Market Segmentation Analysis

Understanding it, Doing it, Making it Useful

Chaganti Venkatarami Reddy, Subesha Sasmal, Parth Sukla

Market Segmentation Analysis by Clustering

"*Systematic sampling is like random sampling, except the members are not chosen totally randomly. You can choose members at regular intervals here*"

Contents

Github : Chaganti Reddy/Market Segmentation

1 Introduction

1.1 What is market Segmentation?

Market Segmentation is a decision-making tool for the marketing manager in the crucial task of selecting a target market for a given product and designing an appropriate marketing mix. Market Segmentation is one of the key building blocks of strategic marketing.

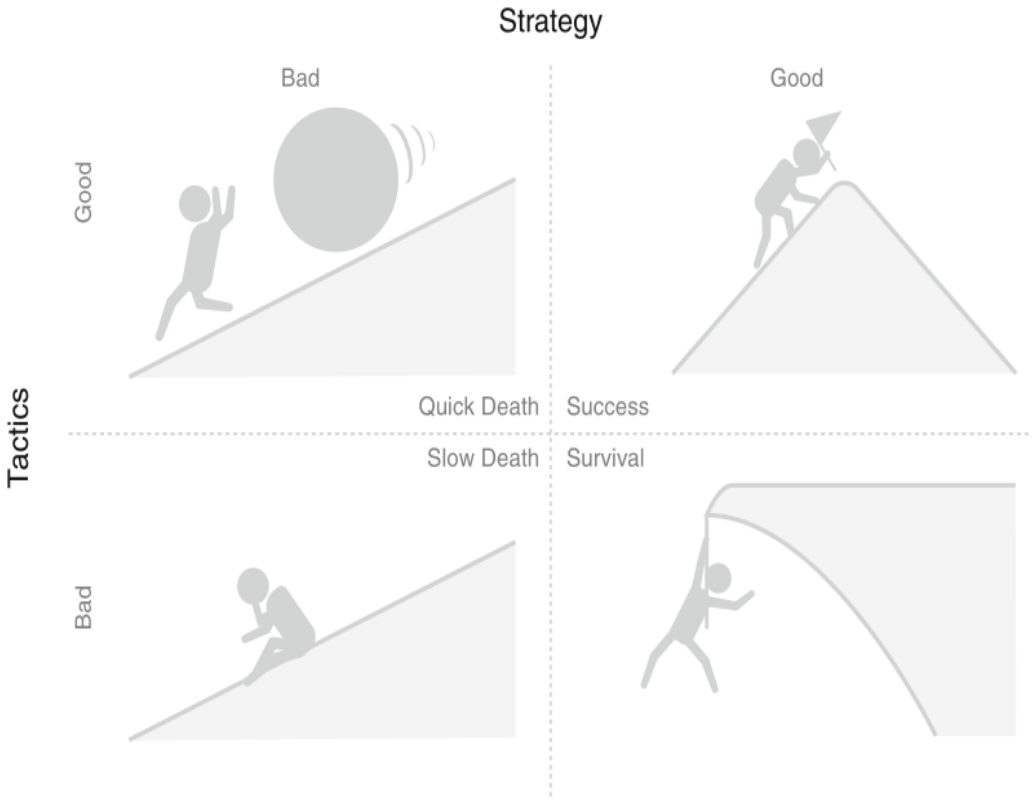


Figure 1: *The assymetry of strategic and tactical marketing*

Conceptually, market segmentation sits between two extreme views that :

- All objects are unique and inviolable
- The population is homogeneous

An ideal market segmentation situation for the simplest case of two product features are given in the below table :

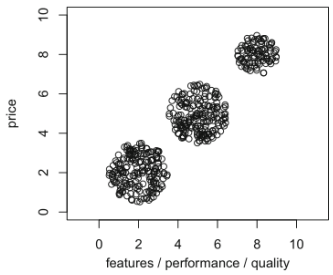
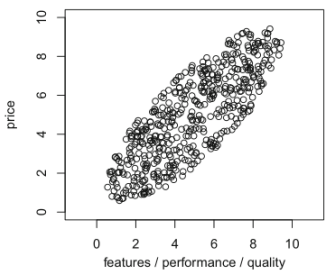
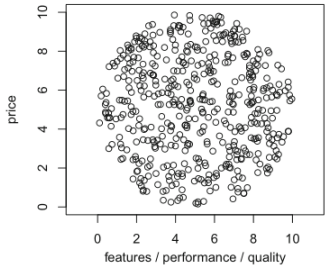
Natural Segmentation	**Reproducible Segmentation**	**Constructive Segmentation**
Cluster Structure in the data	Structure in data, but not cluster structure	No Structure in data
Market segments are received	Market segments are created	Market segments are created
Repeated calculations lead to very similar results (very stable across repeated calculations)	Repeated calculations lead to similar results (reasonable stable across repeated calculations)	Repeated calculations lead to different results (unstable across repeated calculations)
Example data : 	Example data : 	Example data : 

Table 1: *Data driven market segmentation approaches based on data structures*

The x-axis shows the number of desired features of a mobile telephone, and the y-axis shows the price consumers willing to pay. Here three market segments exists :

- A small segment characterized by wanting many mobile telephone features and willing to pay a lot of money.
- A large segment containing consumers who desire the exact opposite (a simple, cheap mobile phone).
- Another large segment in the middle containing members who want a mid-range phone at a mid-range price.

1.2 Benefits of Market Segmentation

Market Segmentation has a number of benefits. At the most general level, market segmentation forces organizations to take stock of where they stand, and where they want to be in future. When implemented well, market segmentation also leads to tangle benefits, including a better understanding of differences between consumers, which improves the match of organizational strengths and consumer needs.

2 Market Segmentation Analysis

2.1 The Layers of Market Segmentation Analysis

Market Segmentation Analysis is the process of grouping consumers into naturally existing or artificially created segments of consumers who share similar product preferences or characteristics.

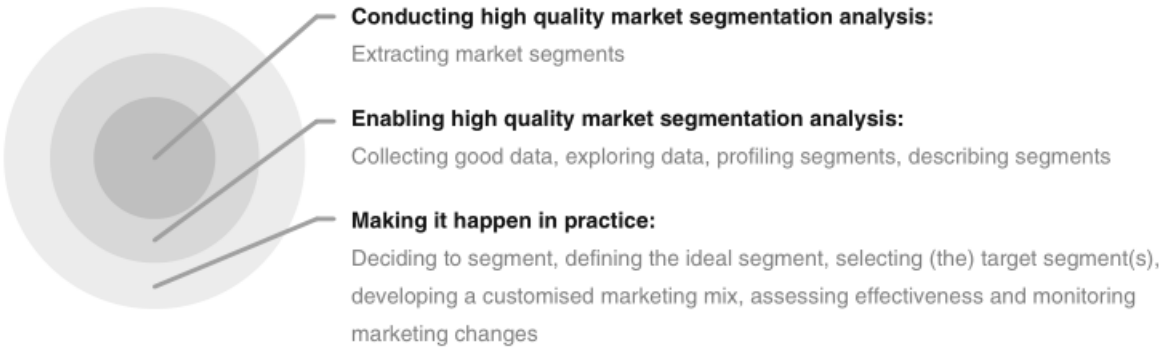


Figure 2: *The layers of market segmentation analysis*

To ensure that the grouping of consumers is one of the highest quality, a number of additional tasks are required, as illustrated in the second layer in Fig-2. All these tasks are still primarily technique in nature. Collecting good data, for example, is critically important. The statical segment extraction process at the core of market segmentation analysis cannot compensate for bad data. The grouping of consumers can always only be as good as the data provided to the segment extraction method. Profiling and describing segments help users to understand each of the segments, and select which one(s) to target. When one or more target segments have been chosen, profiling and describing segments have been chosen, profiling and describing segments inform the development of the customized marketing mix.

If all the tasks in the first (core) and second layer of market segmentation analysis have been implemented well, the result is a theoretically excellent market segmentation solution. But a theoretically excellent market segmentation solution is meaningless unless users can convert such a solution into strategic marketing decisions and tactical marketing action. Therefore, for any market segmentation analysis to be complete, a third layer is required.

2.2 Approaches to Market Segmentation Analysis

No single approach is best when conducting market segmentation analysis. Instead, approaches to market segmentation analysis can be systematized in a number of different ways. We present two systematics here, one uses as its basis the extent to which the organization conducting the market segmentation study is willing or able to make changes to their current approach of targeting the market or a segment of the market.

constraints. The second systematics is based on the nature of the segmentation variable or variables used in the market segmentation analysis.

2.2.1 Based on Organizational Constraints

The quantitative survey-based approach, the creation of segments from existing consumer classifications, and the emergence of segments from qualitative research. These three approaches differ in how radical the resulting change is for the organization. Looking at each one of these approaches in more detail, the segment revolution or quantitative survey-based segmentation approach tends to be seen as the proto-typical market segmentation analysis. The key assumption underlying this approach is that the organization conducting market segmentation analysis is willing and able to start from scratch; to forget entirely about how its marketing was conducted in the past, and commence the segmentation process with a genuinely open mind. The third approach is that of exploratory research pointing to segments. Under this approach, market segments are stumbled upon as part of an exploratory research process possibly being undertaken for a very different purpose initially. In times of big data, such segment mutation may well result from data mining of streams of data, rather than from qualitative research. The same holds for segment evolution.

2.2.2 Based on the choice of Segmentation Variable

A more technical way of systematizing segmentation approaches is to use as a basis the nature of consumer characteristics used to extract market segments. Sometimes one single piece of information about consumers (one segmentation variable) is used. This statistical problem is unidimensional. One example is age. The resulting segments are age groups, and older consumers could be selected as a target segment. In other cases, multiple pieces of information (multiple segmentation variables) about consumers are important. In this case, the statistical problem becomes multidimensional. One example could be consumers expenditure patterns.

Variable	**Dimensions**	**Sample survey questions**
Age	Unidimensional	How old are you?
Gender	Unidimensional	Are you female or male?
Country of region	Unidimensional	Where do you live?
Prior purchase	Unidimensional	Have you booked a cruise trip before?
Benefits sought	Multidimensional	When booking flights online, do you care about • convenience • value for money • speed • ability to compare fares
Motives	Multidimensional	When choosing a vacation, do you want to • rest and relax • explore new things • meet new people • learn about other cultures • get away from everyday routine

Table 2: *Examples of commonly used segmentation variables*

When data-driven segmentation is conducted, the organisation has certain assumptions about the consumer characteristics that are critical to identifying a suitable market segment to target, but does not know the exact profiles of suitable target segments. The aim of data-driven segmentation, therefore, is two fold:

- to explore different market segments that can be extracted using the segmentation variables chosen,
- to develop a detailed profile and description of the segment(s) selected for targeting.

Commonsense and data-driven segmentation are two extremes, the two pure forms of segmentation approaches based on the nature of the segmentation criterion.

2.3 Data Structure and Data-Driven Market Segmentation Approaches

When conducting data-driven market segmentation, data analysts and users of market segmentation solutions often assume that market segments naturally exist in the data.

Such naturally occurring segments, it is assumed, need to merely be revealed and described. In real consumer data, naturally existing, distinct and well-separated market segments rarely exist.

	Commonsense /commonsence segmentation	**Commonsense / data-driven segmentation**	**Data-driven/ commonsence segmentation**	**Data-driven/data-driven segmentation**
Primary Segmentation variable(s)	Commonsense (e.g. age, country of origin)	Commonsense (e.g. age, country of origin)	Data-driven (e.g. expenditures, vacation activities)	Data-driven (e.g. travel motives, expenditures)
Secondary Segmentation variable(s)	Commonsense (e.g. gender, seeking adventure or not)	Data-driven (e.g. travel motives, vacation activities)	Commonsense (e.g. gender, family status)	Data-driven (e.g. vacation activities, information sources used)
Example 1	Young female tourists	Matured aged tourists who play golf, enjoy wine-tastings and fine dancing	Tourists who engage in a large number of activities that attract an entrance fee, such as visiting theme parks and zoos, and who travel with their family.	Tourists who want to learn about culture and local people, and who attend local cultural events and food festivals.
Example 2	Adventure travellers from Australia.	Older tourists who take a holiday to relax, have a change of usual surroundings, and enjoy health / beauty treatments.	Tourists who surf and enjoy the night life of the destination, and who are male.	Tourists who have high expenditures in a wide range of expenditure categories at the destination, and use airline loyalty program mail outs as their key travel information source .

Table 3: *Combinations of segmentation approaches based on the nature of segmentation variables used*

2.3.1 Types of Data-driven Market Segmentation

There are three possible conceptual approaches to data-driven market segmentation

- Natural Segmentation
- Reproducible Segmentation
- Constructive Segmentation

Natural Segmentation

The term natural segmentation reflects the traditional view that distinct market segments exist in the data, and that the aim of market segmentation analysis is to find them. This traditional view is reflected well in the statement that the initial premise in segmenting a market is that segments actually do exist.

Reproducible Segmentation

The term reproducible segmentation refers to the case where natural market segments do not exist in the data. But the data are not entirely unstructured either. Rather, the data contain some structure – other than cluster structure – making it possible to generate the same segmentation solution repeatedly. The ability to repeatedly reveal the same or very similar market segments, makes results of data-driven segmentation studies less random and more reliable. Reliable results represent a stronger basis for long-term strategic segmentation decisions.

Constructive Segmentation

Finally, the term constructive segmentation refers to the case where neither cluster structure nor any other data structure exists, which would enable the data analyst to reproduce similar segmentation solutions repeatedly across applications.

The above data can be referred and easily understand in Table-1

2.4 Market Segmentation Analysis Step-by-Step

Market Segmentation analysis is a ten-step approach. The basic structure is same for both commonsense and data-driven market segmentation:

Figure 3: *Ten steps of Market Segmentation Analysis*

- an organization needs to weigh up the advantages and disadvantages of pursuing a segmentation strategy, and decide whether or not to go ahead .
- Next, the organization needs to specify characteristics of their ideal market segment.
- Only after this preliminary and predominantly conceptual work is finalized, is empirical data collected or compiled from existing sources.
- These data need to be explored before market segments are extracted.
- The resulting market segments are profiled.
- The resulting market segments are described.
- Next is the point of no return where the organization carefully selects one or a small number of market segments to target.
- Based that target choice, a customized marketing mix is developed.
- Upon completion of the market segmentation analysis, the success of implementing a market segmentation strategy needs to be evaluated, and segments need to be continuously monitored for possible changes in size or in characteristics.

  Such changes may require modifications to the market segmentation strategy.

  Although the ten steps of market segmentation analysis are the same for commonsense and data-driven segmentation, different tasks need to be completed for each one of those approaches. Typically, data-driven segmentation requires additional decisions to be made.

3 Ten Steps of Market Segmentation Analysis

3.1 Step-1: Deciding (not) to Segment

3.1.1 Implications of committing to Market Segmentation

Although Market Segmentation has been developed to be a key marketing strategy applied in many organizations, it is not always the best decision to pursue such a strategy. Before investing time and resources in a market segmentation analysis, it is important to understand the implications of pursuing a market segmentation strategy. The key implication is that the organization needs to commit to the segmentation strategy on the long term. Segmenting a market is not free. There are costs of performing research, fielding surveys, and focus groups, designing multiple packages, and designing multiple advertisements and communication messages.

3.1.2 Implementation Barriers

The first group of barriers relates to senior management :

- Lack of leadership
- Pro-active championing
- Commitment and involvement in market segmentation process by senior leadership.

The second group of barriers relates to organizational culture :

- Lack of market or consumer orientation
- Resistance to change and new ideas
- Lack of creative thinking
- Bad communication
- Lack of sharing of information
- Insights across organizational units
- Short term thinking
- Unwillingness to make changes
- Lack of training

Closely linked to these barriers is the lack of a formal marketing function or at least a qualified marketing expert in the organization. The higher the market diversity and the larger the organizations, the more important is a high degree of formalization. the lack of a qualified data manager and analyst in the analyst in the organization can also represent major stumbling blocks.

3.2 Step-2: Specifying the Ideal Target Segment

3.2.1 Segment Evaluation Criterial

The third layer of market segmentation analysis (illustrated in Fig-2) depends primarily on user input. It is important to understand that

- For a market segmentation analysis to produce results that are useful to an organisation,
- User input cannot be limited to either a briefing at the start of the proces or,
- The development of a marketing mix at the end.

In step 2 the organisation must determine two sets of segment evaluation criteria :

- **Knock-out Criteria :** These criteria are the essential, non-negotiable features of segments and the organisation would consider targeting.
- **Attractiveness Criteria :** These criteria are used to evaluate the relative attractiveness of the remaining market segments - those in compliance with the knock-out criteria.

Source	Evaluation Criteria
Day	Measurable, Substantial, Accessible, Sufficiently different, At suitable life-cycle stage
Croft	Large enough, Growing, Competitively advantageous, Profitable, Likely technological changes, Sensitivity to price, Barriers to entry, Buyer or supplier bargaining power, Socio-political considerations, Cyclicality and seasonality, Life-cycle position
Myers	Large enough, Distinguishable, Accessible, Compatible with company
Wedel & Kamakura	Identifiable, Substantial, Accessible, Responsive, Stable, Actionable.
Perreault Jr & McCarthy	Substantial, Operational, Heterogeneous between, Homogeneous within.
Lilien & Rangaswamy	Large enough (market potential, current market penetration), Growing (past growth forecasts of technology change), Competitively advantageous (barriers to entry, barriers to exit, position of competitors), Segment saturation (gaps in marketing), Protectable (patentable products, barriers to entry), Environmentally risky (economic, political, and technological change), Fit (coherence with company's strengths and image), Relationships with other segments (synergy, cost interactions, image transfers, cannibalisation), Profitable (entry costs, margin levels, return on investment)
McDonald & Dunbar	Segment factors (size, growth rate per year, sensitivity to price, service features and external factors, cyclicality, seasonality, bargaining power of upstream suppliers), Competition (types of competition, degree of concentration, changes in type and mix, entries and exits, changes in share, substitution by new technology, degrees and type of integration), Financial and economic factors (contribution margins, capacity utilisation, leveraging factors, such as experience and economies of scale, barriers to entry, or exit), Technological factors (maturity and volatility, complexity, differentiation, patents and copyrights, manufacturing processes), Socio-political factors (social attitudes and trends, laws and government agency regulations, influence with pressure groups and government representatives, human factors, such as unionisation and community acceptance)

...

Continued ···

Dibb & Simkin	Homogeneous, Large enough, Profitable, Stable, Accessible, Compatible, Actionable
Sternthal & Tybout	Influence of company's current position in the market on growth opportunities, Competitor's ability and motivation to retaliate, Competence and resources, Segments that will prefer the value that can be created by the firm over current market offerings, Consumer motivation and goals indicating gaps in marketplace offerings when launching a new company.
west et al.	Differentiable, Measurable, Substantial, Accessible, Actionable.
winer & Dhar	Parsimonious, Large enough, Growing, Competitively advantageous.
Jain	Measurable, Accessible, Substantial, Develops maximum differential in competitive strategy, Preserves competitive advantage, Valid even though imitated.
Kotler & Keller	Measurable, Substantial, Accessible, Differentiable, Actionable, Segment rivalry (competition), Potential entrants, Substitutes, Power of buyers, Power of suppliers, Compatible with company.
Pride et al.	Sales estimates (potential sales for product item, product line, geographical area in the short, medium or long term), Competitive assessment, Cost estimates, Long-term profit opportunities, Financial resources, Managerial skills, Employee expertise, Facilities to compete effectively, Fit with corporate objectives, Legal issues, Conflicts with stakeholders, Technological advances.
Sharp	Measurable, Targetable, Large Enough, Profitable.

Table 4: *Criteria proposed in the literature for the evaluation of the market segments in chronological order*

3.2.2 Knock-Out Criteria

Knock-out criteria are used to determine if market segments resulting from the market segmentation analysis qualify to be assessed using segment attractiveness criteria. The first set of such criteria includes :

- Sustainability
- Measurability
- Accessibility

Additional criteria recommended that fall into Knock-out criteria category :

- **The segment must be homogenous;** members of segment must be similar to one another.
- **The segment must be distinct;** members of segment must be distinctly different from members of other segments.
- **The segment must be large enough;** the segment must contain enough consumers to make it worthwhile to spend extra money on customising the marketing mix for them.
- **The segment must be matching the strengths of the organisation;** the organisation must have the capability to satisfy segment members' needs.
- **Members of the segment must be identifiable;** it must be possible to spot them in the marketplace.
- **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 customized marketing mix accessible to them.

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 need to be specified.

3.2.3 Attractiveness Criteria

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. The attractiveness across all criteria determines whether a market segment.

3.2.4 Implementing a Structured Process

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 along one axis, and organisational competetiveness on the other axis. Factors which constitute both segment attractiveness and organizational 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 the most important for the organisation. 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.

3.3 Step-3: Collecting Data

3.3.1 Segmentation Variables

Empirical data forms the basis of both commonsense and data-driven market segmentation. Empirical data is used to identify or create market segments and – later in the process – describe these segments in detail. The term segmentation variable refers to the variable in the empirical data used in commonsense segmentation split the sample into market segments. In commonsense segmentation, the segmentation variable is typically on single characteristic of the consumers in the sample (as shown in below Fig. . .)

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					

Figure 4: *Gender as a possible segmentation variable in commonsense market segmentation*

Each row in this table represents on consumer, each variable represents one characteristic of that consumer. An entry of 1 in the data set indicates that the consumer has that characteristic. An entry of 0 indicates that the consumer does not have that characteristic. The commonsense segmentation illustrated in Fig-4 uses gender as the

segmentation variable. 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**. They are used to describe the segments in detail.

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 organisation (as shown in below Fig).

Sociodemographics		Travel behaviour	Benefits sought				
gender	age	N° of vacations	relaxation	action	culture	explore	meet people
Female	34	2	1	0	1	0	1
Female	55	3	1	0	1	0	1
Male	87	2	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	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
	descriptor variables		segmentation variables				

Figure 5: *Segmentation variable in data-driven market segmentation*

When commonsense segments are extracted – even if the nature of the segments is known in advance – data quality is critical to both

- assigning each person in the sample to the correct market segment, and
- being able to correctly describe the segments.

The same holds for data-driven market segmentation where data quality deter- mines 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.

3.3.2 Geographic Segmentation

Geographic information is seen as the original segmentation criterion used for the purpose of market segmentation. Typically

- when geographic segmentation is used
- the consumer's location of residence serves as the only criterion to form market segments.

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.

The **key disadvantage** is that living in the same country or area does not necessarily mean that people share other characteristics relevant to marketers, such as benefits they seek when purchasing a product.

Despite the potential shortcomings of using geographic information as the segmentation variable, the location aspect has experienced a revival in international market segmentation studies aiming to extract market segments across geographic boundaries. Such an approach is challenging because the segmentation variable(s) must be meaningful across all the included geographic regions, and because of the known biases that can occur if surveys are completed by respondents from different cultural backgrounds.

3.3.3 Socio-Demographic Segmentation

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; even in times where men are targeted, the female and male segments are treated distinctly differently), baby products (associated with gender), retirement villages (associated with age), tourism resort products (associated with having small children or not).

As is the case with geographic segmentation, socio-demographic segmentation criteria have the advantage that segment membership can easily be determined for every consumer. In some instances, the socio-demographic criterion may also offer an explanation for specific product preferences (having children, for example, is the actual reason that families choose a family vacation village where previously, as a couple, their vacation choice may have been entirely different). 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.

3.3.4 Psychographic Segmentation

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. psychographics was intended as an umbrella term to cover all measures of the mind. Benefit segmentation, is arguably the most popular kind of psychographic segmentation. Lifestyle segmentation is another popular psychographic segmentation approach; it is based on people's activities, opinions and interests.

Psychographic criteria are, by nature, more complex than geographic or socio-demographic criteria because it is difficult to find a single characteristic of a person that will provide insight into the psychographic dimension of interest. As a consequence, most psychographic segmentation studies use a number of segmentation variables, for example: a number of different travel motives, a number of perceived risks when going on vacation.

The psychographic approach has the advantage that it is generally more reflective of the underlying reasons for differences in consumer behaviour. For example, tourists whose primary motivation to go on vacation is to learn about other cultures, have a high likelihood of undertaking a cultural holiday at a destination that has ample cultural treasures for them to explore. 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.

3.3.5 Behavioural Segmentation

Another approach to segment extraction is to search directly for similarities in behaviour or reported behaviour. A wide range of possible behaviours can be used for this purpose, including prior experience with the product, frequency of purchase, amount spent on purchasing the product on each occasion (or across multiple purchase occasions), and information search behaviour.

The key advantage of behavioural approaches is that :

- if based on actual behaviour rather than stated behaviour or stated intended behaviour
- the very behaviour of interest is used as the basis of segment extraction.

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.

3.3.6 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 below.

1. Choice of Variables

   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. 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. Including unnecessary variables can make questionnaires long and tedious for respondents, which, in turn, causes respondent fatigue. 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 included as segmentation variables divert the attention of the segment extraction algorithm away from information critical to the extraction of optimal market segments. Such variables are referred to as **noisy variables or masking variables** and have been repeatedly shown to prevent algorithms from identifying the correct segmentation solution.
2. Response Options

   Answer options provided to respondents in surveys determine the scale of the data available for subsequent analyses. Because many data analytic techniques are based on distance measures, not all survey response options are equally suitable for segmentation analysis. Options allowing respondents to answer in only one of two ways,

   - generate binary
   - 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. Options allowing respondents to select an answer from a range of unordered categories correspond to **nominal variables**.

3. Response Styles

   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. If a bias is displayed by a respondent consistently over time, and independently of the survey questions asked, it represents a response style. A wide range of response styles manifest in survey answers, including respondents' tendencies to use extreme answer options (STRONGLY AGREE, STRONGLY DISAGREE), to use the midpoint (NEITHER AGREE NOR DISAGREE), and to agree with all statements. Response styles affect segmentation results because commonly used segment extraction algorithms cannot differentiate between a data entry reflecting the respondent's belief from a data entry reflecting both a respondent's belief and a response style.

4. Sample Size

   Many statistical analyses are accompanied by sample size recommendations. Not so market segmentation analysis. Fig-6 illustrates the problem any segmentation algorithm faces if the sample is insufficient. The market segmentation problem in this figure is extremely simple because only two segmentation variables are used. Yet, when the sample size is insufficient (left plot), it is impossible to determine which the correct number of market segments is. If the sample size is sufficient, however (right plot) it is very easy to determine the number and nature of segments in the data set.

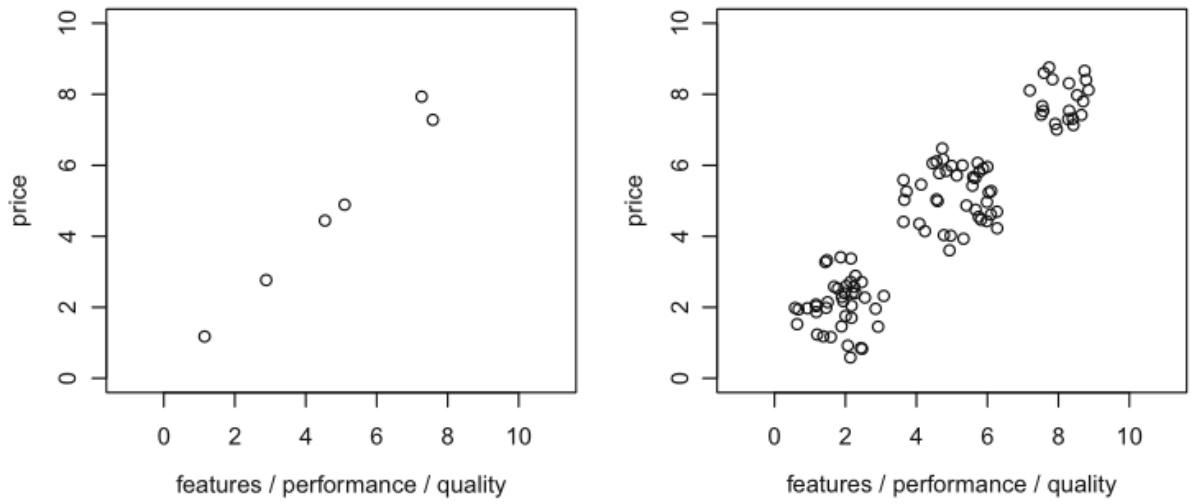


Figure 6: *Illustrating the importance of sufficient sample size in market segmentation analysis*

Viennese psychologist Formann 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.

Qiu and Joe developed a sample size recommendation for constructing artificial data sets for studying the performance of clustering algorithms. According to Qiu and Joe , 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 ($10 \cdot p \cdot k$) where p represents the number of segmentation variables and k represents the number of segments. If segments are unequally sized, the smallest segment should contain a sample of at least ($10 \cdot p$).

In Fig-7, the x-axis plots the sample size (ranging from 10 to 100 times the number of segmentation variables). The y-axis plots the effect of an increase in sample size on the adjusted Rand index. The higher the effect, the better the algorithm identified the correct market segmentation solution.

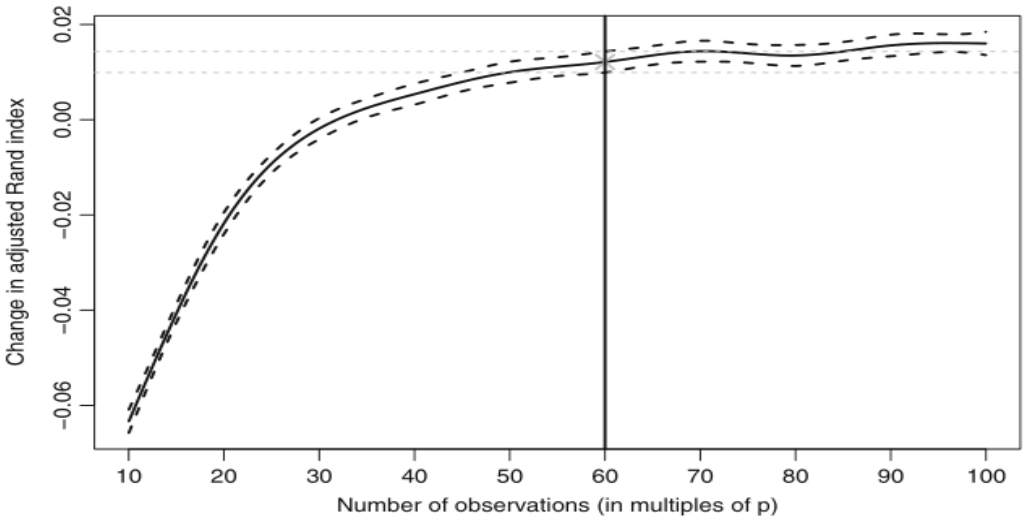


Figure 7: *Effect of sample size on the correctness of segment recovery in artificial data.*

Fig-8 shows the results from this large-scale simulation study using artificial data. Again, the axes plot the sample size, and the effect of increasing sample size on the adjusted Rand index, respectively. As can be seen in Fig-8 larger sample sizes always improve an algorithm's ability to identify the correct market segmentation solution.

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). The recommendation by Dolnicar et al. is to ensure the data contains at least 100 respondents for each segmentation variable. Results from this study also highlight the importance of collecting high-quality unbiased data as the basis for market segmentation analysis.

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:

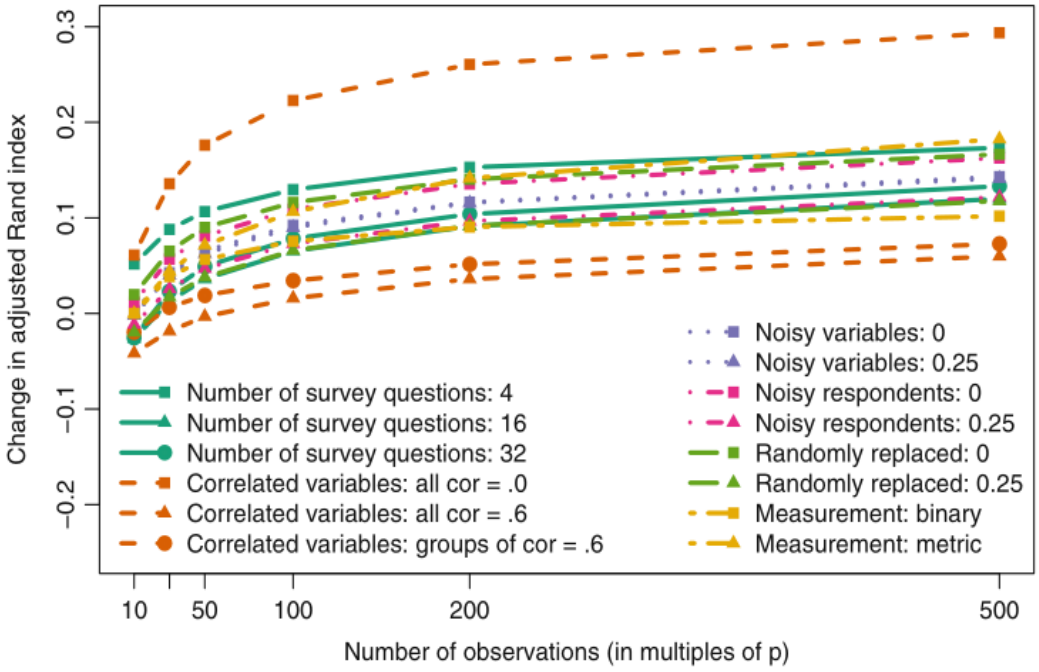


Figure 8: *Sample size requirements in dependence of market and data characteristics.*

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

3.3.7 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. The strength of such data lies in the fact that they represent actual behavior of consumers, rather than statements of consumers about their behavior 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. Another advantage is that such data are usually automatically generated and

- if organisations are capable of storing data in a format that makes them easy to access

- no extra effort is required to collect data.

The danger of using internal data is that it may be systematically biased by over-representing existing customers. What is missing is information about other consumers the organisation may want to win as customers in future, which may differ systematically from current customers in their consumption patterns.

3.3.8 Data from Experimental Studies

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. For example, they can be the result of tests how people respond to certain advertisements. The response to the advertisement could then be used as a segmentation criterion. Experimental data can also result from choice experiments or conjoint analyses. The aim of such studies is to present consumers with carefully developed stimuli consisting of specific levels of specific product attributes. Consumers then indicate which of the products – characterized by different combinations of attribute levels – they prefer. 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.

3.4 Step-4: Exploring Data

3.4.1 A First Glimpse at Data

After data collection, exploratory data analysis cleans and – if necessary – preprocesses the data. This exploration stage also offers guidance on the most suitable algorithm for extracting meaningful market segments. At a more technical level, data exploration helps to

- identify the measurement levels of the variables;
- investigate the univariate distributions of each of the variables and,
- assess dependency structures between variables.

3.4.2 Data Cleaning

The first step before commencing data analysis is to clean the data. This includes checking if all values have been recorded correctly, and if consistent labels for the levels of categorical variables have been used. For many metric variables, the range of plausible values is known in advance. For example, age (in years) can be expected to lie between 0 and 110. It is easy to check whether any implausible values are contained in the data, which might point to errors during data collection or data entry. Similarly, levels of 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. Any other values are not permissible, and need to be corrected as part of the data cleaning procedure. All these process we can see in code in case study which is in github of attached link at the end.

3.4.3 Descriptive Analysis

Being familiar with the data avoids misinterpretation of results from complex analyses. Descriptive numeric and graphic representations provide insights into the data. Statistical software packages offer a wide variety of tools for descriptive analysis. In R, we obtain a numeric summary of the data with command summary(). This command returns the range, the quartiles, and the mean for numeric variables. For categorical variables, the command returns frequency counts. The command also returns the number of missing values for each variable. Helpful graphical methods for numeric data are histograms, boxplots and scatter plots. Bar plots of frequency counts are useful for the visualisation of categorical variables. Mosaic plots illustrate the association of multiple categorical variables. We explain mosaic plots in Step 7 where we use them to compare market segments. Histograms visualise the distribution of numeric variables. They show how often observations within a certain value range occur. Histograms reveal if the distribution of a variable is unimodal and symmetric or skewed. 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. Usu-

ally, 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.

3.5 Step-5: Extracting Segments

3.5.1 Grouping Consumers

Data-driven market segmentation analysis is exploratory by nature. Consumer data sets are typically not well structured. Consumers come in all shapes and forms; a two-dimensional plot of consumers' product preferences typically does not contain clear groups of consumers. Rather, consumer preferences are spread across the entire plot. The combination of exploratory methods and unstructured consumer data means that results from any method used to extract market segments from such data will strongly depend on the assumptions made on the structure of the segments implied by the method. The result of a market segmentation analysis, therefore, is determined as much by the underlying data as it is by the extraction algorithm chosen. Segmentation methods shape the segmentation solution.

Many segmentation methods used to extract market segments are taken from the field of cluster analysis. In that case, market segments correspond to clusters. As pointed out by Hennig and Liao, 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. It is also important to understand how different algorithms impose structure on the extracted segments. One of the most illustrative examples of how algorithms impose structure is shown in Fig-9. In this figure, the same data set – containing two spiralling segments – is segmented using two different algorithms, and two different numbers of segments. The top row in Fig-9 shows the market segments obtained when running k-means cluster analysis with 2 (left) and 8 segments (right), respectively. As can be seen, k-means cluster analysis fails to identify the naturally existing spiral-shaped segments in the data. This is because k-means cluster analysis aims at finding compact clusters covering a similar range in all dimensions.

The bottom row in Fig-9 shows the market segments obtained from single linkage hierarchical clustering. This algorithm correctly identifies the existing two spiralling segments, even if the incorrect number of segments is specified up front. This is because the single linkage method constructs snake-shaped clusters. When asked to return too many (8) segments, outliers are defined as micro-segments, but the two main spirals are still correctly identified. k-means cluster analysis fails to identify the spirals because it is designed to construct round, equally sized clusters. As a consequence, the k-means algorithm ignores the spiral structure and, instead, places consumers in the same market segments if they are located close to one another (in Euclidean space), irrespective of the spiral they belong to.

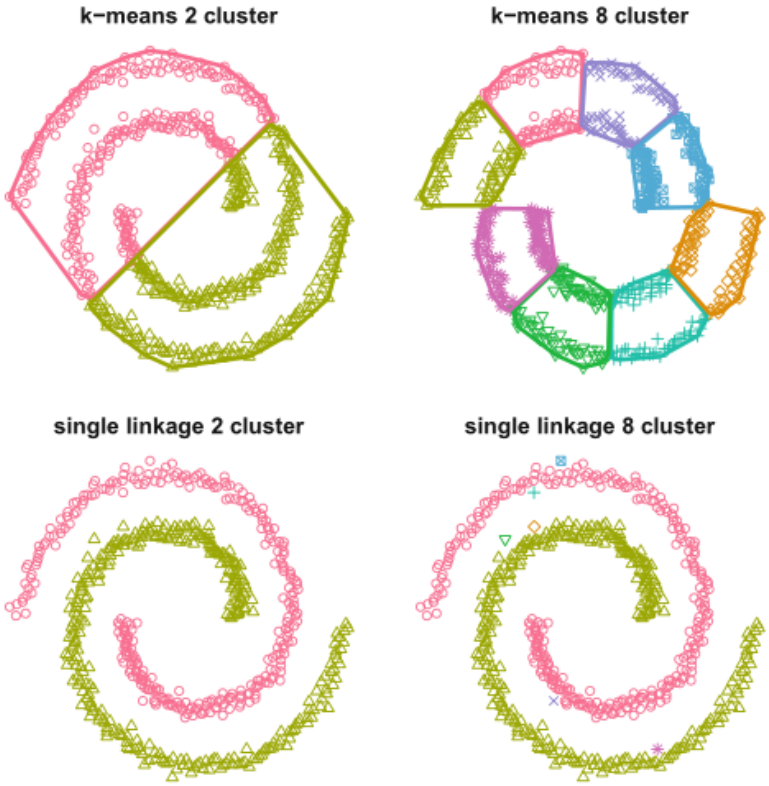


Figure 9: *k-means and single linkage hierarchical clustering of two spirals*

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

Table 5: *Data set and segment characteristics informing extraction algorithm selection*

3.5.2 Distance-Based Methods

Consider the problem of finding groups of tourists with similar activity patterns when on vacation. A fictitious data set is shown in Tab-6. It contains seven people indicating

the percentage of time they spend enjoying BEACH, ACTION, and CULTURE when on vacation. Anna and Bill only want to relax on the beach, Frank likes beach and action, Julia and Maria like beach and culture, Michael wants action and a little bit of culture, and Tom does everything. Market segmentation aims at grouping consumers into groups with similar needs or behaviour, in this example: groups of tourists with similar patterns of vacation activities. Anna and Bill have exactly the same profile, and should be in the same segment. Michael is the only one not interested in going to the beach, which differentiates him from the other tourists. In order to find groups of similar tourists one needs a notion of similarity or dissimilarity, mathematically speaking: a distance measure.

	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

Table 6: *Artificial data set on tourist activities: percentage of time spent on three activities*

1. Distance Measures

   Tab-6 is a typical 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):

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

   The vector corresponding to the i-th row of matrix $\mathbf{X}$ is denoted as $x_i = (x_{i1}, x_{i2}, \ldots, x_{ip})'$ in the following, such that $\mathfrak{X} = \{x_1, x_2, \ldots, x_p\}$ 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 $x_7 = (50, 20, 30)'$. Numerous approaches to measuring the distance between two vectors exist; several are used routinely in cluster analysis and market segmentation. A distance is a function $d(\mathring{\mathrm{u}}, \mathring{\mathrm{u}})$ with two arguments: the two vectors $\mathbf{x}$ and $\mathbf{y}$ between which

the distance is being calculated. The result is the distance between them (a non-negative value). A good way of thinking about distance is in the context of geography. If the distance between two cities is of interest, the location of the cities are the two vectors, and the length of the air route in kilometres is the distance. But even in the context of geographical distance, other measures of natural distance between two cities are equally valid, for example, the distance a car has to drive on roads to get from one city to the other.

3.5.3 Hierarchical Methods

Hierarchical clustering methods are the most intuitive way of grouping data because they mimic how a human would approach the task of dividing a set of n observations (consumers) into k groups (segments). If the aim is to have one large market segment (k = 1), the only possible solution is one big market segment containing all consumers in data X. At the other extreme, if the aim is to have as many market segments as there are consumers in the data set (k = n), the number of market segments has to be n, with each segment containing exactly one consumer. Each consumer represents their own cluster. Market segmentation analysis occurs between those two extremes.

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.

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.

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). They are nested because the partition with k + 1 groups (segments) is obtained from the partition with k groups by splitting one of the groups.

3.5.4 Partitioning Methods

Hierarchical clustering methods are particularly well suited for the analysis of small data sets with up to a few hundred observations. For larger data sets, dendrograms are hard to read, and the matrix of pairwise distances usually does not fit into computer memory. For data sets containing more than 1000 observations (consumers), clustering methods creating a single partition are 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. For a data set including information about 1000 consumers, for example, the agglomerative hierarchical clustering algorithm would have to calculate $(1000 \times 999)/2 = 499{,}500$ distances for the pairwise distance matrix between all consumers in the data set.

A partitioning clustering algorithm aiming to extract five market segments, in contrast, would only have to calculate between 5 and 5000 distances at each step of the iterative or stepwise process (the exact number depends on the algorithm used). 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

1. **k-Means and k-Centroid Clustering**

   The most popular partitioning method is k-means clustering. Within this method, a number of algorithms are available. R function kmeans() implements the algorithms by Forgy, Hartigan and Wong, Lloyd and MacQueen. These algorithms use the squared Euclidean distance. A generalisation to other distance measures, also referred to as k-centroid clustering, is provided in R package flexclust. Let X = {x1, . . . , xn} be a set of observations (consumers) in a data set. Parti- tioning clustering methods divide these consumers into subsets (market segments) such that consumers assigned to the same market segment are as similar to one another as possible, while consumers belonging to different market segments are as dissimilar as possible. The representative of a market segment is referred to in many partitioning clustering algorithms as the centroid. For the k-means algorithm based on the squared Euclidean distance, the centroid consists of the column-wise mean values across all members of the market segment. The data set contains observations (consumers) in rows, and variables (behavioural information or answers to survey questions) in columns. The column-wise mean, therefore, is the average response pattern across all segmentation variables for all members of the segment. The following generic algorithm represents a heuristic for solving the optimi- sation problem of dividing consumers into a given number of segments such that consumers are similar to their fellow segment members, but dissimilar to members of other segments. This algorithm is iterative; it improves the partition in each step, and is bound to converge, but not necessarily to the global optimum.

Figure 10: *Simplified visualisation of the k-means clustering algorithm*

2. **"Improved" k-Means**

   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. Using randomly drawn consumers is suboptimal because it may result in some of those randomly drawn consumers being located very close to one another, and thus not being representative of the data space. Using starting points that are not representative of the data space increases the likelihood of the k-means algorithm getting stuck in what is referred to as a *local optimum*. A local optimum is a good solution, but not the best possible solution. 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. **Hard Competitive Learning**

   Hard competitive learning, also known as learning vector quantisation, differs from the standard k-means algorithm in how segments are extracted. Although hard competitive learning also minimises the sum of distances from each consumer contained in the data set to their closest representative (centroid), the process by which this is achieved is slightly different. 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.

4. **Neural Gas and Topology Representing Networks**

   A variation of hard competitive learning is the neural gas algorithm. Here, not only the segment representative (centroid) is moved towards the randomly selected consumer. Instead, also the location of the second closest segment repre-

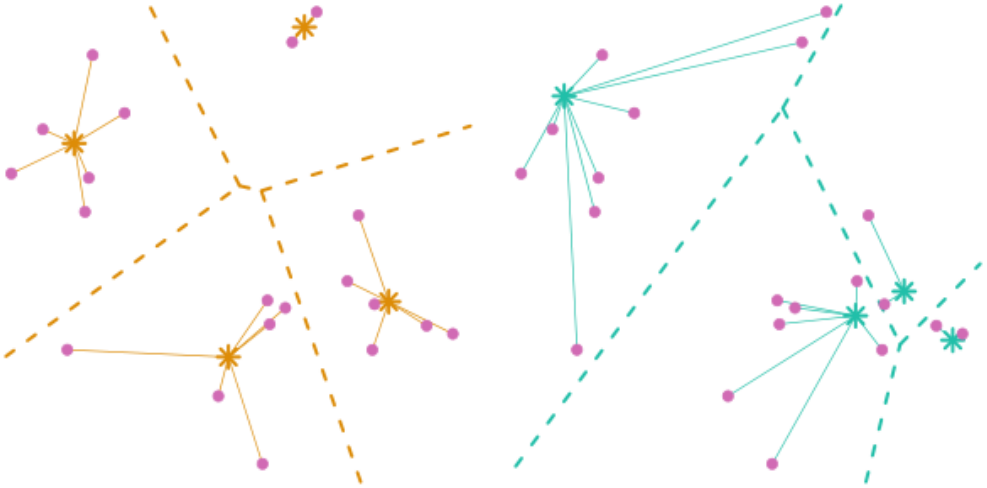


Figure 11: *Examples of good (left) and bad (right) starting points for k-means clustering*

sentative (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. Neural gas clustering can be performed in R using function cclust(x, k, method = "neuralgas") from package flexclust. An application with real data is presented.

5. **Self-Organising Maps**

   Another variation of hard competitive learning are self-organising maps, also referred to as self-organising feature maps or Kohonen maps. Self-organising maps position segment representatives (centroids) on a regular grid, usually a rectangular or hexagonal grid. Examples of grids are provided in Fig-12. The self-organising map algorithm is similar to hard competitive learning: a single random consumer is selected from the data set, and the closest representative for this random consumer moves a small step in their direction. In addition, representatives which are direct grid neighbours of the closest representative move in the direction of the selected random consumer. The 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. What changes over the many repetitions, however, is the extent to which the representatives are allowed to change. The adjustments get smaller and smaller until a final solution is reached. The advantage of self-organising maps over other clustering algorithms is that the numbering of market segments is not random. Rather, the numbering aligns with the grid along which all segment representatives (centroids) are positioned. The price paid for this advantage is that the sum of distances between segment members and segment representatives can be larger than for other clustering algorithms. The reason is that the location of representatives cannot

be chosen freely. Rather, the grid imposes restrictions on permissible locations. Comparisons of self-organising maps and topology representing networks with other clustering algorithms, such as the standard k-means algorithm, as well as for market segmentation applications are provided.

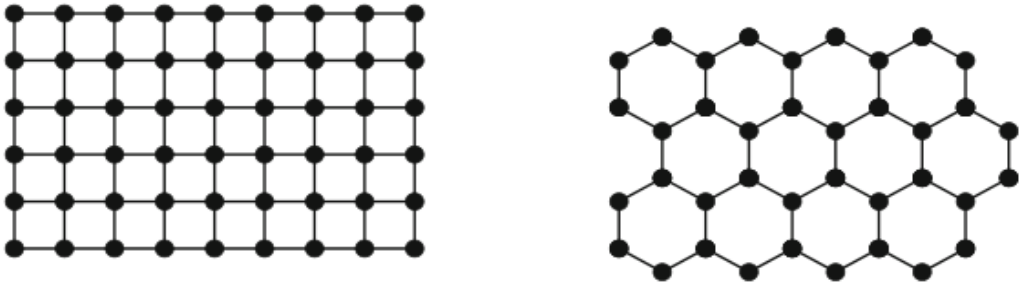


Figure 12: *Rectangular (left) and hexagonal (right) grid for self-organising maps*

3.6 Step-6: Profiling Segments

3.6.1 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. The situation is quite different in the case of data-driven segmentation: users of the segmentation solution may have decided to extract segments on the basis of benefits sought by consumers. Yet – until after the data has been analysed – the defining characteristics of the resulting market segments are unknown. Identifying these defining characteristics of market segments with respect to the segmentation variables is the aim of profiling. Profiling consists of characterising the market segments individually, but also in comparison to the other market segments. If winter tourists in Austria are asked about their vacation activities, most state they are going alpine skiing. Alpine skiing may characterise a segment, but alpine skiing may not differentiate a segment from other market segments.

3.6.2 Traditional Approaches to Profiling Market Segments

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

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

Such tables are hard to interpret, and it is virtually impossible to get a quick overview of the key insights. This is illustrated by Fig-13. Fig-13 shows the mean values of the segmentation variables by segment (extracted from the return object using parameters(vacmot.k6)), together with the overall mean values. Because the travel motives are binary, the segment means are equal to the percentage of segment members engaging in each activity.

Fig-13 provides the exact percentage of members of each segment that indicate that each of the travel motives matters to them. 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.

	Seg. 1	Seg. 2	Seg. 3	Seg. 4	Seg. 5	Seg. 6	Total
Rest and relax	83	96	89	82	98	96	90
Change of surroundings	27	82	73	82	87	77	67
Fun and entertainment	7	71	81	60	95	37	53
Free-and-easy-going	12	65	58	45	87	75	52
Not exceed planned budget	23	100	2	49	84	73	51
Life style of the local people	9	29	30	90	75	80	46
Good company	14	59	40	58	77	55	46
Excitement, a challenge	9	17	39	57	76	36	33
Maintain unspoilt surroundings	9	10	16	7	67	95	30
Cultural offers	4	2	5	96	62	38	28
Luxury / be spoilt	19	24	39	13	89	6	28
Unspoilt nature/natural landscape	10	10	13	15	69	64	26
Intense experience of nature	6	8	9	21	50	58	22
Cosiness/familiar atmosphere	11	24	12	7	49	25	19
Entertainment facilities	5	25	30	14	53	6	19
Not care about prices	8	7	43	19	29	10	18
Everything organised	7	21	15	12	46	9	16
Do sports	8	12	13	10	46	7	14
Health and beauty	5	8	10	8	49	16	12
Realise creativity	2	2	3	8	29	14	8

Figure 13: *Six segments computed with the neural gas algorithm for the Australian travel motives data set. All numbers are percentages of people in the segment or in the total sample agreeing to the motives*

Using Fig-13 as the basis of interpreting segments shows that the defining characteristics of segment 2, for example, are: being motivated by rest and relaxation, and not wanting to exceed the planned travel budget. Also, many members of segment 2 care about a change of surroundings, but not about cultural offers, an intense experience of nature, about not caring about prices, health and beauty and realising creativity. Segment 1 is likely to be a response style segment because – for each travel motive – the percentage of segment members indicating that a travel motive is relevant to them is low (compared to the overall percentage of agreement).

3.7 Step-7: Describing Segments

3.7.1 Using Visualisation to Describe Market Segments

A wide range of charts exist for the visualisation of differences in descriptor variables. Here, we discuss two basic approaches suitable for nominal and ordinal descriptor variables (such as gender, level of education, country of origin), or metric descriptor variables (such as age, number of nights at the tourist destinations, money spent on accommodation). Using graphical statistics to describe market segments has two key advantages: it simplifies the interpretation of results for both the data analyst and the user, and integrates information on the statistical significance of differences, thus avoiding the over-interpretation of insignificant differences. As Cornelius put it: Graphical representations . . . serve to transmit the very essence of marketing research results. The same authors also find – in a survey study with marketing managers – that managers prefer graphical formats, and view the intuitiveness of graphical displays as critically important.

3.7.2 Testing for Segment Differences in Descriptor Variables

Simple statistical tests can be used to formally test for differences in descriptor variables across market segments. The simplest way to test for differences is to run a series of independent tests for each variable of interest. The outcome of the segment extraction step is segment membership, the assignment of each consumer to one market segment. Segment membership can be treated like any other nominal variable. It represents a nominal summary statistic of the segmentation variables. Therefore, any test for association between a nominal variable and another variable is suitable. The association between the nominal segment membership variable and another nominal or ordinal variable (such as gender, level of education, country of origin) is visualised using the cross-tabulation of both variables as basis for the mosaic plot. The appropriate test for independence between columns and rows of a table is the $\mathfrak{X}^2$-test.

3.8 Step-8: Selecting the Target Segments

3.8.1 The Targeting Decision:

Step 8 is where the rubber hits the road. Now the big decision is made: which of the many possible market segments will be selected for targeting? Market segmentation is a strategic marketing tool. The selection of one or more target segments is a long-term decision significantly affecting the future performance of an organisation. This is when the flirting and dating is over; it's time to buy a ring, pop the question, and commit. After a global market segmentation solution has been chosen – typically at the end of Step 5 – a number of segments are available for detailed inspection. These segments are profiled in Step 6, and described in Step 7. In Step 8, one or more of those market segments need to be selected for targeting. The segmentation team can build on the outcome of Step 2. During Step 2, knock-out criteria for market segments have been agreed upon, and segment attractiveness criteria have been selected, and weighed to reflect the relative importance of each of the criteria to the organisation. Optimally, the knock-out criteria have already been applied in previous steps. For example, in Step 6 market segments were profiled by inspecting their key characteristics in terms of the segmentation variables. It would have become obvious in Step 6 if a market segment is not large enough, not homogeneous or not distinct enough. It would have become obvious in Step 7 – in the process of detailed segment description using descriptor variables – if a market segment is not identifiable or reachable. And in both Steps 6 and 7, it would have become clear if a market segment has needs the organisation cannot satisfy. Imagine, for example, that the BIG SPENDING CITY TOURIST emerged as one of the very distinct and attractive segments from a market segmentation analysis, but the destination conducting the analysis is a nature based destination in outback Australia. The chances of this destination meeting the needs of the highly attractive segment of BIG SPENDING CITY TOURIST are rather slim.

Once this is done, the attractiveness of the remaining segments and the relative organisational competitiveness for these segments needs to be evaluated. In other words, the segmentation team has to ask a number of questions which fall into two broad categories:

- Which of the market segments would the organisation most like to target? Which segment would the organisation like to commit to?
- Which of the organisations offering the same product would each of the segments most like to buy from? How likely is it that our organisation would be chosen? How likely is it that each segment would commit to us?

Answering these two questions forms the basis of the target segment decision.

3.9 Step-9: Customizing the Marketing Mix

3.9.1 Implications for Marketing Mix Decisions

Marketing was originally seen as a toolbox to assist in selling products, with marketers mixing the ingredients of the toolbox to achieve the best possible sales results. In the early days of marketing, Borden postulated that marketers have at their disposal 12 ingredients: product planning, packaging, physical handling, distribution channels, pricing, personal selling, branding, display, advertising, promotions, servicing, fact finding and analysis. Many versions of this marketing mix have since been proposed, but most commonly the marketing mix is understood as consisting of the 4Ps: Product, Price, Promotion and Place. Market segmentation does not stand independently as a marketing strategy. Rather, it goes hand in hand with the other areas of strategic marketing, most importantly: positioning and competition. In fact, the segmentation process is frequently seen as part of what is referred to as the *segmentation-targeting-positioning* (STP) approach. 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 organisation can take to ensure that their product is perceived as distinctly different from competing products, and in line with segment needs).

Figure 14: *How the target segment decision affects marketing mix development*

Figure illustrates how the target segment decision – which has to be integrated with

other strategic areas such as competition and positioning – affects the development of the marketing mix. For reasons of simplicity, the traditional 4Ps model of the marketing mix including Product, Price, Place and Promotion serves as the basis of this discussion. Be it twelve or four, each one of those aspects needs to be thoroughly reviewed once the target segment or the target segments have been selected.

3.9.2 Product

One of the key decisions an organisation needs to make when developing the product dimension of the marketing mix, is to specify the product in view of customer needs. Often this does not imply designing an entirely new product, but rather modifying an existing one. Other marketing mix decisions that fall under the product dimension are: naming the product, packaging it, offering or not offering warranties, and after sales support services. The market segments obtained for the Australian vacation activities data set using biclustering present a good opportunity for illustrating how product design or modification is driven by target segment selection. Imagine, for example, being a destination with a very rich cultural heritage. And imagine having chosen to target segment 3. The key characteristics of segment 3 members in terms of vacation activities are that they engage much more than the average tourist in visiting museums, monuments and gardens also like to do scenic walks and visit markets. They share both of these traits with some of the other market segments. Like most other segments, they like to relax, eat out, shop and engage in sightseeing. In terms of the product targeted at this market segment, possible product measures may include developing a new product. For example, a MUSEUMS, MONUMENTS & MUCH, MUCH MORE product (accompanied by an activities pass) that helps members of this segment to locate activities they are interested in, and points to the existence of these offers at the destination during the vacation planning process. Another opportunity for targeting this segment is that of proactively making gardens at the destination an attraction in their own right.

3.10 Step-10: Evaluation and Monitoring

3.10.1 Ongoing Tasks in Market Segmentation

Market segmentation analysis does not end with the selection of the target segment, and the development of a customised marketing mix. As Lilien and Rangaswamy state segmentation must be viewed as an ongoing strategic decision process. Haley elaborates as follows: The world changes . . . virtually the only practical option for an intelligent marketer is to monitor his or her market continuously. After the segmentation strategy is implemented, two additional tasks need to be performed on an ongoing basis:

- The effectiveness of the segmentation strategy needs to be evaluated. Much effort goes into conducting the market segmentation analysis, and customising the marketing mix to best satisfy the target segment's needs. These efforts should result in an increase in profit, or an increase in achievement of the organisational mission. If they did not, the market segmentation strategy failed.
- The market is not static. Consumers change, the environment, and actions of competitors change. As a consequence, a process of ongoing monitoring of the market segmentation strategy must be devised. This monitoring process can range from a regular review by the segmentation team, to a highly automatised data mining system alerting the organisation to any relevant changes to the size or nature of the target segment.

3.10.2 Evaluating Success of the Segmentation Strategy

The aim of evaluating the effectiveness of the market segmentation strategy is to determine whether developing a customised marketing mix for one or more segments did achieve the expected benefits for the organisation. In the short term, the primary desired outcome for most organisations will be increased profit. For non for profit organisations it may be some other performance criterion, such as the amount of donations raised or number of volunteers recruited. These measures can be monitored continuously to allow ongoing assessment of the segmentation strategy. In addition, taking a longer term perspective, the effectiveness of targeted positioning could be measured. For example, a tracking study would provide insight about how the organisation is perceived in the market place. If the segmentation strategy is successful, the organisation should increasingly be perceived as being particularly good at satisfying certain needs. If this is the case, the organisation should derive a competitive advantage from this specialised positioning because the target segment will perceive it as one of their preferred suppliers.

3.10.3 Segment Evolution

Segments evolve. Like any characteristic of markets, market segments change over time. The environments in which the organisation operates, and actions taken by competi-

tors change. Haley, the father of benefit segmentation, says that not following-up a segmentation study means sacrificing a substantial part of the value it is able to generate. Haley proceeds to recommend a tracking system to ensure that any changes are identified as early as possible and acted upon. Haley refers to the tracking system as an early warning system activating action only if an irregularity is detected. Or, as Cahill puts it: keep testing, keep researching, keep measuring. People change, trends change, values change, everything changes. A number of reasons drive genuine change of market segments, including: evolution of consumers in terms of their product savviness or their family life cycle; the availability of new products in the category; and the emergence of disruptive innovations changing a market in its entirety.

A similar approach is used by Oliveira and Gama (2010). In their framework, the following taxonomy is used for changes in segments over time:

- Birth: a new segment emerges.
- Death: an existing segment disappears.
- Split: one segment is split up.
- Merge: segments are merged.
- Survival: a segment remains almost unchanged.

Github : Chaganti Reddy/Market Segmentation