

Market-Segmentation-Analysis

Step-1: Deciding (not) to Segment

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.

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. 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. 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.

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.

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

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. 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.

Step-2: Specifying the Ideal Target Segment

2.1 Segment Evaluation Criteria

The third layer of market segmentation analysis 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 process 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, Cyclicalities 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)
Dibb and Simkin	Homogeneous, Large enough, Profitable, Stable, Accessible, Compatible, Actionable
Sternthal and 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.	Large enough, Sufficient purchasing power, Characteristics of the segment, Reachable, Able to serve segment effectively, Distinct, Targetable with marketing programs
Winer and 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 and 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

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 member's 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 needs to be specified.

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.

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 competitiveness 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.

Step-3: Collecting Data

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

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
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					

Fig: Gender as a possible segmentation variable in commonsense market segmentation

Each row in this table represents one 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 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 organization.

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 : 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 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.

3.2 Segmentation Criteria

The term segmentation criterion is used here in a broader sense than the term segmentation variable. The term segmentation variable refers to one measured value, for example, one item in a survey, or one observed expenditure category. The term segmentation criterion relates to the nature of the information used for market segmentation. It can also relate to one specific construct, such as benefits sought. The 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.

3.2.1 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.

3.2.2 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.

3.2.3 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 psycho-graphic 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 behavior. 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.

3.2.4 Behavioural Segmentation

Another approach to segment extraction is to search directly for similarities in behaviour or reported behavior. A wide range of possible behaviors 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 behavioral 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 behavioral 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.

4.1 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.

4.1.1 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.

4.1.2 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**.

4.1.3 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 overtime, and independently of the survey questions asked, it represents a response style. A wide range of response styles manifest in survey answers, including respondent's 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.

4.1.4 Sample Size

Many statistical analyses are accompanied by sample size recommendations. Not so market segmentation analysis. 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.

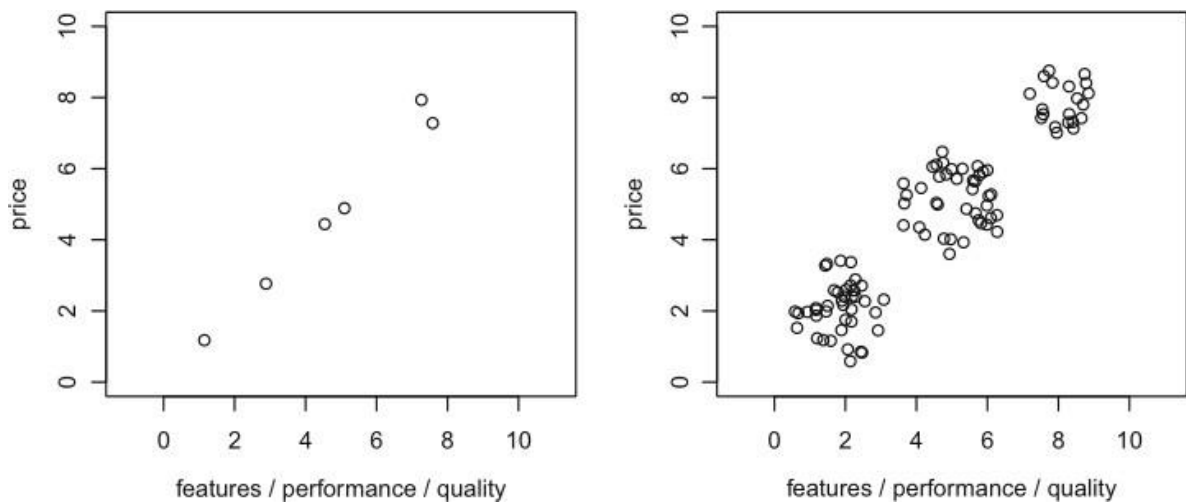


Fig: Illustrating the importance of sufficient sample size in market segmentation analysis

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 $(10 \cdot p)$

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.

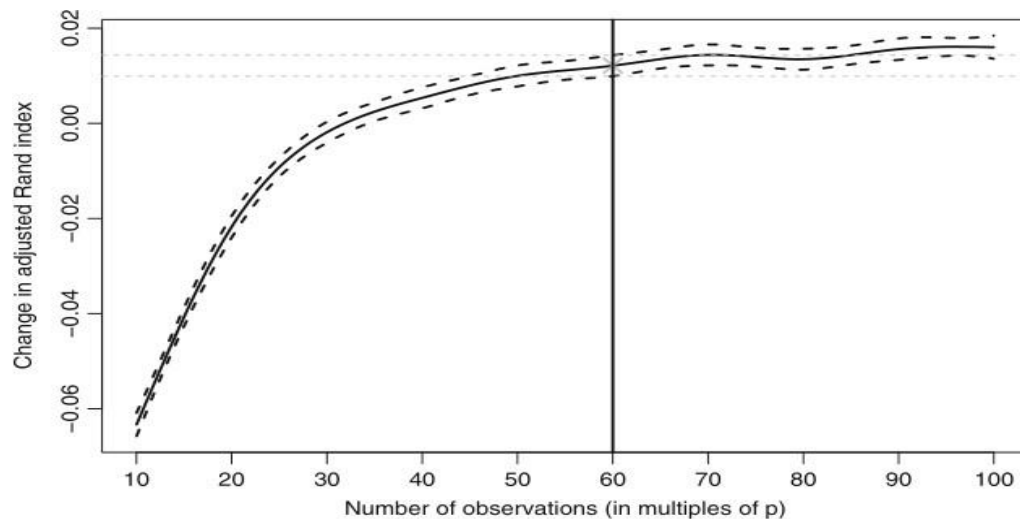


Fig: Effect of sample size on the correctness of segment recovery in artificial data.

Fig 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 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:

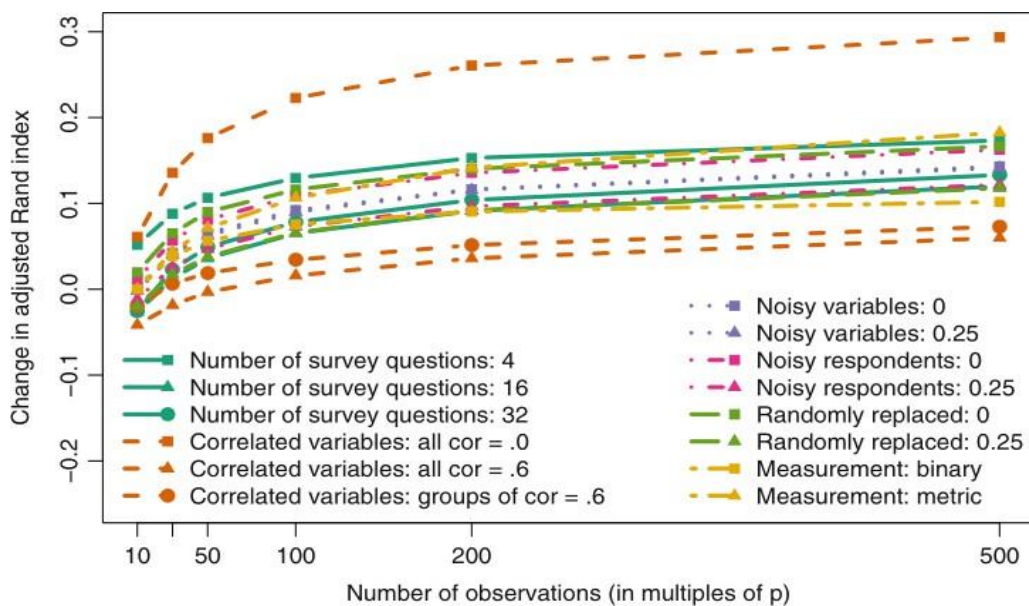


Fig: Sample size requirements in dependence of market and data characteristics.

- 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).

5.1 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.

6.1 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.

Step-7: Describing Segments

7.1 Developing a Complete Picture of Market Segments

Describing segments is similar to the profiling step. The only difference is that the variables being inspected have not been used to extract market segments. Market segments are described using additional information available about segment members. If committing to a target segment is like a marriage, profiling and describing market segments is like going on a number of dates to get to know the potential spouse as well as possible in an attempt to give the marriage the best possible chance, and avoid nasty surprises down the track.

The segment description step uses additional information, such as segment members' age, gender, past travel behaviour, preferred vacation activities, media use, use of information sources during vacation planning, or their expenditure patterns during a vacation. These additional variables are referred to as descriptor variables.

7.2 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.

7.2.1 Nominal and Ordinal Descriptor Variables

When describing differences between market segments in one single nominal or ordinal descriptor variable, the basis for all visualisations and statistical tests is a cross-tabulation of segment membership with the descriptor variable.

The easiest approach to generating a cross-tabulation is to add segment membership as a categorical variable to the data frame of descriptor variables. The following python command gives the number of females and males across market segments:

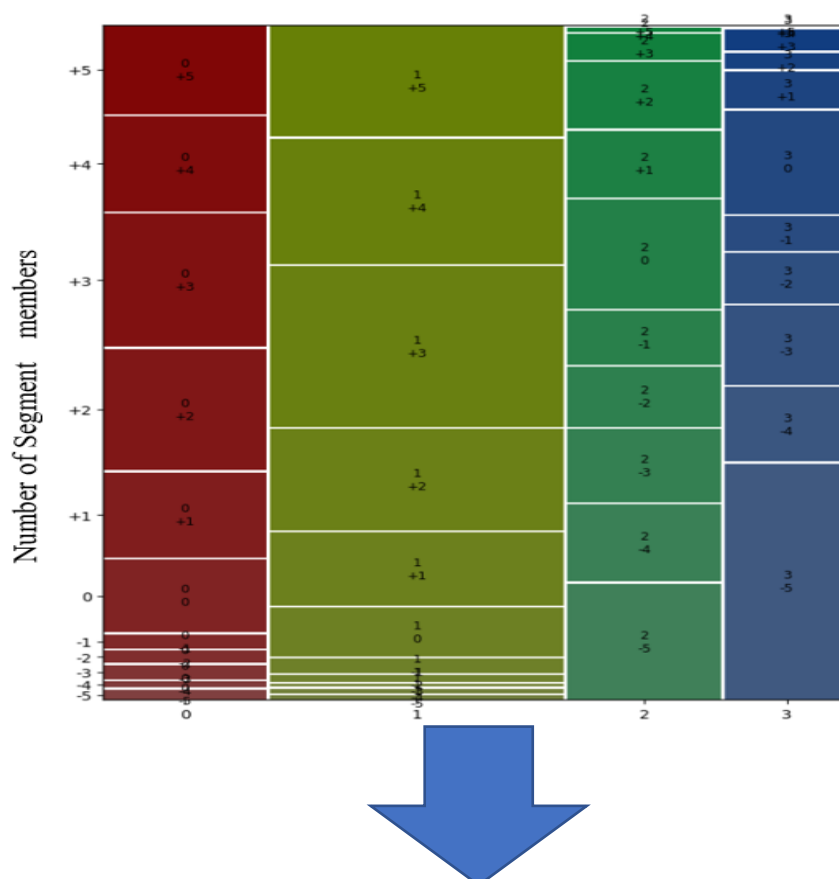
```
crosstab_gender = pd.crosstab(df['Segment_number'],df['Gender'])
```

crosstab_gender

	Gender	Female	Male
Segment_number			
	0	154	169
	1	349	231
	2	179	125
	3	106	140

A visual inspection of this cross-tabulation suggests that there are no huge gender differences across segments. A solution is to draw the bars for women and men next to one another rather than stacking them (not shown). The disadvantage of this approach is that the absolute sizes of the market segments can no longer be directly seen on the y-axis. The mosaic plot offers a solution to this problem.

The mosaic plot also visualises cross-tabulations. The width of the bars indicates the absolute segment size:



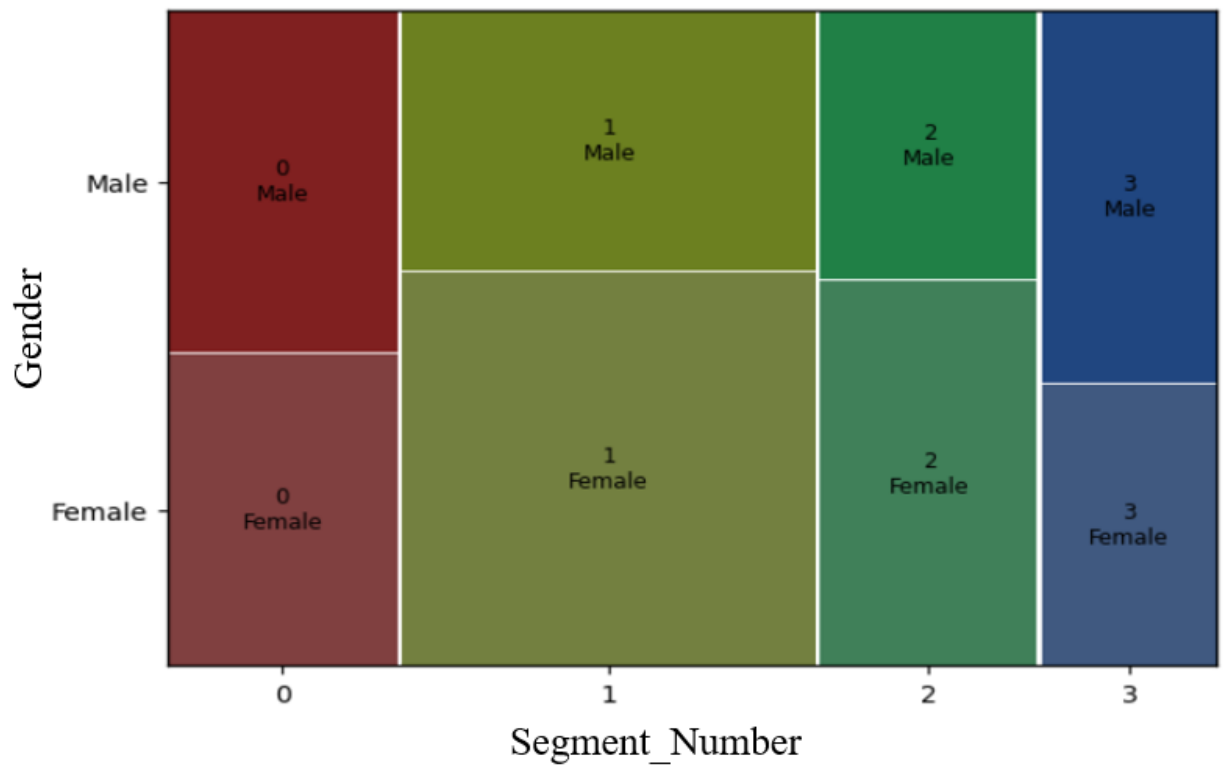


Fig: Comparison of a stacked bar chart and a mosaic plot for the cross-tabulation of segment membership and gender for the mcdonalds data set

We can gain additional insights by using a parallel box-and-whisker plot; it shows the distribution of the variable separately for each segment. We create this parallel box-and-whisker plot for age by market segment in python with the following command:

```
sns.boxplot(x="Segment_number", y="Age", data=df)
```

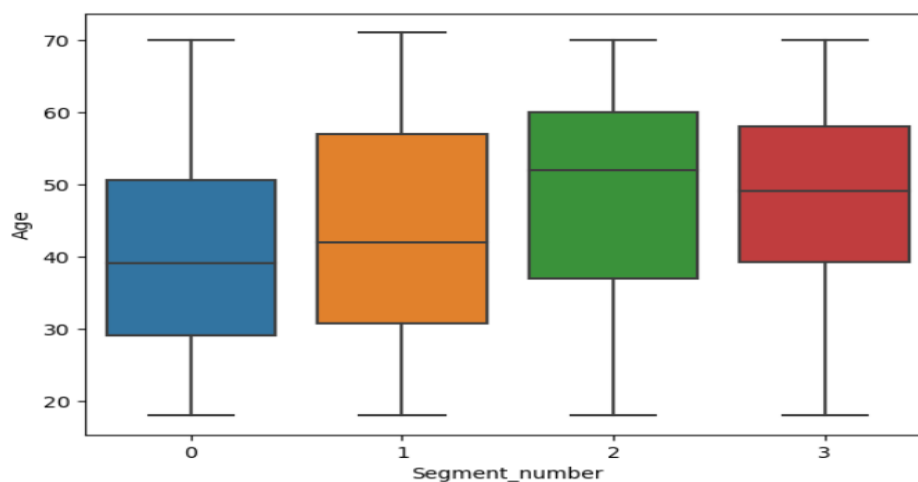


Fig: Parallel box-and-whisker plot of age by segment for the mcdonalds data set

Name: J Narasimha

References

Beane TP, Ennis DM (1987) Market segmentation: a review. *Eur J Mark* 21(5):20–42

McDonald M, Dunbar I (1995) Market segmentation: a step-by-step approach to creating profitable market segments. Macmillan, London

Croft M (1994) Market segmentation. Routledge, London

Day GS (1984) Strategic market planning. West Publishing Company, Minnesota
Dibb S, Simkin L (2008) Market segmentation success: making it happen! Routledge, New York

Jain SC (2012) Marketing: planning and strategy. Cengage Learning Australia, South Melbourne

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B* 57:289–300

Kotler P, Keller KL (2012) Marketing management. Pearson Education, Paris

Tynan AC, Drayton J (1987) Market segmentation. *J Mark Manag* 2(3):301–335

Breiman L, Friedman J, Stone CJ, Olshen RA (1984) Classification and regression trees. Chapman and Hall/CRC, New York

Winer R, Dhar R (2011) Marketing management. Pearson, London

Cornelius B, Wagner U, Natter M (2010) Managerial applicability of graphical formats to support positioning decisions. *Journal für Betriebswirtschaft* 60(3):167–201

Dolnicar S, Leisch F (2008) An investigation of tourists' patterns of obligation to protect the environment. *J Travel Res* 46:381–391

Fox J (2003) Effect displays in R for generalised linear models. *J Stat Softw* 8(15):1–27