

Summary to step 6: Profiling Segments

Profiling is required when data-driven market segmentation is used. Not necessary for common-sense segmentation. Profiling consists of characterizing the market segments individually, but also in comparison to the other market segments. Correct profiling makes good strategic marketing decisions.

For profiling traditional approaches include high-level summaries and large tables for each segment. Such tables were hard to interpret. As a very high number comparison across the table elements needed to do it.

Segment profiling with visualizations:

Getting insights into segment characteristics even in the case of very big data, plots and graphs makes it much easier to comprehend. They also make it easier to assess the usefulness of a market segmentation solution. The process of segmenting data always leads to a large number of alternative solutions. Selecting one of the possible solutions is a critical decision. Visualizations of solutions assist the data analyst and user with this task.

This step can be divided into these 2 sub-steps:

1. Identifying defining characteristics of market segmentation: A good way to understand the defining characteristics of each segment is to produce a *segment profile plot*. The segment profile plot shows how each segment differs from the others. The segment profile plot is a so-called *panel plot*. For each segment, the segment profile plot shows the cluster centers (centroids, representatives of the segments). Heat maps and hierarchical clustering can be used to represent these segments more elaborative.
2. Assessing Segment Separation: The segment separation plot depicts – for all relevant dimensions of the data space – the overlap of segments. Segment separation plots are very simple if the number of segmentation variables is low, but become complex as the number of segmentation variables increases. For this cluster plot, heat maps and pcas are used extensively.

Summary to Step 7: Describing 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. Rather, in this Step 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 behavior, preferred vacation activities, media use, use of information sources during vacation planning, or expenditure patterns during a vacation. These additional variables are referred to as *descriptor variables*.

Visualizations to Describe Market Segments:

1. Nominal and Ordinal Descriptor Variables:

When describing differences between market segments in one single nominal or ordinal descriptor variable, the basis for all visualizations and statistical tests is a cross-tabulation of segment membership with the descriptor variable. For comparisons, the mosaic plot offers a better suitable approach than using cross-tabulation tables. Easier to get proper proportions of segment variables in each of the segments.

2. Metric Descriptor Variables

Testing for segment Differences in descriptor Variable:

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 appropriate test for independence between columns and rows of a table is the χ^2 -test. Using the p-value approach in hypothesis null hypothesis can be rejected as they are the same. The association between segment membership and metric variables is visualized using parallel boxplots. ANOVA can also be used to differentiate two groups.

Predicting Segments from Descriptor Variables:

Another way of learning about market segments is to try to predict segment membership from descriptor variables. To achieve this, we use a regression model with the segment membership as the categorical dependent variable, and descriptor variables as independent variables. Generalized linear models are also used for accommodating large distributions.

Linear models that can be used

1. Binary Logistic Regression
2. Multinomial Logistic Regression
3. Tree based methods