Team Leader: Pranav Pimpalkar

Name: Priya_

Assigned Task: STEP 8 (Selecting the Traget Segment)

Market Segmentation Analysis

(Common for All)

Step 1: Deciding (not) to Segment

Based on the "Market Segmentation Analysis" PDF, here is the detailed evaluation for deciding whether to proceed with segmentation for McDonald's:

1. **Assess the Need for Segmentation**

Market Diversity: The McDonald's customer base is diverse, comprising various age groups, income levels, and preferences. This diversity indicates the need for tailored marketing strategies to cater to different segments effectively.

Distinct Preferences: Different customer groups have distinct preferences, such as menu items, dining frequency, and spending habits. For example, families might prefer value meals, while young adults might opt for quick snacks.

2. **Evaluate Benefits and Costs**

Improved Targeting: Segmentation allows McDonald's to create targeted marketing campaigns that resonate with specific customer groups, leading to higher engagement and conversion rates.

Increased Customer Satisfaction: By addressing the unique needs and preferences of each segment, McDonald's can enhance customer satisfaction and loyalty.

Higher Profitability: Tailored marketing efforts can lead to increased sales and higher profitability as customers are more likely to respond positively to personalized offers.

Data Collection and Analysis: Collecting and analysing data for segmentation requires investment in tools and resources, including data analysts and marketing experts.

Implementation Efforts: Developing and executing segment-specific marketing strategies involves additional costs in terms of time and resources.

Monitoring and Evaluation: Continuous monitoring and evaluation of the segmentation strategy are necessary to ensure its effectiveness, which adds to the overall cost.

3. **Identify Implementation Barriers**

Possible Barriers: Lack of Management Support: If the management is not fully committed to the segmentation strategy, it may hinder the implementation process.

Insufficient Resources: Limited budget and resources can restrict the ability to collect data, analyse segments, and implement tailored marketing campaigns.

Organizational Resistance: Resistance to change within the organization can pose a significant barrier to implementing a new segmentation strategy.

4. **Make the Decision**

Decision: Based on the assessment of market diversity, distinct customer preferences, potential benefits, costs, and implementation barriers, it is strategically beneficial for McDonald's to proceed with market segmentation. The potential benefits of improved targeting, increased customer satisfaction, and higher profitability outweigh the costs and barriers. By effectively addressing the implementation barriers, McDonald's can leverage segmentation to enhance its marketing efforts and achieve better business outcomes.

Summary:

- Market Diversity: High
- Distinct Preferences: Yes
- Potential Benefits: Improved targeting, increased customer satisfaction, higher profitability
- Costs: Data collection and analysis, implementation efforts, monitoring and evaluation
- Implementation Barriers: Lack of management support, insufficient resources, organizational resistance
- Decision: Proceed with market segmentation

Step 2: Specifying the Ideal Target Segment:

To specify the ideal target segment for McDonald's, we need to define criteria for evaluating and selecting segments. Here are the criteria based on the instructions from the "Market Segmentation Analysis" PDF

1. Define Segment Evaluation Criteria

- Size: Number of customers in the segment.
- Growth Potential: Expected growth rate.
- Profitability: Potential revenue and profit.
- Accessibility: Ease of reaching the segment.
- Compatibility: Alignment with McDonald's brand and capabilities.

2. Establish Knock-Out Criteria

- Minimum Size: Must have a sufficient number of customers.
- Minimum Revenue Potential: Must have a sufficient revenue potential.

3. Identify Attractiveness Criteria

- High Spending: Higher average spending per visit.
- Frequent Visits: High visit frequency.
- Strong Loyalty: Strong brand loyalty and low churn rates.
- Low Competition: Less competition from other brands.

4. Implement a Structured Evaluation Process

- Scoring System: Rank segments based on evaluation criteria.
- Weighting: Assign importance to each criterion.
- Aggregate Scores: Calculate scores to identify promising segments

Summary on specifying the ideal Target Market:

- Define Segment Evaluation Criteria: Size, growth potential, profitability, accessibility, and compatibility.
- Establish Knock-Out Criteria: Minimum size and revenue potential.
- Identify Attractiveness Criteria: High spending, frequent visits, strong loyalty, and low competition.
- Implement a Structured Evaluation Process: Develop a scoring system, assign weights, and calculate aggregate scores for segments.

Step 3: Collecting Data:

For this project, we need to gather data from relevant sources. The primary data source for the McDonald's case study is the provided dataset. This dataset includes various customer information that can be used for segmentation analysis.

Data-Set: McDonald's dataset

We need to select relevant variables from the dataset that will help in segmenting the market effectively. Key variables include:

- Demographic information like age and gender.
- Behavioural data such as purchase history and frequency of visits.
- Psychographic data including preferences and interests.

Collect Data:

We load the dataset and prepare it for analysis. Here's how to do it in Python:

```
In [1]: import pandas as pd

        # Load the McDonald's dataset
        url = 'https://homepage.boku.ac.at/leisch/MSA/datasets/mcdonalds.csv'
        mcdonalds = pd.read_csv(url)

        # Display basic information about the dataset
        print(mcdonalds.info())
        print(mcdonalds.describe())
```

```
<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
```

To ensure the data is accurate and usable, we need to clean it. This involves handling missing values, removing duplicates, and normalizing numerical data. Then, we can any missing values by filling them with 0. We remove any duplicate rows to ensure each entry is unique. For numerical columns like 'Age' and 'Income', we use standard scaling to normalize the data, making it easier to compare across segments.

In [2]:
```
# Handle missing values by filling with 0
mcdonalds.fillna(0, inplace=True)

# Remove duplicate rows if any
mcdonalds.drop_duplicates(inplace=True)

# Normalizing numerical data, if applicable
from sklearn.preprocessing import StandardScaler

# Assuming 'Age' and 'Income' are numerical columns to be normalized
scaler = StandardScaler()
if 'Age' in mcdonalds.columns and 'Income' in mcdonalds.columns:
    mcdonalds[['Age', 'Income']] = scaler.fit_transform(mcdonalds[['Age', 'Income']])

# Save the cleaned data to a new file
mcdonalds.to_csv('cleaned_mcdonalds_data.csv', index=False)

# Display the first few rows of the cleaned data
print(mcdonalds.head())
```

```
  yummy convenient spicy fattening greasy fast cheap tasty expensive healthy  \
0    No        Yes    No       Yes     No  Yes    Yes    No       Yes      No
1   Yes        Yes    No       Yes    Yes  Yes    Yes   Yes       Yes      No
2    No        Yes   Yes       Yes    Yes  Yes     No   Yes       Yes     Yes
3   Yes        Yes    No       Yes    Yes  Yes    Yes   Yes        No      No
4    No        Yes    No       Yes    Yes  Yes    Yes    No        No     Yes

  disgusting Like  Age      VisitFrequency  Gender
0         No   -3   61  Every three months  Female
1         No   +2   51  Every three months  Female
2         No   +1   62  Every three months  Female
```

Summary of Step 3

- We load the McDonald's dataset and examine its structure.
- We clean the dataset by handling missing values, removing duplicates, and normalizing numerical data.
- The cleaned data is then saved for further analysis.

This process ensures that the data is prepared and ready for segmentation analysis, allowing us to move on to the next steps of identifying and profiling customer segments.

Assigned Task:

Step 8: Selecting the Target Segment

Selecting target segments involves evaluating and choosing the most promising segments to focus marketing efforts on. This process typically involves assessing the attractiveness and fit of each segment with the company's strategic goals.

Steps to Perform

- Evaluate Segment Attractiveness
- Create a Scoring System
- Prioritize Segments
- Select Target Segments

1. Evaluate Segment Attractiveness:

```
from scipy.cluster.hierarchy import dendrogram, linkage
from statsmodels.graphics.mosaicplot import mosaic
```

In [ ]:
```
from rpy2.robjects.packages import importr
from rpy2.robjects import Formula, pandas2ri
import rpy2.robjects as ro
import plotly.express as px
import warnings
warnings.filterwarnings('ignore')
```

In [3]:
```
# Load the McDonald's dataset from a CSV file
mcdonalds = pd.read_csv('https://homepage.boku.ac.at/leisch/MSA/datasets/mcdonalds.csv')
print(mcdonalds.columns)
```

```
Index(['yummy', 'convenient', 'spicy', 'fattening', 'greasy', 'fast', 'cheap',
       'tasty', 'expensive', 'healthy', 'disgusting', 'Like', 'Age',
       'VisitFrequency', 'Gender'],
      dtype='object')
```

In [4]:
```
# Print the dimensions of the dataset
print(mcdonalds.shape)
```

```
(1453, 15)
```

In [5]:
```
# Display the first 3 rows of the dataset
print(mcdonalds.head(3))
```

```
  yummy convenient spicy fattening greasy fast cheap tasty expensive healthy  \
0    No        Yes    No       Yes     No  Yes   Yes    No       Yes      No
1   Yes        Yes    No       Yes    Yes  Yes   Yes   Yes       Yes      No
2    No        Yes   Yes       Yes    Yes  Yes    No   Yes       Yes     Yes

  disgusting Like  Age      VisitFrequency  Gender
0         No   -3   61  Every three months  Female
1         No   +2   51  Every three months  Female
2         No   +1   62  Every three months  Female
```

2. Create a Scoring System

We created a scoring system based on the columns available. For simplicity, we focussed on VisitFrequency, Like, and Age.

Steps to Encode and Normalize Data:

- Extract Numerical Values from Like Column: Remove non-numeric characters.

- Encode Categorical Variables: Convert VisitFrequency and Gender into numerical values.
- Normalize Numerical Variables**:** Normalize the Like and Age columns to bring all features to a comparable scale.

In [73]:
```
mcdonalds['Like'] = LabelEncoder().fit_transform(mcdonalds['Like'])
Like = mcdonalds.groupby('cluster_num')['Like'].mean()
Like = Like.to_frame().reset_index()
Like
```

Out[73]:

	cluster_num	Like
0	0	5.109890
1	1	6.741667
2	2	3.868668
3	3	3.787975

In [74]:
```
#Gender
from sklearn.preprocessing import LabelEncoder
mcdonalds['Gender'] = LabelEncoder().fit_transform(mcdonalds['Gender'])
Gender = mcdonalds.groupby('cluster_num')['Gender'].mean()
Gender = Gender.to_frame().reset_index()
Gender
```

Out[74]:

	cluster_num	Gender
0	0	0.414835
1	1	0.579167
2	2	0.392120
3	3	0.525316

In [75]:
```
#Calculating the mean
#Visit frequency
mcdonalds['VisitFrequency'] = LabelEncoder().fit_transform(mcdonalds['VisitFrequency'])
visit = mcdonalds.groupby('cluster_num')['VisitFrequency'].mean()
visit = visit.to_frame().reset_index()
visit
```

Out[75]:

	cluster_num	VisitFrequency
0	0	2.730769
1	1	2.691667
2	2	2.602251

3. Prioritize Segments

I have normalized the scores and create a composite score to rank the segments.

Steps to Calculate Composite Score:

- Assign Weights to Each Criterion: Define the importance of each criterion.
- Calculate Composite Score**:** Use the weighted sum of the criteria.

4. Select Target Segments

In [75]:
```
weights = {
    'VisitFrequency': 0.4,
    'Like': 0.4,
    'Age': 0.2
}

# Calculate composite score
mcdonalds['Score'] = (
    mcdonalds['VisitFrequency'] * weights['VisitFrequency'] +
    mcdonalds['Like'] * weights['Like'] +
    mcdonalds['Age'] * weights['Age']
)

# Sort by score
mcdonalds = mcdonalds.sort_values(by='Score', ascending=False)

# Display the top segments
print(mcdonalds.head(10))
```

```
     yummy convenient spicy fattening greasy fast cheap tasty expensive  \
1409   Yes         Yes    No       Yes     No  Yes    No   Yes       Yes
1104    No          No    No       Yes    Yes  Yes    No    No       Yes
1124    No         Yes    No       Yes    Yes  Yes   Yes    No        No
643    Yes         Yes    No       Yes     No  Yes   Yes   Yes        No
226     No         Yes   Yes       Yes    Yes  Yes    No   Yes       Yes
1316    No         Yes   Yes       Yes    Yes  Yes    No   Yes       Yes
881     No         Yes    No       Yes    Yes  Yes   Yes    No        No
624    Yes         Yes    No        No     No  Yes   Yes   Yes        No
1197    No         Yes    No       Yes    Yes  Yes    No    No       Yes
522    Yes         Yes    No        No     No  Yes   Yes   Yes        No

     healthy disgusting  Like  Age  VisitFrequency  Gender  cluster_num  Score
1409      No         No    10   70               4       1            3   19.6
1104      No        Yes     9   70               5       0            1   19.6
1124      No         No     8   70               5       0            0   19.2
643       No         No     8   70               5       0            2   19.2
226       No         No     9   67               5       0            3   19.0
1316      No        Yes     9   67               5       1            1   19.0
881       No        Yes     9   67               5       1            0   19.0
624      Yes         No    10   69               3       0            2   19.0
1197      No        Yes     9   66               5       0            1   18.8
522      Yes         No    10   68               3       1            2   18.8
```

Summary

- Evaluate Segment Attractiveness: Use VisitFrequency, Like, and Age as criteria.
- Create a Scoring System: Assign weights to each criterion and calculate a composite score.
- Normalize and Calculate Scores: Ensure all features are on a comparable scale.
- Select Top Segments: Choose segments with the highest scores for targeted marketing efforts.
- This process ensures you systematically evaluate and select the most promising target segments for McDonald's using the dataset provided.

Githhub link:

https://github.com/priya99karn/Market-segment-Analysis.git