Market Target Segmentation Analysis

Team-Lahari Shastri

- Pawar Mayur Dattatray
- Rayudu Mounika
- Krisi UmangKumar Doshi
- Rajesh Kannan

What is Market Segmentation?

It is the process of dividing a broad consumer or business market, normally consisting of existing and potential customers, into sub-groups of consumers based on shared characteristics. Consumer characteristics are very important for market segmentation. Approaches to market segmentation analysis: There are two systematics, one uses as its basis the extent to which the organization is willing or able to make changes to their current approach of targeting the market or a segment of the market. The second systematics is based on the nature of the segmentation variable.

There is a ten-step approach to market segmentation analysis which are as follows:

Step 1: Deciding (not) to Segment

1.1 Implications of Committing to Market Segmentation

The key implication is that the organization needs to commit to the segmentation strategy in the long term. It must be systematically and continuously communicated and reinforced at all organizational levels and across all organizational units. There should be some changes in market operations in order to improve profits by using market segmentation.

Changes include the development of new products, the modification of existing products, changes in pricing and distribution channels used to sell the product, as well as all communications with the market.

1.2 Implementation barriers

Successful market segmentation can be implemented in organization's by removing the barriers that impede the strategy for market segmentation. Below are the barriers,

- Lack of leadership in the market segmentation process by senior management. Without definite actions by senior management for segmentation reviews, it is impossible to implement strategies to all levels.
- Organizational culture, if the organization is not ready for the modifications, unwillingness to make changes and office policies that prevent the successful implementation of market segmentation.
- Lack of training and qualified marketing experts in the organizations.
- Objective restrictions faced by the organization, including lack of financial resources, or the inability to make the structural changes required.
- Not having clarified the objectives of the market segmentation exercise.
- Lack of structured processes to guide the team through all steps of the market segmentation process.

Step 2: Specifying the Ideal Target Segment

2.1 Segment Evaluation Criteria

The organization must determine two sets of segment evaluation criteria. Members of the segmentation team need to select which of these criteria they want to use to determine how attractive potential target segments are. The following two criteria are as follows:

Knock-out criteria: These criteria are the essential, non-negotiable features of segments that the organization would consider targeting. This will automatically eliminate some of the available market segments.

The selected market segment must be

- **Homogeneous**: Members in the segment similar to one another.
- **Distinct**: Members in other segments distinctly differ from one another.
- **Large enough**: Segment should contain enough members.
- **Identifiable**: Members of the segment must be identifiable

Attractiveness criteria: These criteria are first negotiated by segmentation teams, and then applied to determine the overall relative attractiveness of each market segment– those in compliance with the knock-out criteria. Rather, each market segment is rated; it can be more or less attractive with respect to a specific criterion.

2.2 Implementing a Structured Process

- Following a structured process when assessing market segments is beneficial.
- The most popular structured approach for evaluating market segments in view of selecting them as target markets is the use of a segment evaluation plot, showing segment attractiveness on one axis and organizational competitiveness on the other axis.
- A large number of possible criteria has to be investigated before agreement is reached on which criteria are most important for the organization.
- The core team should propose a solution and report their choices to the advisory committee.
- There is a huge benefit in selecting the attractiveness criteria for market segments at this early stage in the process.
- At the end of this step, the market segmentation team should have a list of approximately six segment attractiveness criteria.

Step 3: Collecting Data

3.1 Segmentation Variables:

The term segmentation variable refers to one measured value, for example, one item in a survey, or one observed expenditure category. Here we use segmentation variables throughout the process of market segmentation. Empirical data forms the basis of both commonsense and data-driven market segmentation. The difference between commonsense and data-driven market segmentation is that data-driven market segmentation is based not on one, but on multiple segmentation variables.

In data-driven market segmentation some of the variables are considered as segmentation variables in the data set and others are descriptor variables. Descriptor variables are used to describe the segments in detail.

3.2 Segmentation Criteria:

The term segmentation criterion relates to the nature of the information used for market segmentation. The most common segmentation criteria are:

- Geographic
- Socio-demographic
- Psychographic and
- Behavioral

Geographic Segmentation

It is seen as the original segmentation criterion used for the purpose of market segmentation. Geographic locations of customers serve to form geographic segmentation for market segments. Knowing the customers geographic locations helps to understand the customer needs.

Socio-Demographic Segmentation

It includes age, gender, income, and education. This segment can be very useful in some industries. For example: luxury goods, cosmetics, baby products, retirement villages and tourism resort products. Segmentation membership can be easily determined for every customer. It does not represent a strong bias for market segmentation, suggesting that the values, tastes, preferences are more useful because they are more influential in terms of consumers buying decisions.

Psychographic Segmentation

This refers to grouping of people's psychological criteria such as their beliefs, preferences, interests, aspirations, or benefits sought when purchasing a product. This segment basically is complex by nature, since it is more difficult to find a single characteristic of a person that will provide insight into the psychographic dimension of interest.

Behavioural Segmentation

Behavioural patterns or similarities can be used to identify the customer needs. Such behaviours like frequency of purchase, amount spent on every purchase on each occasion or multiple occasions, information search behavior and prior experience with the product. Behaviors variables are reported as superior to geographic variables since it shows the true interest of customers.

3.3 Survey

Survey data is cheap and easy to collect, making it a feasible approach for any organization. A few key aspects that need to be considered when using survey data are as follows:

- Choice of Variables
- Response Options
- Response Styles
- Sample Size

3.4 Data from Internal Sources

Increasingly organizations have access to substantial amounts of internal data that can be harvested for the purpose of market segmentation analysis. Typical examples are scanner data available to grocery stores, booking data available through airline loyalty programs, and online purchase data. If organizations are capable of storing data in a format that makes them easy to access, then no extra effort is required to collect data.

3.5 Data from Experimental Studies

The source of data that can form the basis of market segmentation analysis is experimental data. Experimental data can result from field or laboratory experiments. All these experiments help to identify the consumers interest on products which shows the characteristics of different combinations of attribute levels. This information can be used as a segmentation criterion.

Step 4: Exploring Data

After data collection, exploratory data analysis cleans and pre-process the data. This exploration stage also offers guidance on the most suitable algorithm for extracting meaningful market segments.

Data exploration helps in following things:

(1) identify the measurement levels of the variables
(2) investigate the univariate distributions of each of the variables
(3) assess dependency structures between variables.

To illustrate data exploration using real data, we have used a travel motives data set. This dataset contains 20 travel motives reported by 1000 Australian residents in relation to their last vacation. The dataset can be downloaded from the following link.

The link for dataset: http://www.marketsegmentationanalysis.org/

We read the dataset and we store it into a data frame.

```
R> vac <- read.csv("vacation.csv", check.names = FALSE)
```

We can also inspect and learn about the column names, size of the dataset.

```
R> colnames(vac)
 [1] "Gender"
 [2] "Age"
 [3] "Education"
 [4] "Income"
 [5] "Income2"
 [6] "Occupation"
 [7] "State"
 [8] "Relationship.Status"
 [9] "Obligation"
[10] "Obligation2"
[11] "NEP"
[12] "Vacation.Behaviour"
[13] "rest and relax"
[14] "luxury / be spoilt"
[15] "do sports"
[16] "excitement, a challenge"
[17] "not exceed planned budget"
[18] "realise creativity"
[19] "fun and entertainment"
[20] "good company"
[21] "health and beauty"
[22] "free-and-easy-going"
[23] "entertainment facilities"
[24] "not care about prices"
[25] "life style of the local people"
[26] "intense experience of nature"
[27] "cosiness/familiar atmosphere"
[28] "maintain unspoilt surroundings"
[29] "everything organised"
[30] "unspoilt nature/natural landscape"
[31] "cultural offers"
[32] "change of surroundings"
```

Below code generates the summary of the dataset. We have chosen 'gender, 'age', 'income' and 'income2' columns from the above data set.

```
R> summary(vac[, c(1, 2, 4, 5)])
    Gender          Age                        Income
 Female:488   Min.   : 18.00   $30,001 to $60,000   :265
 Male  :512   1st Qu.: 32.00   $60,001 to $90,000   :233
              Median : 42.00   Less than $30,000    :150
              Mean   : 44.17   $90,001 to $120,000  :146
              3rd Qu.: 57.00   $120,001 to $150,000: 72
              Max.   :105.00   (Other)              : 68
                               NA's                 : 66
    Income2
 <30k    :150
 >120k   :140
 30-60k  :265
 60-90k  :233
 90-120k:146
 NA's    : 66
```

4.1 Data Cleaning

This is the important step before connecting data analysis that is to clean the data. Levels of categorical variables can also be checked to ensure they contain only permissible values. Categorical variables can be checked to ensure they contain only permissible values. For example, gender typically has two values in surveys: female and male. Unless the questionnaire did offer a third option, only those two should appear in the data. Cleaning data using code, requires time and discipline, but makes all steps fully documented and reproducible. After cleaning the data set, we save the corresponding data frame.

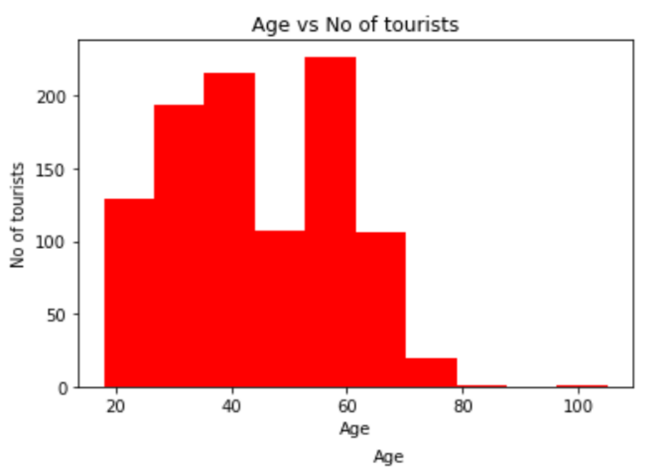
4.2 Descriptive Analysis

Descriptive numeric and graphic representations provide insights into data. Graphical method for numeric data is histograms, box plots, and scatter plots. Bar plots of frequency counts are useful for the visualization of categorical variables. Histograms visualize the distribution of numeric variables. They show how often observations within a certain value range occur.

To obtain a histogram, we first need to create categories of values. We call this binning. The bins must cover the entire range of observations, and must be adjacent to one another. Usually, they are of equal length. Once we have created the bins, we plot how many of the observations fall into each bin using one bar for each bin. We plot the bin range on the x-axis, and the frequency of observations in each bin on the y-axis.

In the vacation dataset we create visualization by introduction of library called matplotlib

```
Import matplotlib.pyplot as plt
plt.hist(vacation1['Age'],bins=10,histtype='barstacked',color='red')
plt.xlabel('Age')
plt.ylabel('No of tourists')
plt.title('Age vs No of tourists')
plt.show()
Ou
```



The boxplot is the most common graphical visualization of unimodal distribution in statistics. The simplest version of a boxplot compresses a data set into minimum, first quartile, median, third quartile and maximum.

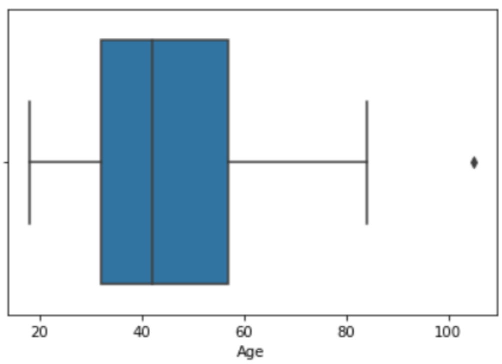
```
vacation1['Age'].describe()
```

```
count    1000.000000
mean       44.168000
std        14.539228
min        18.000000
25%        32.000000
50%        42.000000
75%        57.000000
max       105.000000
```

```
import seaborn as sns

sns.boxplot(vacation1['Age'])
```

Output:



4.3 Pre-processing

1.Categorical variable

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. 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.

2.Numerical variables

Numerical data have the meaning as a measurement or they are a count. For example, one of the segmentation variables is binary (with values 0 or 1 indicating whether or not a tourist likes to dine out during their vacation), and a second variable indicates the expenditure in dollars per person per day (and ranges from zero to $1000), a difference in spend per person per day of one dollar is weighted equally as the difference between liking to dine out or not. To balance the influence of segmentation variables on segmentation results, variables can be standardized. Standardizing variables means transforming them in a way that puts them on a common scale.

4.4 Principal Components Analysis

PCA converts/transforms multivariate data sets containing different variables to a new data set with variables which are basically referred to as principal components which are ordered by importance. The first variable contains most of the variability, the second component contains the second most variability, and so on.

PCA basically works off the covariance or correlation matrix of several numeric variables. Correlation matrix should be used only if the data ranges are different. In most cases, the transformation obtained from principal components analysis is used to project high-dimensional data into lower dimensions for plotting purposes. In this case, only a subset of principal components is used, typically the first few because they capture the most variation. The first two principal components can easily be inspected in a scatter plot. More than two principal components can be visualized in a scatter plot matrix.

As we have taken the Australian travel motives data set. The below code will generate PCA and also check whether the data is standardized. The standard deviations reflect the importance of each principal component.

```
R> vacmot.pca <- prcomp(vacmot)

R> vacmot.pca


Standard deviations (1, .., p=20):
 [1] 0.81 0.57 0.53 0.51 0.47 0.45 0.43 0.42 0.41 0.38
[11] 0.36 0.36 0.35 0.33 0.33 0.32 0.31 0.30 0.28 0.24
```

We can also see the rotation matrix, specifying how to rotate the original data matrix to obtain the principal components.

```
Rotation (n x k) = (20 x 20):
                                   PC1     PC2     PC3
rest and relax                  -0.063  0.0120  0.1345
luxury / be spoilt              -0.109  0.3932 -0.1167
do sports                       -0.095  0.1456 -0.0456
excitement, a challenge         -0.277  0.2227 -0.2103
not exceed planned budget       -0.286 -0.1561  0.5831
realise creativity              -0.110 -0.0122 -0.0153
fun and entertainment           -0.279  0.5205  0.0865
good company                    -0.284 -0.0097  0.1291
health and beauty               -0.140  0.0509  0.0039
free-and-easy-going             -0.317  0.0575  0.2445
entertainment facilities        -0.118  0.3207  0.0050
not care about prices           -0.049  0.2397 -0.2988
life style of the local people  -0.353 -0.2672 -0.3982
intense experience of nature    -0.241 -0.2133 -0.0763
cosiness/familiar atmosphere    -0.132 -0.0133  0.2017
maintain unspoilt surroundings  -0.307 -0.3361  0.0052
everything organised            -0.092  0.1649  0.0780
unspoilt nature/natural landscape -0.269 -0.1831 -0.0556
cultural offers                 -0.260 -0.1160 -0.4282
change of surroundings          -0.259  0.0919  0.1043
```

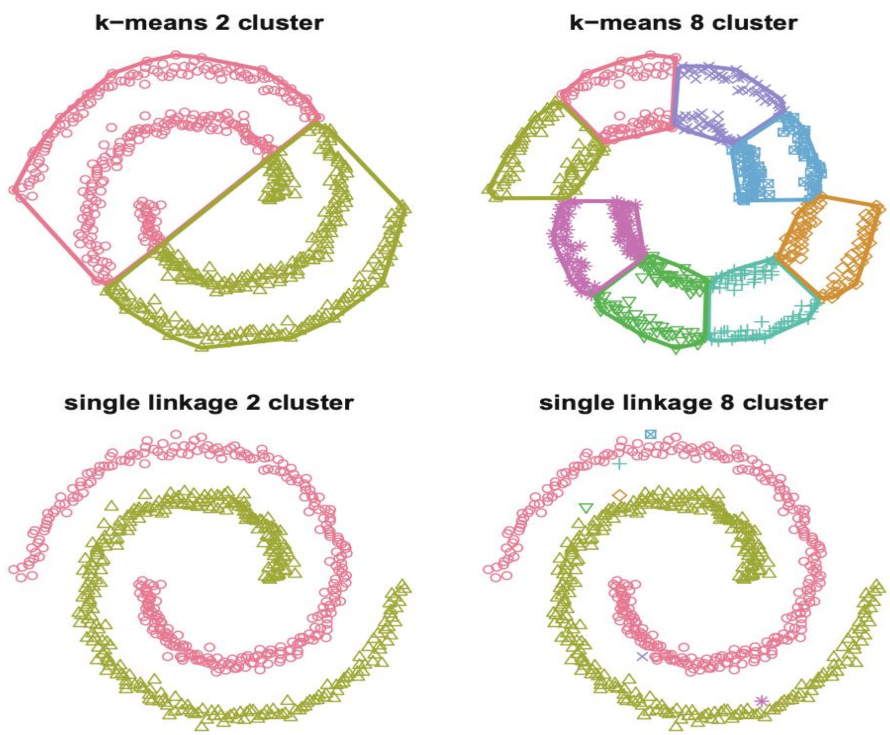
Sometimes principal component analysis is used for the purpose of reducing the number of segmentation variables before extracting market segments from the consumer data. This approach also achieves a reduction in dimensionality, but still works with the original variables collected.

Step 5: Extracting Segments

5.1 Grouping Consumers

Consumers come in all shapes and forms. The combination of exploratory methods and unstructured consumer data results depend on the assumptions made on the structure of the segments implied. The result of a market segmentation analysis is determined by the extraction algorithm chosen.

One of the best examples is in the figure below. It contains two spiraling segments which are segmented using two different algorithms and two different numbers of segments. The top row shows the market segments obtained when running k-means cluster analysis. K-means cluster fails to identify the naturally occurring spiral-shaped segments in the data. The bottom of the below figure shows the market segments obtained from single linkage hierarchical clustering.



The algorithm correctly identifies the spiraling segments since the single linkage method constructs snake-shaped clusters. Outliers are defined as micro-segments, but the two spirals are correctly identified where in k-means clusters ignore the spiral structure and fail to identify the spirals because it is designed to construct round, equally sized clusters.

5.2 Distance-Based Methods

Distance-based methods are used to find the similarity or distance between observations (consumers) and try to find groups of similar observations (market segments). Consider the problem of finding groups of tourists with similar activity patterns when on vacation. Data set is shown below.

	beach	action	culture
Anna	100	0	0
Bill	100	0	0
Frank	60	40	0
Julia	70	0	30
Maria	80	0	20
Michael	0	90	10
Tom	50	20	30

It contains tourists and their actions. Market segmentation aims at grouping consumers into groups with similar needs or behavior. A distance measure is used to find out the similarity or dissimilarity.

5.3 Distance Measures

Using the above dataset, each row represents an observation (tourist) and every column represents a variable (a vacation activity). This can be represented as $n \times p$ matrix where n is number of observations (rows) and p is the number of variables (columns):

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}$$

The vector corresponding to the i-th row of matrix **X** is denoted as **x**i = (xi1,xi2,...,xip)′ in the following, such that X = {**x**1,**x**2,...**x**p} is the set of all observations. In the example above, Anna's vacation activity profile is vector **x**1 = (100, 0, 0) ′ and Tom's vacation activity profile is vector **x**7 = (50, 20, 30)′. Distance (*d*) is calculated between x and y.

$$d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$$

A second criterion is that the distance of a vector to itself and only to itself is 0:

$$d(\mathbf{x},\mathbf{y})=0 \Leftrightarrow \mathbf{x}=\mathbf{y}.$$

In addition, most distance measures fulfill the so-called triangle inequality:

$$d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}).$$

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:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{j=1}^{p} (x_j - y_j)^2}$$

Manhattan or absolute distance:

$$d(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{p} |x_j - y_j|$$

Asymmetric binary distance*:* applies only to binary vectors, that is, all xj and yj are either 0 or 1.

$$d(\mathbf{x}, \mathbf{y}) = \begin{cases} 0, & \mathbf{x} = \mathbf{y} = \mathbf{0} \\ (\#\{j|x_j = 1 \text{ and } y_j = 1\})/(\#\{j|x_j = 1 \text{ or } y_j = 1\}) \end{cases}$$

Euclidean distance is the most common distance measure in market segmentation analysis. Euclidean distance is between two direct straight lines where Manhattan distance is between two points that need to be used to get from one point to another.

We can calculate the distance in standard R function as dist(). Using the vacation dataset

R> data("annabill", package = "MSA")

Then, we can calculate the Euclidean distance between all tourists with the following command:

R> D1 <- dist(annabill)
R> round(D1, 2)

	Anna	Bill	Frank	Julia	Maria	Michael
Bill	0.00					
Frank	56.57	56.57				
Julia	42.43	42.43	50.99			
Maria	28.28	28.28	48.99	14.14		
Michael	134.91	134.91	78.74	115.76	120.83	
Tom	61.64	61.64	37.42	28.28	37.42	88.32

The distance between Anna and Bill is zero since they have identical vacation activity profiles. The distance between Michael and all other people in the data set is substantial because Michael does not go to the beach where most other tourists spend a lot of time.

Manhattan distance – which is also referred to as absolute distance – is very similar to Euclidean distance for this data set:

R> D2 <- dist(annabill, method = "manhattan")
R> D2

```
        Anna Bill Frank Julia Maria Michael
Bill       0
Frank     80   80
Julia     60   60    80
Maria     40   40    80    20
Michael  200  200   120   180   180
Tom      100  100    60    40    60     140
```

In the R package cluster, function daisy calculates the dissimilarity matrix between observations contained in a data frame. If the variables are metric, the results are the same as for dist:

R> library("cluster")
R> round(daisy(annabill), digits = 2)

```
Dissimilarities :
          Anna   Bill  Frank  Julia  Maria Michael
Bill      0.00
Frank    56.57  56.57
Julia    42.43  42.43  50.99
Maria    28.28  28.28  48.99  14.14
Michael 134.91 134.91  78.74 115.76 120.83
Tom      61.64  61.64  37.42  28.28  37.42   88.32

Metric :  euclidean
Number of objects : 7
```

5.4 Hierarchical Methods

Hierarchical clustering refers to grouping of data that how a human would approach the task of dividing a set of n observations (consumers) into k groups (segments). If k = 1, one big market segment containing all consumers in data X. If k = n, then each segment contains exactly one consumer where each consumer represents their own cluster. Market segmentation analysis occurs between those two extremes.

One such clustering is called Divisive hierarchical clustering where Complete dataset X splits into two market segments and it further splits into two segments till each consumer has their own market segment.

Agglomerative hierarchical clustering approaches differently. Each consumer represents their own market and two market segments closest to one another merged until the complete data set forms one large market segment.

Distance between groups of observations (segments) can be determined as follows:

Single linkage: distance between the two closest observations of the two sets, where x and y are consumers.

$$l(\mathcal{X}, \mathcal{Y}) = \min_{\mathbf{x}\in\mathcal{X}, \mathbf{y}\in\mathcal{Y}} d(\mathbf{x}, \mathbf{y})$$

Complete linkage: distance between the two observations of the two sets that are farthest away from each other.

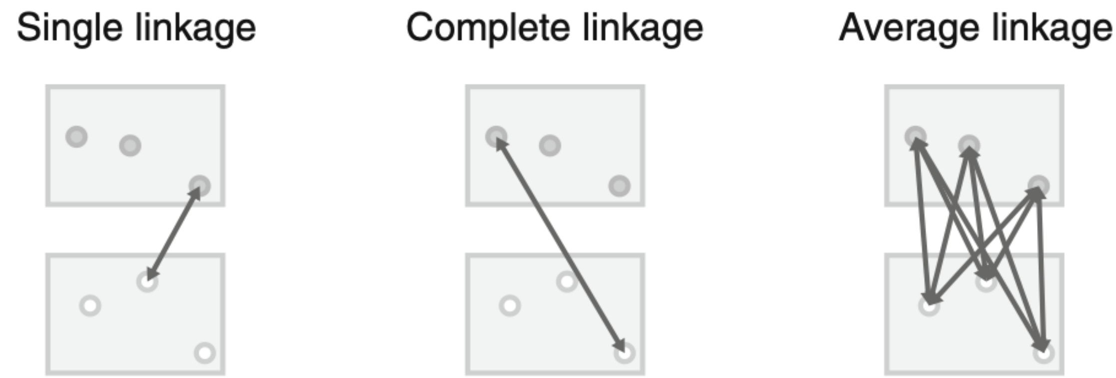
$$l(\mathcal{X}, \mathcal{Y}) = \max_{\mathbf{x}\in\mathcal{X}, \mathbf{y}\in\mathcal{Y}} d(\mathbf{x}, \mathbf{y})$$

Average linkage: mean distance between observations of the two sets.

$$l(\mathcal{X}, \mathcal{Y}) = \frac{1}{|\mathcal{X}||\mathcal{Y}|} \sum_{\mathbf{x}\in\mathcal{X}} \sum_{\mathbf{y}\in\mathcal{Y}} d(\mathbf{x}, \mathbf{y}),$$

where $|X|$ denotes the number of elements in X.

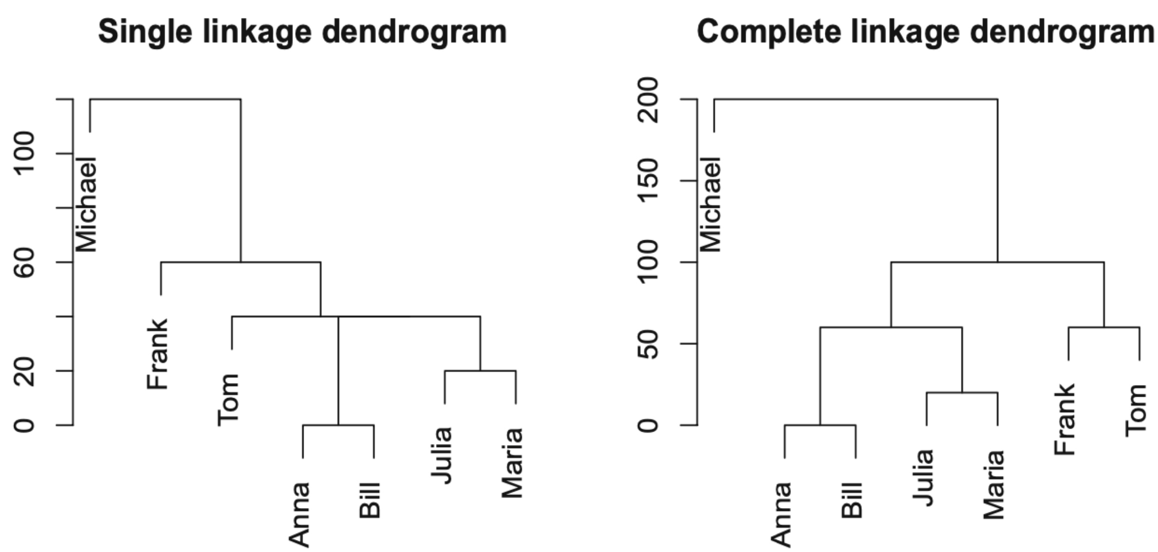
Below illustration shows the linkage method. Clustering in general are exploratory techniques. Different combinations can reveal different features of the data. Linkage technique uses closest neighbor or next neighbor to calculate distance.



A clear illustration for clustering can be done using dendrogram. A dendrogram is a tree diagram. It contains the root and leaves of a tree, where root represents the one-cluster (one market segment contains all consumers) and leaves represent single observations (consumers). Branches represent hierarchy of market segments.

The height of the branch point refers to the distance between the clusters. Higher branches point to more distinct market segments. Dendrograms are often recommended as a guide to select the number of market segments.

An illustration using vacation dataset is given below.



5.5 Partitioning Methods

Hierarchical clustering methods are used to analyse small datasets with up to a few hundred observations. Dendrograms are used for larger datasets since they are hard to read and the pairwise distances do not fit into computer memory. Single partition is suitable for data sets containing more than 1000 observations (consumers). Only distances between each consumer in the data set and the center of the segments are computed rather than computing all distances.

k-Means and k-Centroid Clustering

The most popular partitioning method is k-means clustering and a number of algorithms are available. R function k-means () implements the algorithm. These algorithms use the squared Euclidean distance. A generalization to other distance measures, also referred to as k-centroid clustering, is provided in R package flexclust.

Let X = {**x**1, . . . , **x**n} be a set of observations (consumers) in a data set. This dataset contains observations (consumers in rows, and variables (behavioural information) in columns.

To solve the problem of dividing consumers into a given number of segments, an algorithm is followed and it involves five steps.

1. Specify the desired number of segments k.

2. Select k observations (consumers) randomly from data set *X* and use them as the initial set of cluster centroids C = {**c**1, . . . , **c**k}. Choose five consumers randomly and declare them as representatives.

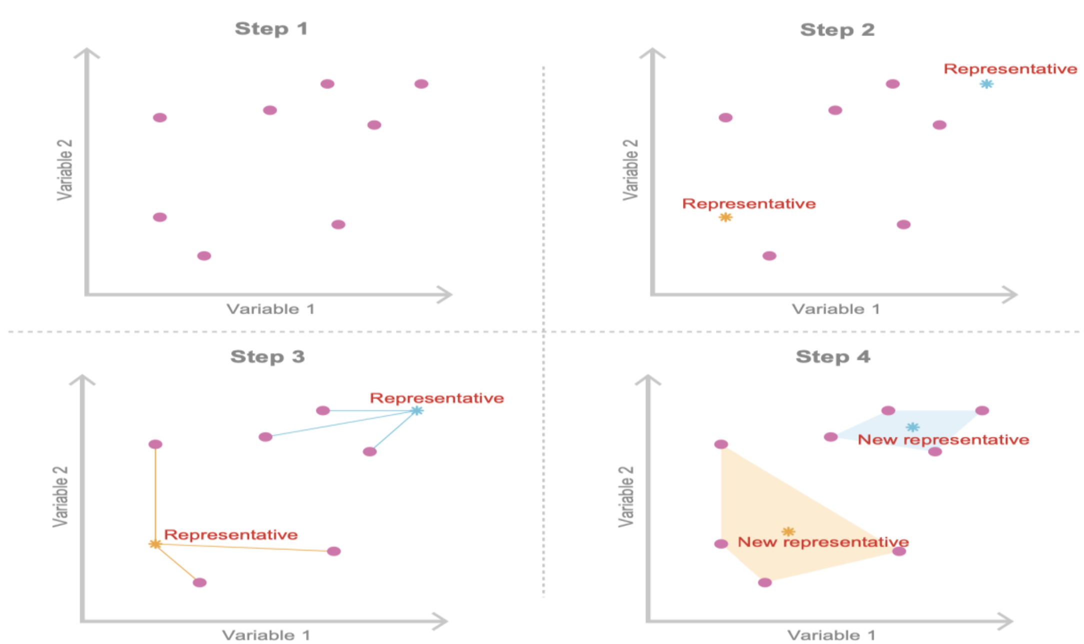
3. Assign each observation ***xi*** to the closest cluster centroid to form a partition of the data,

$$\mathcal{S}_j = \{\mathbf{x} \in \mathcal{X} | d(\mathbf{x}, \mathbf{c}_j) \leq d(\mathbf{x}, \mathbf{c}_h),\ 1 \leq h \leq k\}.$$

4. Fix the cluster membership and minimize the distance from each consumer to the corresponding cluster centroid.

$$\mathbf{c}_j = \arg\min_{\mathbf{c}} \sum_{\mathbf{x} \in \mathcal{S}_j} d(\mathbf{x}, \mathbf{c}).$$

5. Repeat from step 3 until convergence or a pre-specified maximum number of iterations is reached



“Improved” k-means

It is not advisable to draw k-means clustering values randomly as starting points. Using randomly drawn consumers being located too close to one another, is not being representative of the data space. Thus, it is referred to as local optimum where the k-means algorithm gets stuck in the data space where the starting points are not representative of the data.

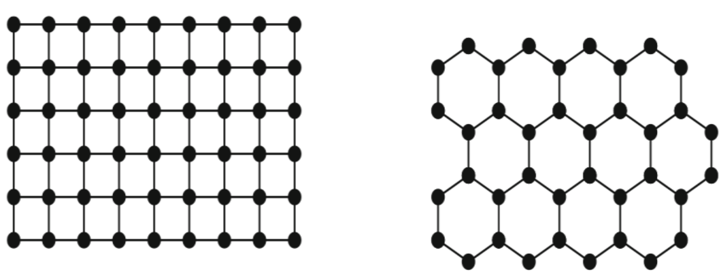
To avoid such problems, drawing starting points which are evenly spread across the entire data space would represent the entire data set in a better way. Steinley and Brusco conducted a simulation using artificial data sets of known structure. Where they concluded that the best approach is to randomly draw many starting points and select the best one.

5.6 Hard Competitive Learning

It is also known as learning vector quantisation, differs from k-means algorithm by randomly picks one consumer and moving this consumer’s closest segment representative a small step into the direction of the randomly chosen consumer.

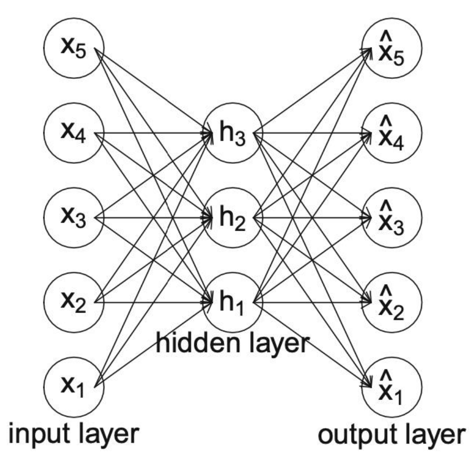
Self-Organising Maps / Kohonen maps

Self-organizing maps are another hard competitive learning in which the self-organizing maps segment representatives (centroids) on a regular grid, usually a rectangular or hexagonal grid.



5.7 Neural Networks

The most popular method from this family of algorithms uses a so-called single hidden layer perceptron.



The network has three layers and each layer takes the data as input. The output layer gives the response of the network. There is a layer in-between input and output as a hidden layer, since it has no connections to the outside of the network. The input layer has a node for every segmentation variable and the three hidden layers are h1, h2 and h3 are weighted linear combinations of the inputs for a non-linear function fj.

$$h_j = f_j\left(\sum_{i=1}^{5} \alpha_{ij} x_i\right)$$

Each weight αij in the formula is depicted by an arrow connecting nodes in the input layer and hidden layer. The fj is chosen such that 0 ≤ hj ≤1, and all hj sum up to one (h1 +h2 +h3 =1)
where coefficients βji correspond to the arrows between hidden nodes and output nodes.

$$\hat{x}_i = \sum_{j=1}^{3} \beta_{ji} h_i,$$

5.8 Hybrid Approaches

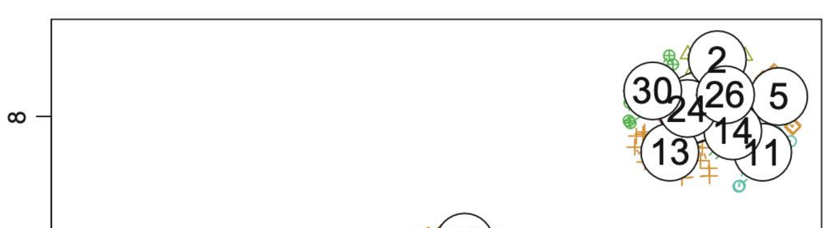
The basic idea behind hybrid segmentation approaches is to first run a partitioning algorithm because it can handle data sets of any size.

Two-Step Clustering

The two steps consist of running a partitioning procedure followed by a hierarchical procedure. The idea can be demonstrated using simple R commands using an artificial dataset which has been used partitioning clustering. Assuming k=30, with much larger than the number of market segments sought.

```
R> set.seed(1234)
R> PF3.k30 <- stepcclust(PF3, k = 30, nrep = 10)
R> plot(PF3.k30, data = PF3)
```

Only 30 clusters are considered from the 500 observations. The choice of the original cluster is not crucial because the primary aim of the first is to reduce the size of the data set by retaining only one representative member of each of the extracted clusters. Such application of cluster methods is often also referred to as vector quantisation.



```
R> PF3.k30.cent <- parameters(PF3.k30)
R> sizes <- table(clusters(PF3.k30))
R> PF3.hc <- hclust(dist(PF3.k30.cent), members = sizes)
```

Upon further proceeding with the codes, cluster centers and segment sizes from the k-means are extracted.

5.9 Bagged Clustering

Bagged clustering also combines hierarchical clustering algorithms and partitioning clustering algorithms, but adds bootstrapping. Bootstrapping can be implemented by random drawing from the dataset with replacement, i.e., the process of extracting segments is repeated many times with randomly drawn (bootstrapped) samples of the data. It has the advantage of making the final segmentation solution less.

5.10 Model-Based Methods

This method does not use similarities or distances to assess which consumers should be assigned to the same market segment. Instead, they assume that the true market segmentation solution which is unknown and has the following two general properties

1. Each market segment has a certain size
2. If a consumer belongs to market segment A, that consumer will have characteristics which are specific to members of market segment A.

Model-based methods use the empirical data to find those values for segment sizes and segment-specific characteristics that best reflect the data. The model-based method used here is called finite mixture models because the number of market segments is finite, and the overall model is a mixture of segment-specific modes. The two properties of the finite mixture model are,

Property 1: Segment membership z of a consumer is determined by the multinomial distribution with segment size π:

$$z \sim \text{Multinomial}(\pi)$$

Property 2: The members of each market segment have segment-specific characteristics. It is captured by the vector θ. Function *f()*, together with θ, capture the specific values *y* for given segment membership z

$$f(y|x, \theta_z).$$

This leads to the following finite mixture model:

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

5.11 Finite Mixtures of Distributions

The simplest case of model-based clustering has no independent variables x, and simply fits a distribution to y. Finite mixtures of distributions basically use the same segmentation variables and a number of pieces of information about consumers. The finite mixture model reduces to

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

Normal Distributions

The most popular finite mixture model is a mixture of several multivariate normal distributions. The multivariate normal distribution can easily model covariance between variables and approximate multivariate normal distributions.

Binary Distributions

Finite mixture of binary distributions (binary data), also referred to as latent class models or latent class analysis are popular. In this case, the p segmentation variables in the vector y are not metric, but binary (either 0 or 1). This model assumes that respondents in different segments have different probabilities of undertaking certain activities.

5.12 Finite Mixtures of Regressions

This method is similar to distance-based clustering methods and, in many cases, result in similar solutions. Finite mixtures of regression models offer a completely different type of market segmentation analysis. It assumes that dependent target variable y can be explained by a set of independent variables x.

These are more complicated than distance-based methods. The complexity makes it very flexible. For binary data, we can use mixtures of binary distributions and for nominal variables, we can use mixtures of multinomial distributions or multinomial logit models. Mixture models also allow to simultaneously include segmentation and descriptor variables.

Algorithms with Integrated Variable Selection

Most algorithms focus only on extracting segments from data. These algorithms assume that each of the segmentation variables contributes in determining the segmentation solution. Variable selection for binary data is more challenging because single variables are not informative for clustering, making it impossible to pre-screen or pre-filter variables one by one.

5.13 Biclustering Algorithms

This algorithm simultaneously clusters both consumers and variables, this exists for any kind of data, including metric and binary. In the simplest case, a bicluster is defined for binary data as a set of observations with values of 1 for a subset of variables. This algorithm which extracts the bicluster follows a sequence of steps.

1. Rearranging rows and columns of data matrix to create a rectangle with identical entries of 1s.The aim is for this rectangle to be as large as possible.
2. Assign the observations falling into this rectangle to one bicluster.
3. Remove from the data matrix the rows containing the consumers who have been assigned to the first bicluster. Once removed, repeat the procedure from step 1 until no more biclusters of sufficient size can be located.

Biclustering is not one single very specific algorithm; rather it is a term describing a family of algorithms differing with respect to the properties of data they can accommodate, the extent of similarity between members of market segments required, and whether individual consumers can be assigned to only one or multiple market segments. Biclustering is particularly useful in market segmentation applications with many segmentation variables.

It also has some other advantages:

- ➢ No data transformation
- ➢ Ability to capture niche markets

Variable Selection Procedure for Clustering Binary Data (VSBD)

This method is based on the k-means algorithm as a clustering method, and Brusco assumes that not all variables available are relevant to obtain a good clustering solution. In particular, the method assumes the presence of masking variables.

The procedure first identifies the best small subset of variables to extract segments. Because the procedure is based on the k-means algorithm, the performance criterion used to assess a specific subset of variables is the within-cluster sum-of-squares. This is the criterion minimized by the k-means algorithm. After having identified this subset, the procedure adds additional variables one by one. The variable added is the one leading to the smallest increase in the within-cluster sum-of-squares criterion. The procedure stops when the increase in within-cluster sum-of-squares reaches a threshold.

Variable Reduction: Factor-Cluster Analysis

The term factor-cluster analysis refers to a two-step procedure of data-driven market segmentation analysis.

In the first step, segmentation variables are factor analyzed. The raw data, the original segmentation variables, are then discarded. In the second step, the factor scores resulting from the factor analysis are used to extract market segments.

Some of the disadvantages of this analysis is as follows:

- Factor analyzing data leads to a substantial loss of information
- Factor analysis transforms data
- Factors-cluster results are more difficult to interpret.

5.14 Data Structure Analysis

Data structure analysis provides valuable insights into the properties of the data. These insights guide subsequent methodological decisions. Most importantly, stability-based data structure analysis provides an indication of whether natural, distinct, and well-separated market segments exist in the data or not.

If there is structure in the data, be it cluster structure or structure of a different kind, data structure analysis can also help to choose a suitable number of segments to extract. There are four different approaches to data structure analysis:

- cluster indices
- gorge plots
- global stability analysis
- segment level stability analysis.

Cluster Indices

Cluster indices provide insight into particular aspects of the market segmentation solution. Which basically, depends on the nature of the cluster index used.

Generally, two groups of cluster indices are distinguished:

- **Internal cluster indices:** These are calculated on the basis of one single market segmentation solution, and use information contained in this segmentation solution to offer guidance.

- **External cluster indices**: The external cluster index measures the similarity between two segmentation solutions. If the correct market segmentation is known, the correct assignment of members to segments serves as the additional input.

Gorge Plots

A simple method to assess how well segments are separated, is to look at the distances of each consumer to all segment representatives. If natural, well-separated market segments are present in the data, we expect the gorge plot to contain many very low and many very high values. This is why this plot is referred to as gorge plot. Optimally, it takes the shape of a gorge with a peak to the left and a peak to the right. For a real market segmentation analysis, gorge plots have to be generated and inspected for every number of segments.

Global Stability Analysis

To assess the global stability of any given segmentation solution, several new data sets are generated using resampling methods, and a number of segmentation solutions are extracted. Resampling methods- It is combined with many repeated calculations using the same or different algorithms which provide critical insight into the structure of the data.

Global stability analysis helps determine which of the concepts applies to any given data set. This acknowledges that both the sample of consumers, and the algorithm used in data-driven segmentation introduce randomness into the analysis. It also provides valuable guidance for selecting the number of segments to extract. However, global stability does not provide information about the stability of each one of the segments individually in the four-segment solution

Segment Level Stability Within Solutions (SLSW)

SLSW measures how often a market segment with the same characteristics is identified across a number of repeated calculations of segmentation solutions with the same number of segments. It is calculated by drawing several bootstrap samples, calculating segmentation solutions independently for each of those bootstrap samples, and then determining the maximum agreement across all repeated calculations.

To assess segment level stability within solutions (SLSW), we use the following R commands:

```
R> PF3.r3 <- slswFlexclust(PF3, PF3.k3)
R> PF3.r6 <- slswFlexclust(PF3, PF3.k6
```

Segment Level Stability Across Solutions (SLSA)

Segment Level Stability Across Solutions (SLSA) serve as indicators of market segments occurring naturally in the data, rather than being artificially created. Natural segments are more attractive to organisations because they actually exist, and no managerial judgment is needed in the artificial construction of segments.

SSLSA can be calculated in combination with any algorithm which extracts segments. However, for hierarchical clustering, segment level stability across solutions will reflect the fact that a sequence of nested partitions is created. If partitioning methods (k-means, k-medians, neural gas, ...) or finite mixture models are used, segmentation solutions are determined separately for each number of segments k.

An algorithm to renumber a series of partitions (segmentation solutions), which is implemented in function relabel() in package flexclust. This function was used to renumber segmentation solutions. We create the SLSA plot using the command slsaplot(PF3.k38) from package flexclust. This plot shows the development of each segment across segmentation solutions with different numbers of segments.

Step 6: Profiling Segments

6.1 Identifying Key Characteristics of Market Segments

The aim of the profiling step is to get to know the market segments resulting from the extraction step. Profiling is only required when data-driven market segmentation is used. For commonsense segmentation, the profiles of the segments are predefined. If, for example, age is used as the segmentation variable for the commonsense segmentation, it is obvious that the resulting segments will be age groups. Therefore, Step 6 is not necessary when commonsense segmentation is conducted.

Identifying these defining characteristics of market segments with respect to the segmentation variables is the aim of profiling. Profiling consists of characterizing the market segments individually, but also in comparison to the other market segments.

6.2 Traditional Approaches to Profiling Market Segments

We use the Australian vacation motives data set. Segments were extracted from this data set using the neural gas clustering algorithm with the number of segments varied from 3 to 8 and with 20 random restarts. We reload the segmentation solution derived and saved on page:

```
R> library("flexclust")
R> data("vacmot", package = "flexclust")
R> load("vacmot-clusters.RData")
```

Data-driven segmentation solutions are usually presented to users (clients, managers) in one of two ways:

1. as high-level summaries simplifying segment characteristics to a point where they are misleadingly trivial, or
2. as large tables that provide, for each segment, exact percentages for each segmentation variable

To identify the defining characteristics of the market segments, the percentage value of each segment for each segmentation variable needs to be compared with the values of other segments or the total value provided in the far-right column.

6.3 Segment Profiling Visualization

Graphics are particularly important in exploratory statistical analysis (like cluster analysis) because they provide insights into the complex relationships between variables. In times of big and increasingly bigger data, visualization offers a simple way of monitoring developments over time. So, there is high usage of visualization techniques to make the results of a market segmentation analysis easier to interpret.

Visualizations are useful in the data-driven market segmentation process 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.

Identifying Defining Characteristics of Market Segments

A good way to understand the defining characteristics of each segment 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. The segment profile plot is the direct visual translation of tables; the option is to order segmentation variables by similarity of answer patterns. We can achieve this by clustering the columns of the data matrix:

```
R> vacmot.vdist <- dist(t(vacmot))
R> vacmot.vclust <- hclust(vacmot.vdist, "ward.D2")
```

The t() around the data matrix vacmot transposes the matrix such that distances between columns rather than rows are computed. Next, hierarchical clustering of the variables is conducted using Ward's method. The segment profile plot is a also called panel plot. For each segment, the segment profile plot shows the cluster centers (centroids, representatives of the segments).

Marker variables are defined as variables which deviate by more than 0.25 from the overall mean. The heat map suggests that it took less effort to find the information required to answer the question. It is therefore well worth spending some extra time on presenting results of a market segmentation analysis as a well-designed graph. Good visualizations facilitate interpretation by managers who make long-term strategic decisions based on segmentation results. Such long-term strategic decisions imply substantial financial commitments to the implementation of a segmentation strategy.

6.4 Assessing Segment Separation

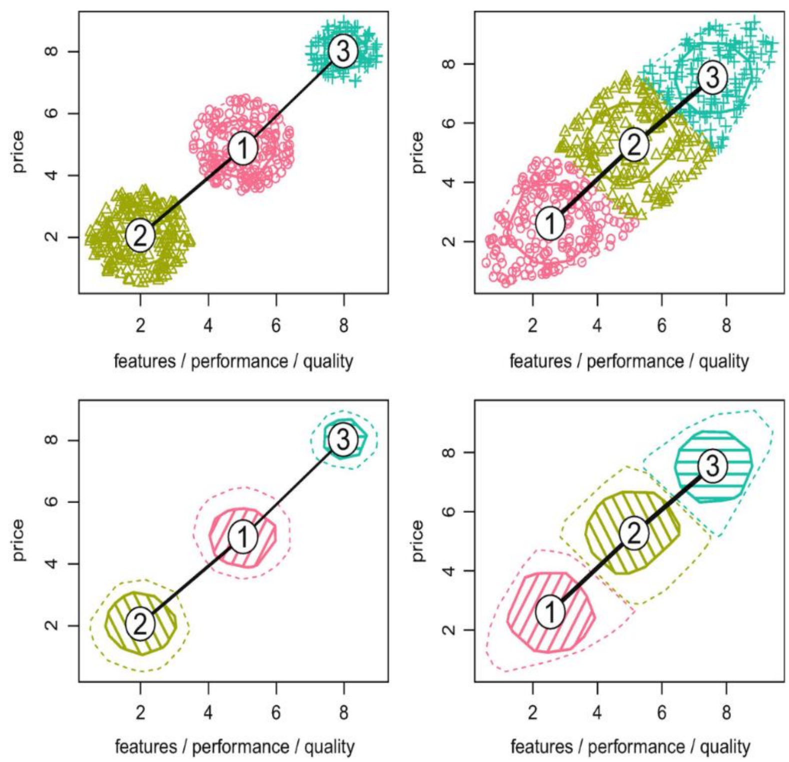
Segment separation can be visualized in a segment separation plot. The segment separation plot depicts – for all relevant dimensions of the data space - the overlap of segments. Segment separation plots are very simple if the number of segmentation variables is low, but become complex as the number of 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.

The two dimensions representing the segmentation variables can be directly plotted. This is not possible if 20-dimensional travel motives data serve as segmentation variables. In such a situation, the 20-dimensional space needs to be projected onto a small number of dimensions to create a segment separation plot. We can use a number of different projection techniques, including some which maximize separation, and principal components analysis. We calculate principal components analysis for the Australian travel motives data set with the following command:

R> vacmot.pca <- prcomp(vacmot)

This provides the rotation applied to the original data when creating our segment separation plot. We use the segmentation solution obtained from neural gas, and create a segment separation plot for this solution:

R> plot(vacmot.k6, project = vacmot.pca, which = 2:3,
+ xlab = "principal component 2",
+ ylab = "principal component 3")



R> projAxes(vacmot.pca, which = 2:3

Due to the overlap of market segments (and the sample size of n = 1000), the plot is messy and hard to read. Modifying colours (argument col), omitting observations (points = FALSE), and highlighting only the inner area of each segment (hull.args = list(density = 10), where density specifies how many lines shade the area) leads to a cleaner version) :

R> plot(vacmot.k6, project = vacmot.pca, which = 2:3,
+ col = flxColors(1:6, "light"),
+ points = FALSE, hull.args = list(density = 10),
+ xlab = "principal component 2",
+ ylab = "principal component 3")

```
R> projAxes(vacmot.pca, which = 2:3, col = "darkblue",
+ cex = 1.2)
```

It is hard to interpret, because natural market segments are not present. This difficulty in interpretation is due to the data, not the visualization. And the data used for this plot is very representative of consumer data.Each segment separation plot only visualizes one possible projection. So, for example, the fact that segments 1 and 5 in this particular projection overlap with other segments does not mean that these segments overlap in all projections. However, the fact that segments 6 and 3 are well-separated in this projection does allow the conclusion – based on this single projection only – that they represent distinctly different tourists in terms of the travel motives.

Step 7: Describing Segments

7.1 Developing a Complete Picture of Market Segment

Understanding the variations in segmentation variables across market segments is key to market segmentation. Early in the market segmentation analysis process, segmentation variables are selected conceptually in Step 2 (specifying the ideal target) and empirically in Step 3 (Collecting data).

We can use either descriptive statistics or inferential statistics to investigate variations between market segments with regard to descriptor variables. Segment descriptions are easier to understand when they are visualized.

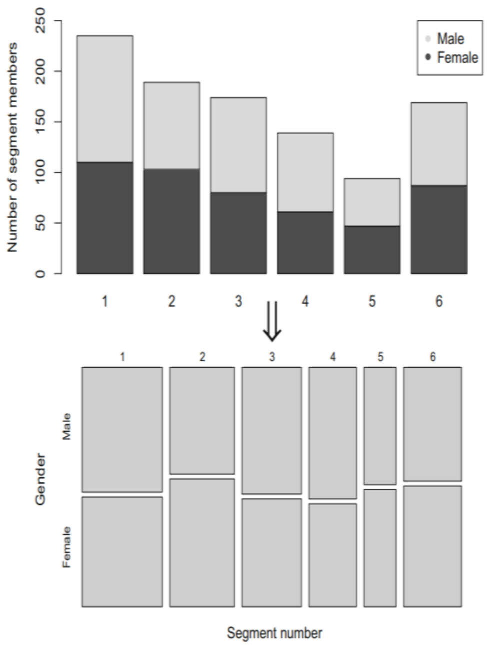
7.2 Using Visualizations to Describe Market Segments

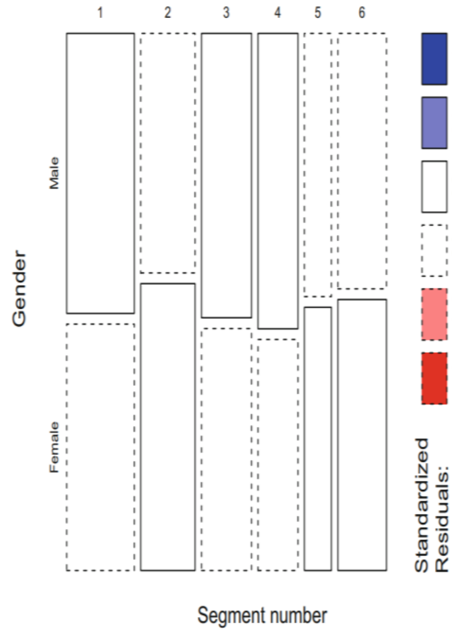
Two major benefits of using graphical statistics to represent market segmentation are that it makes it easier for users and data analysts to understand the results. Additionally, it prevents over interpreting small differences and the statistical importance of differences.

Nominal and Ordinal Descriptor Variables

All visualizations and statistical testing are based on a cross-tabulation of market sectors. Many descriptor variables are present in the data frame vacmotdesc for the Australian travel motives data set. The data set automatically loads these descriptors as well. We require the segment membership for each respondent in order to describe market segments.

Mosaic plots can incorporate components of inferential statistics and visualize tables with more than two descriptor variables. Cellular colors can be used to indicate areas where observed and expected frequencies diverge, if the variables are assumed to be independent. The absolute segment size is shown by the width of the bars. The ratio of men to women in each segment is shown by the height of the rectangles. Each cell in the table has an area that is proportional to its size.



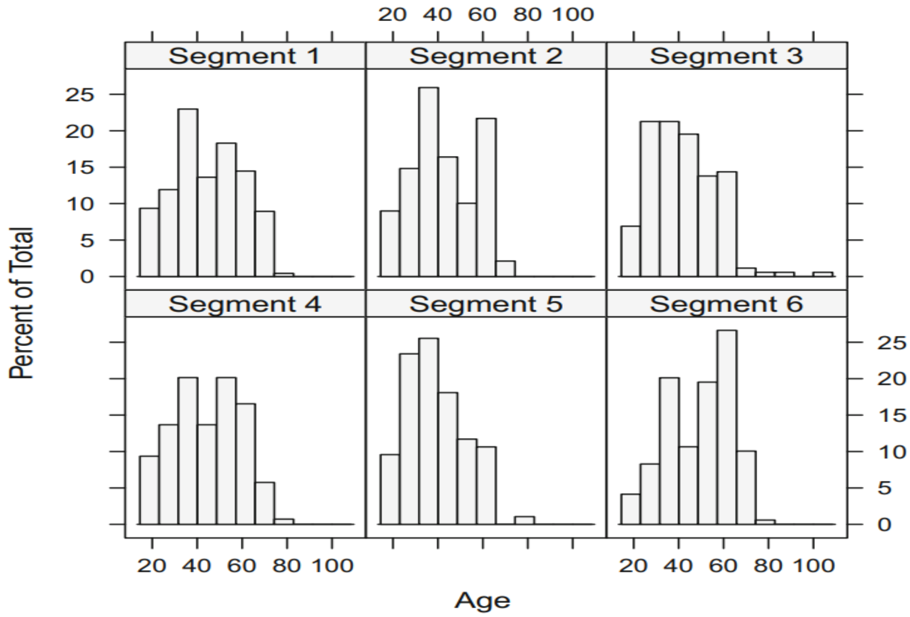


In Second figure all-white cells show that there are no appreciable gender distribution differences among the six market sectors taken from the Australian travel reasons data set. In general, there are about equal numbers of male and female tourists in each sector. Those in Segment 4 (column 4 in Fig.), who are driven by cultural offers and curious about the locals, make more money. Few people with very high earnings make up segment 6, which is for people who love the outdoors.

Figure uses a mosaic plot to visually represent the cross-tabulation of segment membership and declared moral commitment to safeguard the environment. There are much fewer people in Segment 3 (whose members want entertainment) who fall under the category of having a high moral commitment.

7.3 Metric Descriptor Variables

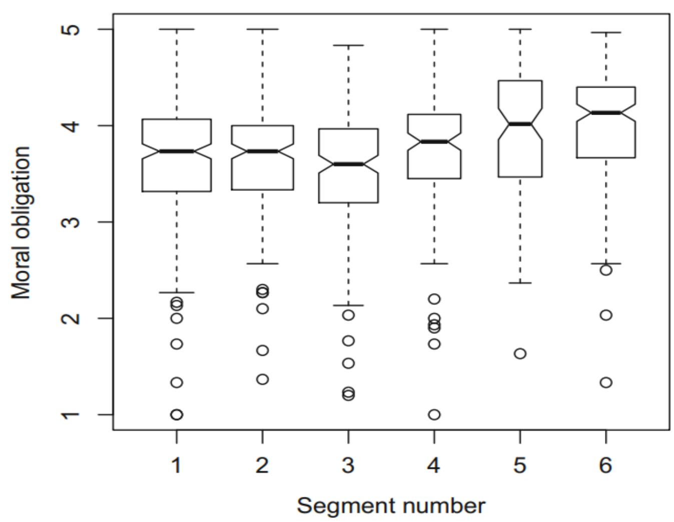
Conditional versions of the majority of common R plots are available in the R packages lattice (Sarkar, 2008) and ggplot2 (Wickham, 2009). Using metric descriptor variables, conditional charts are a good tool for illustrating variations across market groups. Using the following command, the R package lattice generated the segment profile plot in Section 8.3.1 and produced a parallel box-and-whisker plot for age by market segment: R> box plot(Age C6, data = vacmotdesc, + xlab = "Segment number", ylab = "Age"), where xlab and ylab are options that allow the axis labels to be customized. The final histograms are displayed in Figs.



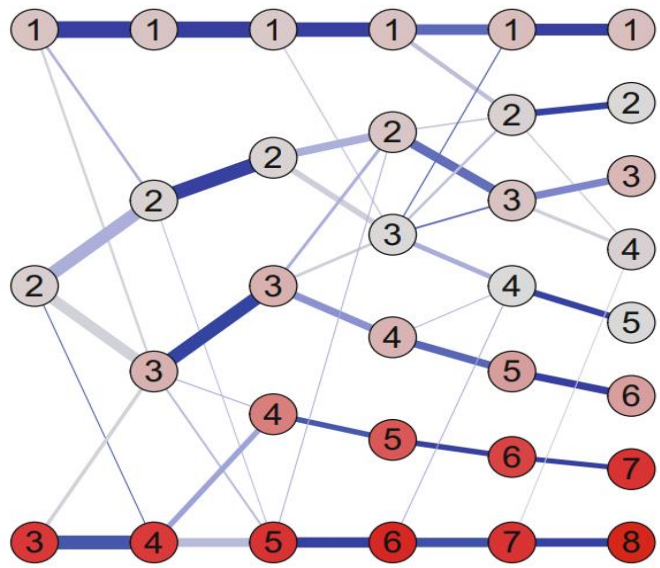
A similar box-and-whisker plot may be made with the following command to acquire further insights:

```
R> boxplot(Obligation ~ C6, data = vacmotdesc,
+   varwidth = TRUE, notch = TRUE,
+   xlab = "Segment number",
+   ylab = "Moral obligation")
```

Segment 5 is shown to be the smallest and have the narrowest box. The largest section is segment 1, and individuals of segment 6 have the highest moral commitment to preserve the environment. These two segments' notches are very different from one another, yet all segments include some outliers who go below the minimum acceptable standard of moral behavior. To follow the evolution of the value of a metric descriptor variable over a range of market situations, one can employ a modified version of the segment level stability across solutions (SLSA) plot.



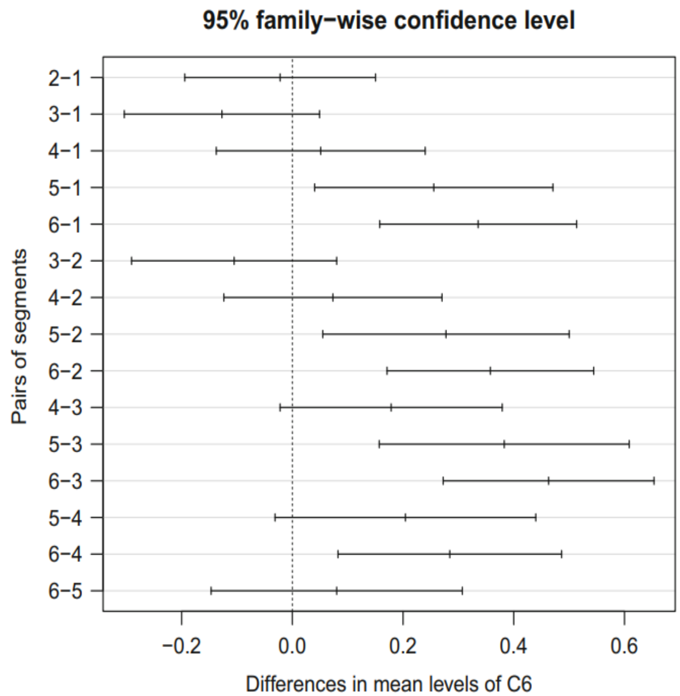
The modified segment level stability across solutions (SLSA) plot depicted below Fig. Uses various colors to represent each segment's mean moral duty to preserve the environment. The segment indicated as a potentially appealing market segment (nature-loving travelers interested in the local inhabitants) has a high sense of moral duty, followed by the segment classified as possessing responses biased towards acquiescence (yes saying) (segment 5 in the six-segment solution). The edges' coloring corresponds to the numeric SLSA value, with dark blue edges denoting high stability values and light grey edges denoting poor stability values.



7.4 Testing for Segment Differences in Descriptor Variables

Formally testing for variations in descriptive variables across market segments can be done using straightforward statistical tests. P-values that are low (usually less than 0.05) are interpreted as indicating that there are differences in the gender distribution between segments. The relationship between the nominal segment membership variable and other variables (such as gender, educational attainment, and country of origin) is visualized using the cross-tabulation of both variables as basis for the mosaic plot.

Parallel box plots are used to visualize the relationship between segment membership and metric variables (such as age, the number of nights spent visiting tourist locations, and the amount of money spent on lodging). The null hypothesis for the Kruskal-Wallis rank sum test is that the median is the same across all segments. Yet, the study of variance fails to pinpoint the various parts. If segment 1 is compared to segment 2 or segment 2, the p-value for the t-test is the same. The benefit of this output is that it condenses the findings into a manageable package.



In this situation, there is only one possible conclusion: all segment means are equal. If at least one of the reported P-values is below the significance level, adjusting the p-values allows for the rejection of the null hypothesis that the segment means are all the same. Fig. shows that segments 1, 2, 3, and 4 do not significantly differ from one another in terms of moral duty. The moral imperative to act sustainably is substantially greater in segments 5 and 6 than it is in the other market categories (with the only exception of segments 4 and 5 not differing significantly).

7.5 Predicting Segments from Descriptor Variables

Prediction models are based on regression analysis, which holds that independent variables, or repressors', x1,..., xp, can predict a dependent variable, y. Assuming that function f () is linear and that y follows a normal distribution with a mean f (x1,...,xp) and variance 2, the most fundamental regression model is the linear regression model. The dependent variable AGE is shown on the left side of the formula interface used to specify regression models. Segment membership C6 is a factor in the data frame and a categorical variable with six categories. The formula interface fits a regression coefficient for each category and appropriately understands categorical variables. According to the linear regression model, changes brought about by adjustments to one independent variable are independent of changes in the absolute levels of the other independent variables.

This is included by generalized linear models by including a connection function, g(). The link function expands the range denoted by from the mean value of y given by to an infinite range. A linear function can then be used to model this altered value. For the dependent variable, we can utilize the normal, Poisson, binomial, and multinomial distributions. We address two special situations of generalized linear regression in the sections that follow.

$$g(\mu) = \eta = \beta_0 + \beta_1 x_1 + \ldots + \beta_p x_p.$$

Binary Logistic Regression

Given a consumer's age and moral responsibility score, Glm() in R builds generalized linear models to estimate the likelihood that the consumer belongs to section 3. The ratio of the likelihood of success to the likelihood of failure is known as the log odds of success. Since the likelihood of belonging to segment 3 decreases as moral obligation rises, it is possible that adding moral obligation to the logistic regression model does not significantly affect model fit. Moral responsibility may be abandoned without noticeably worsening model fit thanks to the fitted model's much improved model fit.

According to the test done for the metric variable AGE, removing the categorical variable OBLIGATION would result in a significant decline in model fit. By putting a dot on the right side of the data input, we may expand the binary logistic regression model's independent variables. A model that is overfit may result from NA. Model selection techniques eliminate unnecessary variables because overfitting models overstate the impact of independent variables. The step function conducts a stepwise analysis to determine whether removing or including an independent variable enhances model fit. Three variables are present in the final model that was chosen: EDUCATION, NEP, and VACATION. AIC balanced goodness-of-fit with a model complexity penalty is evaluated. The distributions of the anticipated probability for segment 3's users are displayed in parallel boxplots.

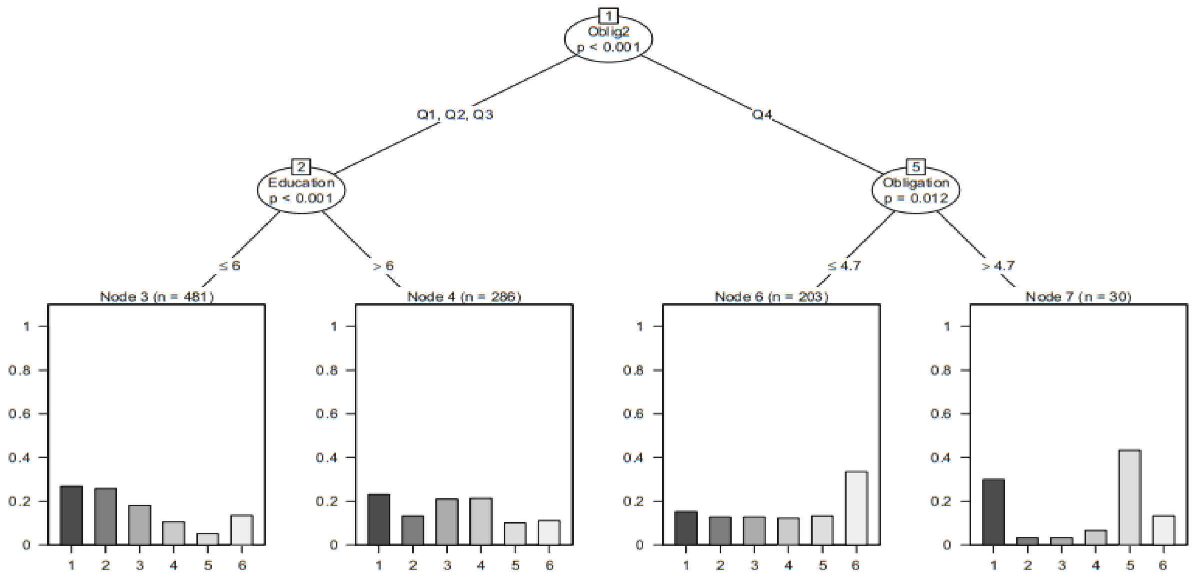
Multinomial Logistic Regression

When the dependent variable y is categorical and is supposed to follow a multinomial distribution with the logistic function as the link function, multinomial logistic regression may construct a model that predicts each segment simultaneously. To fit a model, utilize R's multinom() and Anova() functions. With the exception of segment 1, each segment's regression coefficients are included in the fitted model (the baseline category). The coefficients show how the independent variable's change will affect the log chances. The best-fitting model is the one that performs worst in AIC if an independent variable is either eliminated or added. This is determined by applying function step() to a fitted model.

No customers are anticipated to be in segment 4 in the left panel of Fig. whereas the majority of respondents are predicted to be in segment 1 in the right panel. We depict the expected probability using the function all Effects to make it easier to understand the estimated impacts. Each segment's anticipated probability is displayed separately, and C6 = 1 denotes that the panel includes predictions for segment 1. The right panel displays how, for a customer of average age, the estimated segment membership probability varies with moral duty levels. The likelihood that a respondent is from segment 6 is higher for those with the lowest moral duty score on the Q1 scale and lower for those with higher moral obligation scores.

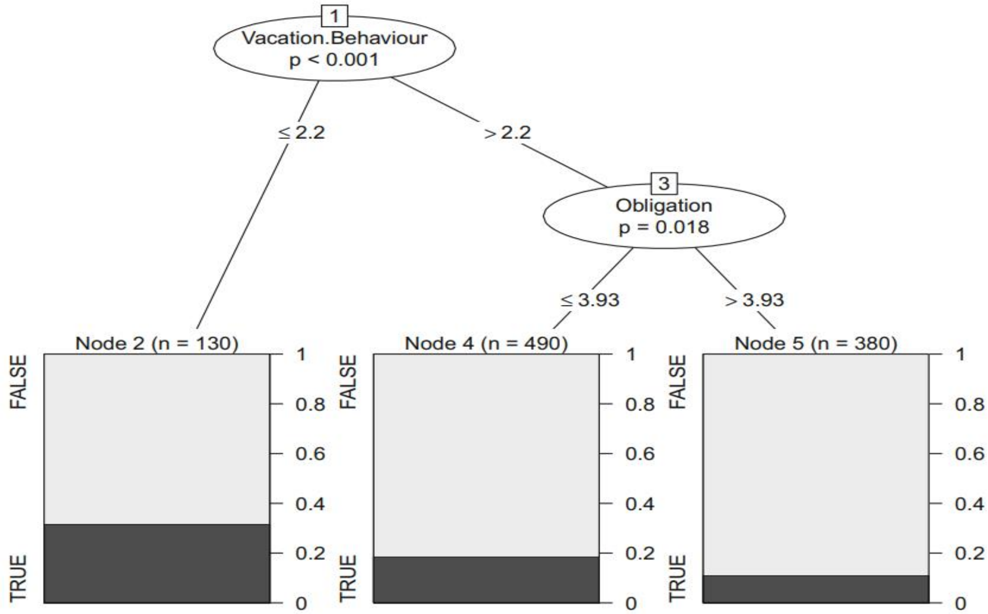
Tree-Based Methods

CARTs are a different modeling strategy that may be used to forecast binary or categorical dependent variables. They include benefits such as variable selection, simplicity in interpretation, and the inclusion of interaction effects. The dividing into two or more groups, the selection criterion, the halting condition, and the outcome prediction vary amongst tree-constructing algorithms. VACATION. Consumer classification is based on two independent variables: BEHAVIOR and OBLIGATION. It is simpler to



understand the classification tree when it is plotted using plot(tree63). Using OBLIGATION as the independent variable, Node 3 is further divided, placing customers in either Node 4 or Node 5. The group with the lowest mean score for practicing ecologically responsible behaviour while on holiday had the largest percentage of segment 3 members, according to stacked bar plots.

The results demonstrate that the categorical variable expressing moral duty is the first splitting variable (OBLIGATION2). Node 11 has responders with a NEP score of at least 4 and the highest attainable value for moral duty. By utilizing EDUCATION as the dividing variable, node 2 is split into nodes 3 and 4. With 481 responders, Node 3 is a terminal node, while Node 4's estimated segment membership is 1. 286 respondents makeup Node 5, and 77% of them do not belong to Segment 1.With a moral responsibility category value of Q4, Node 6 comprises respondents with a moral obligation value of 47 or below. There are 203 responders in Node 7, and 67% of them are not from Segment 6.Node 7 has high proportions of members of segments 1 and 5, but only low proportions of other segments.



Step 8: Selecting the Target Segment(s)

Now the question is: which of the many possible market segments will be selected for targeting?

- At the end of step 5 we have chosen global market segmentation, now we have a number of segments for detailed inspection. These segments are profiled in Step-6, and described in Step-7.
- The aim of step-8 is selecting one or more of those market segments for targeting.
- The first task in Step-8 is to ensure that all the market segments that are still under consideration to be selected as target markets have well and truly passed the knock-out criteria test.
- Once this is done, the attractiveness of the remaining segments and the relative organizational competitiveness for these segments needs to be evaluated.

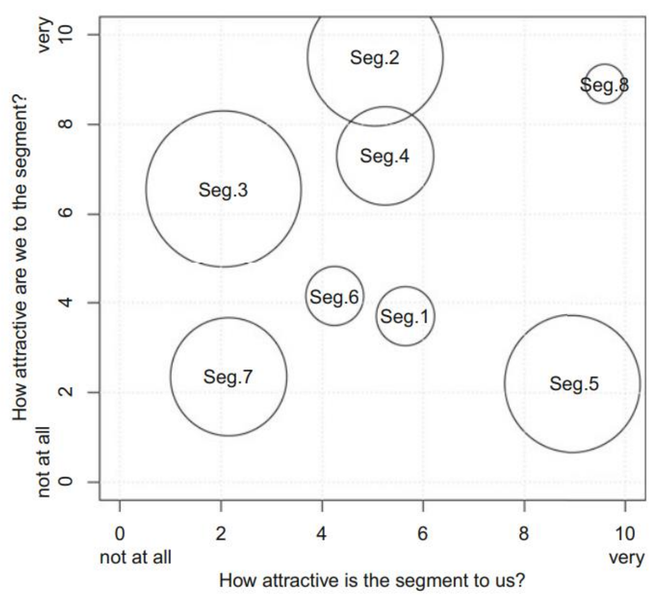
8.1 Market Segment Evaluation

Many books recommend the use of a decision matrix

- Boston matrix
- General Electric / McKinsey matrix
- directional policy matrix
- McDonald four-box directional policy matrix
- market attractiveness-business strength matrix

- The aim of these matrices is to select one or a small number for targeting.
- The axis of matrix is segment attractiveness, and relative organizational competitiveness.

- It is necessary for users to return to their specifications of what an ideal target segment looks like for them. The ideal target segment was specified in Step 2. Step 2 resulted in a number of criteria of segment attractiveness, and weights quantifying how much impact each of these criteria has on the total value of segment attractiveness.
- The location of each market segment in the segment evaluation plot is then computed by multiplying the weight of the segment attractiveness criterion with the value of the segment attractiveness criterion for each market segment.
- Each segment is given a rating from 1 to 10 from step 6 and step 7, for each segment, the rating is multiplied with the weight, and all weighted attractiveness values are added.
- The same procedure is followed for relative organizational competitiveness.
- A Criteria should be selected that may include attractiveness of the product to the segment in view of the benefits segment members seek, suitability of the current price to segment willingness or ability to pay; availability of distribution channels to get the product to the segment.
- The selection of criteria is based on the weights by both segment attractiveness criteria and organizational competitiveness, criteria are agreed upon, next they are weighted, then each segment is rated, and finally the values are multiplied and summed up. After summing up, the more weighted one will be considered.
- Here we took the axes: How attractive is the segment to us? And how attractive are we to the segment?



- Typically bubble size indicates profit.
- Segment 8 is excellent because it is highly attractive to the organization, and views the organization's offer as highly attractive.

Step 9: Customizing the Marketing Mix

Customizing the marketing mix is an essential aspect of any successful marketing strategy. Customization involves adapting the four Ps of the marketing mix (product, price, promotion, and place) to meet the unique needs of a specific target market.

9.1 Implications for Marketing Mix Decisions

- Marketing mix decisions refer to the set of controllable variables that a company uses to influence consumers' perceptions and buying behaviour.
- In the early days of marketing, the marketers had at their disposal 12 ingredients: product planning, packaging, physical handling, distribution channels, pricing, personal selling, branding, display, advertising, promotions, servicing, fact finding and analysis.
- The segmentation-targeting-positioning approach postulates a sequential process. The process starts with market segmentation (the extraction, profiling, and description of segments), followed by targeting (the assessment of segments and selection of a target segment), and finally positioning (the measures an organization can take to ensure that their product is perceived as distinctly different from competing products, and in line with segment needs).
- Viewing market segmentation as the first step in the segmentation-targeting positioning approach is useful because it ensures that segmentation is not seen as independent from other strategic decisions.
- The selection of one or more specific target segments may require the design of new, or the modification or re-branding of existing products (Product), changes to prices or discount structures (Price), the selection of suitable distribution channels (Place), and the development of new communication messages and promotion strategies that are attractive to the target segment (Promotion).

- Most commonly the marketing mix is understood as consisting of the 4Ps:
    - Product
    - Price
    - Promotion
    - Place

- **Product:**
  The product is the core component of the marketing mix. The company needs to ensure that the product meets the needs and wants of the target market. The product should be designed, packaged, and

positioned in a way that is attractive to the target audience. Marketing mix decisions related to product include product features, design, packaging, branding, and quality.

- **Price:**
  Pricing is another crucial element of the marketing mix. The company must set a price that is competitive but still profitable. Marketing mix decisions related to price include pricing strategies, discounting, and bundling.

- **Promotion:**
  Promotion refers to the communication activities that the company uses to reach its target market. Marketing mix decisions related to promotion include advertising, personal selling, sales promotion, public relations, and direct marketing.

- **Place (Distribution)**:
  Place refers to the distribution channels that the company uses to get its products to the target market. Marketing mix decisions related to place include channel selection, channel management, and logistics.

- In today's digital age, marketing mix decisions are also influenced by the rise of digital marketing and e-commerce. Companies need to consider how to integrate digital channels into their marketing mix decisions.
- Overall, the implications for marketing mix decisions are significant, and companies need to carefully consider each variable to create a successful marketing mix that resonates with their target audience.

GitHub Repository code for McDonald's case study:

Lahari Shastri
- **http://www.github.com/LahariShastri/McDonalds-case-study**

Pawar Mayur Dattatray
- https://www.github.com/mayurpawar24/Market-Segmentation-Analysis-McDonald-s-Case-Study-

Rayudu Mounika
- https://www.github.com/mourayu/Mcdonalds.git

Krisi Umangkumar Doshi
- https://github.com/KrisiDoshi/R-to-Python.git

Rajesh Kannan

- https://www.github.com/rajesh-chaitanyaa