MARKET SEGMENTATION ANALYSIS

**TEAM Sarfaraz:**

**Mohd Sarfaraz Faiyaz (Leader):** <https://github.com/bblackheart013/fyenn_labs_market_segmentation_case_study>

**Daksh Agiwal:**

<https://github.com/daksh-025/feynnlabsinternship/blob/main/Feynn%20labs%20report%20-task%201.docx>

**Pratyush Pao:**

<https://github.com/pickleprat/Feynn-Labs>

**Kaustubh Kishor Welde:**

<https://github.com/kaustubhhxxiii/Market-Segmentation>

**Manish C:**

<https://github.com/Manishc1996/Feynn-Labs>

## **Abstract**

Market division assumes a basic part in both the key and strategic showcasing arranging process. In the brilliant course of action, market division assists with recognizing and focus on track portions that the association needs to zero in on. By grasping the necessities, inclinations, and ways of behaving of various client gatherings, an association can foster an essential bearing that is customized to the requirements of those gatherings. For instance, to venture into another market, market division can assist with recognizing which client fragments in that market are generally appealing and worth seeking after.

In the strategic advertising plan, the market division assists with directing the showcasing moves that the association will make to arrive at its objective clients. By understanding the remarkable necessities and inclinations of various sections, an association can foster designated advertising messages, advancements, and item contributions that are explicitly intended to resound with each fragment. For instance, assuming an association is focusing on two distinct sections with various necessities, it could foster two different promoting messages that address the advantages that each portion is searching for.

In general, market division is a basic apparatus for associations to grasp their clients and foster successful showcasing techniques and plans that are customized to the extraordinary requirements and inclinations of various client gatherings even more likely. By understanding the subtleties of each portion, associations can even more actually match their proposals to the veritable requirements and wants of buyers, eventually driving more prominent consumer loyalty and business achievement.

## **Step 1: Deciding (not) to Segment**

## *Implications of Committing to Market Segmentation*

Market division has been created to be a key promoting procedure applied in numerous associations; it isn't generally the best choice to seek after such a methodology. Prior to putting time and assets in a market division examination, understanding the ramifications of seeking after a market division strategy is significant.

The key ramification is that the association needs to focus on the division methodology on the long haul. Market division is a marriage, not a date. The obligation to advertise division remains forever inseparable with the eagerness and capacity of the association to roll out significant improvements (McDonald and Dunbar 1995) and ventures. As Cahill (2006) puts it: It isn't allowed to Fragment a market. There are expenses of carrying out the analysis, handling reviews, and center gatherings, planning numerous bundles, and planning different promotions and correspondence messages. Cahill prescribes not to portion except if the normal expansion in deals is adequate to legitimize executing a division methodology, expressing that one of the clichés of division procedure is that utilizing the plan must be more productive than promoting without it, net of the cost of creating and utilizing the actual plan.

In view of the significant ramifications of such a drawn-out hierarchical responsibility, the choice to examine the capability of a market division methodology should be made at the most noteworthy chief level and should be efficiently and consistently imparted and built up at every single authoritative level and across every hierarchical unit.

*Implementation Barriers*

The main gathering of hindrances connects with senior administration. Absence of administration, favorable to dynamic supporting, responsibility, and contribution in the market division process by senior authority subverts the progress of market division. As McDonald and Dunbar state: There can be no question that except if the CEO sees the requirement for a division survey, comprehends the cycle and shows a functioning revenue in it, it is practically outside the realm of possibilities for a senior promoting chief to carry out the determinations in a significant manner.

A second gathering of boundaries connects with hierarchical culture. Absence of market or customer direction, protection from change and groundbreaking thoughts, absence of imaginative reasoning, terrible correspondence and absence of sharing of data and bits of knowledge across hierarchical units, transient reasoning, reluctance to cause changes and workplace issues to have been recognized as forestalling the fruitful execution of market division .Croft (1994) fostered a short poll to survey the degree to which an absence of market direction in the authoritative culture might address an obstruction to the effective execution of market division.

## **Step 2: Specifying the Ideal Target Segment**

## *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 organization – 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. Rather, the user needs to be involved in most stages, literally wrapping around the technical aspects of market segmentation analysis.

*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 was suggested by Kotler (1994) and includes substantiality, measurability, and accessibility (Tynan and Drayton 1987). Kotler himself and several other authors have since recommended additional criteria that fall into the knock-out criterion category (Wedel and Kamakura 2000; Lilien and Rangaswamy 2003; McDonald and Dunbar 2012):

* + - The segment must be homogeneous; members of the segment must be like one another.
    - The segment must be distinct; members of the 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 customizing the marketing mix for them.
    - The segment must match the strengths of the organization; the organization 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 needs to be specified.

*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 attractive with respect to a specific criterion. The attractiveness across all criteria determines whether a market segment is selected as a target segment in Step 8 of market segmentation analysis.

## **Step 3: Collecting Data**

Segmentation is the process of dividing a larger market into smaller groups of consumers or organizations with similar needs, preferences, and behaviors. The purpose of segmentation is to better understand and meet the needs of specific customer groups, which can lead to more effective marketing strategies and better business outcomes.

Segmentation typically involves analyzing customer data, such as demographics, psychographics, and behavioral information, to identify patterns and group customers with similar characteristics into distinct segments. For example, a clothing retailer might segment its market based on age, gender, and style preferences, and develop targeted marketing campaigns and product offerings for each segment.

There are many ways to segment a market, and the specific approach used will depend on the business and its goals. Some common segmentation approaches include demographic segmentation (based on characteristics such as age, gender, and income), psychographic segmentation (based on attitudes, values, and lifestyle factors), behavioral segmentation (based on purchase history and other behaviors), and geographic segmentation (based on location and other geographic factors).

Overall, segmentation is a key component of effective marketing strategy, as it allows organizations to better understand and meet the needs of their customers, ultimately leading to more satisfied customers and improved business performance.

*Segmentation Variables*

Since your marketing campaigns aren’t going to be relevant to everyone on your contact list, it’s a good idea to build audience segments based on needs, shared beliefs, online behavior.

Marketing variables help you split an audience into segments by providing you with possible categories to group your contacts into. The 4 main types of market segmentation include demographic, geographic, psychographic, and behavioral–which we’ll cover more in depth in the next section.

An example of a marketing segmentation variable would be age. So, you may use popular slang and colloquial language in your campaigns if a particular segment is made up of Gen Z teens. You can also leverage well-known influencers and celebrities to attract more like-minded consumers to your brand.

*Segmentation Criteria*

**Demographic segmentation** enables you to understand who your [target audience](https://mailchimp.com/marketing-glossary/target-audience/) is, which is critical for building customer personas. A customer persona is essentially a profile used to represent your target market according to the data obtained from segmentation. Segmenting based on personas can provide 90% of companies with better knowledge about their audience.

This type of segmentation is also an excellent place to start if you want to understand your audience and if you’re just learning the ropes of segmentation since it’s easy to use. Demographic segmentation variables include:

* + - Age
    - Gender
    - Religion
    - Income level
    - Size of household
    - Occupation
    - Education
    - Marital status

Example: Say you run a wine company and much of your audience is married. In this case, you may promote a special wine bundle for couples to these contacts.

**Geographic segmentation** tells you where your audience is located. Like demographic segmentation, categorizing your contacts according to geographic location is straightforward. Here are a few geographic variables you can use to divide your audience:

* + - Location (Includes factors such as zip code, city, state, and country)
    - Culture
    - Time zone
    - Language
    - Climate
    - Population density

Example: If your target market is in an area where the dominant language is Spanish, you can send marketing postcards in Spanish to inform potential customers of your next big in-store event.

**Psychographic segmentation** enables businesses to segment their contacts based on psychological traits that influence shopping. Variables include:

* + - Attitudes
    - Values
    - Social status
    - Lifestyle
    - Personality
    - Opinions

While it can be difficult to segment audiences using this approach, doing so results in highly effective marketing campaigns. This is because psychographic segmentation provides insight into why consumers buy certain products.

Example: A brand that focuses on plant-based food products may have an audience segment of vegans, vegetarians, or pescatarians and meat eaters who wish to consume less meat.

**Behavioral segmentation** refers to a type of market segmentation in which you group your audience based on consumer behavior, allowing you to see how customers interact with your business. With behavioral segmentation, you can see what your contacts are doing on your website, determine which ones engage with your brand the most, and identify patterns to plan.

Variables of behavioral segmentation include:

* + - Purchasing behavior
    - Stage in the customer journey
    - Occasion or timing
    - Benefits sought
    - [Customer loyalty](https://mailchimp.com/resources/build-brand-loyalty/)
    - Customer satisfaction
    - Engagement

Example: You may send special deals and discount codes to those who signed up for your loyalty program but have not made a purchase yet.

*Should you use segmentation variables?*

Leveraging segmentation variables to create subcategories for your audience can help you make more focused marketing strategies and tailor you’re messaging according to the needs of your contacts.

It can be tempting to build marketing campaigns with generic, non-personalized content; after all, it’s easy and fast. However, this type of content can be vague and irrelevant to your target audience. But you can enhance your messaging by understanding your customers via applicable segmentation variables.

*Use resources more wisely.*

Instead of sending an email campaign to everyone in your contacts or developing an ad for your entire audience, you can focus on a smaller pool of people and obtain better results. This can help you spend less money or valuable resources and ensure they’re used effectively.

*Develop a stronger marketing strategy.*

Your marketing campaign can encompass different strategies, such as email automation or digital advertisements. With segmentation, you can ensure that your content is relevant and tailored to your audience–regardless of which type of tactic you use.

*Boost brand loyalty.*

Encouraging clients to be loyal to your brand can increase repeat sales and help you build a community. As such, your audience may be more likely to recommend your brand to people they know and advocate for your business.

*Ensure consistent branding.*

Segmentation is also valuable for ensuring the consistency of your messaging since it can help you stay on brand. So, no matter who your segment is, they’ll see a consistent marketing campaign that aligns with your organization.

Look at the following steps if you’re interested in using segmentation variables today.

* *Determine your objectives.*

The first step is to determine your objectives. What are you hoping to achieve by segmenting your audience? For example, do you want to raise brand awareness in areas where you have a physical retail store established or strengthen your relationship with your most loyal clients?

* *Verify the viability of your segments.*

It’s a good idea to verify the viability of your segments, as not all segments will align with your marketing strategy or overall business goals. It’s important to ensure the viability of your segments to prevent wasting resources. For your segment to be successful, ask yourself whether it’s stable, accessible, and competitive.

* *Use segmentation tools.*

You can use [segmentation tools](https://mailchimp.com/features/segmentation/) to organize your audience into subgroups that make sense for your business. With Mailchimp, you can also leverage pre-built market segments to ensure the success of your engagement or buying behavior strategy. These types of tools save you valuable time and money.

* *Develop your marketing segmentation strategy.*

You’ll choose between a concentrated or differentiated approach when [developing your](https://mailchimp.com/resources/beginners-guide-to-segmentation/) [marketing segmentation strategy](https://mailchimp.com/resources/beginners-guide-to-segmentation/). Concentrated marketing strategies will only focus on one segment, whereas differentiated marketing strategies will focus on catering to multiple customer segments. So if you want to target more than one segment, a strategy that revolves around differentiated marketing may be more suitable.

* *Create more effective marketing campaigns with market segmentation.*

Market segmentation allows you to split your audience into subgroups, providing an inside look into who your audience is, where they’re located, why they make purchasing decisions, and how they interact with your brand or website.

Segmentation also makes it easier to [create effective campaigns](https://mailchimp.com/marketing-platform/) that are tailored to the specific needs of your target market. When planning your strategy, you’ll choose between demographic, geographic, psychographic, or behavioral variables.

While segmenting your contacts can eat away at your time, our segmentation tools make it easy to build segments and target customers based on their activity, traits, and behaviors. Now, you can create more personalized campaigns that align with your segmentation strategy in less time and enhance your marketing efforts.

**Step 4: Exploring Data**

Data cleaning, also known as data preprocessing, is an essential step in machine learning. It refers to the process of identifying and correcting errors, inconsistencies, and missing data in the dataset before it is used for analysis or training a machine learning model.

The following are some common steps in data cleaning:

1. Handling missing data: Missing data can be imputed using various techniques such as mean, median, mode, or using machine learning algorithms to predict the missing values.

2. Removing duplicate records: Duplicates records can lead to bias in the analysis, and hence it's crucial to remove them from the dataset.

3. Outlier detection and removal: Outliers are data points that lie far away from the other data points and can skew the analysis or model training. Hence, it's essential to detect and remove outliers.

4. Handling inconsistent data: Inconsistent data can arise due to human error, data entry errors, or merge errors. These errors can be corrected manually or using machine learning algorithms.

5. Data transformation: Data transformation techniques such as normalization, scaling, or encoding can be applied to make the dataset more suitable for analysis or model training.

6. Data cleaning helps to improve the quality of data and leads to better analysis or machine learning models.

**Categorical Encoding**

Categorical encoding is the process of converting categorical data into numerical data, which can be used as input for machine learning algorithms. Categorical data is a type of data that includes categories or labels that are not numerical in nature. For example, colors, names, or types of products.

There are various techniques used for categorical encoding, including one-hot encoding, label encoding, binary encoding, and target encoding.

One-hot encoding is a technique where each category is converted into a binary vector, where only one element is 1 and the others are 0. Label encoding is a technique where each category is assigned a unique numerical value. Binary encoding is a technique where categories are converted into binary digits, and target encoding is a technique that involves encoding the categories based on their relationship with the target variable.

The choice of categorical encoding technique depends on the nature of the data and the machine learning algorithm used for analysis or prediction. Categorical encoding is an important step in data preprocessing, as it can significantly affect the performance of the machine learning model.

**PCA [Principal Component Analysis]**

PCA (Principal Component Analysis) is a dimensionality reduction technique that is used to reduce the number of variables in a dataset while retaining as much information as possible. It works by identifying the most important features or components in the data and transforming the data into a new coordinate system that represents these components.

PCA can prevent overfitting by reducing the complexity of the model. Overfitting occurs when a model is too complex and fits the training data too closely, which can lead to poor generalization of new or unseen data. PCA reduces the number of features used in the model, which can help to reduce the complexity of the model and prevent overfitting.

PCA works by identifying the principal components in the data, which are the directions in which the data varies the most. These principal components are orthogonal to each other and represent the most important information in the data. By projecting the data onto these principal components, PCA can reduce the dimensionality of the data while retaining most of the information.

**Step 5: Extracting Segments**

Whether or not distinct non-intersecting clusters are formed or not depends on what type of segmenting variables the organizational units have chosen for the data. If there exists a segmenting pattern in the underlying data, and the segmenting variables are chosen according to classes that cause this pattern, then the segments elected will be distinct. However, if there are noisy parameters that are chosen, then the segments will not be as segregated as you’d expect them to be.

* Clustering is the most common method used for segmenting.
* K means clustering won't always work, for example, in spiral datasets. It identifies grouped clusters based on regions.
* A better performance on spiral shapes can be obtained by the Single Linkage Hierarchical clustering algorithm.
* It's important to note that this doesn't mean that the Single Linkage Algorithm is better, only an indication that different algorithms work better on different types of data.

If consumer data is well-structured and well-separated, distinct market segments exist, and tendencies of different algorithms matter less. However, if the data is not well-structured, the tendency of the algorithm influences the solution substantially. In such situations, the algorithm will impose a structure that suits the objective function of the algorithm.

*Types of methods:*

Distance-based methods: These methods use the proximity of data points to determine clusters. Examples include:

* + K means clustering
  + Hierarchical clustering
  + DBSCAN (Density-Based Spatial Clustering of Applications with Noise)
  + OPTICS (Ordering Points to Identify the Clustering Structure)

Model-based methods: These methods formulate a concise stochastic model to extract segments. Examples include:

* + Gaussian Mixture Models (GMM)
  + Latent Class Analysis (LCA)
  + Finite Mixture Models (FMM)

Important to note that no single best algorithm exists. It’s all trial, error, and logical inferences drawn from what you know about the dataset that help you choose the most appropriate algorithm that yields the best results to form those segments.

Binary Segmentation: Segmentation involved with using binary segmentation variables. Binary segmentation involves using binary segmentation variables to segment data. Two methods of binary segmentation include:

* + Symmetric binary segmentation: where ones and zeros have equivalent values.
  + Asymmetric binary segmentation: where 1 is more valuable for clustering, but 0 doesn't indicate anything. For example, a person being interested in horseback riding is more useful information to the tourism industry than a person who does not like horseback riding.Top of Form

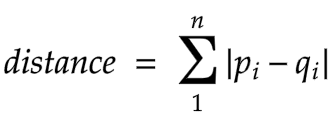
*Distance based methods:*

1. Euclidean distance: Calculates the straight-line distance between two points in space. Assumes that all dimensions of the data are equally important. It is the most commonly used distance metric.

Diagram

Description automatically generated

1. Manhattan distance: Calculates the distance between two points in a grid based on the sum of the absolute differences of their coordinates. Ignores the diagonal distance between two points. Is often used when dealing with sparse datasets.



1. Asymmetric binary distance: Calculates the distance between two binary vectors based on the number of elements that differ between them. Can be used when comparing objects that do not have a natural notion of distance.

Text

Description automatically generated

Euclidean and Manhattan distances:

* Both treat all dimensions of data equally.
* They take a sum over all dimensions of squared or absolute differences.
* Data needs to be standardized if dimensions are not on the same scale.

Hierarchical Clustering:

* The minimum number of segments that can be formed is 1, with the entire data set being part of one whole super cluster.
* The maximum number that can be formed is N, where N is the number of possible available data points, with each observation being its own segment.
* Clustering algorithms aim to find the number of segments K that lies between the two extremes.
* Two types: Divisive clustering and Agglomerative clustering
* Both result in a sequence of nested partitions

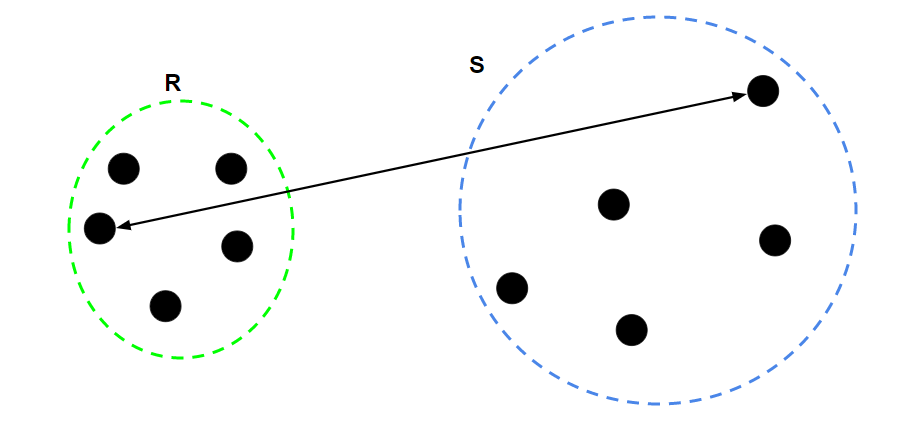
*Three types of linkages:*

* Single linkage: Distance between the closest points belonging to two different sets of clusters.

A picture containing diagram

Description automatically generated

* Complete linkage: Distance between two farthest points of the sets.



* Average linkage: Average distance between all the points in the two sets.

Chart

Description automatically generated with medium confidence

These are the three types of distances that hierarchal clustering algorithms use. You can either have Single linkage, Complete linkage, or Average linkage distances.

To be noted that this is an example of Agglomerative clustering. The step for conducting agglomerative clustering is as follows:

The first step of course is declaring all the individual consumers as Individual clusters. The idea is to converge into one single supercluster after the process is complete. After a dendrogram is ready we can decide upon the ideal number of market segments by observing the charts.

After declaring all the consumers as individual clusters we compute all possible distances from a point say xi with the rest of the clusters, (which at the time is n – 1, assuming the number of observations is n) and then pick the point with the smallest distance with the point xi. This process is then repeated for all the points and we end up with n/2 clusters after the first iteration.

Now how the distance between clusters is computed is based upon the type of method you choose. If you choose a singly linked method, then the distances computed are from two closest points from two different clusters. If you choose Complete linked clustering, then it’s two of the farthest points.

A picture containing diagram

Description automatically generated

Diagram

Description automatically generated with medium confidence

Hierarchical clusters are expressed through dendrograms: A dendrogram is a tree diagram. The root of the tree represents the one-cluster solution where one market segment contains all consumers. The leaves of the tree are the single observations and branches in-between correspond to the hierarchy of market segments formed at each step of the procedure. Dendrograms are often recommended as a guide to select the number of market segments, but rarely provide guidance of this nature because the data sets underlying the analysis are not well-structured enough.

*Partitioning Methods*

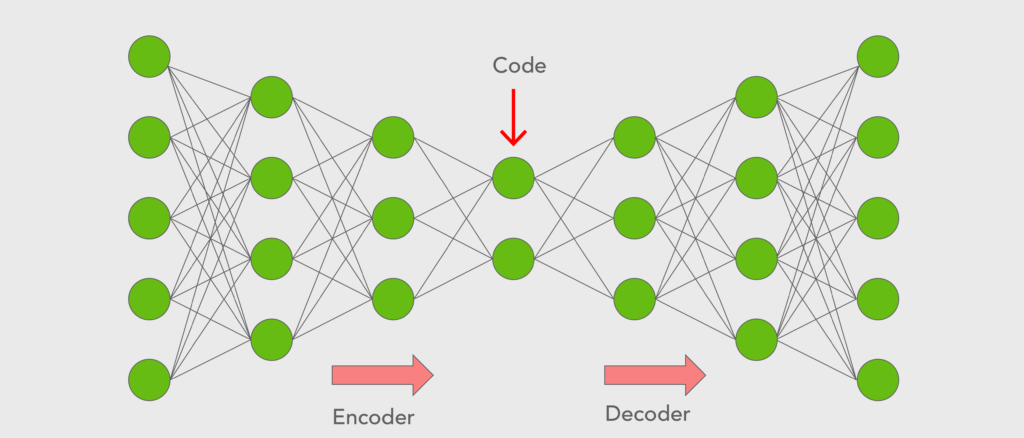
K Means: An Iterative method used to determine the optimal k Market segments, where K is predetermined by the users. In this process the input K is provided by the analyst and the algorithm elects K random data points from the set of records, which are declared as centroids. In the next step, each data record’s Euclidean distance is computed from all the individual centroids and the centroid which is the closest to the record is assigned a segment belonging to that centroid. After repeating this for every single data point, the centroids are re-assigned by the Mean of the elected segments, and the entire process is repeated again, until the co-ordinates of the Segment centroids do not change.

K-Medians is algorithm which uses the same procedure but instead of Euclidean distance it uses Manhattan distance and instead of re-assigning the segment centroids with the means of the newly updated segments it assigns the Medians.

The Neural gas Algorithm: The Neural Gas Algorithm doesn’t just adjust the primary centroid towards the consumer but also adjusts the secondary centroid by moving that closer to the user as well. This adjustment is known as Neural Gas Algorithm.

Before we delve into Model based clustering one more method that can be used for identifying familiar market segments that should be discussed is Autoencoders. Autoencoders are neural networks that are trained to predict the input without using the trivial multiply by one solution. One common question that crosses people’s minds is why would anyone ever do that? Predicting the input is useless because we already have the input.

The answer lies in the architecture of the Neural Network.



Autoencoders consist of an Encoder and a Decoder. As you already know they predict the input variables using the input variable without the Trivial Multiply by one solution. The Encoder part converts the input numbers into smaller coded data which can be interpreted by the decoder to predict the input variables again. This technique is usually used for Image data to compress and decompress large image files or to extract useful features from the images and eliminate the noise.

But one useful attribute of autoencoders is that they can clear noise from your data. Which means that if you provide it with a data point that doesn’t behave like the rest of the data it will have large RECONSTRUCTION ERROR, making it an outlier. This type of technique can therefore be used in Clustering algorithms or for Anomaly detection.

*Model Based Methods:*

The model-based methods are based on Two primary assumptions.

* 1. There exists a certain number of perfect market segments, each of which has their own distribution.
  2. Each segment has a specific set of characteristics that makes it distinct compared to the rest of them.

Model based learning uses these two primary assumptions to compute the probability of a data point belonging to all the possible segments, and the one where the data point is most likely to belong to, is considered as the segment from which the record is coming from.

It assumes a type of distribution for the segments and then uses maximum likelihood to estimate the mean and co-variance of the segment that maximizes the probability of getting a particular data record. After iteratively identifying the correct mean and Co-variance the model is trained and can be used to predict based on the probability of belonging to a certain segment.

Examples are GMM, VGMM and DPGMM.

Two types of Models based learning algorithms:

* 1. Finite Mixture of distributions: In finite mixture of distribution, instead of assuming one single distribution for all of the segments, it assumes that all the segments have their own individual distribution different from the rest. It uses the expected maximum (EM) algorithm to determine the co-variance and the mean of the distribution and after assignment classifies the records based on clusters.
  2. Finite Mixture of Regressions: In this method we use something called a response variable and assume that the data comes from a different type of regression model. Each segment follows its own regression model of course. The data points that can provide the best predictions for the response variables in a segment are all clumped together into the same segment. The data points that increase oppose the co-variance are clumped out of the cluster.

Algorithms with Integrated Variable selection:

Feature selection is a tedious process but obtaining the proper features for segmentation is crucial to reduce noise. It becomes simple, when the segmentation variables are metric to elect features and disregard the irrelevant ones, but when it comes to binary variables it becomes difficult to choose the useful ones and eliminate the noise. To counter this issue, we have the following two algorithms.

1. Biclustering: Binary data is harder to deal with than metric data due to the reasons aforementioned. The 1’s in the segmentation variable might be useful information but the zeroes could act as noise as they don’t convey any relevant information. To counter this biclustering algorithm was introduced.

The biclustering algorithm is specifically designed to cluster binary segmentation variables, It does it by applying matrix consumer switching operations on the rows of a dataset to clump a maximum number of ones in the top left corner of the dataset.

Once the data records are clumped together the records at the top left corner are assigned to be in the same cluster and removed from the matrix. Then the same process Is repeated on the rest of the data frame, to provide with similar clusters.

1. Variable Segmentation for based Clustering for Binary data (VSBD) algorithm: Just like the biclustering algorithm, variable segmentation for binary data algorithm removes noise from the existing data. It does it by iteratively computing the inertia of the clustering algorithm and finding the set of columns that have the least inertia.

It does this by first electing a subset of the entire dataset to reduce computational complexity, and elects the best combination of V Segmentation variables from all possible combinations. Once it has that, it keeps adding new columns to the dataset that reduces the inertia of the clustering, ending in the best subset of columns that give us clusters. The K means cluster of course takes a K value for input but to compute the value of K, Lance ratio or Rakowski’s ratio can be used.

**Step 6: Profiling Segments**

*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. 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. Profiling consists of characterizing the market segments individually, but also in comparison to the other market segments.

At the profiling stage, we inspect a number of alternative market segmentation solutions. This is particularly important if no natural segments exist in the data, and either a reproducible or a constructive market segmentation approach has to be taken. In the following sections we discuss traditional and graphical statistical approaches to segment profiling. Graphical statistical approaches make profiling less tedious, and thus less prone to misinterpretation.

*Traditional Approaches to Profiling Market Segments*

Data-driven segmentation solutions are usually presented to users (clients, managers) in one of two ways: (1) as high-level summaries simplifying segment characteristics to a point where they are misleadingly trivial, or (2) as large tables that provide, for each segment, exact percentages for each segmentation variable.

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.

Sometimes – to deal with the size of this task – information is provided about the statistical significance of the difference between segments for each of the segmentation variables. This approach, however, is not statistically correct. Segment membership is directly derived from the segmentation variables, and segments are created in a way that makes them maximally different, thus not allowing standard statistical tests to assess the significance of differences.

*Segment Profiling Visualization*

Graphics are particularly important in exploratory statistical analysis (like cluster analysis) because they provide insights into the complex relationships between variables. In addition, in times of bigger and increasingly bigger data, visualization offers a simple way of monitoring developments over time. So, recommend the use of visualizations techniques to make the results of a market segmentation analysis easier to interpret.

*Identifying Defining Characteristics of Market Segments*

A good way to understand the defining characteristics of each segment is to produce a segment profile plot. The segment profile plot shows – for all segmentation variables – how each market segment differs from the overall sample. The segment profile plot is the direct visual translation of tables option is to order segmentation variables by similarity of answer patterns.

*Assessing Segment Separation*

Segment separation can be visualized in a segment separation plot. 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. But even in such complex situations, segment separation plots offer data analysts and users a quick overview of the data situation, and the segmentation solution.

The two dimensions representing the segmentation variables can be directly plotted. This is not possible if 20-dimensional travel motives data serve as segmentation variables. In such a situation, the 20- dimensional space needs to be projected onto a small number of dimensions to create a segment separation plot.

# **Step 7: Describing Segments**

# *Developing a Complete Picture of Market Segments*

Segment profiling is about understanding differences in segmentation variables across market segments. Segmentation variables are chosen early in the market segmentation analysis process: conceptually in Step 2 (specifying the ideal target segment), and empirically in Step 3 (collecting data). Segmentation variables form the basis for extracting market segments from empirical data. Step 7 (describing segments) is like the profiling step. The only difference is that the variables being inspected have not been used to extract market segments. Rather, in Step 7 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 few dates to get to know the potential spouse as well as possible to give the marriage the best possible chance and avoid nasty surprises down the track. As van Raaij and Verhallen state: segment should be further described and typified by crossing them with all other variables, i.e., with psychographic, demographic, and socio-economic variables, media exposure, and specific product and brand attitudes or evaluations.

*Using Visualizations to Describe Market Segments*

A wide range of charts exist for the visualization 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).

*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.

**Step 8: Target Segments**

As per the above step, detailed information analysis takes place and retrieve ideal segment to improve organization marketing plan. Organizations decide which segment is crucial to take place as per investment required and long-term plan. Selecting the target segments is a crucial step in market segmentation analysis. Once a business has identified potential segments and collected data to understand their needs and behaviour, it can choose which segments to target with its marketing efforts.

The selection of the target segments should be based on several factors, including the size of the segment, its growth potential, profitability, accessibility, and compatibility with the business's capabilities and resources.

The size of the segment is important because it determines the potential customer base for the business. A larger segment may provide more opportunities for growth and revenue. Growth potential refers to the potential for the segment to expand or grow over time. Businesses should consider whether the segment is likely to grow or shrink in the future.

Profitability is another important factor. A profitable segment can generate revenue and profits for the business. Accessibility refers to the ability of the business to reach and communicate with the consumers in the segment. If a segment is difficult to reach, it may not be an ideal target for the business.

Compatibility refers to the fit between the segment's needs and the business's capabilities and resources. If a segment's needs are not compatible with the business's capabilities or resources, the business may struggle to serve the segment effectively.

Ultimately, businesses should choose the target segments that align with their goals and capabilities. By focusing on a specific segment, businesses can tailor their marketing efforts to meet the needs and preferences of that segment, leading to increased sales and customer loyalty.

**Step 9: Customizing the Marketing Mix**

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 the segmentation targeting is referred to as positioning approach. To best ensure maximizing the benefits of a market segmentation strategy, it is important to customize the marketing mix to the target segment. The selection of one or more specific target segments may require the design of new, or the modification or rebranding of existing products (Product), changes to prices or discount structures (Price), the selection of suitable distribution channels (Place), and the development of new communication messages and promotion strategies that are attractive to the target segment (Promotion).

Customizing the marketing mix is an essential step in market segmentation analysis, where businesses tailor their marketing efforts to meet the needs and preferences of specific consumer segments. The marketing mix comprises four key elements: product, price, promotion, and place. By customizing these elements for each segment, businesses can create a targeted and effective marketing strategy. The product element involves developing products that meet the specific needs and preferences of each segment. For instance, a business may offer different product features, packaging, or branding for different segments. The price element involves setting prices that are attractive to each segment. For instance, a business may offer discounts or promotions to price-sensitive segments or premium prices to segments that value quality over price.

The promotion element involves creating advertising and promotional campaigns that appeal to each segment. For instance, a business may use different advertising channels or messaging to reach different segments.

The place element involves distributing products through channels that are accessible to each segment. For instance, a business may use different distribution channels or locations to reach different segments.

By customizing the marketing mix for each segment, businesses can create a targeted and effective marketing strategy that meets the specific needs and preferences of each segment. This can lead to increased sales, customer loyalty, and overall business success.