

Market Segment Analysis

Market Segment Analysis

- Market segment analysis purpose is to match customers needs and desires.
- Marketing planning is a sequence of strategic marketing and tactical marketing.
- Strategic Marketing focuses on organisation long-term goals whereas tactical marketing focuses on short-term goals of long-term goals marketing.
- Strategic marketing states where the organisation wants to go and why? Whereas tactical marketing contains instructions on what to be done to get there.
- SWOT analysis outlines what the supplier is particularly suitable to offer customers.
- Market research investigates consumer needs and desires.

Step 7 Extracting Segments

Grouping Segments

- Consumer data is not structured and grouped.
- Many segmentation methods used to extract market segments are taken from the field of cluster analysis.
- It is important to select correct extraction methods for the dataset.

Distance-based methods

- Uses a particular notion of similarity or distance between observations and tries to find a group of similarity observations(market segmentation).
- In order to find groups we need a distance measure.
- Distance measure is a data matrix where each row represents an observation and every column represents a variable.
- It is represented by the $n \times p$ matrix.
- Euclidean distance is the most common distance used in the market segmentation.
- The asymmetric binary distance does not use all dimensions of the vector.
- It only uses dimensions where at least one of the two vectors has a value of 1. It is asymmetric because it treats 0's and 1's differently.

Hierarchical methods

- The most intuitive way of grouping data.
- Approach the task of dividing a set of n observations into k groups.
- Divisive hierarchical clustering splits the data into 2 market segments at the initial step, then each of the segments splits into two segments. This process continues until each consumer has their own market segment.
- Agglomerative hierarchical clustering is the opposite of divisive clustering where a group of segments into one single market segment.
- Every time a hierarchical clustering algorithm is applied to the same data set the nested partitions are obtained.
- Underlying is one of the methods, it is the measure of distance between groups of observations.
- Measure is determined by,
 - A distance measured $d(x,y)$ between groups of observations.
 - Linkage method
 - Linkage method generalises how given distance between pairs of observations, distance between groups of observations are obtained.

Partitioning methods

- Suitable for a dataset containing more than 1000 observations.
- K-means clustering
 - Most popular partitioning method.
 - This algorithm uses squared euclidean distance.
- Partitioning method divides the set of observations assigned of the same market segmentation as similar as possible to one another and while consumers belonging to different market segments are as different as possible to one another. This is known as centroid.

Hard Competitive learning

- Hard competitive learning is also known as learning vector quantisation. It differs from the k-means algorithm in how many segments are extracted.
- Minimises the sum of distance from each consumer contained in the data set to their centroid.

- Choses one random variable and moves this to the closest segment representative, a small step into the direction of the randomly chosen consumer.
- **Neural gas and Topology representing networks**
 - A variation of hard competitive learning
 - Instead of the centroid the location of the second closest segment(centroid) is adjusted towards the randomly selected consumer.
 - A further extension of neural gas clustering is topology representing networks.
- **Self-organising maps**
 - Another variation of hard competitive learning.
 - Referred as self-organising feature maps or Kohonen maps.
 - Self-organising maps position segment representatives on a regular grid, usually a rectangular or hexagonal grid.

Two-step clustering

- IBM SPSS implemented a procedure referred to as two-step clustering (SPSS 2001). The two steps consist of running a partitioning procedure followed by a hierarchical procedure.
- The procedure has been used in a wide variety of application areas, including internet access types of mobile phone users.

Bagged Clustering

- Bagged clustering combines hierarchical clustering algorithms and partitioning clustering algorithms.
- Adds bootstrapping.
- Bootstrapping can be implemented by random drawing from the data set with replacement.
- The process of extracting segments is repeated many times with randomly drawn samples of the data.
- Bootstrapping has the advantage of making the final segmentation solution less dependent on the exact people contained in consumer data.

Model-based methods

- Can predict impact on academics and practitioners, next to conjoint analysis.

- Mixture models will prove to be the most influential methodological development spawned by marketing problems to date.
- Clustering methods, model-based segment extraction methods do not use similarities or distances to assess which consumers should be assigned to the same market segment.
- Instead, they are based on the assumption that the true market segmentation solution.
- Two general properties:
 - Each market segment has a certain size.
 - if a consumer belongs to market segment A, that consumer will have characteristics which are specific to members of market segment A.

Finite mixtures of distribution

- The simplest case of model-based clustering has no independent variables x , and simply fits a distribution to y .
- **Normal distribution:**
 - Most popular finite model distribution.
 - Mixture of several multivariate normal distributors.
 - For example, physical measurements on humans like height, arm length, leg length or foot length are almost perfectly modelled by a multivariate normal distribution.
- **Binary distribution:**
 - Finite mixtures of binary distributions and also referred to as latent class models or latent class analysis.
 - The mixture model assumes that respondents in different segments have different probabilities of undertaking certain activities.

Finite Mixtures of Regressions

- Finite mixtures of regressions are useful for segmenting data based on dependent and independent variables.
- A common example involves identifying travel motives by analyzing behavioral data from specific populations, like Australian tourists.
- Techniques like EM (Expectation-Maximization) algorithms are used for parameter estimation in these models.

Extensions and Variations

- Variations of finite mixture models adapt to different types of data, such as binary or multinomial distributions.
- Extensions enhance model flexibility and accommodate dependencies among variables.

Algorithms with Integrated Variable Selection

- This section explores advanced algorithms that incorporate variable selection into the clustering process.
- **Biclustering Algorithms**
 - Biclustering identifies coherent clusters simultaneously in rows and columns of a data matrix.
 - Examples include Australian vacation activities, illustrating biclustering effectiveness in real-world scenarios.

Variable Selection Procedure for Clustering Binary Data (VSBD)

- VSBD focuses on selecting relevant binary variables for segmentation.
- It helps streamline the clustering process by reducing data complexity.

Variable Reduction: Factor-Cluster Analysis

- Combines factor analysis and clustering to simplify data while retaining meaningful patterns.
- Useful for creating interpretable segments from complex datasets.

Data Structure Analysis

- Examines methods for evaluating the validity and stability of segmentation results.
- **Cluster Indices**
 - Internal and external indices measure the quality of clusters.
 - Internal indices focus on compactness and separation, while external indices use known labels for validation.

- **Gorge Plots**
 - Visual tools that provide insights into the clustering structure, helping identify the optimal number of clusters.
- **Global Stability Analysis**
 - Evaluates the robustness of clustering solutions across the entire dataset.
- **Segment Level Stability Analysis**
 - Focuses on the stability of individual segments within or across different clustering solutions.

Profiling Segments

- **-Identifying Key Characteristics:**
 - Profiling is critical for interpreting and characterizing segments, particularly for data-driven segmentation.
 - **Challenges:**
 - Managers often find data-driven segmentation challenging to interpret, citing issues like lengthy and contradictory reports or insufficient summaries.
- **Traditional Profiling Approach:** These involve tabular presentations of data that summarize segment characteristics, often leading to challenges in quick interpretation.
- **Visualization in Profiling:** Visualization techniques enhance interpretability by providing intuitive insights into segment characteristics.

Describing Segments

- **Developing a Comprehensive View:** This involves using additional descriptor variables not used in the segmentation process to describe segments.
- **Importance of Descriptor Variables:** These include demographic, psychographic, and behavioral details that provide deeper insights.
- **Visualization Techniques:** Tools like mosaic plots and boxplots are used to visualize and understand differences between segments.
- **Testing and Predicting:** Statistical methods like logistic regression and tree-based methods help test differences and predict segment memberships.

➤ **Checklist for Profiling and Describing Segments:**

- Ensure data reliability and accuracy.
- Use effective visual tools for clear communication.
- Validate segment definitions using robust statistical methods.

Github link

https://github.com/123uditha/Mcdonals_MSA/blob/main/MCdonals%202.ipynb