**Badakala Yaswanth  
Summarization of Case Study - Fast Food**

**14/01/2025**

***Task 1: Summarization***

***Introduction:***This case study explores market segmentation analysis through empirical data, focusing on McDonald’s brand image perceptions.

***Key Points:***

1. Purpose of Data Collection: The data was originally collected to compare the validity of different survey answer formats in brand image research.
2. Sources of Data: Descriptions of the data are referenced from studies by Dolnicar and Leisch (2012), Dolnicar and Grün (2014), and Grün and Dolnicar (2016), and sections are available in the MSA package.
3. Objective for McDonald’s: The study aims to identify consumer segments with distinct perceptions of McDonald’s, guiding targeted marketing strategies.
4. Marketing Strategy Options:
   * Focus on segments with positive perceptions to strengthen them.
   * Address segments with negative perceptions by understanding and modifying the key drivers of these negative views.

The case study demonstrates the importance of market segmentation in tailoring marketing messages and strategies based on consumer perception data.

***Summarization of Step 1: Deciding (Not) to Segment:***

***Introduction:***  
This section explores the strategic decision-making process McDonald’s faces regarding whether to pursue market segmentation or adopt a universal marketing approach.

***Key Points:***

1. ***Universal Market Approach****:*
   * McDonald’s could decide to treat the market as a homogenous entity, catering to all consumers uniformly.
   * This approach assumes that there is no significant value in understanding or addressing systematic differences in consumer perceptions or preferences.
   * It simplifies marketing efforts, as the same message and products are presented to the entire market, leveraging McDonald’s broad market power.
2. ***Segmented Market Approach****:*
   * Alternatively, McDonald’s could acknowledge that even with significant market power, there exists systematic heterogeneity among consumers.
   * This approach involves conducting detailed market segmentation analysis to uncover distinct consumer segments with varying perceptions and preferences.
   * By understanding these differences, McDonald’s can implement a differentiated marketing strategy, tailoring messages and offerings to specific segments.
   * This strategy aims to maximize market potential by addressing the unique needs and perceptions of each segment, whether by strengthening positive perceptions or addressing negative ones.

***Conclusion:***  
The choice between a universal or segmented market strategy hinges on McDonald’s assessment of the value in leveraging consumer diversity. A segmented approach could provide targeted insights and competitive advantages by customizing marketing efforts to diverse consumer needs and perceptions.

***Summarization of Step 2: Specifying the Ideal Target Segment:***

***Introduction:***This section outlines the factors McDonald’s management should consider when selecting the ideal target market segment for its differentiated marketing strategy.

***Key Points:***

1. ***Knock-out Criteria for Target Segment Selection****:*
   * **Homogeneity**: The segment members should be similar to one another in a key characteristic.
   * **Distinctiveness**: The segment must be clearly differentiated from other segments based on a key characteristic.
   * **Size**: The segment should be large enough to justify the investment in a customized marketing strategy.
   * **Fit with McDonald’s Strengths**: The target segment must be open to fast food, aligning with McDonald’s offerings.
   * **Identifiability**: There should be a way to identify members of the target segment among the general population.
   * **Reachability**: Effective communication and distribution channels must be in place to specifically target this segment.
2. ***Segment Attractiveness Criteria****:*
   * The most straightforward approach is to target a segment that already has a positive perception of McDonald’s, frequently eats out, and enjoys fast food.
   * However, McDonald’s could also choose to focus on segments that currently have negative perceptions of the brand. Understanding the reasons behind these negative views can help McDonald’s alter perceptions and grow its market share in these segments.
3. ***Attractiveness Criteria Used in This Case Study****:*
   * The data primarily reflects brand image, and McDonald’s management will use the criteria of **liking McDonald’s** and **frequent visits to McDonald’s** as key indicators for segment attractiveness.
   * These criteria are crucial for the target segment selection, as outlined in Step 8, where they will help in focusing marketing efforts on the most promising segments.

***Conclusion:***McDonald’s needs to carefully define its target segments based on homogeneity, distinctiveness, size, fit, identifiability, and reachability, while also considering segment attractiveness in terms of consumer perception and behaviour. This ensures that resources are focused on the most viable market segments.

***Summarization of Step 3: Collecting Data:***

***Introduction:***  
This section outlines the data collection process for understanding consumer perceptions of McDonald’s, focusing on various attributes.

***Key Points:***

1. **Data Set Composition**:
   * The data set includes responses from 1453 adult Australian consumers.
   * Consumers were asked about their perceptions of McDonald’s based on the following attributes: **YUMMY**, **CONVENIENT**, **SPICY**, **FATTENING**, **GREASY**, **FAST**, **CHEAP**, **TASTY**, **EXPENSIVE**, **HEALTHY**, and **DISGUSTING**.
2. ***Survey Response Format****:*
   * For each attribute, respondents provided a **YES** or **NO** response, indicating whether they believe McDonald’s possesses that attribute.
   * This binary response format helps categorize consumer opinions on these specific traits.
3. ***Demographic Information****:*
   * In addition to brand perception data, respondents provided basic demographic information, including **AGE** and **GENDER**.
4. ***Limitations and Additional Data****:*
   * The current data set lacks detailed consumer behavior information, such as dining-out habits and media consumption patterns.
   * For a more comprehensive segmentation study, additional data about these aspects would be essential to build richer and more precise consumer segments.

***Conclusion:***The data collection focuses on consumers' perceptions of McDonald’s based on specific attributes, with basic demographic details. More detailed behavioral data would enhance the segmentation analysis and help create a fuller picture of consumer segments.

**Detailed Summarization of Step 4: Exploring Data*:***

***Introduction:***  
This step focuses on exploring the McDonald’s data set to understand its structure and key characteristics, essential for segmentation analysis.

***Key Points:***

1. ***Initial Data Overview:***
   * The dataset comprises responses from **1453 adult Australian consumers**.
   * Variables inspected include consumer perceptions and demographic details such as **AGE** and **GENDER**.
   * Variables related to McDonald’s attributes are: **YUMMY**, **CONVENIENT**, **SPICY**, **FATTENING**, **GREASY**, **FAST**, **CHEAP**, **TASTY**, **EXPENSIVE**, **HEALTHY**, **DISGUSTING**, along with **VisitFrequency** and **Like**.
2. ***Sample Data Inspection:***
   * **First three rows of the dataset** showcase different combinations of YES/NO responses for each attribute.
   * Example responses include:
     + **Respondent 1**: Believes McDonald’s is **not YUMMY**, **CONVENIENT**, **not SPICY**, **FATTENING**, **not GREASY**, **FAST**, **CHEAP**, **not TASTY**, **EXPENSIVE**, **not HEALTHY**, **not DISGUSTING**; they **dislike McDonald’s** (rating of -3), aged **61**, and visit every three months.
     + **Respondent 2**: Thinks McDonald’s is **YUMMY**, **CONVENIENT**, **not SPICY**, **FATTENING**, **GREASY**, **FAST**, **CHEAP**, **TASTY**, **EXPENSIVE**, **not HEALTHY**, **not DISGUSTING**; they **like McDonald’s** (rating of +2), aged **51**, and visit every three months.
     + **Respondent 3**: Views McDonald’s as **not YUMMY**, **CONVENIENT**, **SPICY**, **FATTENING**, **GREASY**, **FAST**, **not CHEAP**, **TASTY**, **EXPENSIVE**, **HEALTHY**, **not DISGUSTING**; they **like McDonald’s** (rating of +1), aged **62**, and visit every three months.
3. ***Data Conversion for Analysis:***
   * The **YES/NO** responses are unsuitable for numeric analysis; thus, they are converted to binary format (YES = 1, NO = 0).
   * Example conversion:
     + **YUMMY**: 55% YES (1), 45% NO (0)
     + **CONVENIENT**: 91% YES, 9% NO
     + **SPICY**: 9% YES, 91% NO
     + **FATTENING**: 87% YES, 13% NO
     + **GREASY**: 53% YES, 47% NO
     + **FAST**: 90% YES, 10% NO
     + **CHEAP**: 60% YES, 40% NO
     + **TASTY**: 64% YES, 36% NO
     + **EXPENSIVE**: 36% YES, 64% NO
     + **HEALTHY**: 20% YES, 80% NO
     + **DISGUSTING**: 24% YES, 76% NO
4. ***Principal Components Analysis (PCA):***
   * PCA is conducted to reduce dimensionality and identify key perceptual dimensions.
   * **First two principal components** capture about 50% of the data variance.
   * **Perceptual Map**:
     + Attributes like **CHEAP** and **EXPENSIVE** highlight the **price dimension**.
     + Clusters of attributes indicate perceptions: **FATTENING**, **DISGUSTING**, **GREASY** contrast with **FAST**, **CONVENIENT**, **HEALTHY**, **TASTY**, **YUMMY**.
5. ***Visualization Insights:***
   * **Figure A.1** (Perceptual Map): Shows how consumers group attributes, illustrating distinct perceptions of McDonald’s based on key factors like price and general sentiment.
   * Respondents’ perceptions are distributed along the **CHEAP vs. EXPENSIVE** axis and the **positive vs. negative** sentiment axis.

***Conclusion:***The exploratory analysis reveals significant insights into consumer perceptions of McDonald’s. The conversion of qualitative data into quantitative form allows for detailed segmentation analysis, highlighting the importance of price perception and general sentiment in shaping consumer opinions. These insights are foundational for further segmentation and targeting strategies.

***Summarization of Step 5: Extracting Segments:***

***Introduction:***Step 5 focuses on extracting market segments using various statistical techniques, providing insights into distinct consumer groups.

***Key Points:***

1. ***Subdivisions of Extraction Techniques:***
   * **Using k-Means:** Standard k-means clustering is applied to identify between two to eight segments, ensuring consistency with multiple random restarts.
   * **Using Mixtures of Distributions:** Finite mixtures of binary distributions are used for segmentation, accommodating binary data structure.
   * **Using Mixtures of Regression Models:** This method leverages regression models for segmenting consumers based on their response patterns.
2. ***k-Means Analysis:***
   * The data is analyzed for a range of segment numbers, from two to eight, without a clear "elbow" in the scree plot, indicating no optimal number of segments through this method alone.
   * Stability-based analysis is suggested to ensure reliability and reproducibility of the segmentation solution.
3. ***Mixtures of Distributions:***
   * Mixtures of binary distributions are employed to model the data, capturing heterogeneity among consumers.
   * This approach is detailed with steps to fit the mixture model and assess its performance through various criteria like AIC, BIC, and ICL.
4. ***Mixtures of Regression Models:***
   * Regression models are used to identify segments, particularly useful when data involves continuous variables or covariates influencing segmentation.
   * Steps involve fitting the regression mixture model and interpreting the segments based on underlying patterns.

***1. Scree Plot (Fig. A.2):***

* **Purpose:** Displays the sum of within-cluster distances for different numbers of clusters, helping determine the optimal number of segments.
* **Observation:** No distinct "elbow" is visible, meaning there's no clear optimal number of segments. The gradual decrease suggests that additional clusters only marginally improve the segmentation​.

***2. Perceptual Map:***

* **Purpose:** Visualizes consumer perceptions along major dimensions such as price and general sentiment.
* **Observation:** Attributes like **CHEAP** and **EXPENSIVE** are crucial in evaluating McDonald’s, while other attributes cluster into positive (e.g., **FAST**, **CONVENIENT**) and negative (e.g., **FATTENING**, **DISGUSTING**) perceptions​.

***3. Cluster Separation Plot:***

* **Purpose:** Depicts how well different clusters are separated based on their data.
* **Details:** The plot includes scatter plots of observations colored by segment membership and cluster hulls indicating segment boundaries. Neighbourhood graphs display similarity between segments, with thicker lines representing greater similarity​​.

***4. Gorge Plots:***

* **Purpose:** Illustrates the distribution of similarity values within each segment.
* **Observation:** High similarity values indicate consumers close to the segment centroid, while low values suggest a wider spread within the segment. This helps in assessing the homogeneity of each segment​​.

***5. Segment Profile Plot:***

* **Purpose:** Shows how each market segment differs from the overall sample for all segmentation variables.
* **Details:** This plot helps in identifying defining characteristics of each segment, providing a visual translation of the segmentation table​​.

These graphs collectively provide insights into the data's structure, helping identify optimal segment numbers and understanding the characteristics and differences between segments. They are crucial for making informed decisions in the market segmentation process.

***Conclusion:***

elaborates on multiple methodologies to extract market segments, each providing unique insights into consumer behaviour and preferences. The techniques range from simple clustering to complex model-based methods, ensuring a comprehensive approach to segmentation.

***Detailed Explanation of Step 6: Profiling Segments:***

***Introduction:***Step 6 involves profiling the identified segments to understand their defining characteristics, using various visualizations and statistical techniques.

**Key Points:**

1. ***Purpose of Profiling:***
   * Profiling provides a deeper understanding of each segment by examining characteristics that differentiate one segment from another.
   * This step is essential for developing targeted marketing strategies.
2. ***Hierarchical Clustering (Fig. 8.1):***
   * **Purpose:** Groups segmentation variables to illustrate how closely related different attributes are within the data.
   * **Observation:** Variables that are frequently associated with similar consumer responses are clustered together, showing relationships between attributes like **TASTY**, **YUMMY**, and **FAST**.
3. ***Segment Profile Plot (Fig. 8.2):***
   * **Purpose:** Visualizes how each segment differs from the overall sample for various attributes.
   * **Details:** Each line represents a segment, showing the deviation of segment responses from the total average. Peaks and troughs indicate the attributes where the segment over- or under-performs relative to the overall population.
4. ***Eye Tracking Heat Maps (Fig. 8.3):***
   * **Purpose:** Demonstrates how different presentations of segmentation results influence user interpretation.
   * **Observation:** Heat maps show areas of focus, indicating which formats (e.g., traditional table vs. segment profile plot) are more intuitive and informative for users.
5. ***Segment Separation Plot (Fig. 8.4):***
   * **Purpose:** Shows the distinctiveness of each segment by plotting them in a reduced-dimensional space.
   * **Details:** The separation between clusters indicates how well-defined and distinct the segments are. Overlapping clusters suggest less clear differentiation.
6. ***Segment Separation Using Principal Components (Figs. 8.5 & 8.6):***
   * **Purpose:** Utilizes principal components to plot segments, enhancing the visualization of segment separation.
   * **Observation:** Clear separation along principal components shows well-defined segments, aiding in understanding which attributes drive the differences.

***Conclusion:***Profiling segments through these visualizations allows for a detailed understanding of the unique characteristics and distinctions of each segment. This detailed profiling aids in crafting precise marketing strategies tailored to the specific needs and preferences of each segment.

***Task 2: Converting the R code to Python code:***

**Summary for McDonald's Case Study Conversion from R to Python**

**1. Data Loading and Preprocessing:**

* **R:** In R, data is loaded using read.csv() and cleaned using na.omit(). Data is then preprocessed to transform categorical values ("Yes", "No") into numeric (1, 0).
* **Python Conversion:**
  + Used pandas.read\_csv() to load the dataset.
  + Categorical values ("Yes"/"No") are converted to binary values (1/0) using .astype(int).

**Python Code:**

mcdonalds = pd.read\_csv('mcdonalds.csv')

MD\_x = mcdonalds.iloc[:, :11].values # Selecting relevant columns

MD\_x = (MD\_x == "Yes").astype(int) # Binary encoding

**2. K-Means Clustering:**

* **R:** The R code applies the kmeans() function from base R, where it computes the optimal number of clusters using the "elbow" method (WCSS) and fits the data to the model.
* **Python Conversion:**
  + Used KMeans() from sklearn to perform clustering.
  + A scree plot was plotted to identify the optimal number of clusters based on WCSS.

**Python Code:**

from sklearn.cluster import KMeans

wcss = []

for i in range(1, 11):

kmeans = KMeans(n\_clusters=i, random\_state=42)

kmeans.fit(MD\_x)

wcss.append(kmeans.inertia\_)

plt.plot(range(1, 11), wcss, marker='o')

plt.title('K-means Scree Plot')

plt.xlabel('Number of Clusters')

plt.ylabel('WCSS')

plt.show()

**3. Segment Stability Evaluation:**

* **R:** Stability of segments is typically measured by silhouette scores or Adjusted Rand Index (ARI) to assess the quality and stability of clusters.
* **Python Conversion:**
  + Stability scores were calculated using the silhouette\_score function from sklearn.
  + The Adjusted Rand Index (ARI) was used as a proxy for evaluating segment stability.

**Python Code:**

from sklearn.metrics import silhouette\_score

stability\_scores = []

for i in range(2, 11):

kmeans = KMeans(n\_clusters=i, random\_state=42)

labels = kmeans.fit\_predict(MD\_x)

stability\_scores.append(silhouette\_score(MD\_x, labels))

sns.boxplot(data=[stability\_scores])

plt.title('Global Stability Boxplot')

plt.show()

**4. Visualization of Segment Stability and Distribution:**

* **R:** Uses ggplot2 or base R plotting functions to visualize clusters and their stability. Mosaic plots are generated to show relationships between segments and categorical features.
* **Python Conversion:**
  + Visualizations like silhouette score plots, and mosaic plots were created using seaborn, matplotlib, and statsmodels.
  + Mosaic plots were used to show the distribution of segments with respect to customer preferences (e.g., 'Like', 'Gender').

**Python Code:**

from statsmodels.graphics.mosaicplot import mosaic

mosaic\_data = pd.crosstab(k4\_labels, mcdonalds['Like'])

mosaic(mosaic\_data.stack(), title='Segment Number vs Like')

plt.show()

**5. Decision Tree for Segment Analysis:**

* **R:** In R, decision trees are created using the rpart package to understand the features influencing a particular cluster or segment.
* **Python Conversion:**
  + In Python, DecisionTreeClassifier from sklearn was used to train a decision tree and visualize the decision boundaries for a given segment (e.g., Segment 3).

**Python Code:**

from sklearn.tree import DecisionTreeClassifier, plot\_tree

mcdonalds['Segment\_3'] = (k4\_labels == 3).astype(int)

X = mcdonalds[['Like.n', 'Age', 'VisitFrequency', 'Gender']]

y = mcdonalds['Segment\_3']

tree = DecisionTreeClassifier(max\_depth=3)

tree.fit(X, y)

plot\_tree(tree, feature\_names=X.columns, class\_names=['Not Segment 3', 'Segment 3'], filled=True)

plt.title('Decision Tree for Segment 3')

plt.show()

**6. Code Optimization and Final Reporting:**

* **R:** Outputs are typically printed in tables or plots to show results such as the cluster centers, evaluation metrics, and segment stability.
* **Python Conversion:**
  + Python results are visualized through matplotlib, seaborn, and statsmodels. Final results like segment probabilities are calculated and displayed.

**Python Code (for Segment Probabilities):**

def segment\_probabilities(data, cluster\_labels, original\_df):

n\_clusters = len(set(cluster\_labels))

segment\_probs = []

for cluster in range(n\_clusters):

segment\_data = data[cluster\_labels == cluster]

probs = segment\_data.mean(axis=0)

segment\_probs.append(probs)

prob\_df = pd.DataFrame(segment\_probs, index=range(n\_clusters), columns=original\_df.columns[:11])

return prob\_df

kmeans\_probs = segment\_probabilities(MD\_x, k4\_labels, mcdonalds)

print(kmeans\_probs)

**Notes:**

* The steps from **data preprocessing** to **clustering** and **visualization** have been adapted from R to Python using relevant Python libraries such as pandas, sklearn, seaborn, and matplotlib.
* The focus was to maintain the overall methodology while adapting it to Python syntax and libraries.

***Code Repository for McDonald's Case Study Conversion:***

The full Python code for converting McDonald's case study from R to Python can be found on my GitHub repository:

***Google Collab link:*** <https://colab.research.google.com/gist/BadakalaYashwanth/ad062eeb212334fda1046e6a91a1cf61/clustering-project.ipynb>

***GitHub Link:*** [Clustering-Projects Repository](https://github.com/BadakalaYashwanth/Clustering-Projects) or https://github.com/BadakalaYashwanth/Clustering-Projects