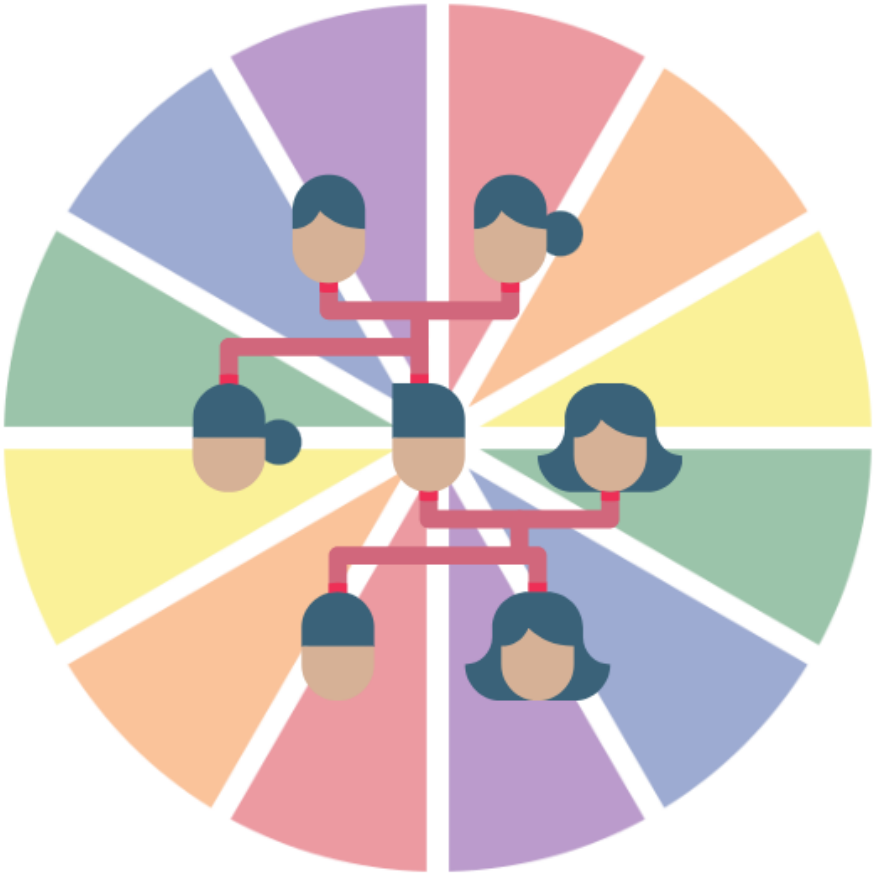
A Study on Market Segmentation

-Raima Joseph

-Feynn Labs Internship Project 2



Abstract

The following report is a detailed study of the understanding of Market Segmentation (MS).

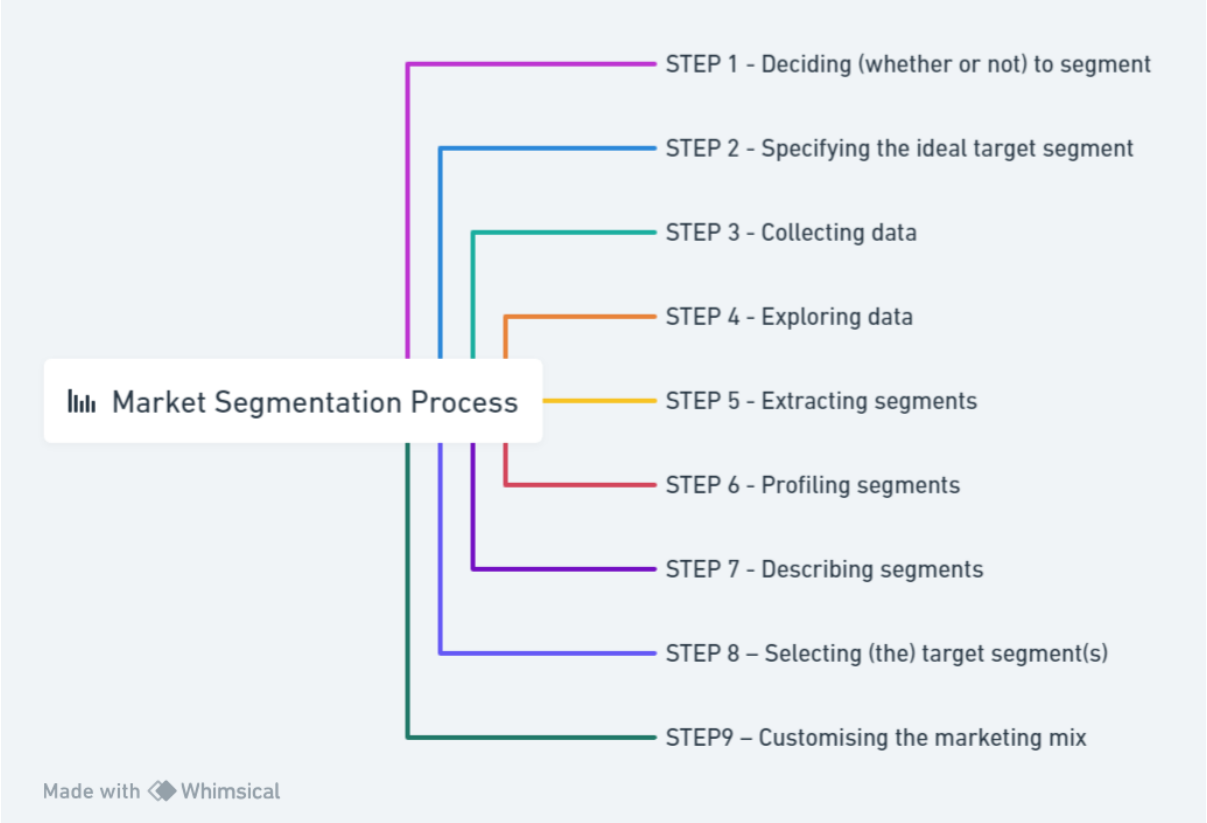
A step-wise report on what is MS and the various steps to achieve MS for a specific company. A case study is conducted using the Python language on McDonald's Market Segmentation in addition to the study.

Background

To match the needs of the consumers of a product or service, it is crucial to understand the audience that gets attracted by the product or service. This match is what makes the marketing planning process of an organisation. The simplest definition for the same is that market segmentation is the process of creating a decision-making tool for managers to divide a large audience into smaller targets for designing and deploying appropriate services. There are various kinds of segmentation that are done based on certain factors. By leveraging the power of data science and machine learning, we can better understand this segmentation and implement it in real-life cases.

MS in general has many benefits, including customer satisfaction and feedback-based improvement based on each segment. Market segmentation analysis is a task that needs a thorough understanding of the targets and deliverables of a product. There exists a ten-step approach to market segmentation analysis; from deciding whether or not to segment to evaluating each market segment. This report is meant to study the ten steps in detail and implement the same in a case study.

Here is a mindmap for the essential steps in market segmentation analysis:



These are steps which can be grouped as commonsense and data-driven segmentation and the gap between the two is bridged by using data science.

1. **Why to segment? What's the ideal target? How to collect data?**

   Step 1: Deciding
   It is not always an ideal choice to segment the market. The implications of market segmentation are crucial to understand. Certain questions that make up a checklist as barriers for an organisation can help in deciding if market segmentation is needed.
   Additionally, market segmentation is not a free-of-cost affair and the amount of research and experimental trials conducted expends time and money. Changes is the services or products offered by an organisation may change its internal and external features which in turn change the segmentation.

Step 2: Determining the ideal targets

The members that make up the segmentation team have criteria to check to determine the potential targets. The team must also determine the importance of each criterion used for distinguishing between the potential target audience. A knock-out criteria is used to determine the relevance of each criterion; including the homogeneity and size of the segments and how each segment differs from the other. The most popular approach is to use an evaluation plot to assess the attractiveness of each segment. The segmentation team is responsible for determining the values of the plot. The criteria to assess attractiveness, however, depends on the perspective of different organisations. In general, this step should result in six criteria for selecting the ideal targets.

Step 3: Collection of data.

In market segmentation, the empirical data creates the market segments. This step talks about a 'segmentation variable' which uses specific characteristics to divide a sample into segments. The general splitting of data occurs by dividing the audience on the basis of gender. 'Descriptor variables' provide extra descriptions regarding the samples of each segment, such as personal characteristics in data.

The source of this data can come from surveys, observations, experimental studies or quantitative analysis. However, before extracting or collecting data, the criteria for selecting the target (Step 3) should be mentioned.

This step also mentions the types of segmentation available:

1. Geographic segmentation: Categorizes audience on the basis of their residence location. Here, access to common media for communication is easy but it does not specifically reflect shared characteristics such as personal preferences.
2. Scio-Demographic segmentation: This kind includes criteria such as age, gender, education and income. This is more useful where preferences for products or services are highly crucial. However, the reason behind the change in consumer behaviour cannot be specified.
3. Psychographic segmentation: Clusters audience in terms of their belief or psychological preferences in lifestyle. This is more useful to get

consumer-based insights for product enhancement. The only issue is in the fact that this depends mainly on customer engagement.

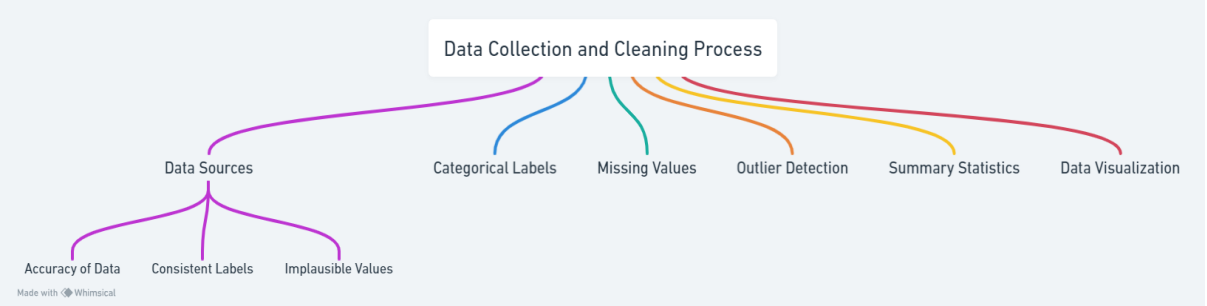
4. Behavioural segmentation: This segments the customers based on observed and actual behaviours, which includes how they reacted in the past regarding the same product.

Using these means to clearly define an objective, leads us to choosing data sources and collection methods. For a survey-based collection of data, the choice of variables and responses is vital to emphasise. For a technical case study, Python comes into play for analysis and segmentation only once we have the data collected in hand. In all cases of collection of data, the data quality and sufficiency play a key role in market segmentation.

2. How to explore and utilise the data?

1. Once the data is obtained, data cleaning is a crucial step to ensure that the data is accurate and does not have discrepancies that may generate miscalculations. We check for the correctness and ordering of the data. For more complex analysis, descriptive analysis is needed to find statistics such as minimum, maximum or mean values. These factors help in handling missing values.

   Visualisations in terms of plots and charts help in the pictorial representation of data.



2. Common pre-processing procedures include merging levels and converting them into numerics. When the scaling can be reduced to an

equal factor, the values can be transformed into numerical values. Hence, binary classification answers are easily transformed into numerical values.

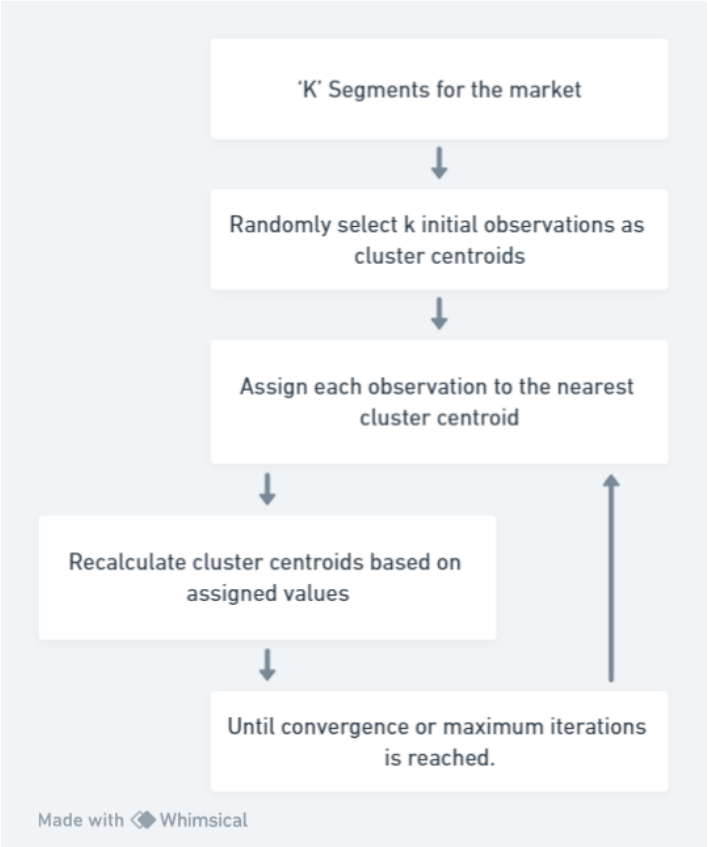
3. It is as necessary as it is common to standardise variables. Values must have a mean of 0 and a standard deviation of 1. Python offers a StandardScaler() from the Scikit-Learn module to transform the data accordingly.
4. Principal Component Analysis: A tool for simplifying complex multivariate datasets into sets of uncorrelated variables called principal]pla components. This is done to reduce dimensionality in the dataset to pre-process and visualise the data.

   These processes ensure clean and consistent data which can easily fit into a machine-learning model, or in this case, for market segmentation analysis.

3. Methods to extract market segments? How do we select which segment to target?

In the given reference material, distance-based methods have been mentioned between the habits or choices of the consumers and grouping has been done based on similarity.

K-Means Algorithm aims to divide a dataset into clusters based on similar characteristics. Here, the clusters would be market segments and the centroids would be the representatives for each segment. The algorithm tries to iteratively update the centroids such that convergence is achieved.


'K' Segments for the market
Randomly select k initial observations as cluster centroids
Assign each observation to the nearest cluster centroid
Recalculate cluster centroids based on assigned values
Until convergence or maximum iterations is reached.
Made with Whimsical

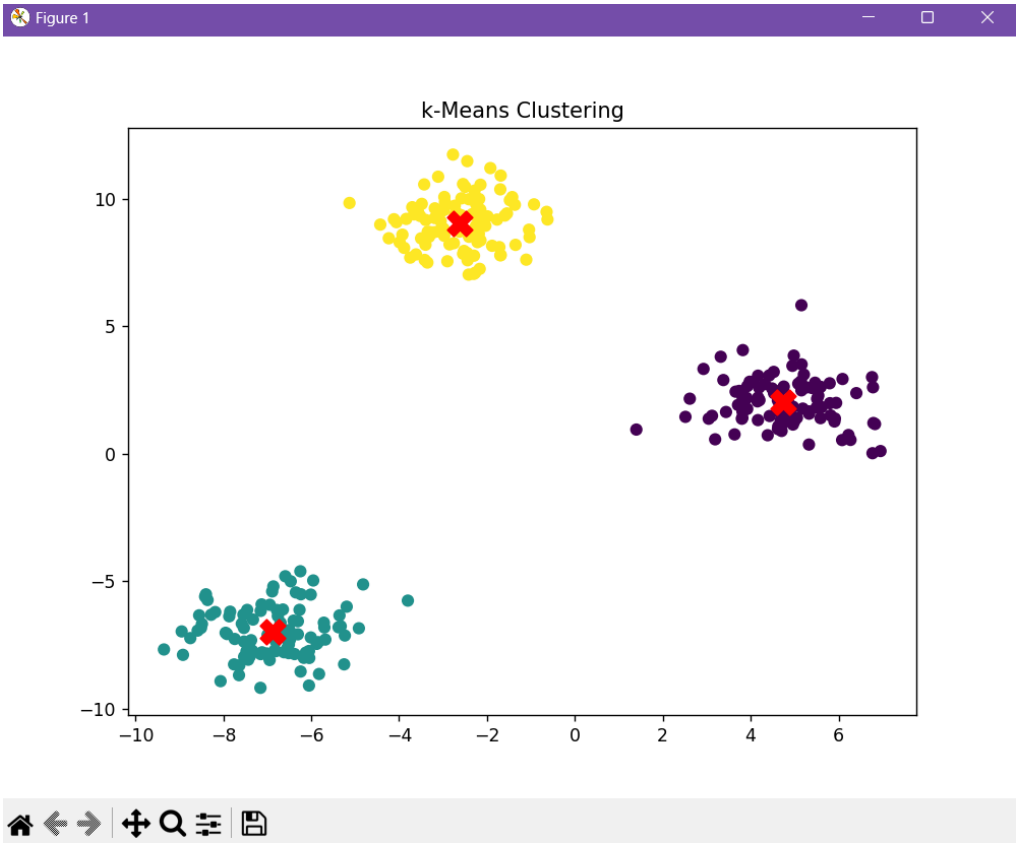
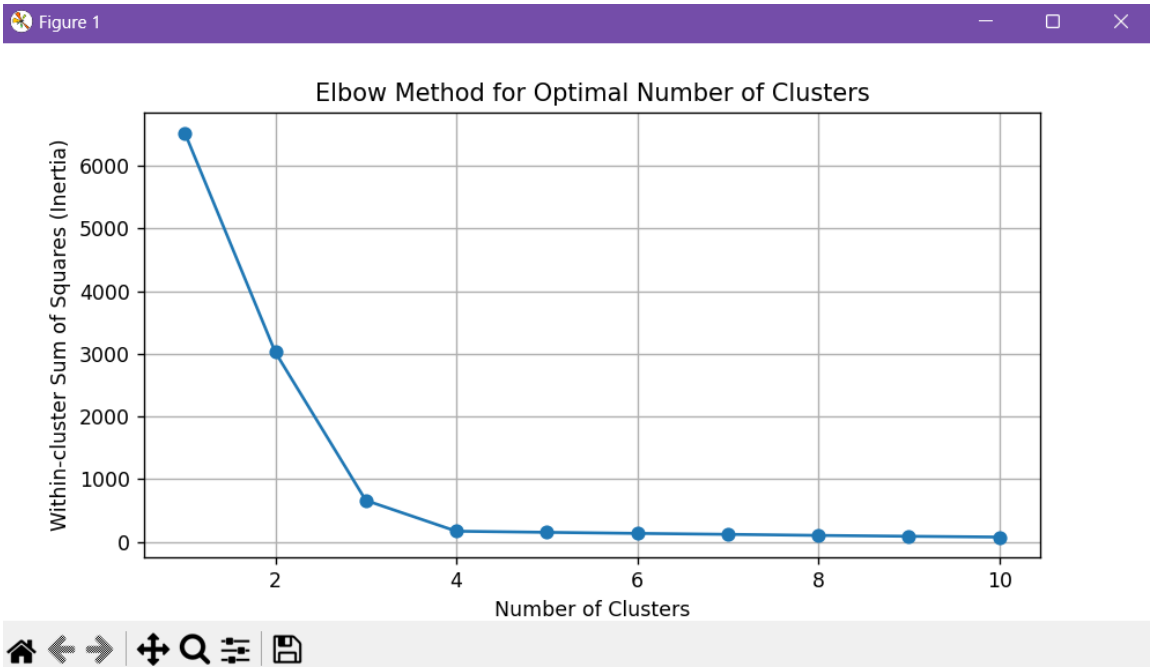


Figure 1
k-Means Clustering
10
5
0
−5
−10
−10
−8
−6
−4
−2
0
2
4
6

When it comes to selecting the number of segments, or ‘k’, a popular method called the ‘elbow method’ is employed. It works by trying various k-numbered clusters and calculating the variation for each. Plotting each K value with its variation gives an ‘elbow’ shape and the point at which the ‘joint’ occurs forming a slope gives the optimal number of clusters.



Another method is mentioned which combines k-means with hierarchical clustering. This method is done to reduce the dataset.

Once clusters have been formed, we move on to profiling. It refers to the process of identifying the crucial features of the market segments based on the segmentation variables. This provides a higher insight into the people’s behaviour and helps in comparing other segments. Visualisations, be it charts or graphical, have been recommended and used in earlier case studies to interpret segmentation analysis.

In order to select the target segment(s), knock-out criteria must be reviewed along with the attractiveness of each segment. The choice of target segment(s) is based on its ability to gain more customers and organisational competitiveness.

The target segment should result in the organization's goals, capabilities and resources.

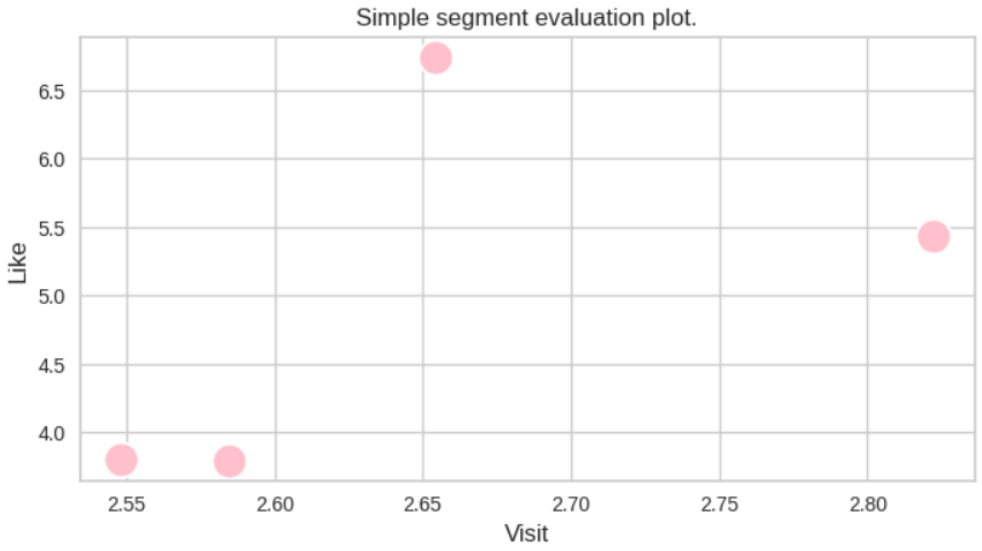
A brief review of the Case Study

In the case study conducted on the Fast Food (McDonald's) Dataset, steps 1-3 have been studied and the dataset was imported for analysis. The data was checked for preprocessing and cleaning and passed for step 4, which was EDA. Barplots, Histograms, Correlation Heatmaps and a Biplot have been covered in order to compare and contrast the relevance of each feature. Along with that, PCA was conducted to find the relevant features.

As mentioned in the report, steps have been taken to perform the case study with maximum inspiration. For step 5, segments were extracted using the above-reviewed elbow method via K-means clustering. Visualising and studying the resulting clusters was then taken place.

As for step 6, profiling of the segments for identifying crucial features in the clusters was done using bar plots and violin plots for different features. Step 7 was to describe the segments based on Gender, Like-ness and Visit count. Lastly, step 8 was conducted to select the target segments similar to the case study done in the report.

The Github link to the Project Repository is as follows:
https://github.com/raima-j/Feynn-Labs-Internship.git



This is the final segment evaluation as shown in the report and replicated for individual case study.

Conclusion

In this report, a summary of market segmentation and its fundamental steps has been drafted based on the book “Market Segmentation Analysis” and external research work. The main points captured are what is the basic definition of market segmentation and how do we technically analyse it. Market segmentation enhances the precision of targeting efforts by dividing the audience or the large mass of customers into clusters based on similarity. This can result in an improved return on investment.
In order to get these clusters, the report highlights the K-Means clustering algorithm, an unsupervised machine learning method, to categorise customers based on similarities. An elbow method is used to determine the optimal number of clusters.
An example is that various e-commerce platforms use K-means to categorise products and segment customers. In conclusion, market segmentation is a vital decision-making process which involves effective machine-learning algorithms.