Project 2 Study Report (Steps 1-4)

Padala Akhil Reddy

Company: Feynn Labs

Batch-8-23-S-B

17-03-2023

Step 1: Deciding(not) to segment.

- Before investing time and resources in a market segmentation strategy, we should understand the implications of pursuing a market segmentation.
- The company should commit to the market segmentation strategy on a long term.
- Potentially required changes include the development of new products, modification of existing products, changes in pricing and distribution channels used to sell the product, as well as all communications with the market.
- Because of the major implications of such a long-term 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.

Implementation Barriers:

There can be various barriers which can impede the successful roll-out of a market segmentation strategy.

- From senior management:
    - Lack of leadership, pro-active championing, commitment and involvement in the market segmentation process by senior leadership.
    - Not making enough resources available, either for initial market segmentation analysis, or for the long-term implementation of the segmentation strategy.
- Organisational culture:

    - Lack of market or consumer orientation.
    - Inertia to adopt new ideas. Lack of creative thinking.
    - Poor communication skills, lack of training.
    - Lack of a qualified marketing experts, data analysts and managers in the organisation.
- Process-related barriers:
    - Lack of good planning.
    - Lack of structured processes to guide the team through all the steps of the segmentation process.
    - Lack of allocation of responsibilities and lack of time.
- Other barriers may include lack of resources or the inability to make required organisational changes.
- 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 market segment.

- The organisation must determine two sets of segment evaluation criteria:
    - Knock-out criteria: Essential, non-negotiable features which the organisation would consider targeting. These are used to determine if market segments resulting from the market segmentation analysis qualify to be assessed using segment attractiveness criteria. Some knock-out criteria are the discussed as follows:
        - The segment must be homogeneous; members of the segment must be similar to one another.
        - 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 customised marketing mix accessible to them.
    - Attractiveness criteria: These criteria are used to evaluate the relative attractiveness of the remaining market segments, those in compliance with the knock-out criteria.
    - Attractiveness criteria are not binary in nature. Segments are generally rated as more or less attractive with respect to a specific criterion.
    - The attractiveness across all criteria determines whether a market segment is selected as a target segment.

Implementing a Structured Process:

- Following a structured process when assessing market segments is beneficial.
- The most popular structured approach to evaluate market segments for selecting them as target markets is to use a segment evaluation plot, showing segment attractiveness along one axis, and organisational competitiveness on the other axis. Both attractivess and competitiveness are evaluated by the segmentation team.
- A core team consisting of representatives of all organisational units must investigate the possible criteria to reach an agreement on which criteria are the most important. Representatives of all organisational units must be included for considering different perspectives on the business of the organisation, and also because all the units in the business are the stakeholders of market segmentation analysis.
- Although there no segments to assess yet on this stage, it is still beneficial to select the attractiveness criteria for market segments at this stage in the process, because knowing what about the market segments matters to the organisation will ensure that this information is captured while collecting data. It also makes the task of selecting a target segment in step 8 much easier because the groundwork is laid before the actual segments are extracted.

- 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.
- Optimally, approval by the advisory committee should
  be sought because the advisory committee contains representatives from multiple
  organisational units bringing a range of different perspectives to the challenge of specifying segment attractiveness criteria.

Step 3: Collecting data.

Segmentation and Descriptor Variables:

- Segmentation variables are used for identifying naturally existing, or for artificially creating market segments which are useful to the organisation.
- Descriptor variables are used to describe the segments in detail. These variables are also important because describing the segments is critical to being able to develop an effective marketing mix targeting the segment.

The difference between commonsense and data-driven market segmentation is that the former involves extracting segments based on only one segmentation variable while the latter involves multiple segmentation variables.

Quality of empirical data is critical to extracting market segments of good quality. Good market segmentation analysis required good data. Correct values in the segmentation variables are necessary to assign each person in the dataset to the right market segment. Correct description makes it possible to develop a customised product, determine appropriate pricing strategy, and select the best distribution channels.

Segmentation criteria:

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.

Generally, the recommendation is to use the simplest possible criteria which works for the business and involves the least cost.

- Geographic Segmentation: The consumer’s location of residence serves as the only criterion to form market segments.
    - Advantages: 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 to reach the selected geographic segments.
    - Disadvantages: 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.
- 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, cosmetics, tourism resort products.
    - Advantages: 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.
    - Disadvantages: In many instances, the socio-demographic criterion is not the cause for product preferences, thus not providing sufficient market insight for optimal segmentation decisions.
- Psychographic Segmentation: Under this segmentation criterion, people are grouped according to psychological criteria, such as their beliefs, interests, preferences, aspirations, or benefits sought when purchasing a product.
    - Advantages: It is generally more reflective of the underlying reasons for differences in consumer behaviour.
    - Disadvantages: Complexity is high of determining segment memberships for consumers. Also, the power of the psychographic approach depends heavily on the reliability and validity of the empirical measures used to capture the psychographic dimensions of interest.
- Behavioural Segmentation: Look for similarities in the behaviour or reported behaviour of consumers. 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.

- Advantages: If based on actual behaviour rather than stated behaviour or stated intended behaviour – the very behaviour of interest is used as the basis of segment extraction. As such, behavioural segmentation groups people by the similarity which matters most.
- Disadvantages: 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.

Data can be collected from different sources, each source of data has its own pros and cons.

Data from Survey Studies:

Most market segmentation analyses are based on data collected from surveys. Survey data is cheap and easy to collect but it is prone to biases which will negatively affect the quality of solutions obtained from the segmentation analyses.

Some importatant aspects to be considered when using survey data are discussed below:

- Choice of variables:
  - All variables relevant to the construct captured by the segmentation criterion need to be included.
  - Unnecessary variables must be avoided. Collecting too many variables will make the survey questionnaires tedious and lead to response fatigue, which leads to lower quality responses. Too many variables will also lead to higher dimensionality of the data, which makes it difficult to extract market segments. These redundant variables are called Noisy variables and can result from not carefully developing the questionnaire.
- 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 the respondent to choose between two options correspond to binary variables.
  - Options allowing the respondent to choose between multiple options correspond to nominal variables.

    - Options allowing respondents to indicate a number, such as age or nights stayed at a hotel, generate metric data. Metric data allow any statistical procedure to be performed (including the measurement of distance), and are therefore well suited for segmentation analysis. But the distance between adjacent answer options is not clearly defined. As a consequence, it is not possible to apply standard distance measures to such data, unless strong assumptions are made.
    - Either metric or binary response options should be provided to respondents if those options are meaningful with respect to the question asked. Using binary or metric response options prevents subsequent complications relating to the distance measure in the process of data-driven segmentation analysis.
    - Although ordinal scales dominate both market research and academic survey research, using binary or metric response options instead is usually not a compromise. If, for example, there is a strong reason to believe that very fine nuances of responses need to be captured, and if capturing those fine nuances does not come at the cost of also capturing response styles, this can be achieved using visual analogue scales.
- Response Styles:
    - Survey data is prone to capturing biases. If a bias is displayed by a respondent consistently over time, and independently of the survey questions asked, it represents a response style.
    - 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.
    - It is critical, therefore, to minimise the risk of capturing response styles when data is collected for the purpose of market segmentation. In cases where attractive market segments emerge with response patterns potentially caused by a response style, additional analyses are required to exclude this possibility.
- Sample Size:
    - If the sample size is sufficient, it is very easy to determine the number and nature of segments in the data set.

    - Increasing the sample size improves the correctness of the extracted segments. Interestingly, however, the biggest improvement is achieved by increasing very small samples. As the sample size increases, the marginal benefit of further increasing the sample size decreases.
    - Larger sample sizes always improve an algorithm's ability to identify the correct market segmentation solution. The extent to which this is the case, however, varies substantially across market and data characteristics.
- It can be concluded that the data used in market segmentation analyses should:
    - Contain all necessary variables and not any redundant or unnecessary variables.
    - 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.
    - Include a sufficient sample size given the number of segmentation variables.

Data from Internal Sources:

- Advantages of using data from internal sources:
    - Data is automatically generated. No extra effort is required to collect it.
    - They represent the actual behaviour of the data.
- Disadvantage:
    - May not have information on potential customers. Contains data mostly about existing customers.

Data from Experimental Studies:

- Obtained from laboratory or field experiments.
- Can also be collected from choice experiments or conjoint analyses.

Step 4: Exploring the data.

After data collection, the next step is to explore the data, understand the data and resolve quality issues if any. Data exploration helps us to identify the measurement

levels of the variables, investigate the univariate distributions of each variable, and assess dependency between variables.

- First load the data loaded and view the basic statistical information about the data, like number of data points, number of columns.
- Check if the values are recorded correctly.
- Check if consistent labels are used.
- Find any implausible values.
- For categorical values, permissible labels must be there. Ex: For gender, unless a third option is given, only male and female labels must be there.
- Look for missing values in the data.

Descriptive numeric and graphic representations provide insights into the data and makes us familiar to the data.

- For numeric data:
    - Histograms, boxplots and scatter plots.
    - Histograms give insight on the distribution of the numeric variables, whether they are unimodal, symmetric or skewed.
    - Boxplots compresses a dataset into a minimum, first quartile, median, third quartile, and maximum. This gives insight into several distributional properties of the dataset assuming unimodality.
- For categorical data:
    - Bar plots of frequency counts and mosaic plots.
- Pre-Processing:
    - Categorical Variables:
        - Merging different categorical variables to have fewer and more relevant categorical variables to work with.
        - Converting categorical variables to numeric variables. Ordinal data can be converted to numeric data if it can be assumed that the distances between the adjacent scale points on the ordinal scale are approximately equal.
        - It is important to consider the uncertainty in distances between the options in the survey response before working with the data. Unless

there is a strong reason for using multi-categorical variables with uncertain distances between the categories, using binary options is preferable.

        - It must be noted that pre-processing will alter the data in some or the other way. But most statistical procedures will work correctly for the converted data if there are only 2 categories.

    - Numeric Variables:

        - Standardise the variables in such a way that all the segmentation variables are in the same range. This should be done because the range of values of a segmentation variable affects its relative influence in distance-based methods in segment extraction.

        - Default standardisation in statistics:

            - $z_i = \frac{x_i - \bar{x}}{s}$
                - Where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
                - $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - x)^2$
                - for the n observations of a variable $x = \{x_1, ..., x_n\}$

- Principal Component Analysis:

    - PCA transforms the given dataset into a new dataset with variables called principal components - which are uncorrelated and ordered by importance.

    - When the ranges of values are different for different variables, PCA involves computation of the correlation matrix of the numeric variables.

    - PCA is used to project high dimensional data into lower dimensions for plotting purposes.

    - If the first few principal components do not explain much of the variance, then it can mean that all the variables are required to derive the insights from the data. In such a case, it is not easy to project the data into lower dimensions because we cannot remove variables, as all variables are important and give valuable information about the data.

    - Sometimes PCA is used for reducing the number of segmentation variables before extracting the market segments from consumer data. This idea is sometimes chosen because this makes it easier for the segment extraction algorithm to extract the segments.

- The shortcomings of this method are that it replaces the original variables with principal components. When we use only some principal components for market segmentation, we are actually looking at the data from a different angle, and use a different space as a basis for extracting the market segments.
- Still, it is useful to use PCA to explore the data and identify the correlated variables. Strongly correlated variables indicate redundancy, so extra variables must be removed to lower the dimensionality of the data.