouette index, Dunn index, and Calinski-Harabasz index. Higher values indicate better clustering solutions with clear boundaries and compact clusters.



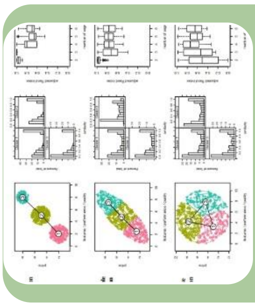
Gorge Plots:

Gorge plots visualize the stability and robustness of clustering solutions. They depict the variation in cluster assignments across multiple iterations or perturbations of the data. Gorge plots consist of stacked bars, where each bar represents a data point, and the height indicates the frequency of cluster membership. A wider gorge plot suggests more stable and reliable segments, while narrower plots indicate instability or ambiguity.



Global Stability Analysis:

Global stability analysis assesses the overall stability of clustering solutions by comparing different iterations or modifications of the data. It provides insights into the consistency and reliability of the segment assignments across variations. Statistical measures, such as Jaccard coefficients or Rand indices, are used to quantify the similarity between different clustering solutions. Higher similarity scores indicate greater stability and consistency.



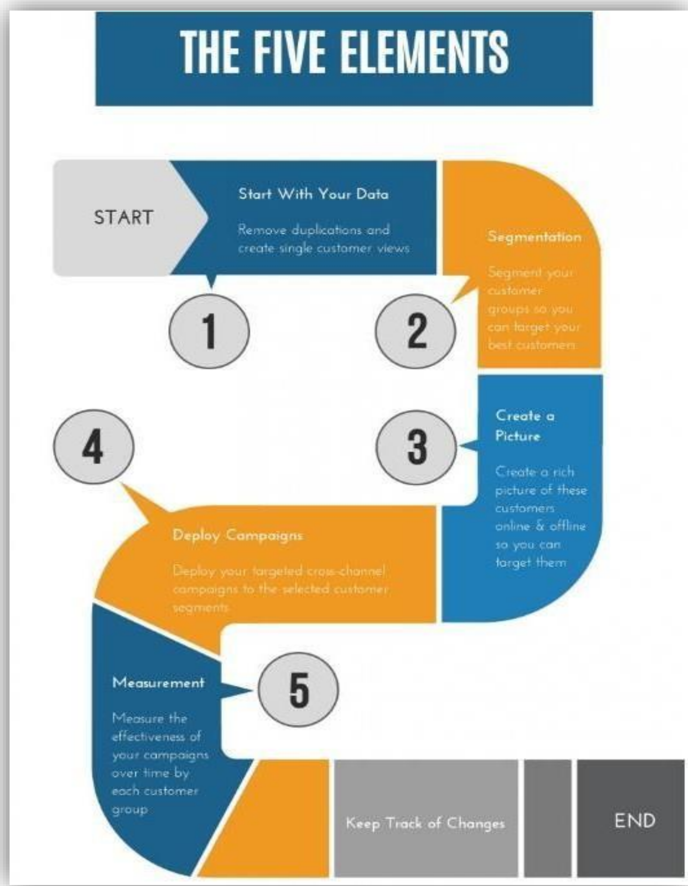
Segment Level Stability Analysis:

Segment level stability analysis evaluates the stability of individual segments within clustering solutions. It focuses on understanding the robustness of segment memberships and boundaries. This analysis involves comparing segment assignments across different iterations or perturbations of the data. Metrics like the Variation of Information (VI) or Adjusted Rand Index (ARI) are used to measure the similarity or agreement between segment assignments. Higher agreement scores indicate more stable and reliable segments.

Cluster indices provide a quantitative assessment of clustering quality, focusing on compactness and separation of clusters. Gorge plots offer a visual representation of stability and variability in cluster assignments. Global stability analysis assesses the overall consistency and similarity of clustering solutions. Segment level stability analysis zooms in on the stability of individual segments within the solutions. Each approach provides unique insights into the reliability and validity of clustering results.

STEP 6 – PROFILING SEGMENTS

Identifying the key characteristics of the market segmentation is the most important process. Profilingthe segments is to know the market segments resulting from the extraction step. Profiling is only required when data driven segmentation is used. Profiling consists of characterizing the market segments individually, but also in comparison to the other market segments. At the profiling stage, we inspect several alternative market segmentation solutions. This is particularly important if no natural segments exist in the data, and either a reproducible or a constructive market segmentation approach must be taken. Good profiling is the basis for correctinterpretation of the resulting segments. Correct interpretation, in turn, is critical to making good strategic marketing decisions.



Visualizations can be used for profiling complex data sets. Graphical representation and tabular representation the most used types. Visualizations are useful in the data-driven market segmentationprocess to inspect, for each segmentation solution, one or more segments in detail. Statistical graphs facilitate the interpretation of segment profiles. They also make it easier to assess the usefulness of a market segmentation solution. The process of segmenting data always leads to many alternative solutions. Selecting one of the possible solutions is a critical decision. Visualizations of solutions assist the data analyst and user with this task. A good way to understand the defining characteristics of eachsegment is to produce a segment profile plot. The segment profile plot shows – for all segmentation variables – how each market segment differs from the overall sample.

Segment separation can be visualized in a segment separation plot. The segment separation plot depicts

– for all relevant dimensions of the data space – the overlap of segments. Segment separation plots arevery simple if the number of segmentation variables is low but become complex as the number of segmentation variables increases. But even in such complex situations, segment separation plots offer data analysts and users a quick overview of the data situation, and the segmentation solution.

STEP 7 – DESCRIBING SEGMENTS

Segment profiling is about understanding differences in segmentation variables across market segments. We can use a wide range of charts to describe the segmentation. When describing differences between market segments in one single nominal or ordinal descriptor variable, the basis forall visualizations and statistical tests is a cross-tabulation of segment membership with the descriptor variable. Simple statistical tests can be used to formally test for differences in descriptor variables acrossmarket segments. The simplest way to test for differences is to run a series of independent tests for each variable of interest. The outcome of the segment extraction step is segment membership, the assignment of each consumer to one market segment. Segment membership can be treated like any other nominal variable. It represents a nominal summary statistic of the segmentation variables.

Therefore, any test for association between a nominal variable and another variable is suitable.

Another way of learning about market segments is to try to predict segment membership from descriptor variables. To achieve this, we use a regression model with the segment membership as categorical dependent variable, and descriptor variables as independent variables. We can use methods developed in statistics for classification, and methods developed in machine learning for supervised learning.

STEP 8 – SELETING THE TARGET SEGMENT(s)

- All market segments under step 8 must comply with the knock-out criteria.
- The team has to consider the following questions in selecting a target segment
    - Which of the market segments would the organisation most like to target? Which segment would the organisation like to commit to?
    - Which of the organisations offering the same product would each of the segments most like to buy from? How likely is it that our organisation would be chosen? How likely is it that each segment would commit to us?
- Use of decision matrix to visualize segment attractiveness

- Segment attractiveness vs relative organisational competitiveness graph s plotted. Segments appear as circles.
- Size of circles represent another criteria of choice relevant to selection
- In step 2, we decided the weight of each criteria on segment attractiveness, this weight is used to plot the graph
- The plot is between 'How attractive is the segment to us?' VS 'How attractive are we to the segment?' (i.e., if the segment gets another organisation with similar services, will they switch?)
- The size of the segment bubble in the graph is based on how profitable that particular segment is to the organisation.
- Refer to page 239-241 for mathematical calculations.

STEP 9 – CUSTOMISING THE MARKET MIX

- Marketing mix consists of 4 Ps, Product, Price, Promotion, Place
- Market segmentation goes together with other areas of marketing, most importantly; positioning and competition
- Segmentation process is seen as part of STP (segmentation-targeting-positioning)
  - Segmentation: the extraction, profiling and description of segments
  - Targeting: the assessment of segments and selection of a target segment
  - Positioning: the measures an organisation can take to ensure that their product is perceived as distinctly different from competing products, and in line with segment needs
- All of the aspects of marketing mix need to be reviewed once target segment(s) have been selected.
- To maximize benefits, the marketing mix are customised to target segments, such as discounts, rebranding, selection of distribution channels, etc.
- The product needs to be specified in view of customer need, by modifying the existing product.
- Other decisions n Product mix are: naming, packaging, warranties, after sales support, etc.
- The price mix includes decisions like: selling price, discount rates, etc.
- Example from case study of Australian vacation is given to determine the expenditure of people of segment 3 taking a vacation.
- The decisions under Place mix are how to distribute the product, should the product be available online or offline, direct selling should be available or a wholesaler should be involved etc.
- The same example is used to determine the means of booking a vacation by segment 3 people in the Australian vacation case study.
- The Promotion mix includes decisions like: developing an advertising message, and finding the most effective way of communicating the message. Public relations, personal selling, sponsorships, etc.
- Using the same example, it was observed that the source of advertisement most used by segment 3 in Australian vacation case study was tourist centre, and most watched TV channel were channel 7 and 9 by the users.
- The checklist is provided to go through the processes of this step thoroughly.

Case Study: Fast Food

Introduction:

The purpose of this case study is to demonstrate market segmentation analysis utilising an empirical dataset gathered for brand image research. The idea, as with McDonald's, is to discover separate consumer categories with differing impressions of the brand. Understanding these differences allows for more focused marketing techniques and message personalization. By identifying key factors and implementing appropriate changes, the focus might be on reinforcing positive perceptions or resolving negative perceptions.

Step 1: Deciding (not) to Segment:

In terms of market segmentation, McDonald's has two options. The first position implies that the brand serves the entire market and hence does not need to investigate systematic distinctions across consumer categories. Alternatively, McDonald's may believe that, despite their market dominance, researching systematic variability among consumers is worthwhile. They can use a differentiated marketing strategy to capitalise on the varying tastes and needs of various segments if they recognise these variances.

Step 2: Specifying the Ideal Target Segment:

McDonald's management must examine various parameters when determining target market segments. The target groups should be homogeneous, distinct from other segments, large enough to justify targeted marketing efforts, and open to fast food consumption. Furthermore, the segments must be recognised and reachable via existing routes. While it may appear that a sector with a positive impression of McDonald's, frequent dining out habits, and a penchant for fast food is evident, McDonald's may also aim to understand and influence attitudes in segments that are currently unfavourable to the brand. The fast-food dataset's limited data will focus on two attractiveness criteria: loving McDonald's and frequent consumption at McDonald's. These parameters will influence the selection of target segments.

Step 3: Collecting Data:

The information utilised in this case study is made up of responses from 1453 adult Australian consumers about their impressions of McDonald's across a variety of qualities. These characteristics, such as "yummy," "convenient," "spicy," "fattening," and others, were extracted from a prior qualitative study. With a YES or NO response, respondents indicated whether McDonald's possessed each feature. The dataset also includes age and gender information for respondents. Additional data on dining out behaviour and information channel usage would have been obtained in a true market segmentation research to provide a more detailed description of each market segment.

Step 4: Exploring Data:

First we explore the key characteristics of the data set by loading the data set and inspecting basic features such as the variable names, the sample size, and the first three rows of the data:

```
R> library("MSA")
R> data("mcdonalds", package = "MSA")
R> names(mcdonalds)

 [1] "yummy"          "convenient"     "spicy"
 [4] "fattening"      "greasy"         "fast"
 [7] "cheap"          "tasty"          "expensive"
[10] "healthy"        "disgusting"     "Like"
[13] "Age"            "VisitFrequency" "Gender"

R> dim(mcdonalds)

[1] 1453   15

R> head(mcdonalds, 3)

  yummy convenient spicy fattening greasy fast cheap tasty
1    No        Yes    No       Yes     No  Yes   Yes    No
2   Yes        Yes    No       Yes    Yes  Yes   Yes   Yes
3    No        Yes   Yes       Yes    Yes  Yes    No   Yes
  expensive healthy disgusting Like Age      VisitFrequency
1       Yes      No         No   -3  61 Every three months
2       Yes      No         No   +2  51 Every three months
3       Yes     Yes         No   +1  62 Every three months
  Gender
1 Female
2 Female
3 Female
```

The first respondent's perception of McDonald's traits is revealed in the data, coded as vocal YES/NO responses. However, these verbal codes are insufficient for segment extraction and must be converted to numeric binary representation. To accomplish this, the segmentation variables are separated into a separate matrix and translated into TRUE (YES) and FALSE (NO) values. The transformation's accuracy is confirmed by checking the average values of the transformed segmentation variables.

Results from principal components analysis indicate that the first two components capture about 50% of the information contained in the segmentation variables. The following command returns the factor loadings:

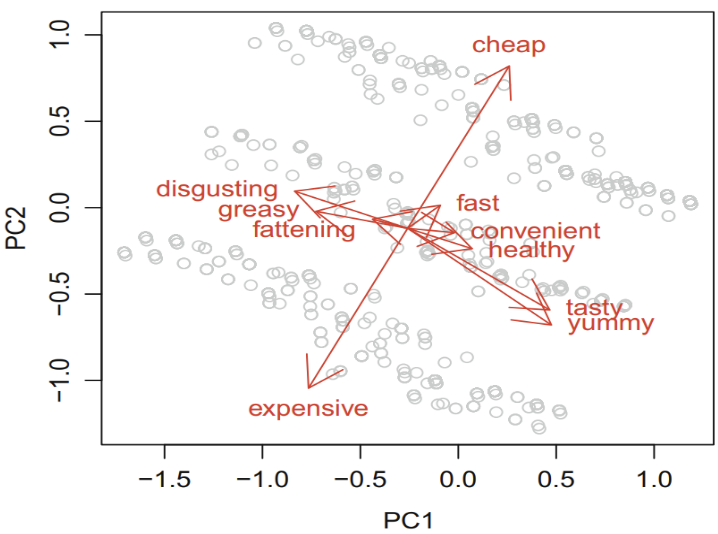
R> print(MD.pca, digits = 1) Standard deviations (1, .., p=11):

[1] 0.8 0.6 0.5 0.4 0.3 0.3 0.3 0.3 0.3 0.2 0.2

```
Rotation (n x k) = (11 x 11):
              PC1   PC2  PC3    PC4    PC5   PC6   PC7
yummy       0.477 -0.36 0.30 -0.055 -0.308  0.17 -0.28
convenient  0.155 -0.02 0.06  0.142  0.278 -0.35 -0.06
spicy       0.006 -0.02 0.04 -0.198  0.071 -0.36  0.71
fattening  -0.116  0.03 0.32  0.354 -0.073 -0.41 -0.39
greasy     -0.304  0.06 0.80 -0.254  0.361  0.21  0.04
fast        0.108  0.09 0.06  0.097  0.108 -0.59 -0.09
cheap       0.337  0.61 0.15 -0.119 -0.129 -0.10 -0.04
tasty       0.472 -0.31 0.29  0.003 -0.211 -0.08  0.36
expensive  -0.329 -0.60 -0.02 -0.068 -0.003 -0.26 -0.07
healthy     0.214 -0.08 -0.19 -0.763  0.288 -0.18 -0.35
disgusting -0.375  0.14 0.09 -0.370 -0.729 -0.21 -0.03
             PC8    PC9   PC10   PC11
yummy       0.01 -0.572  0.110  0.045
convenient -0.11  0.018  0.666 -0.542
spicy       0.38 -0.400  0.076  0.142
fattening   0.59  0.161  0.005  0.251
greasy     -0.14  0.003 -0.009  0.002
fast       -0.63 -0.166 -0.240  0.339
cheap       0.14 -0.076 -0.428 -0.489
tasty      -0.07  0.639 -0.079  0.020
expensive   0.03 -0.067 -0.454 -0.490
healthy     0.18  0.186  0.038  0.158
disgusting -0.17  0.072  0.290 -0.041
```

The principal component analysis loadings show how the basic variables are merged to generate components. In this scenario, the attributes with the largest loadings for major component 2 are CHEAP and EXPENSIVE, indicating that this component represents the price dimension. The data is then projected onto the principle component space, with customers shown as grey dots. In judging McDonald's, the attributes CHEAP and EXPENSIVE are important, while other attributes match with favourable or negative opinions. FATTENING, DISGUSTING, and GREASY are connected with negative impressions, whereas FAST, CONVENIENT, HEALTHY, TASTY, and YUMMY are associated with favourable perceptions. The perceptual chart also shows consumer groups along the EXPENSIVE vs. CHEAP axis.

<u>Figure shows the resulting perceptual map:</u>



Step 5: Extracting Segments

Step 5 is where we extract segments. To illustrate a range of extraction techniques, we subdivide this step into three sections. In the first section, we will use standard k-means analysis. In the second section, we will use finite mixtures of binary distributions. In the third section, we will use finite mixtures of regressions, mixtures of Distribution and K-means clustering.

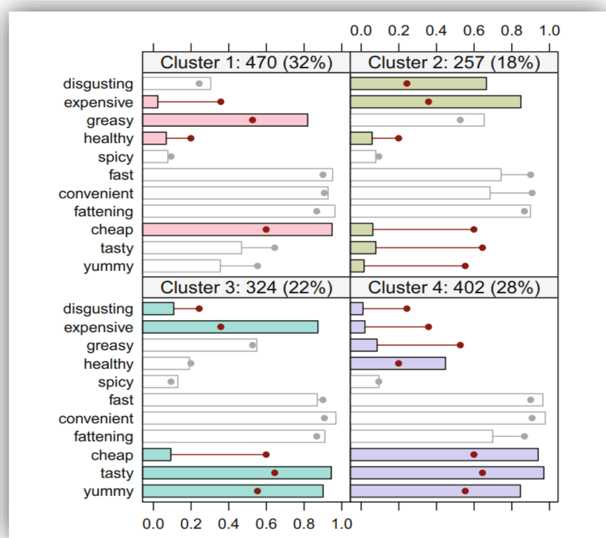
Step 6: Profiling Segments

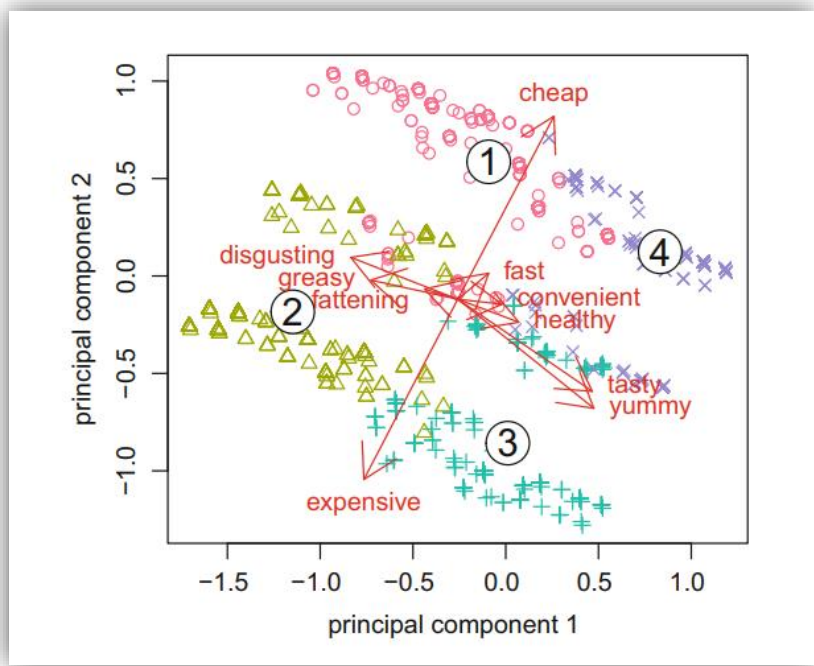
The first step is to create a segment profile plot, which visually presents the key characteristics and differences between the market segments. Hierarchical cluster analysis is used to identify similar attributes, and the ordering of these attributes is used to create the segment profile plot. Marker variables are highlighted to show the percentage of respondents within each segment who associate specific perceptions with McDonald's.

By examining the segment profile plot, McDonald's managers can identify distinct perceptions of each segment. For example, segment 1 perceives McDonald's as cheap and greasy, while segment 2 views it as disgusting and expensive. Segment 3 considers McDonald's expensive but also finds the food tasty and yummy. On the other hand, segment 4 holds positive beliefs about McDonald's, perceiving the food as tasty, yummy, cheap, and somewhat healthy.

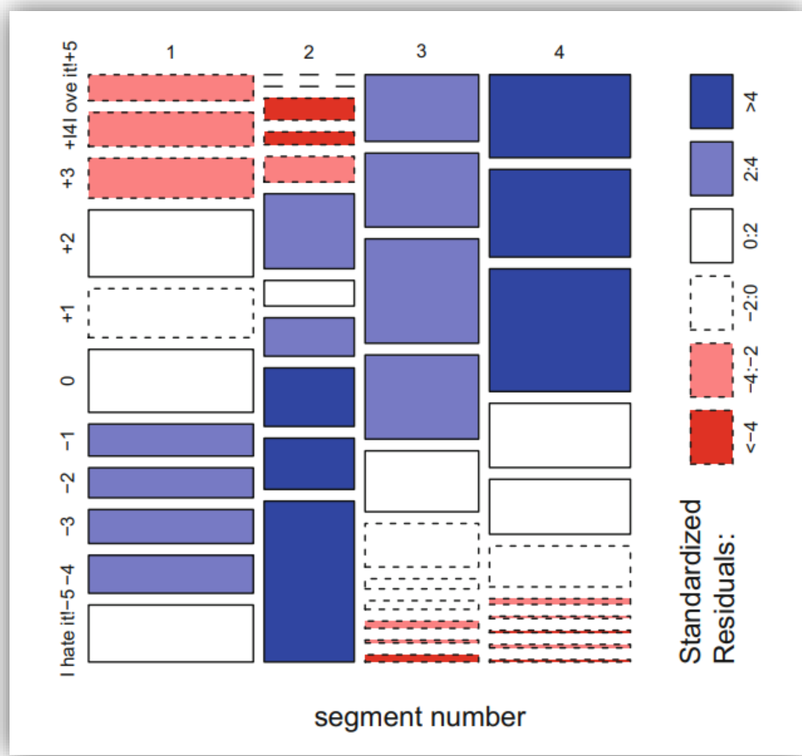
Another visualization called the segment separation plot is introduced, which incorporates principal components analysis. This plot shows the centers of each market segment and colors observations based on segment membership. From this plot, it can be observed that segments 1 and 4 both view McDonald's as cheap, but segment 4 also has positive beliefs, while segment 1 associates it with negative attributes. Segments 2 and 3 disagree on various features, with segment 2 holding a less favorable view compared to segment 3.

By the end of Step 6, McDonald's managers gain a good understanding of the nature of the four market segments based on the information used to create them. However, further exploration and learning about the segments are necessary, which will be the focus of Step 7.





Step 7: Describing Segments



Step 8: Selecting the target Segment(s)

The segment evaluation plot, as shown in Figure A.15, helps McDonald's managers assess the attractiveness of different market segments. Using the provided code in Python, various segment characteristics are analyzed. The mean values of visiting frequency, liking, and the percentage of female consumers are computed for each segment.

To compute the mean visiting frequency for each segment:

visit = mcdonalds.groupby(k4)['VisitFrequency'].mean()

To compute the mean liking for each segment:

like = mcdonalds.groupby(k4)['Like.n'].mean()

To compute the percentage of female consumers for each segment, the GENDER variable needs to be converted to numeric:

female = mcdonalds.groupby(k4)['Gender'].apply(lambda x: (x == 'Female').mean())

Using the obtained values, the segment evaluation plot can be created:

```
import matplotlib.pyplot as plt

plt.scatter(visit, like, s=10 * female, alpha=0.7)
plt.xlim(2, 4.5)
plt.ylim(-3, 3)

for i, txt in enumerate(visit.index):
    plt.text(visit[i], like[i], str(txt))

plt.xlabel('Visiting Frequency')
plt.ylabel('Liking for McDonald\'s')
plt.show()
```

The resulting plot helps identify the positioning of market segments. Segments 3 and 4, located in the attractive quadrant, represent customers who like McDonald's and visit frequently, making them valuable targets for retention. Market segment 2, positioned in the least attractive area, comprises individuals who dislike McDonald's and rarely visit, making them an unattractive segment. Market segment 1, despite negative perceptions and considering the factors of liking and visitation frequency, presents an opportunity for targeted marketing efforts to address their concerns and expand McDonald's customer base. The segment evaluation plot aids McDonald's management in deciding which market segments to focus on in Step 9.

Step 9: Customising the Marketing Mix

In Step 9, McDonald's can focus on segment 3 by introducing the MCSUPERBUDGET line that caters to their price expectations. This approach aims to develop loyal customers who may eventually transition to the regular product range as their income increases. Differentiating product features would prevent cannibalization. Effective communication channels used by segment 3 should be utilized to promote the MCSUPERBUDGET line. Distribution channels would remain the same, but a separate lane could be considered to minimize cannibalization without affecting the main product line.

Step 10: Evaluation and Monitoring

After conducting market segmentation analysis and implementing marketing activities, evaluating the success of the strategy and monitoring the market becomes crucial. Changes within existing segments, such as increased income for segment 3, may necessitate adjustments to product offerings like MCSUPERBUDGET. Additionally, monitoring the larger marketplace is important to identify shifts caused by factors like new competitors entering the market. McDonald's management must stay vigilant and adapt their strategic and tactical marketing approaches accordingly

MacDonald's data set review

- HIMANSHI GARG

  https://github.com/Ghimanshigit03/McDonaldEDA/blob/main/Project.ipynb
- ARYAN KUMAR

  https://github.com/aryan311/mcdonalds_Fynn_labs
- PSADHASIVAM

  https://github.com/Sadhasivam9/Feynn_Labs_Mcdonald-s_case_study.git
- KUSHAGRA BHATNAGAR

  https://github.com/Bhatnagar621/McDCaseStudy