Project-2 : STUDY TASK

- Kavuluru Lakshmi Srinidhi

Contents

**(Click on the topic to go to that section)*

STEP-1: Deciding (not) to Segment

3.1 Implications of Committing to Market Segmentation

1. Before investing time and resources in a market segmentation analysis, it is important to understand the implications of pursuing a market segmentation strategy.
2. The key implication is that the organisation needs to commit to the segmentation strategy on the long term.
3. The commitment to market segmentation goes hand in hand with the willingness and ability of the organisation to make substantial changes.
4. Potentially required 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.
5. These changes, in turn, are likely to influence the internal structure of the organisation, which may need to be adjusted in view of, for example, targeting a handful of different market segments.
6. Because of the major implications of such a long-term organisational commitment, the decision to investigate the potential of a market segmentation strategy must be made at the highest executive level, and must be systematically and continuously communicated and reinforced at all organisational levels and across all organisational units.

3.2 Implementation Barriers

1. The first group of barriers relates to senior management. Lack of leadership, pro-active championing, commitment and involvement in the market segmentation process by senior leadership undermines the success of market segmentation.
2. Senior management can also prevent market segmentation to be successfully implemented by not making enough resources available, either for the initial market segmentation analysis itself, or for the long-term implementation of a market segmentation strategy.
3. A second group of barriers relates to organisational culture. Lack of market or consumer orientation, resistance to change and new ideas, lack of creative thinking, bad communication and lack of sharing of information and insights across organisational units, short-term thinking, unwillingness to make changes and office politics have been identified as preventing the successful implementation of market segmentation.
4. Another potential problem is lack of training. If senior management and the team tasked with segmentation do not understand the very foundations of market segmentation, or if they are unaware of the consequences of pursuing such a strategy, the attempt of introducing market segmentation is likely to fail.
5. Closely linked to these barriers is the lack of a formal marketing function or at least a qualified marketing expert in the organisation. The higher the market diversity and the larger the organisations, the more important is a high degree of formalisation. The lack of a qualified data manager and analyst in the organisation can also represent major stumbling blocks.

6. Another obstacle may be objective restrictions faced by the organisation, including lack of financial resources, or the inability to make the structural changes required.
7. At a more operational level, Doyle and Saunders (1985) note that management science has had a disappointing level of acceptance in industry because management will not use techniques it does not understand (p. 26). One way of counteracting this challenge is to make market segmentation analysis easy to understand, and to present results in a way that facilitates interpretation by managers. This can be achieved by using graphical visualisations (see Steps 6 and 7).
8. Most of these barriers can be identified from the outset of a market segmentation study, and then proactively removed. If barriers cannot be removed, the option of abandoning the attempt of exploring market segmentation as a potential future strategy should be seriously considered.
9. Above all, a resolute sense of purpose and dedication is required, tempered by patience and a willingness to appreciate the inevitable problems which will be encountered in implementing the conclusions.

3.3 Step 1 Checklist

This first checklist includes not only tasks, but also a series of questions which, if not answered in the affirmative, serve as knock-out criteria. For example: if an organisation is not market-oriented, even the finest of market segmentation analyses cannot be successfully implemented.

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.		□
Ensure that the implications of pursuing a market segmentation strategy are fully understood. If they are not: conduct training until the implications of pursuing a market segmentation strategy are fully understood.		□
Put together a team of 2-3 people (segmentation team) to conduct the market segmentation analysis.		□

Task	Who is responsible?	Completed?
Ensure that a marketing expert is on the team.		□
Ensure that a data expert is on the team.		□
Ensure that a data analysis expert is on the team.		□
Set up an advisory committee representing all affected organisational units.		□
Ensure that the objectives of the market segmentation analysis are clear.		□
Develop a structured process to follow during market segmentation analysis.		□
Assign responsibilities to segmentation team members using the structured process.		□
Ensure that there is enough time to conduct the market segmentation analysis without time pressure.		□

STEP-2: Specifying the Ideal Target Segment

4.1 Segment Evaluation Criteria

1. The third layer of market segmentation analysis (illustrated in Fig.2.1) depends primarily on user input.
2. After having committed to investigating the value of a segmentation strategy in Step 1, the organisation has to make a major contribution to market segmentation analysis in Step 2. While this contribution is conceptual in nature, it guides many of the following steps, most critically Step 3 (data collection) and Step 8 (selecting one or more target segments).
3. In Step 2 the organisation must determine two sets of segment evaluation criteria.
    a. One set of evaluation criteria can be referred to as knock-out criteria. These criteria are the essential, non-negotiable features of segments that the organisation would consider targeting.
    b. The second set of evaluation criteria can be referred to as attractiveness criteria. These criteria are used to evaluate the relative attractiveness of the remaining market segments– those in compliance with the knock-out criteria.
4. The literature does not generally distinguish between these two kinds of criteria. Instead, the literature proposes a wide array of possible segment evaluation criteria and describes them at different levels of detail.
5. The shorter set of knock-out criteria is essential. It is not up to the segmentation team to negotiate the extent to which they matter in target segment selection.
6. The second, much longer and much more diverse set of attractiveness criteria represents a shopping list for the segmentation team. 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 segmentation team also needs to assess the relative importance of each attractiveness criterion to the organisation.
7. Where knock-out criteria automatically eliminate some of the available market segments, attractiveness criteria are first negotiated by the team, and then applied to determine the overall relative attractiveness of each market segment in Step 8.

4.2 Knock-Out Criteria

1. Knock-out criteria are used to determine if market segments resulting from the market segmentation analysis qualify to be assessed using segment attractiveness criteria.
2. The first set of such criteria was suggested by Kotler (1994) and includes substantiality, measurability and accessibility.
3. Kotler himself and a number of other authors have since recommended additional criteria that fall into the knock-out criterion category:
    a. The segment must be ***homogeneous***; members of the segment must be similar to one another.
    b. The segment must be ***distinct***; members of the segment must be distinctly different from members of other segments.
    c. The segment must be ***large enough***; the segment must contain enough consumers to make it worthwhile to spend extra money on customizing the marketing mix for them.
    d. The segment must be ***matching*** the strengths of the organisation; the organisation must have the capability to satisfy segment members’ needs.
    e. Members of the segment must be ***identifiable***; it must be possible to spot them in the marketplace.

   f. The segment must be ***reachable***; there has to be a way to get in touch with members of the segment in order to make the customised marketing mix accessible to them.

4. Knock-out criteria must be understood by senior management, the segmentation team, and the advisory committee. Most of them do not require further specification, but some do. For example, while size is non-negotiable, the exact minimum viable target segment size needs to be specified.

4.3 Attractiveness Criteria

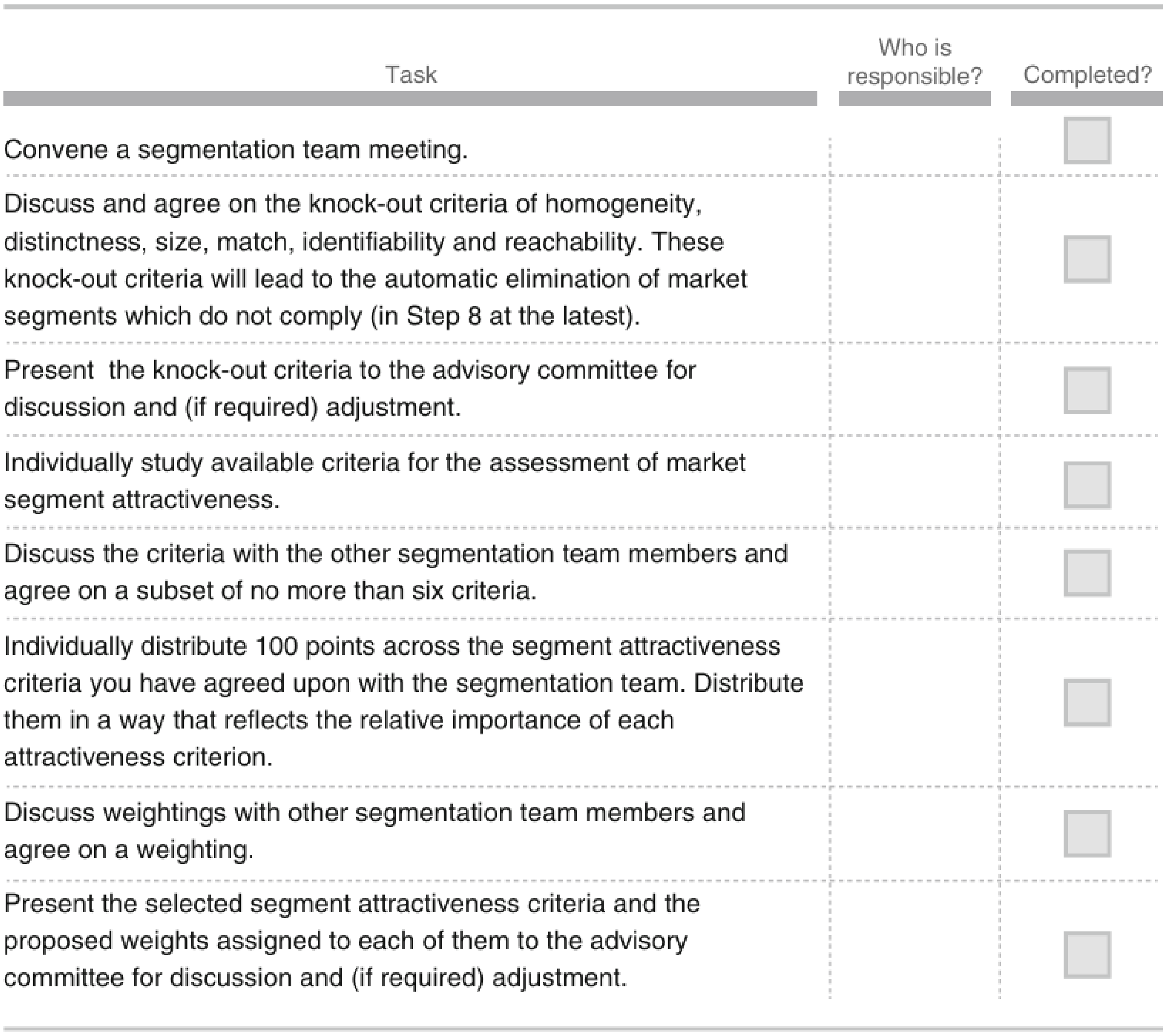
1. In addition to the knock-out criteria, a wide range of segment attractiveness criteria is available to the segmentation team to consider when deciding which attractiveness criteria are most useful to their specific situation.
2. Attractiveness criteria are not binary in nature. Segments are not assessed as either complying or not complying with attractiveness criteria. Rather, each market segment is rated; it can be more or less attractive with respect to a specific criterion.
3. The attractiveness across all criteria determines whether a market segment is selected as a target segment in Step 8 of market segmentation analysis.

4.4 Implementing a Structured Process

1. There is general agreement in the segmentation literature, that following a structured process when assessing market segments is beneficial.
2. 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 (Lilien and Rangaswamy 2003; McDonald and Dunbar 2012) showing segment attractiveness along one axis, and organisational competitiveness on the other axis.
3. The segment attractiveness and organisational competitiveness values are determined by the segmentation team. This is necessary because there is no standard set of criteria that could be used by all organisations.
4. Factors which constitute both segment attractiveness and organisational competitiveness need to be negotiated and agreed upon. To achieve this, a large number of possible criteria has to be investigated before agreement is reached on which criteria are most important for the organization
5. Optimally, this task should be completed by a team of people. If a core team of two to three people is primarily in charge of market segmentation analysis, this team could propose an initial solution and report their choices to the advisory committee– which consists of representatives of all organisational units– for discussion and possible modification.
6. There are at least two good reasons to include in this process representatives from a wide range of organisational units.
   a. First, each organisational unit has a different perspective on the business of the organisation. As a consequence, members of these units bring different positions to the deliberations.
   b. Secondly, if the segmentation strategy is implemented, it will affect every single unit of the organisation. Consequently, all units are key stakeholders of market segmentation analysis.
7. There is a huge benefit in selecting the attractiveness criteria for market segments at this early stage in the process: knowing precisely what it is about market segments that matters to the organisation ensures that all of this information is captured when collecting data (Step 3).
8. It also makes the task of selecting a target segment in Step 8 much easier because the groundwork is laid before the actual segments are on the table.

9. At the end of this step, the market segmentation team should have a list of approximately six segment attractiveness criteria. Each of these criteria should have a weight attached to it to indicate how important it is to the organisation compared to the other criteria.
10. The typical approach to weighting is to ask all team members to distribute 100 points across the segmentation criteria. These allocations then have to be negotiated until agreement is reached.

4.5 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.		□
Discuss weightings with other segmentation team members and agree on a weighting.		□
Present the selected segment attractiveness criteria and the proposed weights assigned to each of them to the advisory committee for discussion and (if required) adjustment.		□

STEP-3: Collecting Data

5.1 Segmentation Variables

1. Empirical data is used to identify or create market segments and later in the process– describe these segments in detail.
2. We use the term **segmentation variable** to refer to the variable in the empirical data used in commonsense segmentation to split the sample into market segments.
3. In commonsense segmentation, the segmentation variable is typically one single characteristic of the consumers in the sample. This case is illustrated in Table 5.1.

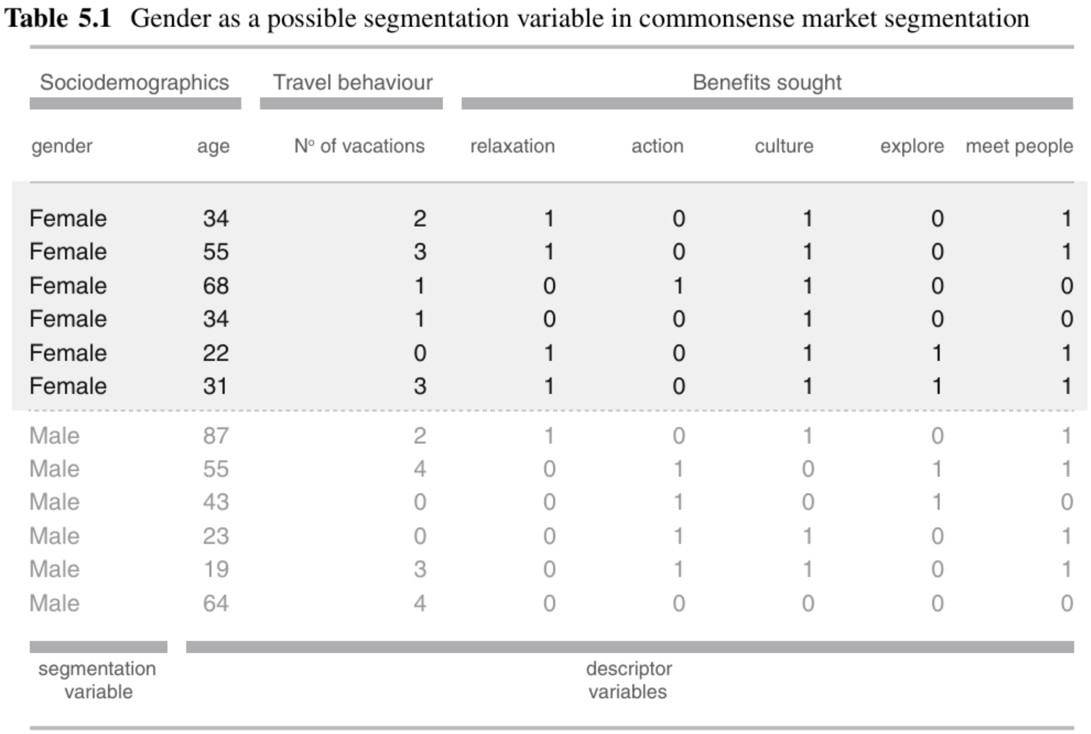


Table 5.1 Gender as a possible segmentation variable in commonsense market segmentation

Sociodemographics		Travel behaviour	Benefits sought				
gender	age	Nº of vacations	relaxation	action	culture	explore	meet people
Female	34	2	1	0	1	0	1
Female	55	3	1	0	1	0	1
Female	68	1	0	1	1	0	0
Female	34	1	0	0	1	0	0
Female	22	0	1	0	1	1	1
Female	31	3	1	0	1	1	1
Male	87	2	1	0	1	0	1
Male	55	4	0	1	0	1	1
Male	43	0	0	1	0	1	0
Male	23	0	0	1	1	0	1
Male	19	3	0	1	1	0	1
Male	64	4	0	0	0	0	0
segmentation variable	descriptor variables						

4. All the other personal characteristics available in the data– in this case: age, the number of vacations taken, and information about five benefits people seek or do not seek when they go on vacation– serve as so-called **descriptor variables**.
5. 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. These segmentation variables serve as the starting point for identifying naturally existing, or artificially creating market segments useful to the organization.
6. When commonsense segments are extracted– even if the nature of the segments is known in advance– data quality is critical to (1) assigning each person in the sample to the correct market segment, and (2) being able to correctly describe the segments. The correct description, in turn, makes it possible to develop a customised product, determine the most appropriate pricing strategy, select the best distribution channel, and the most effective communication channel for advertising and promotion.
7. The same holds for data-driven market segmentation where data quality determines the quality of the extracted data-driven market segments, and the quality of the descriptions of the resulting segments. Good market segmentation analysis requires good empirical data.
8. Empirical data for segmentation studies can come from a range of sources: from survey studies; from observations such as scanner data where purchases are recorded and, frequently, are linked to an individual customer’s long-term purchase history via loyalty programs; or from experimental studies.

9. Optimally, data used in segmentation studies should reflect consumer behaviour.

5.2 Segmentation Criteria

1. Long before segments are extracted, and long before data for segment extraction is collected, the organisation must make an important decision: it must choose which segmentation criterion to use
2. The decision which segmentation criterion to use cannot easily be outsourced to either a consultant or a data analyst because it requires prior knowledge about the market. The most common segmentation criteria are **geographic, socio-demographic, psychographic and behavioural.**
3. the following differences between consumers are the most relevant in terms of market segmentation: profitability, bargaining power, preferences for benefits or products, barriers to choice and consumer interaction effects.

5.2.1 Geographic Segmentation

1. Typically– when geographic segmentation is used– the consumer's location of residence serves as the only criterion to form market segments. While simple, the geographic segmentation approach is often the most appropriate.
2. The key advantage of geographic segmentation is that each consumer can easily be assigned to a geographic unit. As a consequence, it is easy to target communication messages, and select communication channels (such as local newspapers, local radio and TV stations) to reach the selected geographic segments.
3. The key disadvantage is that living in the same country or area does not necessarily mean that people share other characteristics relevant to marketers, such as the benefits they seek when purchasing a product.

5.2.2 Socio-Demographic Segmentation

1. Typical socio-demographic segmentation criteria include age, gender, income and education. Socio-demographic segments can be very useful in some industries. For example: luxury goods (associated with high income), cosmetics (associated with gender), etc.
2. As is the case with geographic segmentation, socio-demographic segmentation criteria have the advantage that segment membership can easily be determined for every consumer.
3. But in many instances, the socio-demographic criterion is not the cause for product preferences, thus not providing sufficient market insight for optimal segmentation decisions.

5.2.3 Psychographic Segmentation

1. When people are grouped according to psychological criteria, such as their beliefs, interests, preferences, aspirations, or benefits sought when purchasing a product, the term psychographic segmentation is used.
2. The psychographic approach has the advantage that it is generally more reflective of the underlying reasons for differences in consumer behaviour.
3. The disadvantage of the psychographic approach is the increased complexity of determining segment memberships for consumers. Also, the power of the psychographic approach depends heavily on the reliability and validity of the empirical measures used to capture the psychographic dimensions of interest.

5.2.4 Behavioural Segmentation

1. Another approach to segment extraction is to search directly for similarities in behaviour or reported behaviour. A wide range of possible behaviours can be used for this purpose,

including prior experience with the product, frequency of purchase, the amount spent on purchasing the product on each occasion (or across multiple purchase occasions), and information search behaviour.

2. The key advantage of behavioural approaches is that– if based on actual behaviour rather than stated behaviour or stated intended behaviour– the very behaviour of interest is used as the basis of segment extraction. As such, behavioural segmentation groups people by the similarity which matters most.
3. But behavioural data is not always readily available, especially if the aim is to include in the segmentation analysis potential customers who have not previously purchased the product, rather than limiting oneself to the study of existing customers of the organisation.

5.3 Data from Survey Studies

1. 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.
2. 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

5.3.1 Choice of Variables

1. Carefully selecting the variables that are included as segmentation variable in commonsense segmentation, or as segmentation variables in data-driven segmentation, is critical to the quality of the market segmentation solution.
2. In data-driven segmentation, all variables relevant to the construct captured by the segmentation criterion need to be included. At the same time, unnecessary variables must be avoided.
3. Including unnecessary variables also increases the dimensionality of the segmentation problem without adding relevant information, making the task of extracting market segments unnecessarily difficult for any data analytic technique. Unnecessary variables are referred to as noisy variables or masking variables.
4. The problem of noisy variables negatively affecting the segmentation solution can be avoided at the data collection and the variable selection stage of market segmentation analysis.
5. Noisy variables can result from not carefully developing survey questions, or from not carefully selecting segmentation variables from among the available survey items.

5.3.2 Response Options

1. Options allowing respondents to answer in only one of two ways, generate binary or dichotomous data. Such responses can be represented in a data set by 0s and 1s. The distance between 0 and 1 is clearly defined and, as such, poses no difficulties for subsequent segmentation analysis
2. Options allowing respondents to select an answer from a range of unordered categories correspond to nominal variables. Nominal variables can be transformed into binary data by introducing a binary variable for each of the answer options.
3. Options allowing respondents to indicate a number, such as age or nights stayed at a hotel, generate metric data. Metric data allow any statistical procedure to be performed (including the measurement of distance), and are therefore well suited for segmentation analysis.

5.3.3 Response Styles

1. Survey data is prone to capturing biases. A response bias is a systematic tendency to respond to a range of questionnaire items on some basis other than the specific item content (i.e., what the items were designed to measure)
2. If a bias is displayed by a respondent consistently over time, and independently of the survey questions asked, it represents a response style.
3. Response styles affect segmentation results because commonly used segment extraction algorithms cannot differentiate between a data entry reflecting the respondent's belief from a data entry reflecting both a respondent's belief and a response style.
4. It is critical, therefore, to minimise the risk of capturing response styles when data is collected for market segmentation. In cases where attractive market segments emerge with response patterns potentially caused by a response style, additional analyses are required to exclude this possibility. Alternatively, respondents affected by such a response style must be removed before choosing to target such a market segment.

5.3.4 Sample Size

1. Viennese psychologist Formann (1984) recommends that the sample size should be at least 2p (better five times 2p), where p is the number of segmentation variables. This rule of thumb relates to the specific purpose of goodness-of-fit testing in the context of latent class analysis when using binary variables. It can therefore not be assumed to be generalisable to other algorithms, inference methods, and scales.
2. According to Qiu and Joe (2015), the sample size should– in the simple case of equal cluster sizes– be at least ten times the number of segmentation variables times the number of segments in the data (10*p*k where p represents the number of segmentation variables and k represents the number of segments).
3. The adjusted Rand index serves as the measure of correctness of segment recovery. The adjusted Rand index assesses the congruence between two segmentation solutions.
4. Increasing the sample size improves the correctness of the extracted segments. Interestingly, however, the biggest improvement is achieved by increasing very small samples.
5. A sample size of at least 60*p is recommended.
6. For a more difficult artificial data scenario Dolnicar et al. (2014) recommend using a sample size of at least 70*p; no substantial improvements in identifying the correct segments were identified beyond this point.
7. Presence of unequally sized segments makes it more difficult for an algorithm to extract the correct market segments.
8. If the variables are not correlated at all, the algorithm has no difficulty extracting the correct segments. If, however, the variables are highly correlated, the task becomes so difficult for the algorithm, that even increasing the sample size dramatically does not help. A small number of noisy variables, on the other hand, has a lower effect.
9. Overall, this study demonstrates the importance of having a sample size sufficiently large to enable an algorithm to extract the correct segments (if segments naturally exist in the data).
10. It can be concluded from the body of work studying the effects of survey data quality on the quality of market segmentation results based on such data that, optimally, data used in market segmentation analyses should
    a. contain all necessary items;
    b. contain no unnecessary items;
    c. contain no correlated items;

d. contain high-quality responses;
e. be binary or metric;
f. free of response styles;
g. include responses from a suitable sample given the aim of the segmentation study; and
h. include a sufficient sample size given the number of segmentation variables (100 times the number of segmentation variables)

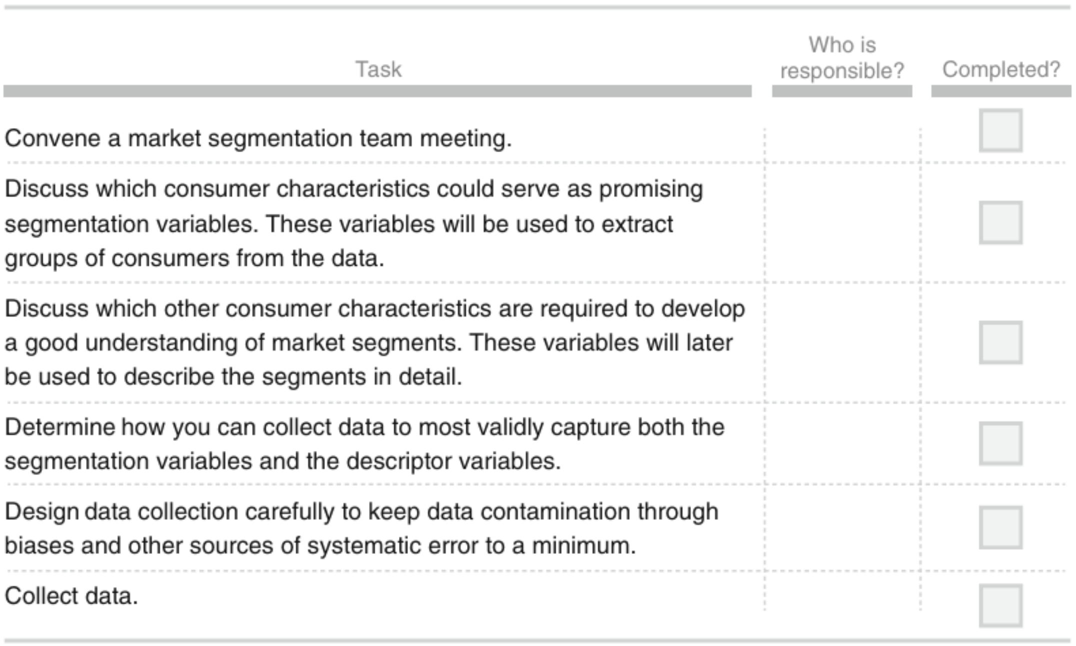
5.4 Data from Internal Sources

1. Increasingly organisations 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.
2. The strength of such data lies in the fact that they represent actual behaviour of consumers, rather than statements of consumers about their behaviour or intentions, known to be affected by imperfect memory, as well as a range of response biases, such as social desirability bias or other response styles.

5.5 Data from Experimental Studies

1. Another possible source of data that can form the basis of market segmentation analysis is experimental data. Experimental data can result from field or laboratory experiments.
2. Conjoint studies and choice experiments result in information about the extent to which each attribute and attribute level affects choice. This information can also be used as a segmentation criterion.

5.6 Step 3 Checklist



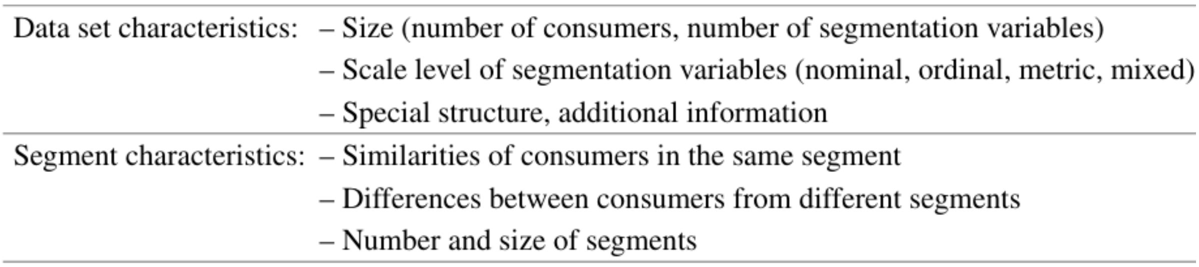
Task	Who is responsible?	Completed?
Convene a market segmentation team meeting.		□
Discuss which consumer characteristics could serve as promising segmentation variables. These variables will be used to extract groups of consumers from the data.		□
Discuss which other consumer characteristics are required 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-5: Extracting Segment (7.1 & 7.2)

7.1 Grouping Consumers

1. The result of a market segmentation analysis is determined as much by the underlying data as by the extraction algorithm chosen. Segmentation methods shape the segmentation solution.
2. Selecting a suitable clustering method requires matching the data analytic features of the resulting clustering with the context-dependent requirements that are desired by the researcher. It is, therefore, important to explore market segmentation solutions derived from a range of different clustering methods.
3. There is no single best algorithm for all datasets. If consumer data is well-structured and well-separated, distinct market segments exist, and the tendencies of different algorithms matter less. If, however, data is not well-structured, the tendency of the algorithm influences the solution substantially. In such situations, the algorithm will impose a structure that suits the objective function of the algorithm.
4. The aim of this chapter is to provide an overview of the most popular extraction methods used in market segmentation, and point out their specific tendencies of imposing structure on the extracted segments.
5. There are two main groups of extraction methods:
   a. Distance-based methods
   b. Model-based methods
6. Distance-based methods use a particular notion of similarity or distance between observations (consumers), and try to find groups of similar observations (market segments).
7. Model-based methods formulate a concise stochastic model for the market segments.
8. In addition to those main two groups of extraction methods, a number of methods exist which try to achieve multiple aims in one step.
9. Data characteristics and expected or desired segment characteristics allow a pre-selection of suitable algorithms to be included in the comparison. Table 7.1 contains the information needed to guide algorithm selection.

Table 7.1 Data set and segment characteristics informing extraction algorithm selection



Data set characteristics:	– Size (number of consumers, number of segmentation variables) – Scale level of segmentation variables (nominal, ordinal, metric, mixed) – Special structure, additional information
Segment characteristics:	– Similarities of consumers in the same segment – Differences between consumers from different segments – Number and size of segments

7.2 Distance-Based Methods

To find groups of similar things with similar activity patterns one needs a notion of similarity or dissimilarity, mathematically speaking: a distance measure.

7.2.1 Distance Measures

1. In a Data matrix each row represents an observation (in this case a tourist), and every column represents a variable (in this case a vacation activity). Mathematically, this can be represented as an n × p matrix where n stands for the number of observations (rows) and p for the number of variables (columns):

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

2. The vector corresponding to the i-th row of matrix X is denoted as xi = (xi1,xi2,...,xip) in the following, such that X ={x1,x2,...xp} is the set of all observations.
3. A distance measure has to comply with a few criteria.
    a. One criterion is symmetry, that is:
       d(x,y) = d(y,x).
    b. A second criterion is that the distance of a vector to itself and only to itself is 0:
       d(x,y) = 0 ⇔ x=y
    c. In addition, most distance measures fulfil the so-called triangle inequality:
       d(x,z) ≤ d(x,y) +d(y,z)
4. Let x = (x1,...,xp) and y = (y1,...,yp) be two p-dimensional vectors. The most common distance measures used in market segmentation analysis are:
    a. Euclidean distance:

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

    b. Manhattan or absolute distance:

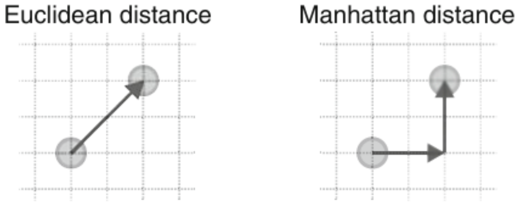
$$d(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{p} |x_j - y_j|$$

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

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

In words: the number of dimensions where both x and y are equal to 1 divided by the number of dimensions where at least one of them is 1.

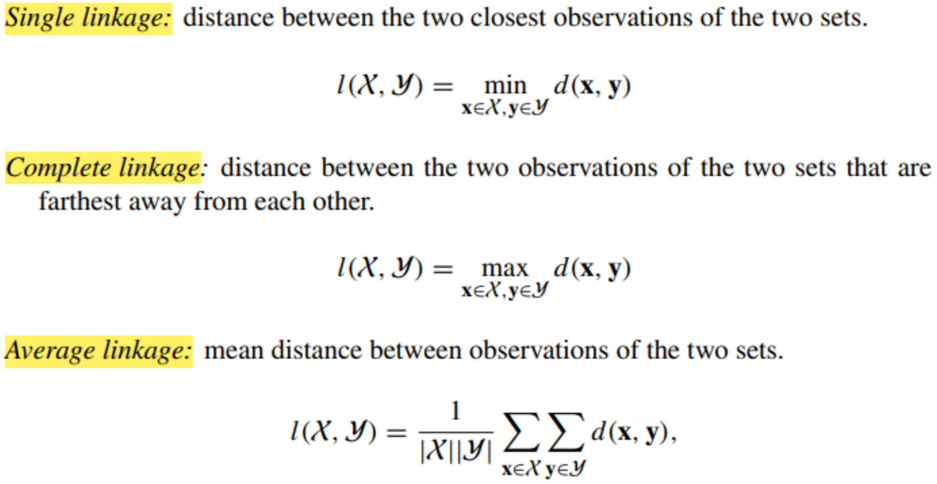
Fig. 7.2 A comparison of Euclidean and Manhattan distance



5. Euclidean distance is the most common distance measure used in market segmentation analysis.
6. Both Euclidean and Manhattan distance treat all dimensions of the data equally; they take a sum over all dimensions of squared or absolute differences.

7.2.2 Hierarchical Methods

1. Hierarchical clustering methods are the most intuitive way of grouping data because they mimic how a human would approach the task of dividing a set of n observations (consumers) into k groups (segments).
2. Divisive hierarchical clustering methods start with the complete data set X and splits it into two market segments in a first step. Then, each of the segments is again split into two segments. This process continues until each consumer has their own market segment.
3. Agglomerative hierarchical clustering approaches the task from the other end. The starting point is each consumer representing their own market segment (n singleton clusters). Step-by-step, the two market segments closest to one another are merged until the complete data set forms one large market segment.
4. Both approaches result in a sequence of nested partitions. A partition is a grouping of observations such that each observation is exactly contained in one group. The sequence of partitions ranges from partitions containing only one group (segment) to n groups (segments)
5. Underlying both divisive and agglomerative clustering is a measure of distance between groups of observations (segments). This measure is determined by specifying (1) a distance measure d(x, y) between observations (consumers) x and y, and (2) a linkage method.
6. Assuming two sets X and Y of observations (consumers), some of the linkage methods for measuring the distance *l*(X, Y) between these two sets of observations:



Single linkage: distance between the two closest observations of the two sets.

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

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

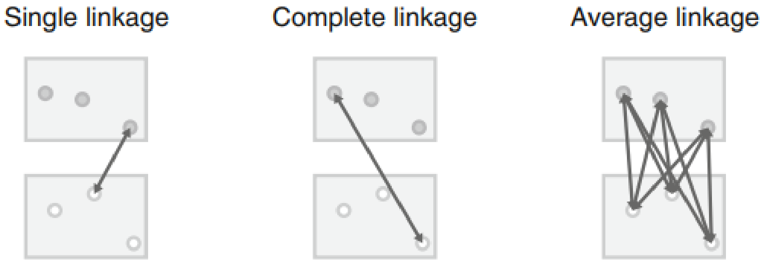


Fig. 7.3 A comparison of different linkage methods between two sets of points

7. The result of hierarchical clustering is typically presented as a dendrogram.
8. A dendrogram is a tree diagram. The root of the tree represents the one-cluster solution where one market segment contains all consumers. The leaves of the tree are the single observations (consumers), and branches in-between correspond to the hierarchy of market segments formed at

each step of the procedure. The height of the branches corresponds to the distance between the clusters. Higher branches point to more distinct market segments

9. At every split into two branches, the left and right branch could be exchanged, resulting in 2n possible dendrograms for exactly the same clustering where n is the number of consumers in the data set.

7.2.3 Partitioning Methods

1. Hierarchical clustering methods are particularly well suited for the analysis of small data sets with up to a few hundred observations.
2. For larger data sets, dendrograms are hard to read, and the matrix of pairwise distances usually does not fit into computer memory clustering methods creating a single partition is more suitable than a nested sequence of partitions. This means that – instead of computing all distances between all pairs of observations in the data set at the beginning of a hierarchical partitioning cluster analysis using a standard implementation – only distances between each consumer in the data set and the centre of the segments are computed.
3. In addition, if only a few segments are extracted, it is better to optimise specifically for that goal, rather than building the complete dendrogram and then heuristically cutting it into segments.

7.2.3.1 k-Means and k-Centroid Clustering

1. The most popular partitioning method is k-means clustering. Within this method, a number of algorithms are available.
2. These algorithms use the squared Euclidean distance. A generalisation to other distance measures, also referred to as k-centroid clustering, is also provided.
3. The Algorithm involves five steps -
    1. Specify the desired number of segments k.
    2. Randomly select k observations (consumers) from data set X and use them as initial set of cluster centroids C = {c1,..., ck}.
    3. Assign each observation xi to the closest cluster centroid to form a partition of the data, that is, k market segments S1,...,Sk where

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

    4. Recompute the cluster centroids (segment representatives) by holding cluster membership fixed, and minimising the distance from each consumer to the corresponding cluster centroid:

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

       For squared Euclidean distance, the optimal centroids are the cluster-wise means, for Manhattan distance cluster-wise medians, resulting in the so-called k-means and k-medians procedures, respectively.

    5. Repeat from step 3 until convergence or a pre-specified maximum number of iterations is reached.
4. The algorithm will always converge: the stepwise process used in a partitioning clustering algorithm will always lead to a solution.
5. The key idea is to systematically repeat the extraction process for different numbers of clusters (or market segments), and then select the number of segments that leads to either the most stable overall segmentation solution, or to the most stable individual segment.

6. The process of selecting random segment representatives is called random initialisation.
7. Specifying the number of clusters (number of segments) is difficult because, typically, consumer data does not contain distinct, well-separated naturally existing market segments.
8. A popular approach is to repeat the clustering procedure for different numbers of market segments (for example: everything from two to eight market segments), and then compare – across those solutions.
9. Both partitions obtained using either hierarchical or partitioning clustering methods are reasonable from a statistical point of view. Which partition is more suitable to underpin the market segmentation strategy of an organisation needs to be evaluated jointly by the data analyst and the user of the segmentation solution using the tools and methods presented.

7.2.3.2 “Improved” k-Means

1. Many attempts have been made to refine and improve the k-means clustering algorithm. The simplest improvement is to initialise k-means using “smart” starting values, rather than randomly drawing k consumers from the data set and using them as starting points.
2. One way of avoiding the problem of the algorithm getting stuck in a local optimum is to initialise it using starting points evenly spread across the entire data space. Such starting points better represent the entire data set.
3. Steinley and Brusco conclude that the best approach is to randomly draw many starting points, and select the best set.

7.2.3.3 Hard Competitive Learning

1. Hard competitive learning, also known as learning vector quantisation (e.g. Ripley 1996), differs from the standard k-means algorithm in how segments are extracted.
2. k-means uses all consumers in the data set at each iteration of the analysis to determine the new segment representatives (centroids). Hard competitive learning randomly picks one consumer and moves this consumer’s closest segment representative a small step into the direction of the randomly chosen consumer.
3. As a consequence of this procedural difference, different segmentation solutions can emerge, even if the same starting points are used to initialise the algorithm. It is also possible that hard competitive learning finds the globally optimal market segmentation solution, while k-means gets stuck in a local optimum (or the other way around).
4. Neither of the two methods is superior to the other; they are just different.

7.2.3.4 Neural Gas and Topology Representing Networks

1. A variation of hard competitive learning is the neural gas algorithm.
2. Here, not only the segment representative (centroid) is moved towards the randomly selected consumer. Instead, also the location of the second closest segment representative (centroid) is adjusted towards the randomly selected consumer. However, the location of the second closest representative is adjusted to a smaller degree than that of the primary representative. Neural gas has been used in applied market segmentation analysis
3. A further extension of neural gas clustering are topology representing networks
4. The underlying algorithm is the same as in neural gas. In addition, topology representing networks count how often each pair of segment representatives (centroids) is closest and second closest to a randomly drawn consumer.

7.2.3.5 Self-Organising Maps

1. Another variation of hard competitive learning are self-organising maps (Kohonen 1982, 2001), also referred to as self-organising feature maps or Kohonen maps.
2. The self-organising map algorithm is similar to hard competitive learning: a single random consumer is selected from the data set, and the closest representative for this random consumer moves a small step in their direction. In addition, representatives which are direct grid neighbours of the closest representative move in the direction of the selected random consumer.
3. The process is repeated many times; each consumer in the data set is randomly chosen multiple times, and used to adjust the location of the centroids in the Kohonen map

7.2.3.6 Neural Networks

1. Auto-encoding neural networks for cluster analysis work mathematically differently than all cluster methods presented so far. The most popular method from this family of algorithms uses a so-called single hidden layer perceptron.
2. The input layer has one so-called node for every segmentation variable.
3. Neural network clustering is an example of a fuzzy segmentation with membership values between 0 (not a member of this segment) and 1 (member of only this segment). Membership values between 0 and 1 indicate membership in multiple segments.

7.2.4 Hybrid Approaches

1. Several approaches combine hierarchical and partitioning algorithms in an attempt to compensate for the weaknesses of one method with the strengths of the other
2. The strengths of hierarchical cluster algorithms are that the number of market segments to be extracted does not have to be specified in advance and that similarities of market segments can be visualised using a dendrogram.
3. The biggest disadvantage of hierarchical clustering algorithms is that standard implementations require substantial memory capacity, thus restricting the possible sample size of the data for applying these methods. Also, dendrograms become very difficult to interpret when the sample size is large.
4. The strength of partitioning clustering algorithms is that they have minimal memory requirements during calculation, and are therefore suitable for segmenting large data sets.
5. The disadvantage of partitioning clustering algorithms is that the number of market segments to be extracted needs to be specified in advance
6. The basic idea behind hybrid segmentation approaches is to first run a partitioning algorithm because it can handle data sets of any size. However, the partitioning algorithm used initially does not generate the number of segments sought. Rather, a much larger number of segments is extracted. Then, the original data is discarded and only the centres of the resulting segments (centroids, representatives of each market segment) and segment sizes are retained and used as input for the hierarchical cluster analysis. At this point, the data set is small enough for hierarchical algorithms, and the dendrogram can inform the decision of how many segments to extract.

7.2.4.1 Two-Step Clustering

1. IBM SPSS (IBM Corporation 2016) implemented a procedure referred to as two-step clustering.
2. The two steps consist of running a partitioning procedure followed by a hierarchical procedure.
3. First we cluster the original data using k-means with k much larger than the number of market segments sought.

4. The choice of the original number of clusters to extract is not crucial because the primary aim of the first step is to reduce the size of the data set by retaining only one representative member of each of the extracted clusters. Such an application of cluster methods is often also referred to as vector quantisation.
5. It cannot be determined from the hierarchical cluster analysis, however, which consumer belongs to which market segment. This cannot be determined because the original data was discarded.
6. What needs to happen in the final step of two-step clustering, therefore, is to link the original data with the segmentation solution derived from the hierarchical analysis.

7.2.4.2 Bagged Clustering

1. Bagged clustering also combines hierarchical clustering algorithms and partitioning clustering algorithms, but adds bootstrapping.
2. Bootstrapping can be implemented by random drawing from the data set with replacement. That means that the process of extracting segments is repeated many times with randomly drawn (bootstrapped) samples of the data. Bootstrapping has the advantage of making the final segmentation solution less dependent on the exact people contained in consumer data.
3. Bagged clustering is suitable in the following circumstances (Dolnicar and Leisch 2004; Leisch 1998):
    - If we suspect the existence of niche markets.
    - If we fear that standard algorithms might get stuck in bad local solutions.
    - If we prefer hierarchical clustering, but the data set is too large.
4. Bagged clustering is an example of an ensemble clustering method. These methods are called ensemble methods because they combine several segmentation solutions into one.
5. Ensembles are also referred to as committees. Every repeated segment extraction using a different bootstrap sample contributes one committee member. The final step is equivalent to all committee members voting on the final market segmentation solution.
6. An additional advantage of bagged clustering – compared to standard partitioning algorithms – is that the two-step process effectively has a built-in variable uncertainty analysis. This analysis provides element-wise uncertainty bands for the cluster centres.