Step 1: Deciding (not) to Segment

The market segmentation is very importance corresponding in any business prospective, In business , market segmentation determines who is in your target market and who is not, When utilizing market segmentation you look at all the people who could buy your product and decide how to break them up into groups that have similar needs, wants or demand characteristics, When you do this, you are able to communicate with different groups using different messages and marketing techniques. Marketers categorize their prospects into customer segments in order to keep their efforts focused and effective. When your prospects are grouped correctly, it's much easier to target specific groups and tailor your efforts for maximum impact. Below are the most common forms of segmentation.

There are some barriers that occur in market segmentation. The first group of barriers relates to senior management. Lack of leadership, pro-active championing, commitment and involvement in the market segmentation process by senior leadership undermines the success of market segmentation.

A second group of barriers relates to organisational culture. Lack of market or consumer orientation, resistance to change and new ideas, lack of creative thinking, bad communication and lack of sharing of information and insights across organisational units, short-term thinking, unwillingness to make changes and office politics have been identified as preventing the successful implementation of market segmentation.

Step 2: Specifying the Ideal Target Segment

In general, the following is true: Your target segment should be homogenous, distinct, highly legit to your marketing mix, recognizable and accessible to everyone.
In this case study, it was a natural choice to select a segment with positive perceptions of McDonald. However, management tried to understand the situation and to know about segments of concern for customers.

Step 3: Collecting Data

The dataset that was used in this report is all about McDonald's customer's feedback as per their orders. Link of the dataset
https://homepage.boku.ac.at/leisch/MSA/datasets/mcdonalds.csv

import pandas as pd
from matplotlib import pyplot as plt
import seaborn as sns
data
=pd.read_csv("https://homepage.boku.ac.at/leisch/MSA/datasets/mcdonalds.csv")
data.head()

	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

print(data.shape)
(1453, 15)

The shape of this dataset is (1453, 15). means the number of rows are 1453 and the number of columns are 15.

print(data.columns)

The names of the columns are followed by:
Index(['yummy', 'convenient', 'spicy', 'fattening', 'greasy', 'fast', 'cheap',
    'tasty', 'expensive', 'healthy', 'disgusting', 'Like', 'Age',
    'VisitFrequency', 'Gender'],
    dtype='object')

data.info()

```
<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
```

From above figure we can see that there are 15 columns, in those columns the column name "Age " only has int64 data type, rather than other columns have object datatype.

Extraction of segments is one of the main step in this case. This step is divided into 3 sub-steps.
1) using K-means analysis.
2) using mixture of binary distributions.
3) using mixture of regressive models.

1. use of K-means:
   Compute solutions for 2 to 8 segments using 10 random constraints using K-means. We don't know the optimal number of market segments in advance, so we extract 2-8 segments. After calculating a set of solutions, you can compare them and choose a solution that extracts segments with similar consumers that differ significantly from members of other segments.

   Plotted screen plot has no noticeable arcs. As the number of market segments increases, the total distance within a market segment slowly decreases. A decrease in value is expected because more market segments automatically

result in smaller segments, resulting in segment members becoming more similar profiling segments. So this is not a good way to represent segments.

A second approach to determining the appropriate number of segments is to use stability-based data structure analysis. A stability-based data structure analysis indicates whether market segments occur naturally within the data or have to be artificially constructed. Global stability is the degree to which the same segmentation solution emerges when the analysis is repeated over and over again using bootstrap samples (randomly drawn subsets) of the data.

2. using mixture of binary distributions:
    Compute the latent class analysis using a finite mixture of binary distributions. A mixed model maximizes the probability of drawing a segment. In doing so, we use the inverse minimization technique used in K-Means for the Euclidean distance. The stepFlexmix() call uses 10
    random restarts of the EM algorithm to extract 2-8 segments (k=2:8).

    Next, record information criteria using custom labels on the y-axis to select the appropriate number of segments. Plot the information criteria values AIC, BIC, and ICL on the Y-axis for different numbers of components (segments) on the X-axis. As you can see, the values of all information criteria decrease sharply until we reach the four components (market segments).

    If the information criteria are strictly applied based on the theory of statistical inference, ICL recommends the extraction of seven market segments by a small margin. BIC also points out seven market segments. AIC values continue to decline across seven market segments, indicating that at least eight components are required to properly fit the data.

3. using mixture of regressive models:
    This sub-step primarily targets a market segment characterized by a similar perception of McD as "love or hate". This segmentation approach allows McDonald's to selectively change key perceptions for specific target segments to increase love and decrease hate.

    uses finite mixtures of linear regression models, also called latent class regression, to extract such market segments. Not all variables are treated the same. Rather, you should specify dependent variables that obtain information predicted using the independent variables. Select whether consumers like or

dislike McDonald's as the dependent variable. The dependent variable contains the response to the statement 'I LIKE MCDONALDS'.

First convert the ordinal variable LIKE to a numeric variable to create a numeric dependent variable. I need a numeric variable to fit a mixture of linear regression models. The categorical variable has 11 levels, from numeric code 1 "I LOVE IT!" (+5) to numeric code 11 "I HATE IT!" (-5). You can then manually construct the model formula for the regression model by entering 11 variable names and separating them with plus signs. Alternatively, you can automate this process with Python. First concatenate the 11 independent variables into one string, separate them with plus signs, then insert the dependent variable (like N). Finally, convert the resulting string to a formula.

Step 6: Profiling Segments

8.1   Identifying Key Characteristics of Market Segments

Just doing the segmentation of the product will not solve the problem entirely. For that we have to move for Profiling segments. For example age is segmentation , and making the groups by targeting the resulting age range, like 15-25. As I explained by suitable example , analysing the data of Mcdonald neatly, making the segmentations based on all segmentation varieties like geographic, Demographic ,Psychological and Behavioural segmentation with proper justification. And as I gave an example of age just like that making the resulting profiling segments which gives the proper result in the business. But this method is not easy to interpret. Have many difficulties with resulting profile segments.

8.2   Traditional Approaches to Profiling Market Segments

Data-driven segmentation solutions are usually presented to clients or managers in one of two ways : (1) Summaries or say explanations which will give the proper idea about the segment .

(2) Large tables that provide, for each segment, exact percentages for each segmentation variable. Such tables are hard to interpret, and it is virtually impossible to get a quick overview of the key insights.

8.3 Segment Profiling with Visualisations

Visualization of the driven data is very useful to understand the data. For the inspection of the market segmentation process visualization is very useful, for each segmentation to understand. They also make it easier to assess the usefulness of a market segmentation solution. The process of segmenting data always leads to a large number of alternative solutions. Selecting one of the possible solutions is a critical decision. Visualisations of solutions assist the data analyst and user with this task.

At the heart of segmentation analysis is the extraction of complete market segments. Now we need to understand what we mean by a k-means solution with four segments. The first step in this direction is to create a segment profile chart. The segment profile chart makes it easy to see key characteristics of each market segment. It also highlights the differences between segments. The second step is to make the plot easy to interpret. Similar attributes should be placed close to each other. This is achieved by computing a hierarchical cluster analysis. And in the final step. Hierarchical cluster analysis applied to attributes identifies the most similar attributes for each attribute.

This makes it easier for McDonald's executives to interpret. We can see that there are four market segments. You can also see the size of each market segment. The smallest segment (Segment 2) accounts for his 18% of consumers and the largest segment (Segment 1) accounts for his 32%. The name of the segmentation variable (attribute) is written on the left side of the chart. Horizontal lines ending in dots indicate the percentage of respondents across the sample who associated each perception with McDonald's. Bars represent the percentage of respondents within each segment who associated each perception with McDonald's.