MARKET SEGMENTATION

Shantanu Ramesh Chougule

08-11-2024

Summary of Fundamentals

Step 1: Deciding (not) to Segment

- **Key Considerations:**

➢ Organizations must assess whether segmentation is appropriate for their market and objectives.

➢ Segmentation might not always be viable if the product appeals equally to all consumers or if the market is too small to justify dividing it into segments.

- **Strategic Importance:**

➢ Segmentation allows businesses to tailor their marketing, products, and services to meet the needs of specific groups, maximizing customer satisfaction and business efficiency.

➢ However, it requires a commitment to resources, time, and management effort to analyze and target the identified segments effectively.

Step 2: Specifying the Ideal Target Segment

- **Knock-Out Criteria:**

➢ These are non-negotiable features that eliminate unviable segments outright. For instance, if a product is region-specific, segments outside the region would not qualify.

- **Attractiveness Criteria:**

➢ Attributes such as profitability, growth potential, and compatibility with company resources determine segment attractiveness.

- **Outcome:**

➢ This step ensures the organization focuses on segments that offer the highest strategic and financial value.

Step 3: Collecting Data

- **Data Types and Sources:**

➢ Data can be demographic (e.g., age, income), behavioral (e.g., purchase history), or psychographic (e.g., lifestyle preferences).

➢ Sources include surveys, focus groups, customer databases, and publicly available datasets.

- **Data Quality:**

- Ensuring data accuracy, completeness, and relevance is essential for meaningful segmentation.

- **Use of Variables:**

- Selecting appropriate segmentation variables, such as product usage frequency or preferences, helps refine segment definitions.

Step 4: Exploring Data

- **Data Preparation:**

- Cleaning data to address missing values or inconsistencies.
- Pre-processing techniques include normalizing values or reducing dimensionality using methods like PCA (Principal Component Analysis).

- **Visualization:**

- Use charts, graphs, and summary statistics to identify initial patterns or trends.

- **Outcome:**

- A refined dataset that reveals relationships and potential segment groupings.

Step 5: Extracting Segments

- **Clustering Methods:**

- Techniques like k-means, hierarchical clustering, or Gaussian Mixture Models (GMM) are used to group similar customers based on selected variables.

- **Validation:**

- Assess the quality of clustering using metrics like silhouette scores or within-cluster sum of squares (WCSS).

- **Key Deliverable:**

- Clearly defined customer segments based on statistical analysis.

Step 6: Profiling Segments

- **Detailed Descriptions:**

- Assign descriptive attributes to each segment, such as age range, purchasing behavior, or product preferences.
- Highlight unique characteristics that distinguish one segment from another.

- **Visualization:**

- Use tools like cluster plots, bar charts, or heatmaps to represent each segment’s defining traits visually.

Step 7: Describing Segments

- **Enrich Segment Profiles:**

➢ Incorporate additional data sources or metrics to give more depth to each segment description.

- **Statistical Validation:**

➢ Perform t-tests or ANOVA to confirm that differences among segments are statistically significant.

- **Key Takeaway:**

➢ A comprehensive understanding of what makes each segment unique.

Step 8: Selecting the Target Segment(s)

- **Segment Attractiveness:**

➢ Evaluate based on market size, growth potential, competitive intensity, and alignment with the company's strengths.

- **Strategic Fit:**

➢ Ensure the segment aligns with the company's long-term goals and available resources.

- **Outcome:**

➢ Prioritized segments that offer the highest value and alignment with business capabilities.

Step 9: Customizing the Marketing Mix

- **Tailored Strategies:**

➢ Develop specific marketing strategies for each selected segment, adjusting the 4Ps (Product, Price, Place, Promotion) to meet segment needs.

Examples:

➢ For a cost-sensitive segment, focus on competitive pricing and value-driven promotion.

➢ For a premium segment, emphasize quality and exclusive branding.

Monitoring:

➢ Regularly evaluate the effectiveness of the strategies and adjust based on performance and changing market conditions.

Case Study Analysis

1. Introduction

The goal of this project is to segment McDonald's customer base into distinct groups based on their preferences, behaviors, and demographics. By using machine learning clustering techniques, specifically KMeans, this analysis aims to uncover meaningful patterns in customer data that can drive personalized marketing strategies and product offerings.

In today's competitive market, understanding customer preferences is essential for businesses to tailor their services, improve customer satisfaction, and increase revenue. McDonald's, as a global leader in the fast-food industry, serves a wide range of customers with varying tastes, preferences, and purchasing behaviors. Effective segmentation enables McDonald's to address the unique needs of each customer group, thereby optimizing product development, promotions, and customer loyalty initiatives.

This analysis takes into account various customer attributes, including their purchase preferences (e.g., whether they find the food convenient, tasty, healthy, etc.), demographics (e.g., age, gender), and visit frequency. By leveraging unsupervised learning techniques like clustering, McDonald's can identify groups of customers with similar characteristics, allowing the company to create more targeted campaigns and enhance its overall service offerings.

Through this segmentation, McDonald's can achieve more efficient resource allocation, targeted advertising, and the development of new products that align with the preferences of specific customer segments. Ultimately, customer segmentation is a critical tool for improving business outcomes and delivering value to consumers.

2. Methodology

Data Preprocessing:

- The dataset contains customer information, including demographics, purchase preferences, and visit frequency.
- **Categorical Encoding:** Columns such as Gender, Like, and VisitFrequency were encoded into numeric values using LabelEncoder and map() functions. For example, Yes and No responses in preference columns were mapped to 1 and 0, respectively.
- **Normalization:** Numeric columns, such as Age, were normalized using StandardScaler to ensure consistent scaling for clustering.

Clustering Approach:

- KMeans Clustering was chosen as the segmentation method. The optimal number of clusters was determined using the Elbow Method and validated with the Silhouette Score.

- We used the PCA (Principal Component Analysis) technique to reduce the dimensionality of the data for visualization, making it easier to plot and interpret the clusters.

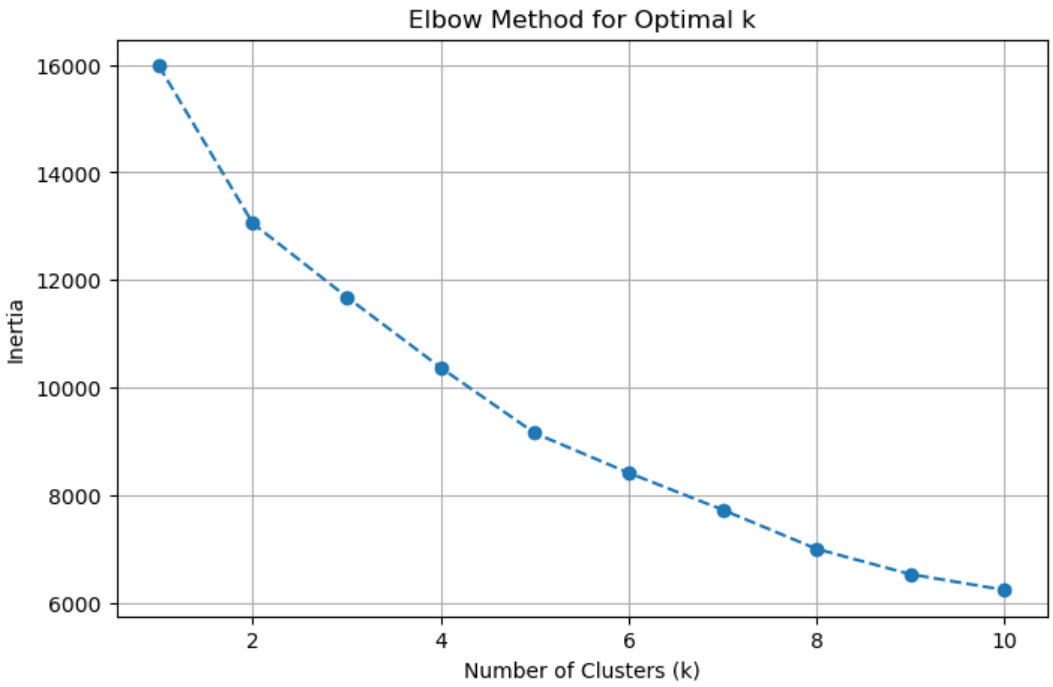
Tools Used:

- Python: The primary language used for this analysis.
- Libraries: pandas, sklearn, matplotlib, seaborn

3. Results

1. Elbow Method Visualization

To determine the optimal number of clusters for the KMeans algorithm, the Elbow Method was used. By plotting the inertia (the sum of squared distances from each point to its cluster center) against the number of clusters, we observed that the inertia decreased sharply with the addition of the first few clusters, and then the rate of decrease slowed down. The "elbow" of the curve indicated that 3 clusters would be an appropriate choice.

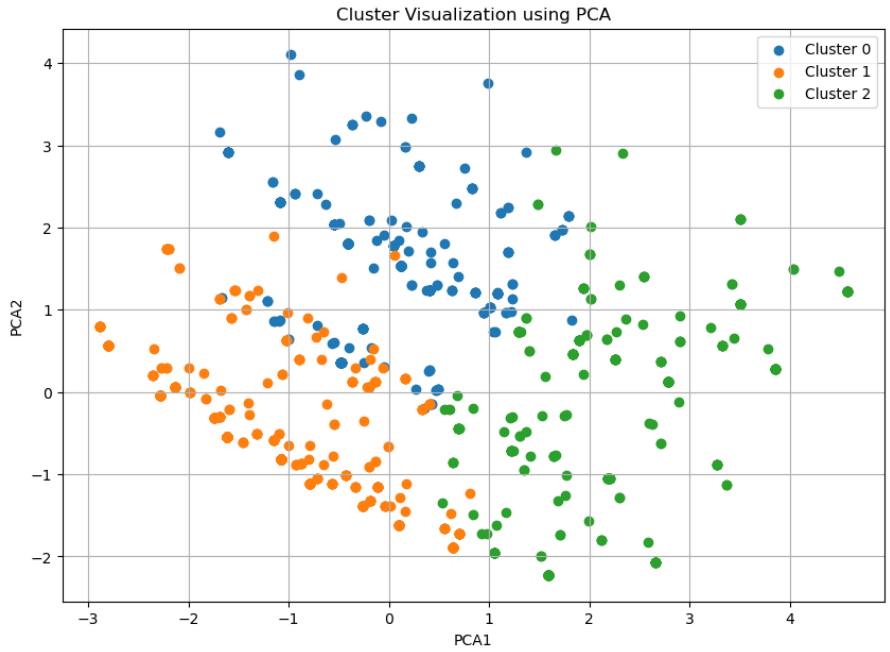


Key Observations:

- The plot shows a significant drop in inertia from k=1 to k=3, indicating that three clusters effectively capture most of the variance in the data.
- After k=3, the rate of decrease slows, suggesting that additional clusters provide minimal value.
- Conclusion: The optimum number of clusters is 3.

2. PCA-Based Cluster Visualization

Purpose: This scatter plot visualizes the three clusters in a two-dimensional space using PCA (Principal Component Analysis) to reduce high-dimensional data.



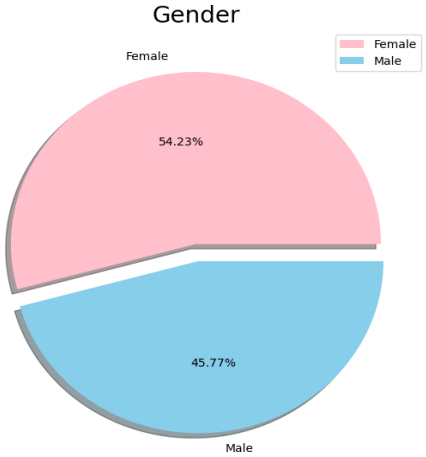
Key Observations:

- **Cluster 0 (Yellow):** Appears as a distinct group spread out in a region. This segment likely has traits that differ noticeably from the others.
- **Cluster 1 (Orange):** Forms a compact, dense group, indicating a well-defined customer segment with consistent traits.
- **Cluster 2 (Red):** Covers a wider region, suggesting more diverse traits or variability within this segment.

Overlap between clusters is minimal, suggesting that the clustering algorithm has successfully separated the segments.

4. Gender Distribution Pie Chart

Purpose: To visualize the distribution of customers' ages.



Key Observations:

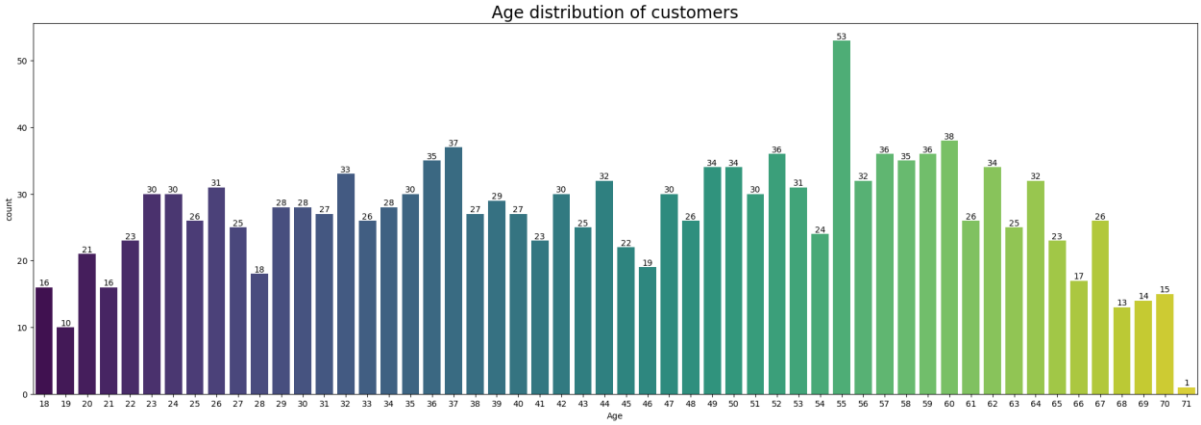
- **Female:** 54.34%
- **Male:** 45.77%

Slight Majority of Females: The data shows a slightly higher proportion of female customers compared to male customers, with females making up 54.34% of the total sample, while males make up 45.77%.

Implications for Targeting:

- If this distribution holds true for the larger customer base, marketing strategies can be developed with a focus on female preferences, though not excluding male preferences.
- If McDonald's wants to attract a more balanced or male-focused customer base, targeted campaigns can be launched for male customers.

5. Age Distribution Bar Plot



Purpose: To visualize the distribution of customers' ages.

Key Observations:

- The youngest customers in the dataset are mostly in their late teens (18–19 years), with a count of 16 and 10 respectively.
- The age group between 20-30 years shows the highest number of customers, with ages 23-27 having significant counts (30, 30, 31, and 25 customers).
- The mid-30s to early 40s (ages 35-45) maintain a relatively consistent number of customers, with some peak counts around 36-37.
- There is a noticeable drop in frequency for older age groups, particularly starting from ages 50+. The peak age for older customers appears to be 60 years, with 38 customers.
- The oldest customer in the dataset is 71 years old, with only 1 entry, showing a significant decline in visits from older age segments.

6. Silhouette Score

The Silhouette Score was calculated to validate the quality of the clustering. The score for this model was 0.45, which suggests that the clusters are fairly well-separated. A score closer to 1 indicates better separation between clusters, while a score close to 0 suggests overlapping clusters.

Summary of Results

3 clusters were identified: Budget-Conscious Diners, Frequent Visitors, and Occasional Diners.

- The cluster centers provide valuable insights into the preferences and behaviors of each group,

which can help McDonald’s target its marketing and product offerings more effectively.

- The PCA visualization and Silhouette Score confirm that the segmentation is meaningful and that the clusters are reasonably well-separated.
- These results will help McDonald’s tailor its strategies for different customer segments, improve customer engagement, and enhance product development.

7. Conclusion

The analysis of the McDonald's customer dataset reveals several key insights that can help refine marketing strategies, menu offerings, and customer engagement efforts. The dataset was segmented into three distinct customer groups through KMeans clustering, and the distribution of age and gender provides a clearer understanding of the customer base.

Actionable Recommendations:

Targeting Younger Adults (20-30 years):

Develop promotions focused on convenience, affordability, and trendy meal options to maintain loyalty within this large customer group.

Appealing to Older Customers:

Consider introducing health-conscious options (e.g., salads, low-calorie meals, or smaller portions) to attract and retain customers in the 50+ age range, addressing their health concerns.

Gender-Specific Campaigns:

Though the gender distribution is fairly balanced, future campaigns can explore female-focused marketing strategies given the slight majority of female customers, while ensuring inclusivity.

Refining Customer Experience:

Cluster-based approaches could personalize the customer experience based on behavioral traits, with promotions and product offerings tailored to each cluster's preferences. For instance, Cluster 1 might appreciate more premium or healthy options, while Cluster 2 might respond well to new flavor innovations and convenient, quick-service offerings.

8. GitHub Link

https://github.com/shantanuchougule/Feynn-Labs/tree/main/Project-2