SUMMARY OF FUNDAMENTALS OF MARKET SEGMENTATION

Team members:

Manish Gholam

S. Charan Raju

Bage Aman Shaikh

Index:

Market Segmentation

Marketing is the activity, set of institutions, and processes for creating, communicating, delivering, and exchanging offerings that have value for customers, clients, partners, and society at large.

1. **What is marketing?**

Marketing strategy is a set of specific ideas and actions that outline and guide decisions on the best or chosen way to create, distribute, promote, and price a product or service (Manage the marketing mix variables).

2. **What is market segmentation?**

At its core, market segmentation is the practice of dividing your target market into approachable groups. Market segmentation creates subsets of a market based on demographics, needs, priorities, common interests, and other psychographic or behavioral criteria used to better understand the target audience.

Benefits of Market Segmentation.

1. Stronger marketing messages: You no longer have to be generic and vague – you can speak directly to a specific group of people in ways they can relate to, because you understand their characteristics, wants, and needs.
2. Targeted digital advertising: Market segmentation helps you understand and define your audience's characteristics, so you can direct your marketing efforts to specific ages, locations, buying habits, interests etc.
3. Developing effective marketing strategies: Knowing your target audience gives you a head start about what methods, tactics and solutions they will be most responsive to.
4. Better response rates and lower acquisition costs: These will result from creating your marketing communications both in ad messaging and advanced targeting on digital platforms like Facebook and Google using your segmentation.
5. Attracting the right customers: Market segmentation helps you create targeted, clear and direct messaging that attracts the people you want to buy from you.
6. Increasing brand loyalty: when customers feel understood, uniquely well served and trusting, they are more likely to stick with your brand.
7. Differentiating your brand from the competition: More specific, personal messaging makes your brand stand out.
8. Identifying niche markets: segmentation can uncover not only underserved markets, but also new ways of serving existing markets – opportunities which can be used to grow your brand.
9. Staying on message: As segmentation is so linear, it's easy to stay on track with your marketing strategies, and not get distracted into less effective areas.

10. Driving growth: You can encourage customers to buy from you again, or trade up from a lower-priced product or service.
11. Enhanced profits: Different customers have different disposable incomes; prices can be set according to how much they are willing to spend. Knowing this can ensure you don't over (or under) sell yourself.
12. Product development: You'll be able to design with the needs of your customers top of mind, and develop different products that cater to your different customer base areas.

3. Types of Market Segmentation:

	Demographic (B2C)	**Firmographic (B2B)**	**Psychographic (B2B/B2C)**	**Behavioral (B2B/B2C)**
Definition	Classification based on individual attributes	Classification based on company or organization attributes	Classification based on attitudes, aspirations, values, and other criteria	Classification based on behaviors like product usage, technology laggards, etc.
Examples	Geography Gender Education Level Income Level	Industry Location Number of Employees Revenue	Lifestyle Personality Traits Values Opinions	Usage Rate Benefit Types Occasion Purchase Decision
Decision Criteria	You are a smaller business or you are running your first project	You are a smaller business or you are running your first project	You want to target customers based on values or lifestyle	You want to target customers based on purchase behaviors
Difficulty	Simpler	Simpler	More advanced	More advanced

Step 1: Deciding (not) to Segment

Marketing segmentation is a long-term work strategy and includes research, fielding surveys, focus groups, designing multiple packages, advertisements, communication. Before deciding to use or not use this method or strategy one must check with constraints and requirements. The scheme must be more profitable than marketing without it, and net of expenses of developing and using the scheme itself.

Barriers:

They are some parameters we need to see before starting up the process of market segmentation. There are some obstacles between successful segmentation implementation.

1. **Senior Management:**
   Lack of leadership in implementation and decision making will cost much to the company as investments are done for these segments. Lack of commitment and involvement will cause loss.

2. **Organizational culture:**
   A proper culture is needed to be developed in the organizing committee. Where there is a lack of consumer orientation, no new ideas, no creative thinking, bad communication, short term thinking, politics and the unwillingness.

3. **Lack of training:**
   Necessary marketing and analytics training is not provided then lack in performance will cause loss to implementation of market segmentation.

4. **Inadequate financial resources:**
   While entering or implementing marketing segmentation we sometimes need to make changes in infrastructure or product details and designing too. Hence, we need financial investments to go further. If not, adequate resources are available then it may cause an obstacle between implementation of segmentation and running them.

Then what's required first?

1. Sense of purpose
2. Dedication
3. Patience
4. Willingness

Step 2: Specifying the Ideal Target Segment

In Step 2 the organization must determine two sets of segment evaluation criteria. One set of evaluation criteria can be referred to as knock-out criteria. These criteria are the essential, non-negotiable features of segments that the organization would consider targeting. The second set of evaluation criteria can be referred to as attractiveness criteria. These criteria are used to evaluate the relative attractiveness of the remaining market segments – those in compliance with the knock-out criteria.

Step 3: Collection Data

While collecting the data we have to look up on these points. The need of the useful data is important while collecting because it plays a major role for accurate results. we have to do market segmentation based on the data driven Technique. In our case McDonald's dataset contains features Age, Like, tasty, Gender etc. The term segmentation variable refers to the one measured value such as Gender in our dataset etc.

Dataset:

	yummy	convenient	spicy	fattening	greasy	fast	cheap	tasty	expensive	healthy	disgusting	Like	Age	VisitFrequency	Gender
0	No	Yes	No	Yes	No	Yes	Yes	No	Yes	No	No	-3	61	Every three months	Female
1	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	+2	51	Every three months	Female
2	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	No	+1	62	Every three months	Female
3	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No	Yes	+4	69	Once a week	Female
4	No	Yes	No	Yes	Yes	Yes	Yes	No	No	Yes	No	+2	49	Once a month	Male
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1448	No	Yes	No	Yes	Yes	No	No	No	Yes	No	Yes	I hate it!-5	47	Once a year	Male
1449	Yes	Yes	No	Yes	No	No	Yes	Yes	No	Yes	No	+2	36	Once a week	Female
1450	Yes	Yes	No	Yes	No	Yes	No	Yes	Yes	No	No	+3	52	Once a month	Female
1451	Yes	Yes	No	No	No	Yes	Yes	Yes	No	Yes	No	+4	41	Every three months	Male
1452	No	Yes	No	Yes	Yes	No	No	No	Yes	No	Yes	-3	30	Every three months	Male

3.1 Segmentation Criteria:

Segmentation involves creating similar groups made up of individuals with identifiable common characteristics. These might be place of residence, how usually they buy our Product, age, expenditure, lifestyle or even how they behave on your Market. The individuals within a same segment are supposed to have the similar expectations and buying on our products and should react in a similar when we kept the offer and when we ask about Feedback. There are four 4 major Segments in which we make a Criteria.

1. **Geographic Criteria**:

This type of segmentation is based on the geolocation, Weather, Day etc.

Ex: The international Clothing Company plans to sells Sweatshirt for its consumers according to their geolocation and local weather. It suggests Sweatshirts that meet the immediate needs of visitors. They see two different offers depending on their local temperature a Sweatshirt rated 100$ where temperature is <-10°. And another jacket Price is 200$ withstand temperatures of -30°.

2. **Socio-Demographic Criteria**:

Demographic segmentation is the most commonly used criteria, this criteria Includes more specific information about us like Gender, Age, Expenditure, salary, Education, in this we can retrieve Information Easily and target the Customers

Ex: Fashion Applications targeting the consumers based on Gender, salary etc. which product has to recommend (If the person was studying in a prestigious college and have a high purchasing from the application then it will recommend costly items compare to another consumers)

3. **Psychographic Criteria**:

When people are segmented based on their Interests, beliefs and preferences these comes under psychographic criteria

Ex: For example, people who are in young or newly married couples' primary motivation to go on vacation to Maldives and enjoy in the beaches these kinds of people have a high likelihood of taking leave on summer.

4. **Behavioral Criteria:**

Behavioral segmentation criteria depend on the way visitors interact with the website. Some data depends on their immediate online behavior and giving positive feedback while other data depends on their past offline behavior or negative feedback.

Ex: The more the user visited the particular page and the time spent on that product etc.

In our case McDonald's market segmentation, we use mostly Geographic and Socio-Demographic data because psychographic criteria much not useful in this case

3.2 Data from Survey Studies:

1) Choice of Variables:

Selecting the right variables that are included commonsense segmentation and data-driven segmentation is critical to the quality of the market segmentation solution. while surveying the data we have to look upon these actions. If we use categorical values for representing gender or scale range from 1 to 10 for feedback on that specific item these comes under Response style. Sample data can be collected through surveys, experimental studies, or the company's internal data on particular item.

Step 4: Exploratory Data Analysis

4.1 A First Glimpse at the Data:

It is an approach to analyze the data using visualization techniques. It is used to discover trends, patterns, or to check our assumptions with the help of statistical summary and graphical representations. From the collected data we must do pre-processing because most of the data is un-structured. This round of investigation also provides recommendations for the best algorithm for extracting useful market segments.

Let's look up McDonald's Dataset:

```
In [5]: df.info()
        df.shape

        <class 'pandas.core.frame.DataFrame'>
        RangeIndex: 1453 entries, 0 to 1452
        Data columns (total 15 columns):
         #   Column          Non-Null Count  Dtype
        ---  ------          --------------  -----
         0   yummy           1453 non-null   object
         1   convenient      1453 non-null   object
         2   spicy           1453 non-null   object
         3   fattening       1453 non-null   object
         4   greasy          1453 non-null   object
         5   fast            1453 non-null   object
         6   cheap           1453 non-null   object
         7   tasty           1453 non-null   object
         8   expensive       1453 non-null   object
         9   healthy         1453 non-null   object
         10  disgusting      1453 non-null   object
         11  Like            1453 non-null   object
         12  Age             1453 non-null   int64
         13  VisitFrequency  1453 non-null   object
         14  Gender          1453 non-null   object
        dtypes: int64(1), object(14)
        memory usage: 170.4+ KB

Out[5]: (1453, 15)
```

4.2 Data Cleaning:

The first step before commencing data analysis is to clean the data. This includes checking if all values have been recorded correctly, and if consistent labels for the levels of categorical variables have been used. For many metric variables, the range of plausible values is known in advance. For example, age (in years) can be expected to lie between 0 and 110. It is easy to check whether any implausible values are contained in the data, which might point to errors during data collection or data entry. In or dataset there is no null values and we use label encoding for converting the categorical data into numerical data

4.3 Descriptive Analysis:

Descriptive techniques often include constructing tables of quantiles and means, methods of dispersion such as variance or standard deviation, and cross-tabulations or "crosstabs" that can be used to carry out many disparate hypotheses. These hypotheses often highlight differences among subgroups

	Age
count	1453.000000
mean	44.604955
std	14.221178
min	18.000000
25%	33.000000
50%	45.000000
75%	57.000000
max	71.000000

4.4 Pre-Processing:

1) Categorical Variables:

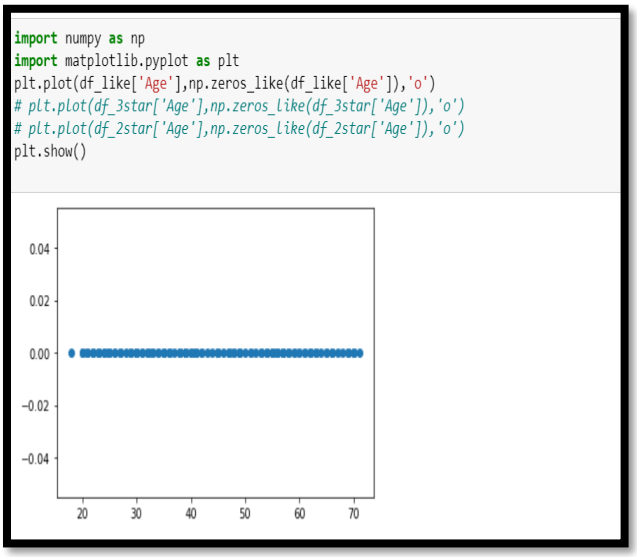
Two pre-processing procedures are often used for categorical variables. One is merging levels of categorical variables before further analysis, the other one is converting categorical variables to numeric ones, if it makes sense to do so. Merging levels of categorical variables is useful if the original categories are too differentiated (too many). Thinking back to the income variables

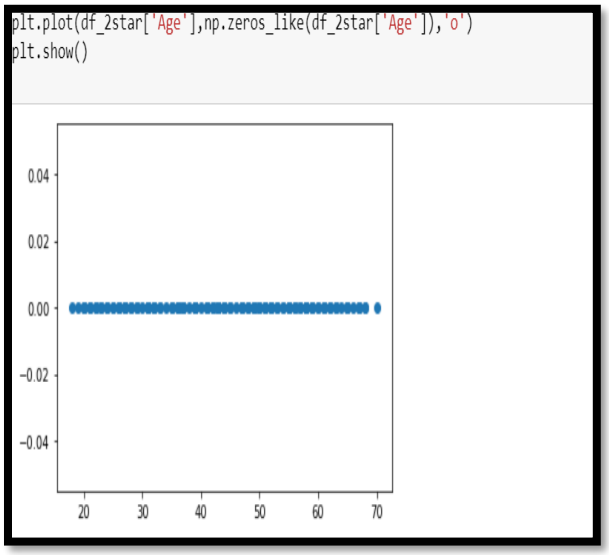
2) **Numerical variables**:

In distance-based methods of segment extraction, the range of values of a segmentation variable determines its relative influence. If one of the segmentation variables is binary (with values 0 or 1 indicating whether or not a customer views on the product of fast food), and a second variable indicates the expenditure in dollars per person per day (with values ranging from zero to $1000), a one-dollar difference in spend per person per day is weighted equally as the difference in liking to dine out or not.

3) Univariate Analysis:

We take one feature and based on that we will try to classify what the output. Suppose in our dataset we took age as feature and classify based how much they are liked. From our data all the persons who gave positive feedback (+4) their age is >20 and the data is fit (overlapped) one guy from age<20 gave positive feedback





4) Bivariate analysis:

Bivariate analysis is slightly more analytical than Univariate analysis. When the data set contains two variables and researchers aim to undertake comparisons between the two data set then Bivariate analysis is the right type of analysis technique.

	yummy	convenient	spicy	fattening	greasy	fast	cheap	tasty	expensive	healthy	disgusting	Age	VisitFrequency
yummy	1.000000	0.254168	0.008500	-0.086769	-0.151239	0.104908	0.104199	0.686970	-0.060765	0.246928	-0.419128	-0.281647	-0.022890
convenient	0.254168	1.000000	0.029000	0.036368	-0.107193	0.244490	0.151925	0.290241	-0.164363	0.099351	-0.341563	-0.070637	-0.036239
spicy	0.008500	0.029000	1.000000	-0.041239	0.054161	0.019673	0.016941	0.060896	0.046306	0.106154	0.033261	0.146261	-0.016897
fattening	-0.086769	0.036368	-0.041239	1.000000	0.319143	0.046557	-0.026326	-0.087257	0.088832	-0.338314	0.145743	-0.139308	0.005570
greasy	-0.151239	-0.107193	0.054161	0.319143	1.000000	-0.056444	-0.072526	-0.162809	0.146544	-0.210944	0.321171	-0.233739	0.006931
fast	0.104908	0.244490	0.019673	0.046557	-0.056444	1.000000	0.250514	0.143589	-0.199668	0.032670	-0.134906	-0.020273	0.013965
cheap	0.104199	0.151925	0.016941	-0.026326	-0.072526	0.250514	1.000000	0.138510	-0.721282	0.129662	-0.124314	0.017108	0.033924
tasty	0.686970	0.290241	0.060896	-0.087257	-0.162809	0.143589	0.138510	1.000000	-0.103771	0.225969	-0.436190	-0.194543	-0.022660
expensive	-0.060765	-0.164363	0.046306	0.088832	0.146544	-0.199668	-0.721282	-0.103771	1.000000	-0.069463	0.198455	-0.071051	-0.010868
healthy	0.246928	0.099351	0.106154	-0.338314	-0.210944	0.032670	0.129662	0.225969	-0.069463	1.000000	-0.173317	0.021425	-0.007774

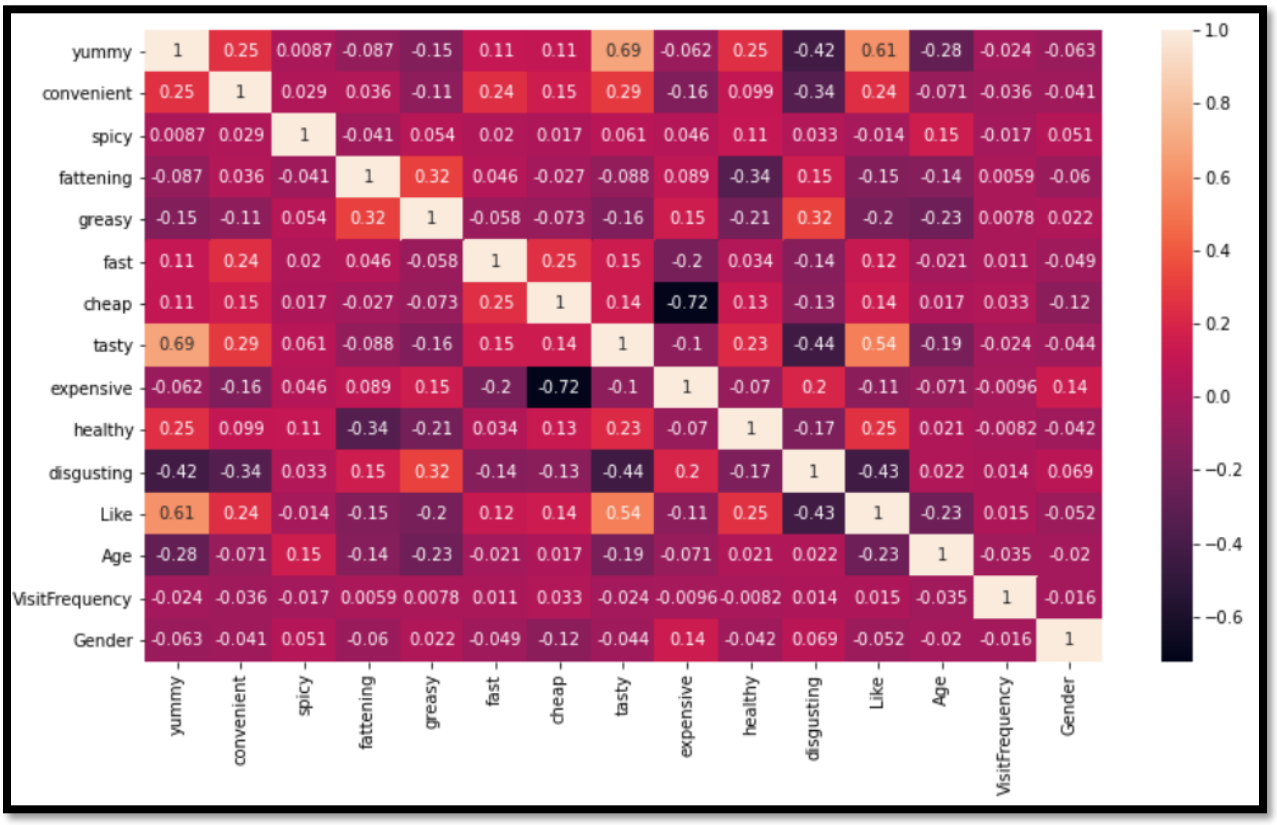
5) Multivariate analysis:

Multivariate analysis is a more complex form of statistical analysis technique and used when there are more than two variables in the data set. Here we can apply PCA to reduce the dimensions

4.5 Principal Component Analysis:

Principal component analysis Principal components analysis (PCA) converts a multivariate data set with metric variables into a new data set with uncorrelated and importance-ordered variables called principal components. The most variability is contained in the first variable (principal component), the second principal component contains the second most variability, and so on. Because principal components analysis generates as many new variables as there were old ones, observations (consumers) retain their relative positions to one another after transformation, and the dimensionality of the new data set remains the same. The data space is essentially unchanged in principal components analysis, but it is viewed from a different perspective. Prior to creating market segments from consumer data, principal components analysis may be used to reduce the number of segmentation factors. This concept is appealing as more variables increase the intricacy of the problem that the segment extraction technique must deal with, making extraction more difficult and requiring larger sample sizes

Heat Map for Visualizing:



From this data we can say that the features yummy and tasty are more contributing to the 'Like' output. And if it is disgusting then more people are giving negative feedback. So, when we are

marketing from next time we make sure that these features yummy, tasty, healthy, convenient are required for better sales

Step 5: Extracting Segments

Step 5 is where we extract segments. To illustrate a range of extraction techniques, we subdivide this step into sections. In the first section, we will use standard k-means analysis.

5.1 Grouping Consumers:

For extracting market segments from data mostly clustering analysis is used with a variety of approaches. Selecting a suitable clustering method requires matching the data analytic features of the resulting clustering with the context-dependent requirements that are desired by the researcher of the market.

Some of Data set and segment characteristics informing extraction algorithm selection given:

Data set characteristics: – Size (number of consumers, number of segmentation variables) – Scale level of segmentation variables (nominal, ordinal, metric, mixed) – Special structure, additional information.

Segment characteristics: – Similarities of consumers in the same segment – Differences between consumers from different segments – Number and size of segments

None of these methods outperform other methods in all situations. Rather, each method has pros and cons.

Distance-based methods use a particular notion of similarity or distance between observations (consumers), and try to find groups of similar observations (market segments). For distance- based methods, the choice of the distance measure depends on the scale level of the data.

Model-based methods formulate a concise stochastic model for the market segments where mainly distributions and probabilistic approach is used. Model based approaches are more probability based considering parameters of segment size and characteristics consumer's probability of getting fit into segment is derived and best solution is provided in end. If the data set contains repeated measurements of consumers over time, for example, an algorithm that takes this longitudinal nature of the data into account is needed. Such data generally requires a model-

based approach. If the data contains purchase histories and price information, and market segments are based on similar price sensitivity levels, regression models are needed. This, in turn, calls for the use of a model-based segment extraction algorithm.

In the case of binary segmentation variables, another aspect needs to be considered. We may want consumers in the same segments to have both the presence and absence of segmentation variables in common, here these variables would be symmetrical (with 0s and 1s treated equally).

Alternatively, we may be concerned about segmentation variables consumers have in common, here these variables would be symmetrical (with only common 1s being of interest). Biclustering uses binary information asymmetrically. Distance-based methods can use distance measures that account for this asymmetry, and extract segments characterized by common 1s.

Data-driven market segmentation analysis is exploratory by nature. Consumer data sets are typically not well structured. Consumers come in all shapes and forms a two-dimensional plot of consumers' product preferences typically does not contain clear groups of consumers. Rather, consumer preferences are spread across the entire plot. The combination of exploratory methods and unstructured consumer data that results from any method used to extract market segments from such data will strongly depend on the assumptions made on the structure of the segments implied by the method. The result of a market segmentation analysis, therefore, is determined as much by the underlying data as it is by the extraction algorithm chosen.

There are few types of **Distance-based Methods:**

5.2 Distance Methods:

1) Euclidean distance

Euclidean Distance represents the shortest distance between two points. Most machine learning algorithms including K-Means use this distance metric to measure the similarity between observations

$$d = ((p_1 - q_1)^2 + (p_2 - q_2)^2)^{1/2}$$

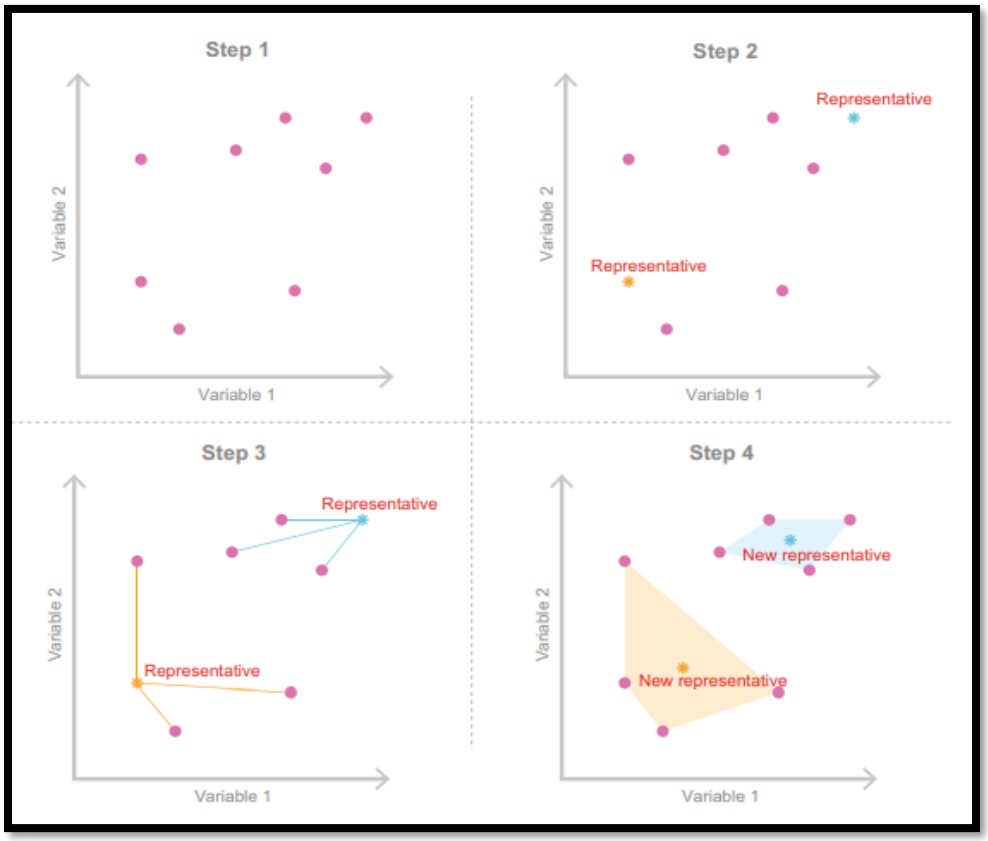
2) Hierarchical measures

Hierarchical clustering methods are the most intuitive way of grouping data because they mimic how a human would approach the task of dividing a set of n observations(consumers) into k groups (segments). If the aim is to have one large market segment (k = 1), the only possible solution is one big market segment containing all consumers in data X.

$$l(\mathcal{X}, \mathcal{Y}) = \frac{1}{|\mathcal{X}||\mathcal{Y}|} \sum_{\mathbf{x} \in \mathcal{X}} \sum_{\mathbf{y} \in \mathcal{Y}} d(\mathbf{x}, \mathbf{y}),$$

3) Partitioning method

Hierarchical clustering methods are particularly well suited for the analysis of small data sets with up to a few hundred observations. For larger data sets, dendrograms are hard to read, and the matrix of pairwise distances usually does not fit into computer memory**.** For data sets containing more than 1000 observations (consumers), clustering methods creating a single partition are more suitable than a nested sequence of partitions. This means that – instead of computing all distances between all pairs of observations in the data set.

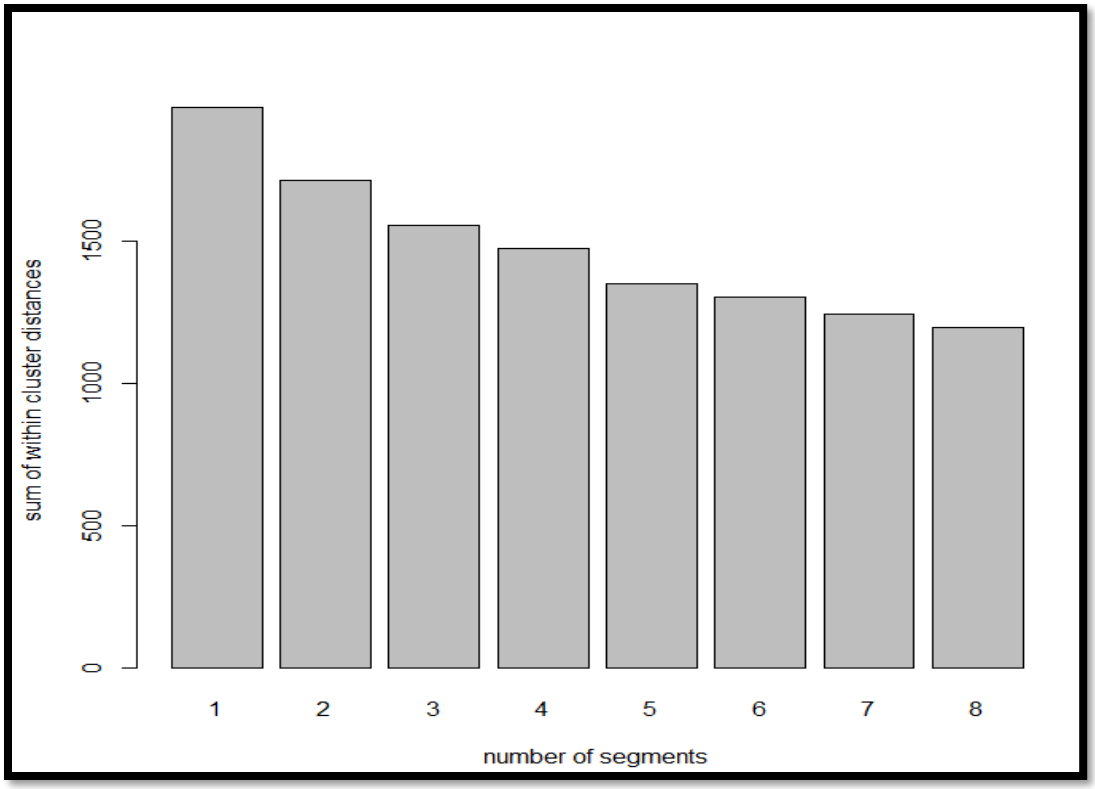


Simplified visualisation of the k-means clustering algorithm

5.3 Using k-Means

We calculate solutions for two to eight market segments using standard k-means analysis with ten random restarts. We then relabel segment numbers such that they are consistent across segmentations.

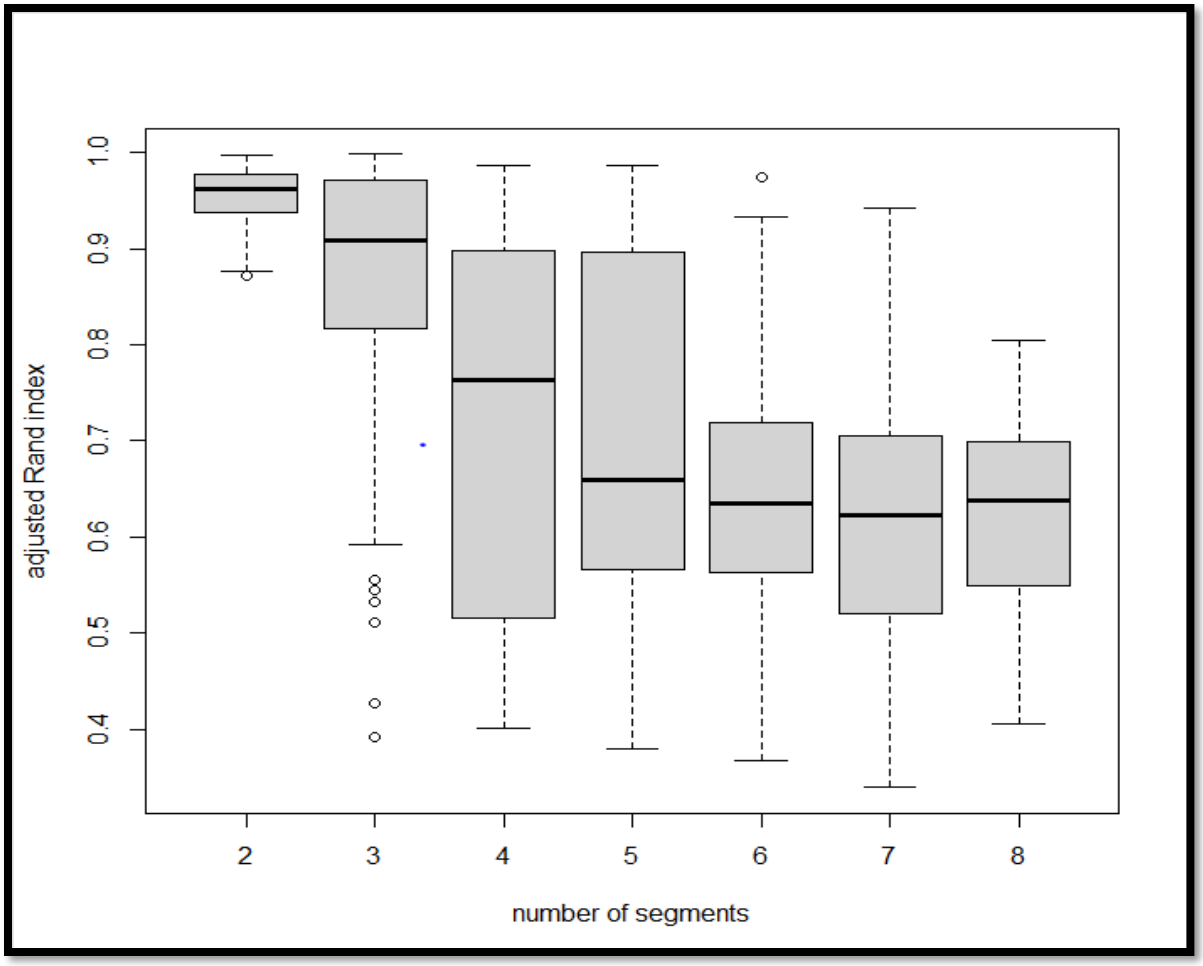
We extract between two and eight segments because we do not know in advance what the best number of market segments is. If we calculate a range of solutions, we can compare them and choose the one which extracts segments containing similar consumers which are distinctly different from members of other segments. We compare different solutions using a scree plot:



Scree plot for the McDonald's data set

The scree plot has no distinct elbow: the sum of distances within market segments drops slowly as the number of market segments increases. We expect the values to decrease because more market segments automatically mean that the segments are smaller and, as a consequence, that segment members are more similar to one another. But the much-anticipated point where the sum of distances drops dramatically is not visible. This scree plot does not provide useful guidance on the number of market segments to extract.

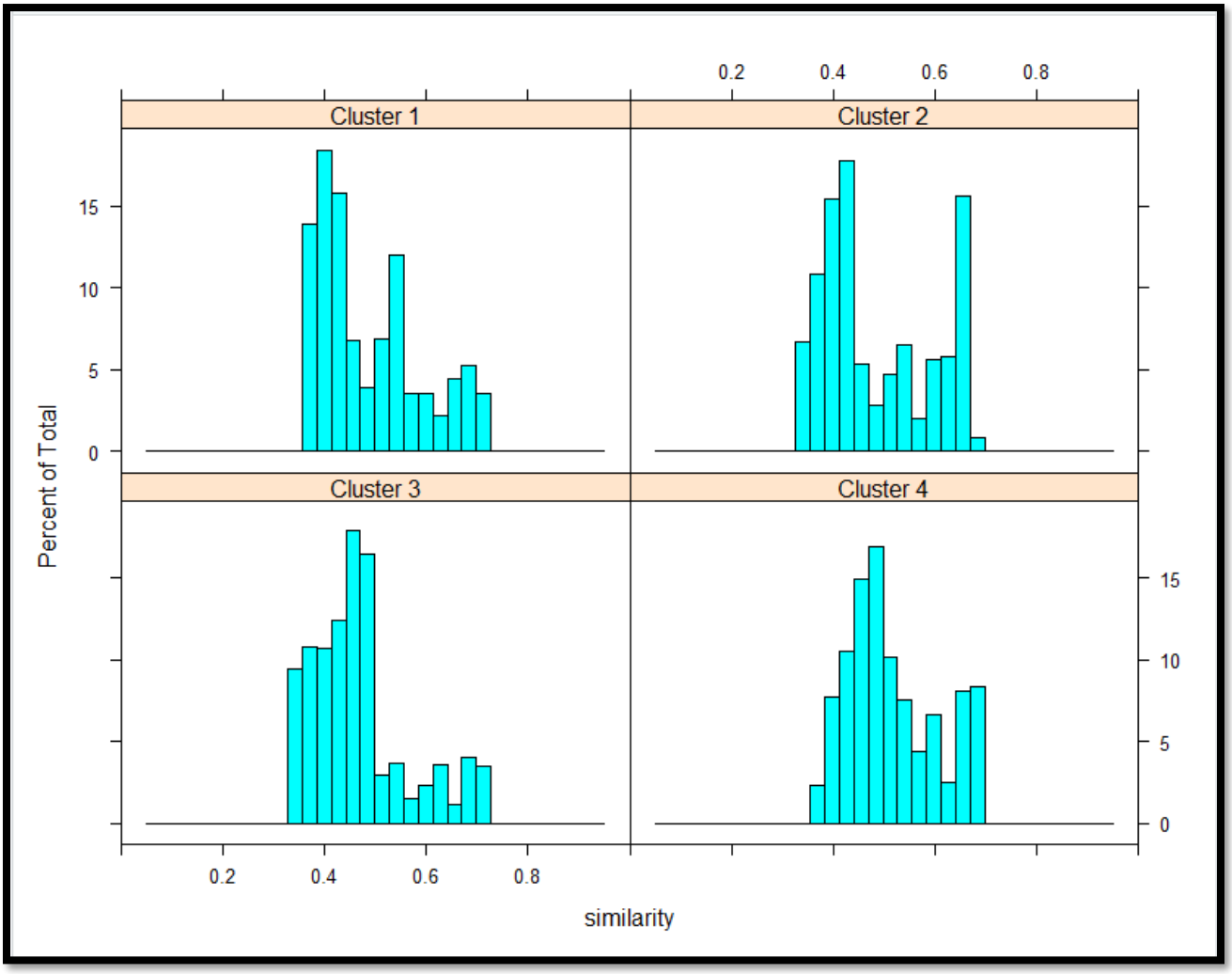
A second approach to determining a good number of segments is to use stability-based data structure analysis. Stability-based data structure analysis also indicates whether market segments occur naturally in the data, or if they have to be artificially constructed. Stability-based data structure analysis uses stability across replications as criterion to offer this guidance. Imagine using a market segmentation solution which cannot be reproduced. Such a solution would give McDonald's management little confidence in terms of investing substantial resources into a market segmentation strategy.



Global stability of k-means segmentation solutions for the McDonald's data set

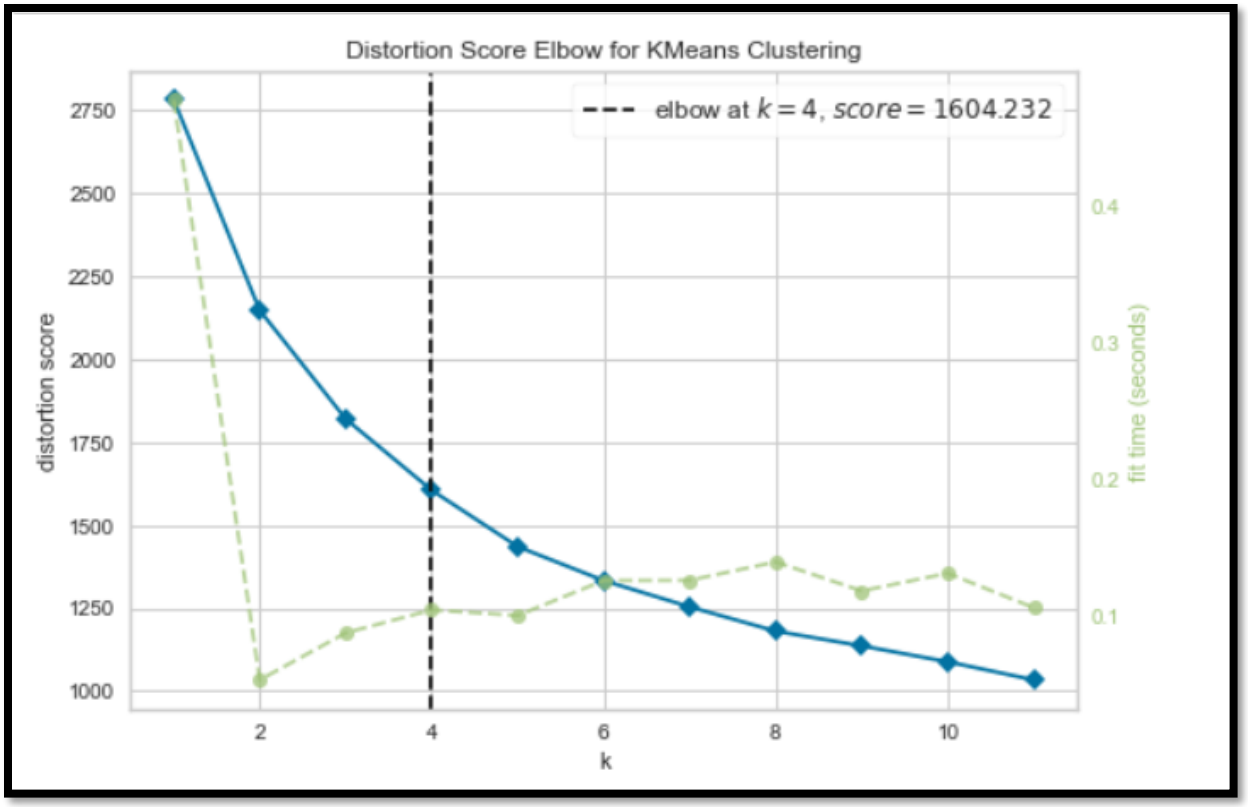
The vertical boxplots show the distribution of stability for each number of segments. The median is indicated by the fat black horizontal line in the middle of the box. Higher stability is better. Inspecting points to the two-, three- and four-segment solutions as being quite stable. However, the two- and three-segment solutions do not offer a very differentiated view of the market.

We gain further insights into the structure of the four-segment solution with a gorge plot:



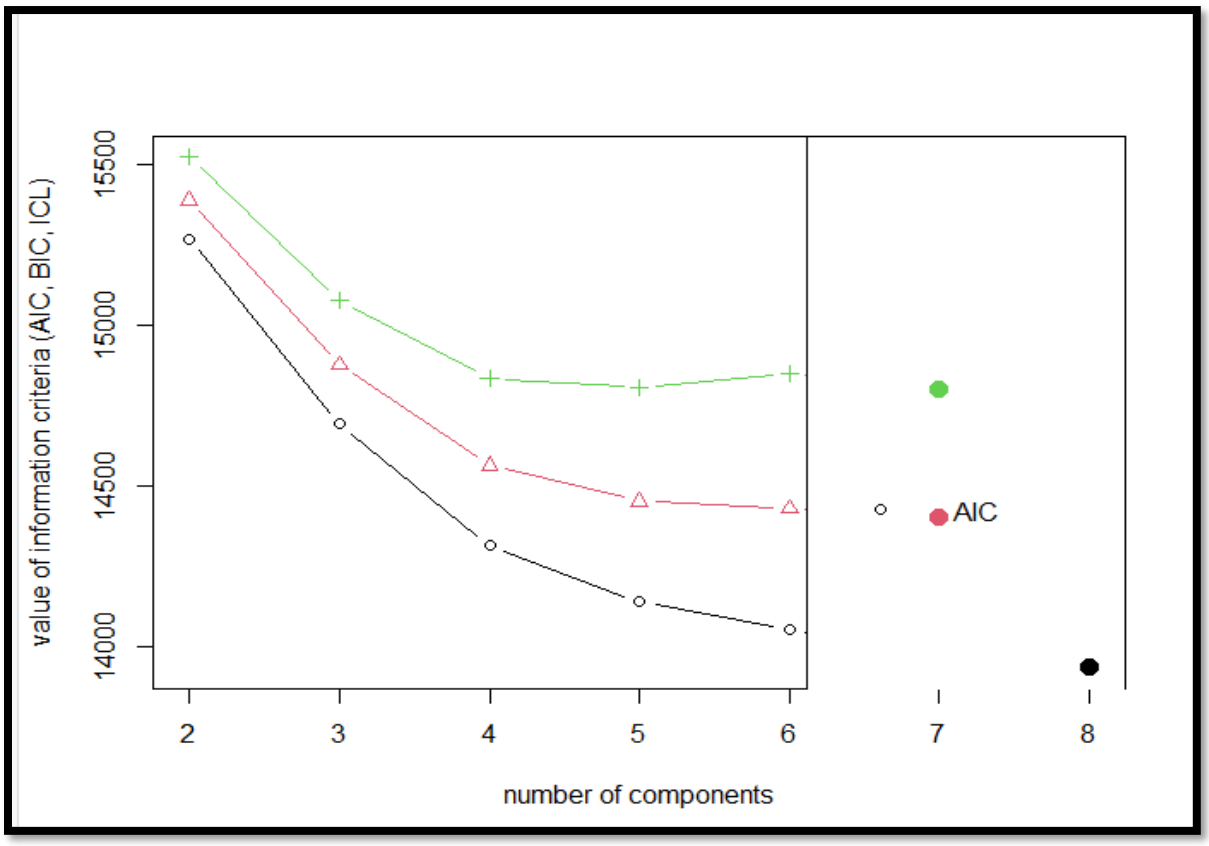
Gorge plot of the four-segment k-means solution for the McDonald's data set

None of the segments is well separated from the other segments, and proximity to at least one other segment is present as indicated by the similarity values all being between 0.3 and 0.7.


Distortion Score Elbow for KMeans Clustering
elbow at $k = 4$, $score = 1604.232$
distortion score
fit time (seconds)
k
2750
2500
2250
2000
1750
1500
1250
1000
0.4
0.3
0.2
0.1
2
4
6
8
10

5.4 Using Mixtures of Distributions:

We calculate latent class analysis using a finite mixture of binary distributions. The mixture model maximizes the likelihood to extract segments (as opposed to minimizing squared Euclidean distance, as is the case for k-means).



Information criteria for the mixture models of binary distributions with 2 to 8 components (segments) for the McDonald's data set

Step 6: Profiling Segments

6.1 Identifying Key Characteristics of Market Segments:

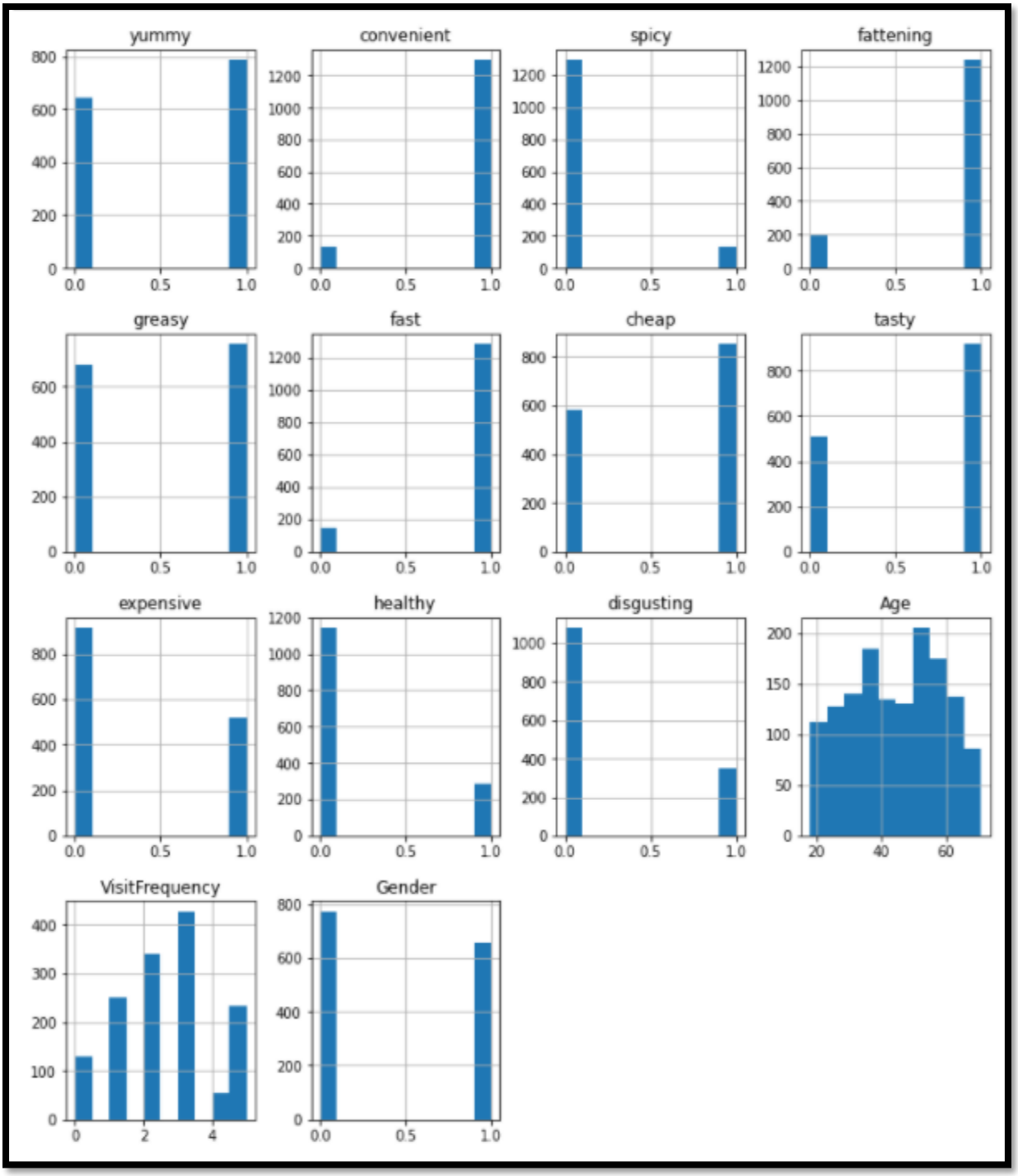
The profiling step aims to get to know the market segments resulting from the extraction step. Profiling is only required when data-driven market segmentation is used. For commonsense segmentation, the profiles of the segments are predefined. If, for example, age is used as the segmentation variable for the commonsense segmentation, it is obvious that the resulting segments will be age groups. Therefore, Step 6 is not necessary when commonsense segmentation is conducted.

At the profiling stage, we inspect several 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. Good profiling is the basis for the correct interpretation of the resulting segments. A correct interpretation, in turn, is critical to making good strategic marketing decisions.

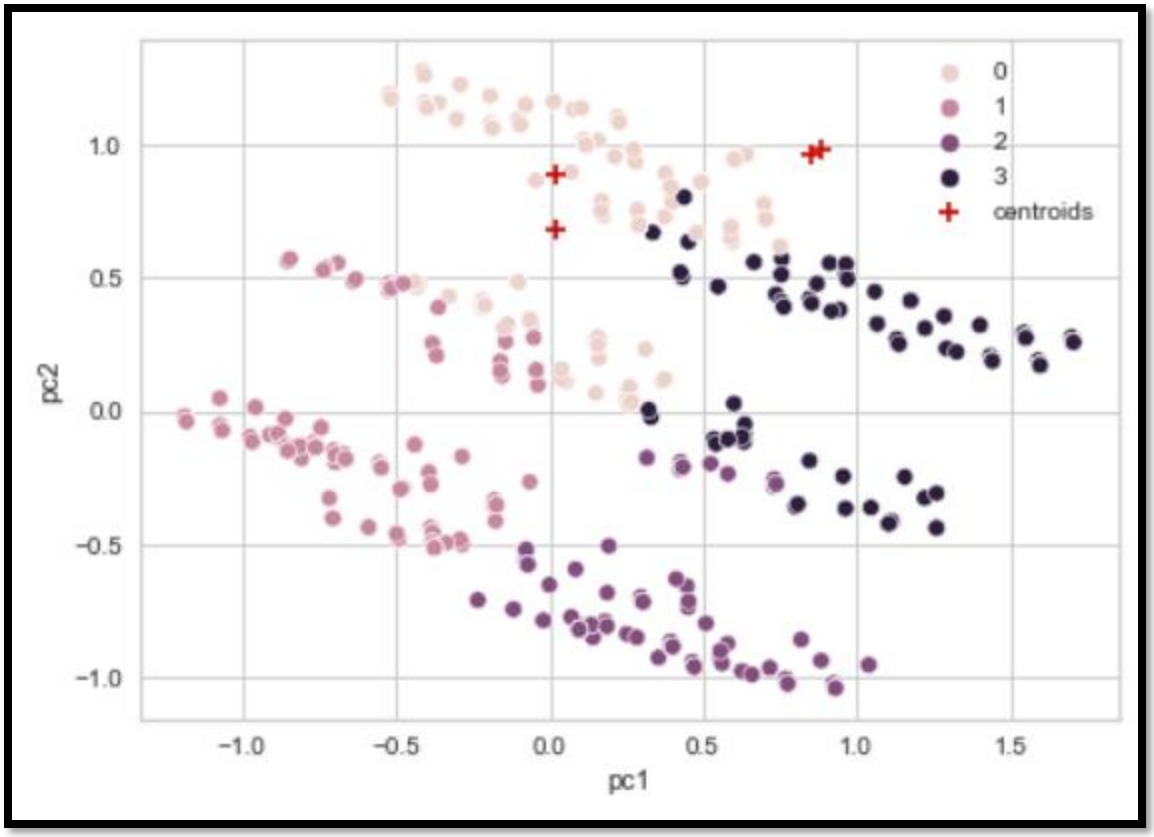
6.2 Segment Profiling with Visualizations

Neither the highly simplified, nor the very complex tabular representation typically used to present market segmentation solutions make much use of graphics, although data visualization using graphics is an integral part of statistical data analysis

Visualizations are useful in the data-driven market segmentation process to inspect, for each segmentation solution, one or more segments in detail.



Histogram for each attribute



Visualizing the cluster for pc1(yummy) and pc2(convenient)

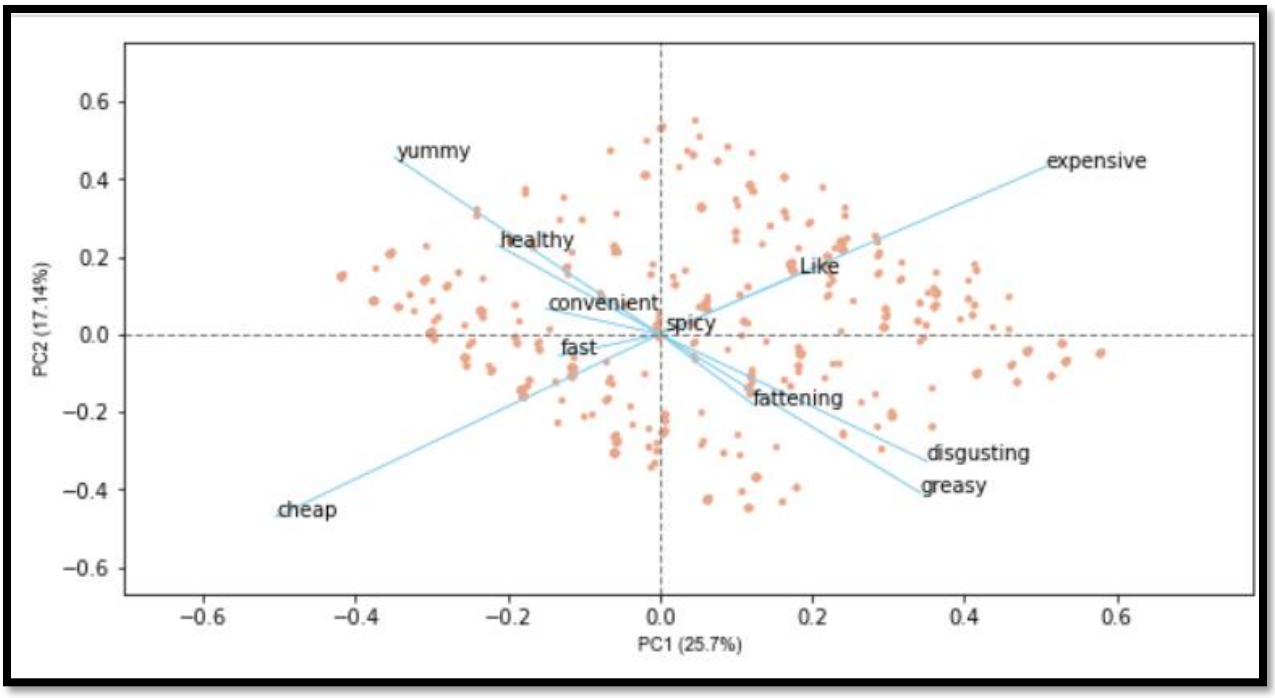
In the above visualization, the color of the small data points represents the cluster to which they have been assigned. The (+) marker are cluster centroids, with their color indicating the cluster that they represent.

first, we choose k in our case k=4, the number of clusters we want to find in the data. Then, the centers of those k clusters, called centroids, are initialized in some fashion.

Of course, there is still the problem of choosing k, the number of clusters. Usually, you don't know beforehand how many clusters the data contains, and usually you can't look at the data directly because it lies in a higher dimension than two or three. (Indeed, if you can look at your data and see obvious clusters as you can here, you may be better off clustering manually). So, in practice, people often try different values of k and see how their results vary.

1) 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.



Segment separation plot using principal components 1 and 2 for the MacDonald's data set

Due to the overlap of market segments, the plot above is messy and hard to read. The plot is still not trivial to assess, but it is easier to interpret than the segment separation plot

Each segment separation plot only visualizes one possible projection. So, for example, the fact that segments like and expensive in this particular projection overlap with other segments does not mean that these segments overlap in all projections. However, the fact that segments cheap and expensive are well-separated in this projection.

Step 7: Describing Segments

7.1 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 similar to 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.

Good descriptions of market segments are critical to gaining detailed insight into the nature of segments. In addition, segment descriptions are essential for the development of a customized marketing mix.

For example, when conducting a data-driven market segmentation analysis using McDonald's data set, profiling means investigating differences between segments concerning the like and dislike. The segment description step uses additional information, such as segment age, gender, visit frequency. These additional variables are referred to as descriptor variables.

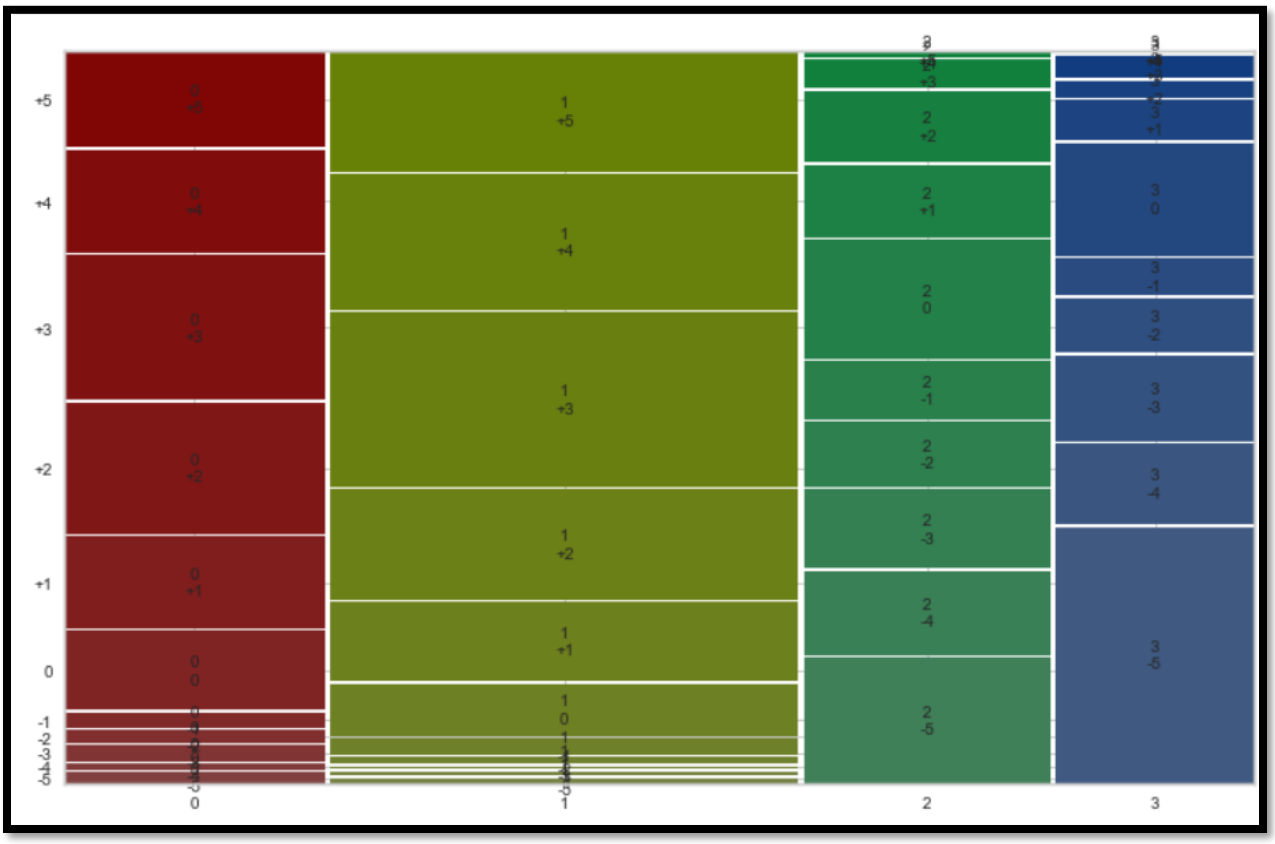
We can study differences between market segments concerning descriptor variables in two ways: we can use descriptive statistics including visualizations, or we can analyze data using inferential statistics. The marketing literature traditionally relies on statistical testing and tabular presentations of differences in descriptor variables. Visualizations make segment descriptions more user-friendly.

7.2 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, or metric descriptor variables.

The fast-food data set is not typical for data collected for market segmentation analysis because it contains very few descriptor variables. Descriptor variables – additional pieces of information about consumers – are critically important to gaining a good understanding of market segments. One descriptor variable available in the fast-food data set is the extent to which consumers love or hate McDonald’s. We cross-tabulate segment membership and the love hate variable

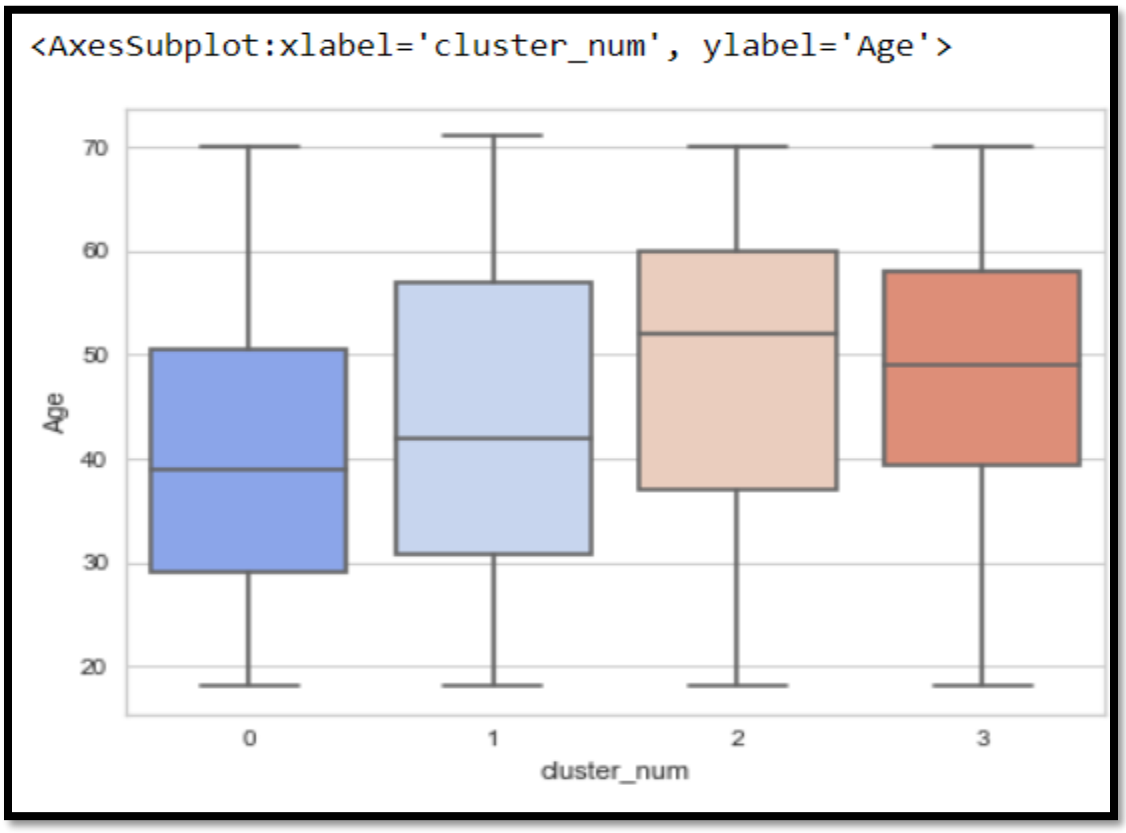
The mosaic plot segment number along the x-axis, and loving or hating McDonald's along the y-axis. The mosaic plot reveals a strong and significant association between those two variables.



mosaic plot for cross-tabulation of segment membership and I LIKE IT for the MacDonald's dataset

The fast-food data contains a few other basic descriptor variables, such as gender and age.

In a box plot, we draw a box from the first quartile to the third quartile. A vertical line goes through the box at the median.



Box plot of age by segment for the McDonald's dataset

plots segments along the x-axis, and age along the y-axis. We see immediately that the notches do not overlap, suggesting significant differences in average age across segments.

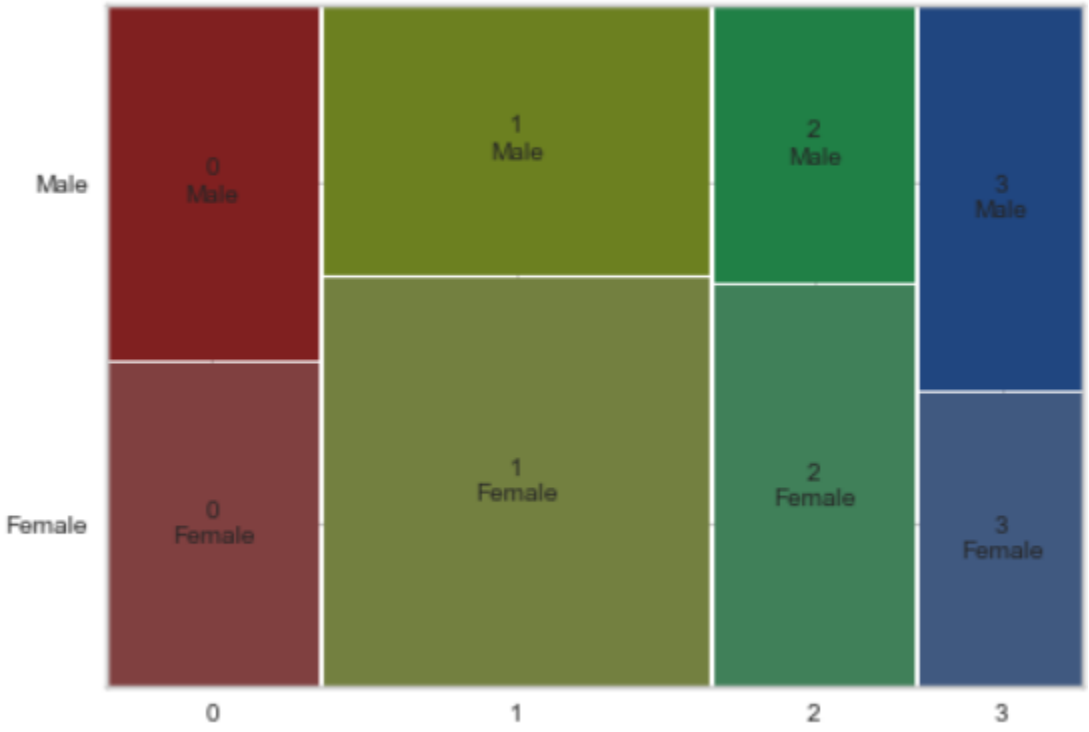
1) Nominal and Ordinal Descriptor Variables:

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.

The easiest approach to generating a cross-tabulation is to add segment membership as a categorical variable to the data frame of descriptor variables.

Gender	0	1
cluster_num		
0	151	164
1	278	190
2	203	141
3	140	164

For every cluster we see male-female are equality included. As in Shop irrespective of geography male female wont different in count considering some exceptions.



Optimally, additional descriptor variables would be available. Of particular interest would be information about product preferences, frequency of eating at a fast-food restaurant, frequency of dining out in general, hobbies and frequently used information sources (such as TV, radio, newspapers, social media). The availability of such information allows the data analyst to develop a detailed description of each market segment. A detailed description, in turn, serves as the basis for tasks conducted in Step 9 where the perfect marketing mix for the selected target segment is designed.

Step 8: Selecting the Target Segment(s)

After the implementation of algorithms or methods for clustering or grouping the consumers import step comes of targeting specific segments to work on depending upon their characteristics. These characteristics include geographic locations, cultures, psychographics properties of people in region, Behavioral characteristics feedback received from consumers, etc. Market segmentation is a strategic marketing tool. The selection of one or more target segments is a long-term decision significantly affecting the future performance of an organization. This is when the flirting and dating is over; it's time to buy a ring, pop the question, and commit.

Need and purpose of Target Segmentation:

A particular market requires particular improvements in terms of service and product. After extracting the optimal number of segments, we decide target segments among them to focus on. These segments will be those where collective action making improvement in product and services will take place. To grow business sales, we need to see consumer's tendencies in segments which we are specialized in facilitating. This helps us to focus on specific groups and make strategies for improving consumer satisfaction with profitable business returns.

Step 9: Customizing the Marketing Mix

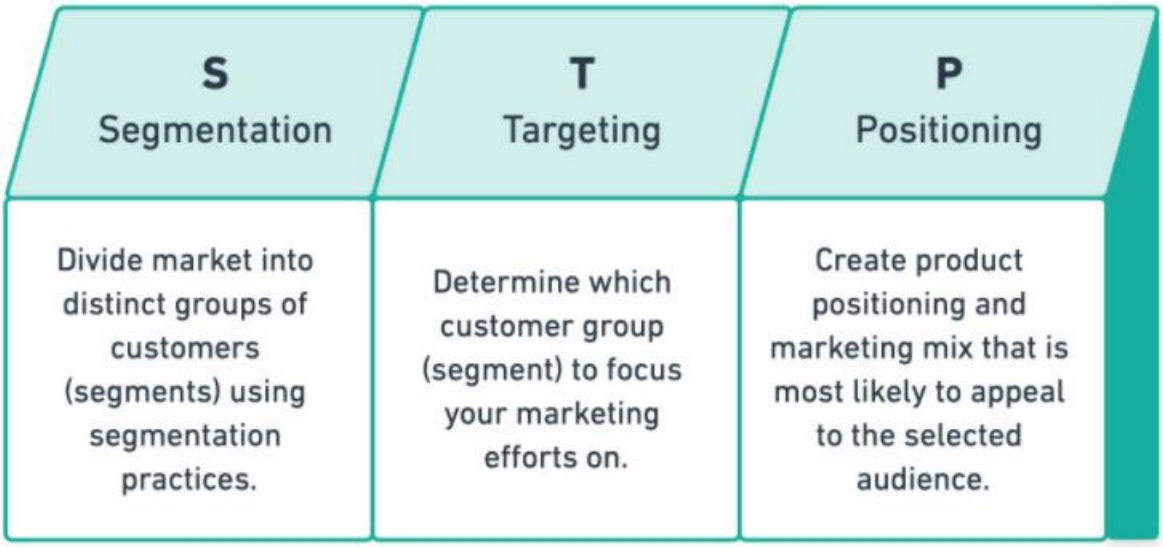
After segmenting the market, finding the target segment and positioning itself, each company needs to come up with an offer. The 5 P's used by McDonald are:

1. Product
2. Place
3. Price
4. People
5. Promotion

1. **Product:**

Product is the physical product or service offered to the consumer. Product includes certain aspects such as packaging, guarantee, looks etc. This includes both the tangible and the non- tangible aspects of the product and service.

How should the company design, manufacture the product so that it enhances the customer experience?



Customizing marketing Mix

McDonald's has intentionally kept its product depth and product width limited. McDonald's studied the behavior of the Indian customer and provided a totally different menu as compared to its international offering. It dropped ham, beef and mutton burgers from the menu. McDonald's success worldwide has been attributed to the ‘’Think Global, Act Local and Sell like a Retailer" philosophy. McDonald's followed this international mantra while opening doors to the Indian subcontinent. With respect to the cultural and traditional sentiments, not only did McDonald's not serve its most popular product The BIG MAC (a beef burger) but also developed an egg-less mayonnaise for the first time in the worldwide system.

To suit the Indian palette, the McAloo Tikki burger, Veg. Pizza McPuff and Chicken McGrill burger were among other offerings that were formulated and introduced using spices favored by Indians. Furthermore, cache restaurant kitchen was designed to maintain separate Vegetarian and Non-Vegetarian food counters.

Reasons for introducing the products

1) Beef which was a taboo in the Indian market was planned to be replaced by filling of Aloo Tikki to match up with the Indian culture.
2) Salad sandwiches were introduced targeting the people who are more health conscious about health.
3) The introduction of Jain salad in the outlet keeps the Jain people's rituals of not cutting ginger, onion, potatoes etc.
4) Seeing the Indian mentality of having wheat chapatis led to the idea to introduce paneer salsa wrap and chicken Mexican wrap.
5) Filet-o-fish was introduced keeping in mind the demand of the sea food lovers having no competitors in the segment before.

2. Place

The place mainly consists of distribution channels and outlets of the company. It is considered very important because the product must be available to the customer at the right place, at the right time and in the right quantity. In the U.S.A nearly 50% of outlets are situated within the distance of 3 minutes.

There is a certain degree of fun and happiness that McDonald's provides to its customers. It provides a value position based on the needs of the customer. McDonald's offers a proper hygienic atmosphere, good ambience and better services.

Now McDonald's have also started offering internet facilities at their outlets, along with a music system through radio, not the normal music but the music which is preferred by the younger generation in order to attract them.

There are also games for children, one example is air hockey. Children play games till their parents spend quality time in McDonald's.

3. Price:

Pricing includes the list price, the discount functions available, the financing options available etc. It should also take into consideration the probable reaction from the competitor to the pricing strategy. This is the most important part of the marketing mix as this is the only part which generates revenue. All the other three are expenses incurred. The price must take into consideration the appropriate demand-supply equation

McDonald's has certain value pricing and bundling strategies such as happy meal, combo meal, family meal etc. to increase overall sales volumes.

This pricing strategy was founded to attract middle class and lower-class people and the effect can

clearly be seen in the consumer base that McDonald's has now.

McDonald's found success in its strategy of Branded Affordability and introduced the Happy Price Menu' of Rs.20/-. Engaging and Memorable campaigns were created to establish the Branded Affordability communication.

4. **People:**

McDonald's understands the value of both its employees and its customers. It understands the fact that a happy employee can serve well and result in a happy customer.

McDonald continuously does Internal Marketing. This is important as it must precede external marketing. This includes hiring, training and motivating able employees. This way they serve customers well and the final result is a happy customer.

The level of importance has changed to be in the following order (the more important people are at the top):

1. Customers
2. Front line employees
3. Middle level managers
4. Front line managers

The punch line:

"I'm loving it" is an attempt to show that the employees are loving their work at McDonald's and will love to serve the customers.

5. **Promotion:**

The promotional activities adopted by McDonald helps to communicate efficiently with the target customers. The diagram gives ideas about the promotion strategy of McDonald's Application of above-mentioned Communication Mix describes the cost that is feasible as perthe consumers

McDonald's corporate used advertising, personal selling, sales promotion, public relations, and direct marketing and became the world's largest leading Burger Empire. These five promotion tools are used by McDonald's to integrate marketing communication programs which allows McDonald's to access the communication channels clearly, consistently and easily transfers messages and products to the target audiences

Sitting the Promotion Mix:

1. Advertising:

An Advertisement is targeted to attract the masses it reaches to a large number of people at a time. Advertising is one of the most important tools for promotion which has various ways of advertisement in that advertising through billboards and media are often used by any of the business enterprises. Consumers mostly perceive goods which are advertised goods, as they assume it is more rightful.

McDonald's also holds the hand of Advertising. There are three main objectives of advertising for McDonald's are to make people aware of an item, feel positive about it and remember it. The right message has to be communicated to the right people through the right medina. McDonald's does its promotion through television, hoardings and bus shelters. They use print ads and the television programs are also an important marketing medium for promotion.

2. Personal Selling:

Personal selling is the most effective tool for building buyers' preference, convenience and actions. Personal interaction allows knowing feedback and adjustments if required. If the organization has a good Relationship with Buyers, they are more attentive towards personal selling.

In personal selling McDonald's employees working in different outlets are the best example of personal interaction, the employees are directly serving the customers so, and the face-to-face communication is easily possible. In the McDonald's outlet there are such stuff which are appointed for personal selling they are the one who perform the activities regarding selling goods to customers.

3. Public Relations:

Highly credible; Very believable; Many forms: news stories, news features. events and sponsorships, etc.; Reaches many prospects missed by means of other forms of promotion: Dramatizes company or product; Often the most under used element in the promotional mix, relatively inexpensive (certainly not 'free' as many people think--there are costs involved)

Public Relations are also an important part of the McDonald's marketing strategy. The restaurant employees play a huge role in interacting with the public. On a day-to-day basis the employees commit themselves to customers and the customers' feelings toward the brand. McDonald's feels that before they communicate with their customers, they need to be aware of what their competitors are communicating, so they can create a beneficial difference between themselves and the competitors

4. **Direct Marketing**

Many forms: Telephone marketing, direct mail, online marketing, etc.: Four distinctive characteristics: Nonpublic, Immediate, Customized, Interactive; Well suited to highly-targeted marketing efforts.

Direct marketing is also one of the efficient tools for promotion. The McDonald's uses tool in the home delivery services in which they directly serve the order to their home. Also, they have a website which are more in preference for direct marketing in that they usually mentioned all the new offers along with the contact number of your nearby outlets.

5. **Sales Promotions:**

Sales promotion activity consists of promoting the business unit through organizing various contests, programs, functions, distribution of free discounts coupons etc. that attracts attention of the customers. Also offers strong purchase incentives, dramatizes offers, boosts sagging sales Stimulates quick response; Short-lived: Not effective at building long-term brand preferences.

McDonald's organizes several sales promoting contests and programs in different retail markets and outlets in which they distribute free discounts coupons. The statue of Mascot McDonald's is always there for any occasions that are also one of the logos of McDonald's.

Conclusion:

In the given dataset we have 15 features out of that we are assumed that 'Like' as a target feature. After the preprocessing and cleaning the data (such as applying label encoding and replacing some of the negative values). We have moved further to a clustering part of how to group customers based on their feedback. As we can see that the features yummy and tasty are more contributing to the 'Like' output. And if it is disgusting then more people are giving negative feedback. And gender has a negative correlation. So, it cannot play a major role and it is not the deciding factor.

when we are marketing from next time, we make sure that these features yummy, tasty, healthy, convenient are required for better sales. After the market segmentation analysis is completed, and all strategic and tactical marketing activities have been undertaken, the success of the market segmentation strategy has to be evaluated, and the market must be carefully monitored on a continuous basis. Changes can occur within existing market segments. But changes can also occur in the larger marketplace, for example, if new competitors enter the market. All potential sources of change have to be monitored in order to detect changes which require McDonald's management to adjust their strategic or tactical marketing in view of new market circumstances

GitHub Link: https://github.com/c7r7/Feynn-Labs