Step 1: Deciding (not) to Segment

Implementing a market segmentation plan necessitates a long-term commitment and significant investment from the organisation. It entails expenses such as research, the creation of different packages and adverts, and the creation of specialised communication messages. Segmentation may necessitate adjustments to items, pricing, distribution routes, and internal organisational structure. Decisions should be made at the highest executive level, and ongoing communication is essential. Before investing in a market segmentation study, it is critical to understand and weigh these effects.

A number of obstacles may stand in the way of successfully implementing a market segmentation strategy. These include the absence of senior management leadership, dedication, and resources, as well as organisational resistance to change and a lack of market orientation. Implementation can also be hampered by a lack of knowledge and experience, limits on goals and resources, inadequate planning, and operational challenges. It takes deliberate detection, removal, or consideration of alternate ways to overcome these hurdles. Implementing market segmentation successfully requires a committed and patient strategy, underpinned by clear communication and comprehension.

Step 1 Checklist

Task	Who is responsible?	Completed?
Ask if the organisation's culture is market-oriented. If yes, proceed. If no, seriously consider not to proceed.		☐
Ask if the organisation is genuinely willing to change. If yes, proceed. If no, seriously consider not to proceed.		☐
Ask if the organisation takes a long-term perspective. If yes, proceed. If no, seriously consider not to proceed.		☐
Ask if the organisation is open to new ideas. If yes, proceed. If no, seriously consider not to proceed.		☐
Ask if communication across organisational units is good. If yes, proceed. If no, seriously consider not to proceed.		☐
Ask if the organisation is in the position to make significant (structural) changes. If yes, proceed. If no, seriously consider not to proceed.		☐
Ask if the organisation has sufficient financial resources to support a market segmentation strategy. If yes, proceed. If no, seriously consider not to proceed.		☐
Secure visible commitment to market segmentation from senior management.		☐
Secure active involvement of senior management in the market segmentation analysis.		☐
Secure required financial commitment from senior management.		☐
Ensure that the market segmentation concept is fully understood. If it is not: conduct training until the market segmentation concept is fully understood.		☐

Chapter 4
Step 2: Specifying the Ideal Target Segment

User feedback is crucial to the third layer of market segmentation analysis at every stage. User involvement should go beyond the initial briefing and the creation of the marketing mix. The organization's conceptual contribution to the study directs actions like data collecting and target segment selection. In Step 2, the organisation chooses two categories of segment evaluation criteria: attractiveness criteria and knock-out criteria, which assess the relative attractiveness of surviving segments. Numerous factors, including size, growth, profitability, accessibility, compatibility, and more, are suggested in the literature for segment appraisal. While attractiveness criteria offer the segmentation team a range of options, knock-out criteria are non-negotiable. These criteria play a crucial role in Step 8 when assessing the overall attractiveness of each market segment.

Which market sectors are acceptable for evaluation utilising segment attractiveness criteria are filtered using knock-out criteria. Aspects like homogeneity, distinctiveness, size, alignment with the organization's strengths, identifiability, and reachability are among these requirements. While distinctiveness emphasises how they differ from members of other segments, homogeneity refers to how similar members are to one another within a segment. Additionally, the sector must be sizable enough to support tailoring the marketing mix, and it must also be compatible with the organization's capabilities. Segment members can be recognised and efficiently contacted if they are identifiable and reachable. Understanding these knock-out criteria is crucial for senior management, the segmentation team, and the advisory committee. Although the majority of the criteria do not need to be further specified, some elements, such as the minimum viable target segment size.

In addition to the knock-out criterion, Table 4.1 offers different segment attractiveness standards that the segmentation team may take into account depending on their particular requirements. Attraction criteria are not binary, in contrast to knockout criteria. Instead, the attractiveness of each market sector for each criterion is evaluated. In Step 8 of the market segmentation study, a market segment is chosen as a target segment based on its overall attractiveness across all criteria.

It is advised to examine market segmentation using an organised method. It is usual to use the segment evaluation plot to display competitiveness and attractiveness. Representatives from various organisational units should negotiate the assessment criteria to reach an agreement. Early attractiveness criteria selection aids in data gathering and target segmentation. The team should have six or fewer criteria, each with a weighted rating of importance. It is advantageous to ask the advisory committee, which has a variety of viewpoints, for approval.

Step 2 Checklist

Task	Who is responsible?	Completed?
Convene a segmentation team meeting.		☐
Discuss and agree on the knock-out criteria of homogeneity, distinctness, size, match, identifiability and reachability. These knock-out criteria will lead to the automatic elimination of market segments which do not comply (in Step 8 at the latest).		☐
Present the knock-out criteria to the advisory committee for discussion and (if required) adjustment.		☐
Individually study available criteria for the assessment of market segment attractiveness.		☐
Discuss the criteria with the other segmentation team members and agree on a subset of no more than six criteria.		☐
Individually distribute 100 points across the segment attractiveness criteria you have agreed upon with the segmentation team. Distribute them in a way that reflects the relative importance of each attractiveness criterion.		☐

Step 3: Collecting Data:

Empirical data is essential for both common sense and data-driven market segmentation. In common sense segmentation, a single characteristic, such as gender, is used to divide the sample into market segments. Other personal characteristics serve as descriptor variables to describe the segments. Data-driven segmentation involves multiple segmentation variables to identify or create market segments. Descriptor variables provide detailed information about the segments.

- Gender as a possible segmentation variable in common sense market segmentation.
- Segmentation variables in data-driven market segmentation.

Both intuitive and data-driven market segmentation depend heavily on the quality of the empirical data. For effective categorization of persons into the appropriate market segment and accurate segmentation, data quality is crucial in common sense segmentation. The creation of specialised items, price plans, distribution avenues, and advertising methods all depend on this knowledge. The same is true for data-driven segmentation, where the calibre of the market segments that are retrieved and their descriptions is determined by the calibre of the data. Numerous sources, including surveys, observations (such scanner data), and experimental research, can provide empirical data for segmentation investigations. It is crucial to investigate various data sources and give priority to those that most accurately depict actual consumer behaviour. Despite being widely utilised, surveys may not be able to capture all behaviours.

Segmentation Criteria:

Before extracting segments and collecting data, organizations need to decide on the segmentation criterion to use. The segmentation criterion refers to the nature of the information used for market segmentation, such as geographic, sociodemographic, psychographic, or behavioural factors. Choosing the appropriate criterion requires market knowledge and cannot be easily outsourced. While various segmentation criteria exist, it is generally recommended to use the simplest approach that works for the product or service. The focus should be on what is effective and cost-efficient rather than what may seem more sophisticated.

- Geographic Segmentation
- Socio-Demographic Segmentation
- Psychographic Segmentation
- Behavioural Segmentation

Data from Survey Studies:

Most market segmentation analyses are based on survey data. Survey data is cheap and easy to collect, making it a feasible approach for any organisation. But survey data – as opposed to data obtained from observing actual behaviour – can be contaminated by a wide range of biases. Such biases can, in turn, negatively affect the quality of solutions derived from market segmentation analysis. A few key aspects that need to be considered when using survey data are discussed

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

Data from Internal Sources

Nowadays, businesses have access to vast amounts of internal data that may be used to analyse market segments. Examples include data from grocery store scanners, reservations from airline loyalty programmes, and information from online purchases. The advantage of this data over self-reported data, which might be skewed by response biases and memory constraints, is that it represents actual customer behaviour. Internal data is also frequently accessible and doesn't need to be collected separately. The bias that could result from over-representing current consumers and underrepresenting potential new customers with diverse consumption patterns, however, is a possible negative.

Step 3 Checklist:

Task	Who is responsible?	Completed?
Convene a market segmentation team meeting.		☐
Discuss which consumer characteristics could serve as promising segmentation variables. These variables will be used to extract groups of consumers from the data.		☐
Discuss which other consumer characteristics are required to develop a good understanding of market segments. These variables will later be used to describe the segments in detail.		☐
Determine how you can collect data to most validly capture both the segmentation variables and the descriptor variables.		☐
Design data collection carefully to keep data contamination through biases and other sources of systematic error to a minimum.		☐
Collect data.		☐

Step 7: Describing Segments

Understanding the variations in segmentation factors across market segments is necessary for segment profiling. Early in the analysis process, segmentation factors are selected as the foundation for identifying market segments. Step 7 uses extra data, referred to as descriptor variables, to describe market segments. For insight and targeted marketing, effective segment descriptions are crucial. Descriptive statistics and visualisations can be used to study differences in descriptor variables.

Managers favour graphical visualisation due to its intuitiveness, which also makes result interpretation easier and combines statistical significance. Compared to tabular output, it increases the efficiency of processing information.

Cross-tabulations are used to visualise and examine differences in descriptor variables across market groups. The data set on Australian travel motivations is presented as an illustration. Segment membership is kept in a helper variable, and the data frame includes descriptor variables. By including segment membership as a categorical variable to the data frame, cross-tabulations are produced. Stack bar charts or mosaic plots can be used to graphically show the cross-tabulations. Mosaic plots can include components of inferential statistics and are particularly helpful for visualising tables with several descriptor variables. In mosaic plots, colour coding draws attention to notable discrepancies between observed and predicted frequencies. The offered examples demonstrate the relationships between segment membership and gender, income, and the ethical responsibility to preserve the environment in Australian travel motivations.

The "lattice" R package (Sarkar, 2008) offers conditional versions of standard R plots, allowing for visualizing differences between segments. It divides data into subsets or panels, and supports histograms and box-and-whisker plots for segment profiles. By combining the word "Segment" with segment numbers, segment names can be displayed in plots. Histograms show age and moral obligation distributions per segment, but assessing differences visually can be challenging. To gain insights, parallel box-and-whisker plots are created, indicating minor age variations across segments. The package facilitates statistical hypothesis testing by customizing the plots with proportional box widths and 95% confidence intervals for medians. Notches in the plot reveal significant differences in moral obligation between segments 3 and 6. A modified segment level stability across solutions (SLSA) plot traces metric descriptor variable values over multiple segmentation solutions, using colors to represent values. The plot consistently shows high moral obligation for a nature-loving tourist segment, followed by an acquiescence-biased segment. The lattice package offers a comprehensive toolkit for visualizing and describing market segments, including conditional plots, histograms, box-and-whisker plots, and SLSA plots, enabling a better understanding of segment differences and characteristics.

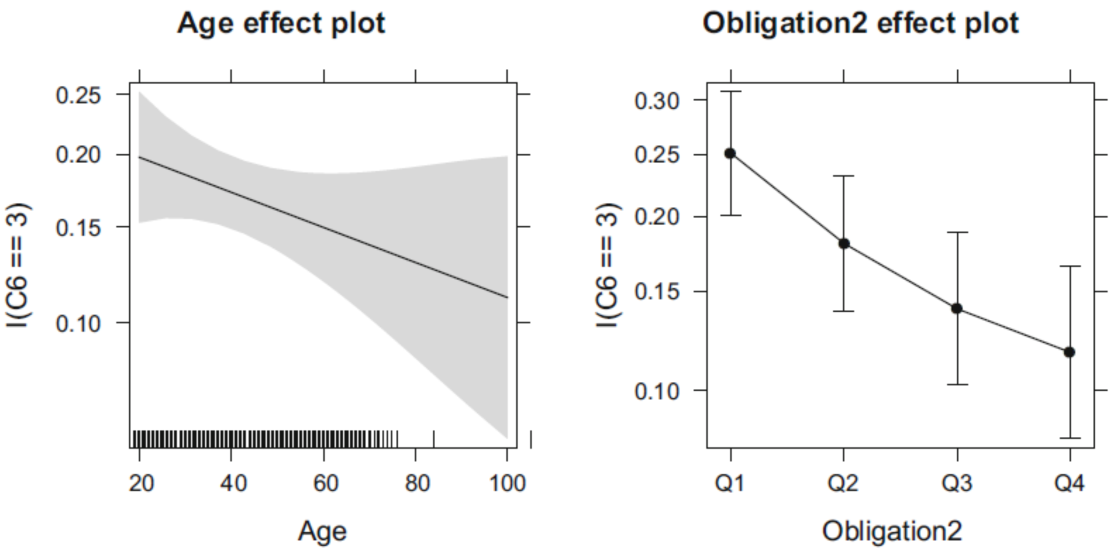
The "lattice" R package (Sarkar, 2008) provides conditional versions of standard R plots for visualizing segment differences. It supports histograms and box-and-whisker plots for segment profiles, with the ability to display segment names. While visually assessing differences in histograms can be challenging, parallel box-and-whisker plots show minor age variations across segments. The package allows for statistical hypothesis testing by customizing plots with proportional box widths and 95% confidence intervals. Notches in the plot reveal significant differences in moral obligation between specific segments. Additionally, a modified segment level stability across solutions (SLSA) plot tracks variable values over multiple segmentation solutions, using colors to represent values. Overall, the lattice package offers a comprehensive toolkit for visualizing and describing market segments, facilitating a better understanding of their differences and characteristics.

The passage discusses using regression models to predict market segment membership based on descriptor variables. Regression analysis is employed with the segment membership as the categorical dependent variable and the descriptor variables as independent variables. This approach tests differences in all descriptor variables simultaneously, providing insights into identifying market segments and determining critical descriptor variables. The passage also introduces the linear regression model, assuming a linear relationship between the dependent variable and independent variables. The mean value of the dependent variable is modelled using a linear function, and regression coefficients represent the mean age difference between segments. Generalized linear models are also mentioned, which can accommodate a wider range of distributions for the dependent variable, allowing for modelling categorical variables. Two special cases of generalized linear models, binary and multinomial logistic regression, are briefly mentioned, where the dependent variable follows a binary or multinomial distribution, respectively, and the link function is the logit function.

$$g(\mu) = \eta = \log\left(\frac{\mu}{1-\mu}\right).$$

R's lm () function fits generalised linear models. A family defines the distribution of the dependent variable and the link function. The family of the Bernoulli distribution with the logit link is binomial (link = "logit") or

Because the logit link is the default, family = binomial (). The binomial distribution is a generalisation of the Bernoulli distribution where the variable y indicates the number of successes out of several independent Bernoulli-distributed trials with the same success probability instead of just taking the values 0 and 1.

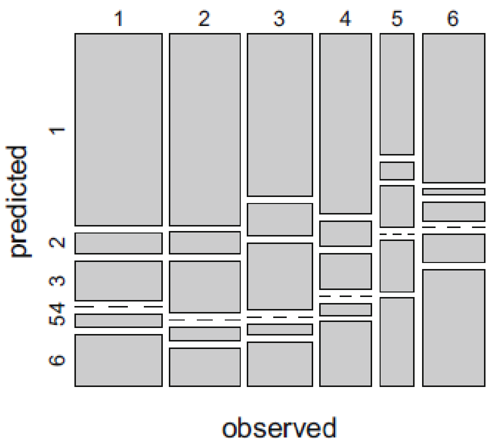


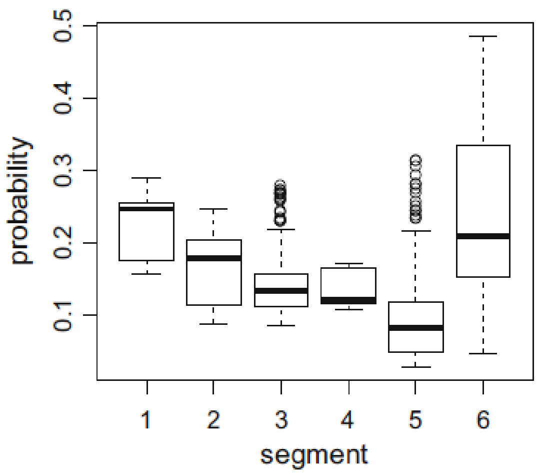
The text describes a table of estimated coefficients and their standard errors, which is part of a regression model's output. Additionally, it offers p-values and test statistics, particularly for the z-test. The z-test contrasts the model that was fitted to data with a model with a regression coefficient of 0. The coefficient is implied to be greater than zero and should be included in the model if the null hypothesis is rejected.

The null hypothesis is not rejected in the case of the variable "AGE," indicating that its removal from the model is not likely to have a major impact on how well it fits the data. Three regression coefficients are fitted to the model to represent the differences between categories Q2, Q3, and Q4 in comparison to category Q1 if the variable "moral obligation" is included. In every test, the entire model is contrasted with a model in which the regression coefficient for a particular category is set to 0. It is impossible to tell from this comparison if the moral responsibility model performs better than the one without it.

The "Anova" function from the "car" package is used to contrast the model with and without moral duty. In order to compare the two models, all regression coefficients for this variable are set to zero in the model when moral responsibility is omitted.

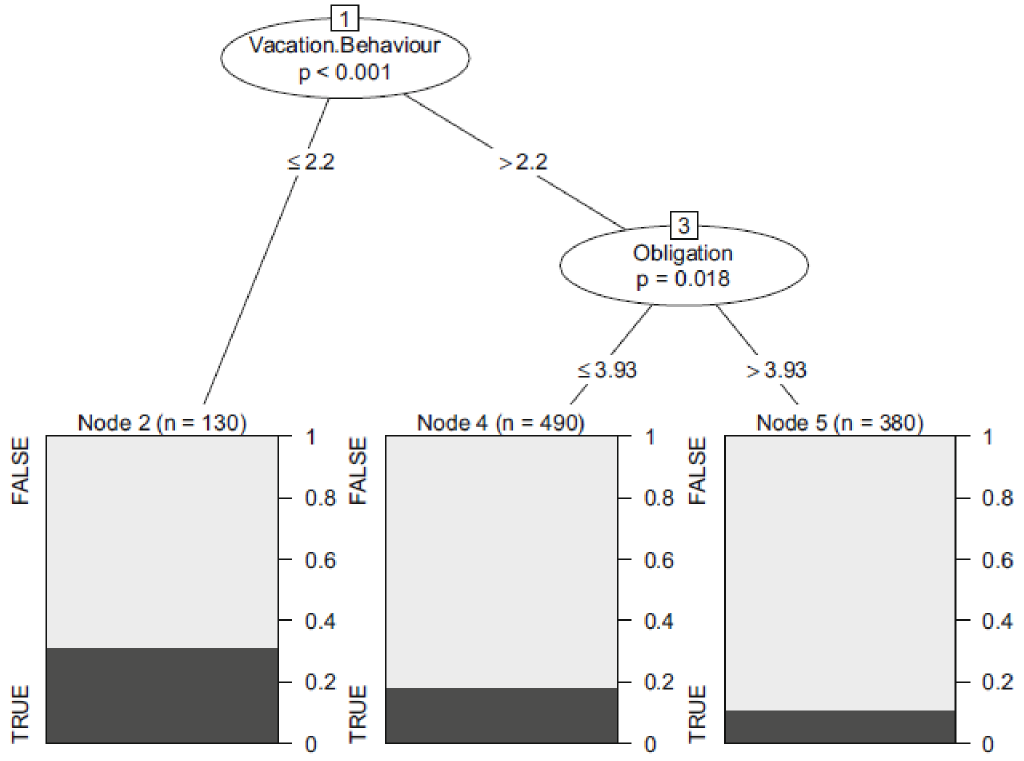
In the section, it is discussed how to handle missing values and prevent regression models from becoming overfit. In order to pick models based on AIC, it introduces the "step" function in R. Three variables are part of the model that was chosen. Comparing predictive performance amongst models and visualising expected probability for various customer segments using parallel boxplots.





The passage talks about multinomial logistic regression, which is applied to simultaneously predict numerous segments. This kind of regression is fitted using the "multinom" function found in the "nnet" package of the R language. The AIC value's decline is used to determine which model to use. By contrasting projected and observed segment memberships, the model's predictive capability is evaluated. The correlation between age, moral obligation, and estimated segment probability is shown in the figures. The likelihood that a respondent falls into segment 6 is lower for younger respondents and higher for older respondents. As people get older, their likelihood of falling into segment 5 reduces. Segment membership probability are also influenced by moral duty values, with higher values indicating a larger likelihood of belonging to segment 6.

Machine learning techniques like classification and regression trees (CART) are used to forecast categorical or binary dependent variables based on independent factors. They provide benefits including variable choice, interpretability, and simplicity in introducing interaction effects. CART performs well when there are many independent factors, but it can produce unpredictable results. The method uses a sequential splitting process to build a tree of nodes, each of which stands for a group of consumers. By navigating the tree depending on the values of the consumer's independent variables, segment membership is predicted. Several R packages and methods, including rpart and partykit, implement the building and visualisation of trees. A conditional inference tree can be fit using the partykit package's ctree() method.



The output shows the results of a classification tree analysis. The first splitting variable is the moral obligation category (OBLIGATION2), which separates the consumers into two nodes (2 and 5). Node 2 consists of consumers with moral obligation values of Q1, Q2, and Q3, while node 7 consists of consumers with a moral obligation value of Q4. Node 2 is further split based on the education level variable (EDUCATION), creating nodes 3 and 4. Node 3 is a terminal node mostly comprising respondents from segment 1, while node 4 is also a terminal node with a predicted segment membership of 1 but containing respondents primarily not from segment 1. Node 5 represents consumers highly morally obliged to protect the environment and is split into nodes 6 and 7 based on the metric version of moral obligation. Node 6 consists of respondents with lower moral obligation scores and primarily belongs to segment 6, while node 7 represents higher moral obligation scores and has a predicted segment membership of 5. The visualization of the tree (plot(tree6)) provides a clear representation of the segmentation and the distribution of segment memberships in each terminal node.

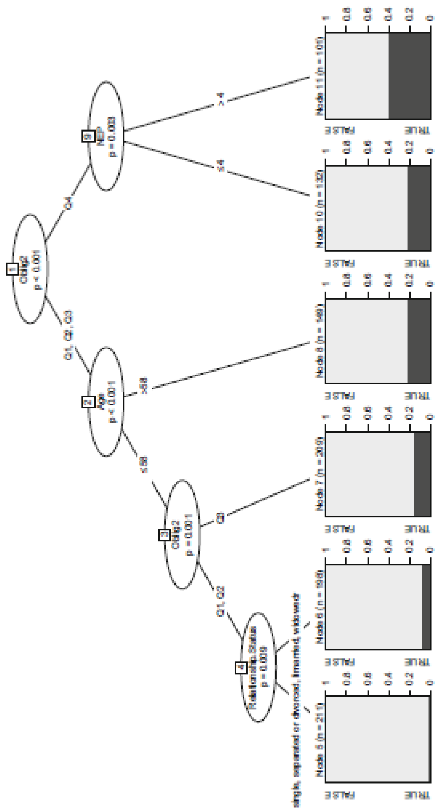


Fig. 9.16 Conditional inference tree using membership in segment 6 as dependent variable for the Australian travel motives data set

The classification tree for segment 6 is more complex than the one for segment 3, with a larger number of inner and terminal nodes. The stacked bar charts in the terminal nodes indicate the purity of each node and the variation in the proportion of segment 6 members within them. Terminal node 11 has the highest proportion of consumers assigned to segment 6, consisting of respondents with the highest moral obligation value and a NEP score of at least 4.

The ctree() function can also be used to fit a tree for categorical dependent variables with more than two categories. The tree plot shows the splits based on the independent variable VACATION.BEHAVIOUR, with consumers above a threshold value directed to the right branch (node 3) and consumers below or equal to the threshold value directed to the left branch (node 2). Node 2 is a terminal node, indicating the proportion of respondents in segment 3. Node 3 is further split using the independent variable OBLIGATION, creating terminal nodes 4 and 5, which are visualized using stacked bar plots showing the proportions of respondents belonging to segment 3.

The parameters of the partykit package can be adjusted using the ctree_control function to influence tree construction, such as restricting nodes considered for splitting, setting the minimum size for terminal nodes, selecting the test statistic for association tests, and determining the minimum criterion value for implementing a split.

As an example, a tree with segment 6 membership as the dependent variable is fitted, ensuring that terminal nodes contain at least 100 respondents (minbucket = 100) and setting the minimum criterion value to 0.99 (equivalent to a p-value smaller than 0.01). Figure 9.16 provides a visualization of this tree.

9.5 Step 7 Checklist

Task	Who is responsible?	Completed?
Bring across from Step 6 (profiling) one or a small number of market segmentation solutions selected on the basis of attractive profiles.		☐
Select descriptor variables. Descriptor variables are additional pieces of information about each consumer included in the market segmentation analysis. Descriptor variables have not been used to extract the market segments.		☐
Use visualisation techniques to gain insight into the differences between market segments with respect to descriptor variables. Make sure you use appropriate plots, for example, mosaic plots for categorical and ordinal descriptor variables, and box-and-whisker plots for metric descriptor variables.		☐
Test for statistical significance of descriptor variables.		☐
If you used separate statistical tests for each descriptor variable, correct for multiple testing to avoid overestimating significance.		☐
"Introduce" each market segment to the other team members to		☐