

FEYNN LABS PROJECT - 2

‘Market Segmentation Analysis - Case Study’

Sowmiya V

Aditya Sangole

Adwait Suhas gore

Koli Nilesh Ravindra

PROBLEM STATEMENT :

McDonald's knows it has different types of customers, but it's not sure exactly how these customers view the brand. This lack of understanding makes it difficult to:

- Target the right customers with marketing campaigns.
- Craft messages that resonate with each customer group.
- Develop products that appeal to specific customer needs.

OBJECTIVE :

- **Identify Market Segments:** To uncover distinct groups of consumers who hold significantly different perceptions of McDonald's brand image. These segments will be based on factors like demographics, psychographics, or behavioral patterns.
- **Refine Marketing Strategies:** By understanding these segments and their unique brand image perceptions, McDonald's can tailor its marketing efforts and messaging to:
 - **Target specific segments:** Focus on attracting and retaining customers who already have a positive perception or address concerns of those with a negative perception.
 - **Craft targeted messaging:** Develop communication that resonates with the values and priorities of each segment, leading to a more effective connection with each customer group.

STEPS OF MARKET SEGMENTATION ANALYSIS :

STEP 1 : Deciding (not) to Segment

STEP 2 : Specifying the Ideal Target Segment

STEP 3 : Collecting Data

STEP 4 : Exploring Data

STEP 5 : Extracting Segments

STEP 6 : Profiling Segments

STEP 7 : Describing Segments

STEP 8 : Selecting (the) Target Segments

STEP 9 : Customizing the Marketing Mix

STEP 1 - DECIDING (NOT) TO SEGMENT :

SOWMIYA V

Long-Term Commitment:

- Market segmentation is a strategic decision requiring long-term commitment from leadership.
- It involves changes in products, pricing, distribution, and communication.
- Organizational structure might need adjustments to focus on segments.

Implementation Barriers:

- Lack of leadership, resources, and involvement can hinder success.
- Resistance to change, poor communication, and short-term thinking can be roadblocks.
- Absence of qualified marketing personnel or data analysts can create hurdles.
- Financial limitations or inability to make structural changes can pose challenges.
- Unclear objectives, poor planning, and time pressure can lead to suboptimal segmentation.

Mitigating Barriers:

- Proactive leadership and clear communication are crucial.
- Assessing the organizational culture for market orientation is essential.
- Training and expertise development can address knowledge gaps.

ADITYA SANGOLE

Implications of Committing to Market Segmentation

- Requires a long-term commitment and substantial changes.
- Involves significant costs for research, surveys, and marketing materials.
- Must be more profitable than not segmenting, accounting for all costs.
- May necessitate developing new products, adjusting pricing, and reorganizing the structure.
- Decisions should be made at the highest executive level and communicated across all organizational levels.

Implementation Barriers in Market Segmentation

- Lack of leadership and commitment from senior management undermines success.
- Insufficient resources for analysis and implementation impede progress.
- Resistance to change, poor communication, and office politics hinder implementation.
- Inadequate understanding of market segmentation by management and team leads to failure.
- Lack of a qualified marketing expert or data manager hinders success.
- Financial constraints and structural limitations limit success.
- Unclear objectives, poor planning, and time pressure obstruct efforts.
- Barriers should be identified and addressed proactively.

ADWAIT GORE

Step 1: Deciding (not) to segment

a) Implications of committing to market segmentation

- Long-term organizational commitment and willingness to make substantial changes
- Investment of money in market research, focus groups, field surveys and advertisements.
- Potentially required changes include change in pricing, new product development and modification of existing products, etc.
- To maximize the benefits focus should be on market segments rather than products.
- Segmentation strategy decision must be made at highest level and continuous communication with all organizational levels and across all units.

b) Implementation barriers

- Lack of leadership and involvement of senior management in segmentation process.
- Lack of creative thinking and resistance to change.
- Lack of financial resources and inability to make structural changes.
- Lack of qualified marketing experts in the organizations.
- Lack of planning or bad planning to exercise the market segmentation and having no clear objectives about it.

STEP 2 - SPECIFYING THE IDEAL TARGET SEGMENT :

SOWMIYA V

Evaluation criteria :

- Knock-Out Criteria : These are essential features a target segment needs (e.g., big enough to be worth targeting).
- Attractiveness Criteria : These help compare potential segments (e.g., how much they spend).

Benefits of Early Definition:

- Focused Data Collection: Knowing what makes a segment attractive ensures relevant data collection in the next step.
- Simplified Target Selection: Defining these criteria early streamlines choosing target segments later.

Structured Selection Process:

- Involves the segmentation team and representatives from different organizational units.
- "Segment evaluation plot" is a popular method, considering both segment attractiveness and organizational competitiveness.

ADITYA SANGOLE

Segment Evaluation Criteria

- User Involvement: User input is crucial throughout the market segmentation process.
- Conceptual Contribution: After Step 1 commitment, the organization guides subsequent steps in Step 2.
- Attractiveness Criteria: These evaluate the relative attractiveness of compliant segments.
- Literature Distinction: The literature provides various evaluation criteria.
- Team Negotiation: The team selects and assesses the importance of attractiveness criteria.
- Segment Evaluation: Attractiveness criteria determine the overall segment attractiveness in Step 8.

Knock-Out Criteria

- Homogeneity: Segments must comprise members who are similar to one another.
- Distinctiveness: Segments must be distinctly different from each other.
- Size Adequacy: Segments must be large enough to warrant customized marketing efforts.
- Organizational Alignment: Segments must match the organization's strengths and capabilities.

Attractiveness Criteria

- A variety of segment attractiveness criteria are available for the segmentation team to consider.
- Attractiveness criteria are not binary; segments are rated based on each criterion's level of attractiveness, collectively determining target segment selection in Step 8.

Implementing a Structured Process

- A structured approach, like the segment evaluation plot, assesses segment attractiveness and organizational competitiveness.
- Factors for both are determined by the segmentation team, involving representatives from diverse organizational units.
- Early development of criteria streamlines data collection and target segment selection.
- The team finalizes about six weighted criteria through negotiation and approval by the advisory committee.

ADWAIT GORE

Segment evaluation criteria

- This is a third layer of market segmentation which depends on user inputs.
- User input should be throughout the process of market segmentation analysis.

Knock-out criteria

It is used to determine if market segments identified in segmentation analysis qualify for assessment using segment attractiveness criteria.

- The segment must be homogenous.
- The segment must be distinct.
- The segment must be large enough and matching the strengths of the organization
- Members of the segment must be identifiable (possible to spot them in the marketplace).
- Segment must be reachable.

The senior management, the segmentation team, and the advisory committee must have clear idea about knock-out criteria.

Attractiveness criteria

- Attractiveness criteria are not binary in nature.
- Each market is rated (more attractive or less attractive) with respect to a specific criterion.
- These attractiveness criteria determines whether a market segment is selected as a target segment in market segmentation analysis.

Implementing a structured process

- Most popular structured approach for evaluating market segments is the use of segment evaluation plot.
- A team of two to three people is sufficient who can propose their initial solution to advisory committee which consists of representatives of all organizational units for discussion and possible modification.
- Approximately six segment attractiveness criteria have to be listed and weighted accordingly by distributing 100 points across the segmentation criteria by all team members to compare with other criteria.

STEP 3 - COLLECTING DATA :

SOWMIYA V

Commonsense Segmentation :

- Uses one variable (e.g., gender) to split the sample into segments
- Relies on pre-existing knowledge about the market
- Easier and faster to implement
- May not be as effective as data-driven segmentation

Data-Driven Segmentation :

- Uses multiple variables to identify natural or create artificial segments
- Requires data collection and analysis
- More complex and time-consuming to implement.

- Can be more effective in targeting specific customer groups.

ADITYA SANGOLE

Segmentation Variables

- Empirical data drives both commonsense and data-driven segmentation.
- Commonsense segmentation uses a single characteristic, like gender, to split the sample into segments.
- Descriptor variables, such as age and vacation preferences, detail segment characteristics.
- Descriptor variables aid in developing effective marketing strategies.
- Data-driven segmentation employs multiple variables to identify or create segments.
- Unlike commonsense segmentation, data-driven segmentation relies on several variables for segment definition.

Segmentation Criteria

- Before data collection for segment extraction, organizations must decide on segmentation criteria, such as geographic, socio-demographic, psychographic, or behavioral.
- Geographic segmentation categorizes consumers based on their location, facilitating targeted communication but may overlook other relevant characteristics.
- Socio-demographic segmentation relies on age, gender, income, and education, providing easy segmentation but limited insight into consumer behavior.
- Psychographic segmentation considers beliefs, interests, and preferences, offering deeper insights into consumer behavior but requires complex measures.
- Behavioural segmentation examines actual consumer behavior, offering direct insight into consumer preferences but may require extensive data collection efforts.

Data from Survey Studies

Market segmentation analyses often use survey data due to its cost-effectiveness and ease of collection. However, this data can be contaminated by biases, negatively impacting the quality of solutions.

- Choice of Variables: Carefully select necessary variables while avoiding unnecessary ones to maintain data quality and prevent respondent fatigue.
- Response Options: Prefer binary or metric response options over ordinal scales to ensure compatibility with statistical procedures and simplify segmentation analysis.
- Response Styles: Minimize response biases like extreme responding to prevent distortion of segmentation results and ensure accurate interpretation of market segments.
- Sample Size: Ensure a sufficient sample size (at least 100 respondents per segmentation variable) for reliable segmentation analysis and robust results.

Data from Internal Source

- Organizations utilize internal data for segmentation, like scanner data for stores or airline loyalty bookings.
- This data reflects real consumer behavior, bypassing biases like social desirability.
- It's advantageous due to automatic generation and accessibility.
- However, it may be biased and miss potential future customer trends.

Data from Experimental Studies

- Experimental data: Results from field or lab tests on consumer responses to ads.
- Choice experiments: Consumers choose preferred products based on various attribute combinations.
- The impact of each attribute helps in segmentation.

ADWAIT GORE

Segmentation variables

- Empirical data forms the basis of both commonsense and data-driven market segmentation.
- It is important to identify or create market segments and also in developing a valid segmentation solution.
- In commonsense segmentation, the segmentation variable is typically one single characteristic of that consumer, while data-driven segmentation relies on multiple variables rather than just one.

- Empirical data can come from wide range of sources such as, survey studies, scanner data, experimental studies, etc.

Segmentation criteria

The term “Segmentation variable” refers to one measured value but “Segmentation Criterion” relates to the nature of the information used for market segmentation.

- 1) **Geographic segmentation:** - In this, the consumer’s location of residence used as the only criterion to form market segments.
- 2) **Socio-Demographic segmentation:** - It includes age, gender, income and education. It can be very useful in some industries.
- 3) **Psychographic segmentation:** - When people are grouped according to psychological criteria, such as their beliefs, interests, preferences, aspirations, or benefits sought when purchasing a product.
- 4) **Behavioural:** - In this method, customer’s prior experience with the product, frequency of purchase, amount spent on purchasing the product on each occasion. This kind of information is used for market segmentation.

Data from survey studies

- 1) **Choice of Variables:** - Careful selection of variables is necessary for the better performance of the algorithm. The unnecessary variables termed as “Noisy variables” which do not contribute any information necessary for the identification correct market segments can negatively impact the segmentation solution if included.
- 2) **Response Options:** - Binary and metric data are preferable as they facilitate distance measurement essential for market segmentation.

- 3) **Response Styles:** - Common responses styles include tendencies to choose extreme options, the midpoint, or to agree with all statements which can distort segmentation results.
- 4) **Sample Size:** - Adequate sample size is required for statistical validity. Too small sample data may not give the clear picture and too large data can lead to unnecessary computation cost and data complexity.

Data from Internal Sources

- Organizations have substantial internal data for market segmentation such as scanner data, online purchase data, etc.
- This data represents actual consumer behaviour.
- A potential drawback is systematic bias by over representing existing customers and lacks information about potential future customers who may have different consumption patterns.

Data from experimental Studies: - Experimental data can result from field or laboratory experiments. Experimental data can also result from choice experiments or conjoint analyses. The aim of such studies is to present consumers with carefully developed stimuli consisting of specific levels of specific product attributes.

STEP 4: EXPLORING DATA

ADITYA SANGOLE

A First Glimpse at the Data

- Purpose: Understand the data, detect patterns, anomalies, and relationships.
- Visualization: Utilize plots (histograms, scatter plots, box plots) for visualizing distributions, correlations, and outliers.
- Pattern and Anomaly Identification: Analyze data for trends, correlations, and irregularities.
- Libraries: Pandas, Matplotlib, Seaborn

- Visualization Tools: Generate plots using Matplotlib and Seaborn.
- Correlation Analysis: Use `df.corr()` to identify correlations among features.

Data Cleaning

Objective: Ensure data accuracy and consistency through thorough cleaning.

Tasks:

- Addressing Missing Data: Determine whether to fill, discard, or flag missing values.
- Eliminating Duplicates: Identify and remove duplicate records.
- Rectifying Errors: Fix data entry inconsistencies or errors.
- Managing Outliers: Identify and potentially eliminate outliers that could skew results.

Tools: Pandas

Procedures:

- Handling Missing Data: Use `df.fillna()` or `df.dropna()` to manage missing values.
- Removing Duplicates: Apply `df.drop_duplicates()`.
- Correcting Errors: Employ column-specific functions for data rectification (e.g., `df['column'].apply(lambda x: ...)`).
- Outlier Management: Detect outliers using statistical methods and consider removal

Descriptive Analysis

- Descriptive Statistics: Using Pandas functions like `df.describe()`, `df.mean()`, and `df.median()` for basic statistical analysis.
- Providing an overview of dataset characteristics including sample size, data types, and missing values presence with tools like `df.info()` and `df.describe()`.

Univariate Analysis:

- Focuses on individual variables, highlighting central tendency (mean, median, mode), dispersion (range, variance, standard deviation), and frequency distributions.
- Visualizing variable characteristics using methods like histograms and bar charts.

Bivariate and Multivariate Analysis:

Examines relationships between variables through techniques like scatter plots, correlation matrices, and cross-tabulations.

Data Visualization:

Utilizes graphical representations like histograms, box plots, scatter plots, and heatmaps to visualize data trends, patterns, and outliers.

Handling Missing Values:

Identifying and addressing missing data through deletion, imputation, or analysis based on present data.

Correlation Analysis:

Computes correlation coefficients to determine the strength and direction of relationships between variables.

Normality Tests:

Uses statistical tests and plots, such as Q-Q plots, to assess if data distribution deviates from a normal distribution.

Pre-Processing

Purpose: To prepare data for analysis by cleaning and transforming it.

Tools Used: Pandas for data manipulation and cleaning.

Processes:

- Handling Missing Values: Identify and address missing data through deletion or imputation using functions like `df.dropna()` or `df.fillna()`.
- Removing Duplicates: Identify and eliminate duplicate records using `df.drop_duplicates()`.
- Correcting Errors: Fix data entry errors or inconsistencies by applying functions to columns using `df['column'].apply()`.
- Dealing with Outliers: Identify and potentially remove outliers that could skew results through statistical methods.

Principal Component Analysis (PCA):

Purpose: To decrease data complexity while preserving the majority of its variability.

What It Entails:

- Dimensionality Reduction: Converting data from a high-dimensional to a lower-dimensional space.
- Principal Component Identification: PCA identifies new axes (principal components) that capture the most variance.
- Data Projection: Mapping data onto these new axes, reducing the number of attributes.

Library Utilized: Scikit-learn

Procedures:

- PCA Implementation: Employing PCA from Scikit-learn.
- PCA Fitting: Applying PCA to the dataset.
- Data Transformation: Projecting the data onto principal components.
- Variance Examination: Analyzing the explained variance.

ADWAIT GORE

1) A first glimpse at the data

- The Exploratory data analysis and data cleaning and pre-processing are important to analyse the trends in the data.

- To understand the trends, we use python libraries like pandas for data cleaning purpose and matplotlib, seaborn for visualization purpose.

2) Data Cleaning

- The first step before doing data analysis is to clean the given data.
- This includes check for null values, duplicate values, using measures of central measures (Mean, Mode, Median) to replace null values, etc.
- Visualize the data to check the outliers [using python command `df. plot (kind = box)`].

3) Descriptive Analysis

- Descriptive numeric and graphic representation provide insights into the data.
- To obtain numeric summary in python use command `df. describe ()`. This gives you the five number summary of data which includes [max, min, Q1(25 percentile), Q2=Median (50% percentile), Q3 (75 percentile), count, standard deviation].
- Histogram, boxplot and scatter plot of numeric data helps to visualize the data and to understand the trends in the data.
- Histogram shows how often observations within a certain value range occur. It also shows that whether data is skewed or normally distributed.
- Boxplot gives the outliers present in the data.

4) Pre-processing:

a) Categorical Variables

- Two pre-processing procedures are often used for categorical variables.
 - 1) merging of categorical levels before further analysis.
 - 2) converting the categorical variables to numeric one. This is done by two methods I) using label encoder II) one hot encoding
 - 3) These encoded columns are now used for further analysis.

b) Numeric Variables

- The range of values of a segmentation variable affects its relative influence in distance-based methods of segment extraction.
- To balance the influence of segmentation variables on segmentation results, variables can be standardized.
- Standardizing variables means transforming them in a way that puts them on a common scale.
- In python's sci-kit learn library there are two scalers 1) MinMaxScaler 2) StandardScaler
- MinMaxScaler converts the data between minimum and maximum value of that particular column that needs to be standardized. On the other hand, StandardScaler converts that numerical column data between 0 and 1.

5) Principal Component Analysis

- It transforms the multivariate data set containing metric variables to a new data set with new variables referred to as principal components.
- These principal components are uncorrelated and ordered by importance.
- Principal components analysis basically keeps the data space unchanged, but looks at it from a different angle.
- Works off the covariance or correlation matrix of several numeric variables.
- If the data ranges are different, the correlation matrix should be used (equivalent to standardizing the data)
- PCA typically used to project high-dimensional data into lower dimension for plotting purposes.

STEP 5: EXTRACTING SEGMENTS

ADITYA SANGOLE

1. Grouping Consumers

In the context of market segmentation, grouping customers entails splitting a bigger market into more homogenous, smaller groups according to specific attributes. These attributes could include behaviors, psychographics, demographics, or other pertinent elements. The objective is to recognize discrete groupings.

of customers who have comparable requirements, tastes, and habits in order to customize marketing tactics for each group.

Businesses may create more relevant and successful marketing strategies for each sector of the consumer base after they have been divided into groups. Enhancing customer happiness and brand loyalty might result from this, as can gaining a deeper comprehension of the distinct requirements and inclinations of different consumer segments.

2. Distance-Based Methods

A set of approaches known as distance-based methods is employed in data analysis and clustering to quantify how similar or different two items are. These techniques are essential to clustering algorithms, which group things that are similar together.

These distance measurements are essential for many different applications, such as similarity-based searches, clustering, and classification. The type of data and the particular needs of the investigation determine which distance metric is best. The effectiveness of the analysis depends on the choice of distance measure, since various metrics may produce different findings.

2.1.Distance Measures

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}$$

Every column in this example represents a variable (a holiday activity), and every

row an observation (a tourist in this case). This may be mathematically represented as a $n \times p$ matrix, where p is the number of variables and n is the number of observations (rows).

The vector corresponding to the i -th row of matrix X is denoted as $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})$ in the following, such that $X = \{x_1, x_2, \dots, x_p\}$ is the set of all observations. In the example above, Anna's vacation activity profile is vector $x_1 = (100, 0, 0)$ and Tom's vacation activity profile is vector $x_7 = (50, 20, 30)$

2.2.Hierarchical Methods

Hierarchical clustering algorithms offer a natural approach to organizing data by simulating how individuals would group a set of observations into distinct segments. In the realm of market segmentation analysis, these algorithms occupy a middle ground between two extremes. Initially, the entire dataset, X

X , is divided into two market segments using divisive hierarchical clustering methods. Subsequently, each segment is further subdivided into two, and this process continues until every customer has their own dedicated market sector.

On the other hand, agglomerative hierarchical clustering approaches the task from a different angle. Initially, each customer represents a unique market segment, forming n

n singleton clusters. Then, through a step-by-step process, the algorithm merges the two closest market sectors until a comprehensive segmentation is achieved.

2.3.Partitioning Method

Hierarchical clustering techniques are notably well-suited for examining small datasets comprising up to several hundred observations. Conversely, for datasets surpassing 1000 observations (consumers), clustering methods that establish a single partition are more appropriate than hierarchical methods involving nested partitions. This entails avoiding the computation of all distances between pairs of observations at the outset of a hierarchical partitioning cluster analysis using a standard implementation.

On the other hand, a partitioning clustering algorithm tailored to identify five market segments would only need to compute between 5 and 5000 distances during each step of the iterative or stepwise process, depending on the algorithm utilized. Furthermore, when aiming to extract only a few segments, it is preferable to optimize specifically for that purpose rather than constructing the complete dendrogram and subsequently segmenting it heuristically.

2.4. Hybrid Approaches

Various methodologies combine hierarchical and partitioning algorithms in an effort to offset the drawbacks of one approach with the advantages of the other. While hierarchical cluster algorithms excel in determining the number of market segments, their major drawback lies in their substantial memory requirements, which limits their applicability to large datasets.

Hybrid segmentation strategies adopt a two-step process. Initially, a partitioning algorithm is employed due to its capability to handle datasets of any size. However, instead of generating the desired number of segments, this algorithm produces a considerably larger number. Subsequently, the original dataset is discarded, retaining only the centroids (representatives of each market segment) and segment sizes. These are then utilized as input for hierarchical cluster analysis. This process reduces the dataset to a manageable size for hierarchical algorithms, allowing the dendrogram to guide the determination of the optimal number of segments.

3. Model-Based Methods

Model-based techniques for segment extraction involve clustering methods that entail fitting probabilistic models to the dataset. Unlike distance-based approaches, which focus on assessing similarity or dissimilarity between data points, model-based methods operate under the assumption that the data originates from a specific underlying probabilistic model. These techniques seek to determine the parameters of the model that best describe the observed data and, in the context of clustering, identify distinct segments or clusters within the dataset.

One widely used model-based clustering technique is the Gaussian Mixture Model (GMM). GMM assumes that the data is generated from a combination of multiple Gaussian distributions. Each Gaussian distribution represents a cluster, and the model estimates the parameters of these distributions, including mean, covariance, and mixing coefficients.

Model-based clustering approaches offer a versatile framework for segment extraction, accommodating more intricate data distributions compared to certain distance-based methods. It is essential to comprehend the assumptions of the selected model and validate the outcomes based on the data's characteristics and the analysis objectives.

4. Algorithms with Integrated Variable Selection

These algorithms operate under the assumption that each segmentation variable contributes to determining the segmentation solution. However, there are instances where segmentation variables are inadequately chosen and may contain redundant or noisy attributes. Preprocessing techniques can help in identifying such variables.

Selecting variables for binary data poses a greater challenge because individual variables may not provide sufficient information for clustering, making it impractical to pre-screen or filter them individually.

In cases where segmentation variables are binary and redundant or noisy variables cannot be identified and eliminated during data preprocessing, it becomes necessary to identify suitable segmentation variables during segment extraction. Several algorithms are designed to extract segments while concurrently selecting appropriate segmentation variables.

5. Data Structure Analysis

Extracting market segments involves an inherently exploratory process, regardless of the segmentation algorithm employed. Traditional validation, aiming for a clear optimality criterion, proves unattainable in this context. Ideally, validation should entail testing various segmentation solutions, targeting different segments, and evaluating their profitability or success in achieving objectives.

Validation in market segmentation typically involves assessing the reliability or stability of solutions through repeated calculations with slight modifications to the algorithm or the data. This approach differs significantly from validation using an external criterion. Throughout this book, we refer to this validation approach as stability-based data structure analysis.

Data structure analysis offers valuable insights into the characteristics of the data, guiding subsequent methodological decisions. Importantly, stability-based data structure analysis indicates whether the data contains natural, distinct, and well-separated market segments.

If such segments exist, they can be easily identified. However, if no clear structure is evident, users and analysts must explore numerous alternative solutions to identify the most relevant segment(s) for the organization. Additionally, if the data exhibits any form of structure, data structure analysis can assist in determining an appropriate number of segments to extract.

STEP - 6: PROFILING SEGMENTS

NILESH KOLI

Identifying Key Characteristics Of Market Segments

- The aim of the profiling step is to get to know the market segment results from the extraction step.
- Profiling is only required when data-driven market segmentation is used.

Example: Age is used as the segmentation variable for the common-sense segmentation, it is obvious that the resulting segment will be age group.

- The situation is quite different in the case of data-driven segmentation: Users of the segmentation solution may have decided to extract segments on the basis of benefits sought by consumers
- Yet until the data has been analyzed the defining characteristics of the resulting market segment are unknown

Identifying these defining characteristics of market segment with respect to the segmentation variable is the aim of profiling

Profiling consists of characterizing the market segments individually, but also in comparison to the other market segment. If asked about their activities, operational activities and market activities are such as

1. Food preparations
2. Customer service
3. Market and supply chain management

At the profiling stage, we inspect a number of alternative market segmentation solutions. This is particularly important if no natural segment exists in data, and either a reproducible or a constructive market segmentation approach has to be taken. Good profiling is the basis for correcting interpretation of the resulting segments. Correct interpretation, in turn, is critical to making good strategic marketing decisions.

- Data-driven market segmentation solutions are not easy to interpret. Managers have difficulties interpreting segmentation results correctly and 70% feel that segmentation analysis is like a black box.

STEP 7: DESCRIBING SEGMENTS

ADWAIT GORE

a) Developing a complete picture of Market Segments

- Segment profiling involves understanding the differences in segmentation variables across market segments.
- Descriptive statistics, visualization, inferential statistics are some methods to understand these differences.

b) Using visualizations to describe market segments

- A wide range of charts are available to visualize the descriptor variable differences.

- Visualization simplifies the interpretation for both data analyst and the user.
- People process graphical data easily as compared to tabular data.

i) Nominal and Ordinal Descriptor Variables

- Use of descriptor variables are required to describe the market and effective segmentation.
- The use of segment plot and segment number will effectively project the different variables that are influencing the segment.

ii) Metric Descriptor Variables

- Use of metric variables with continuous values and trends of the variables is key to segmentation and its success.
- Conditional plots are well-suited for visualising differences between market segments using metric descriptor variables.

c) Testing for Segment Differences in Descriptor Variables

The use of statistical methods and tests are required for checking the success of the segmentation. The testing of model success using f1 score and R2 score will also be useful for successful segmentation.

d) Predicting Segments from Descriptor Variables

- To learn market segments, regression model is used with segment variables as categorical dependent variable and descriptor variables as dependent variables.
- This is done by using supervised machine learning models.
- Regression analysis is the basis of model prediction.

1) Binary Logistic Regression

- This model is used when data has binary descriptor variables.

2) Multinomial Logistic Regression

- Multinomial logistic regression can fit a model that predicts each segment simultaneously. Because segment extraction typically results in more than two market segments, the dependent variable y is not binary.
- The regression coefficients are arranged in matrix form.

3) Tree Based Methods

- Classification and regression trees are a supervised learning technique from machine learning.
- The advantages of classification and regression trees are their ability to perform variable selection, ease of interpretation supported by visualizations, and the straight-forward incorporation of interaction effects.
- The tree approach uses a stepwise procedure to fit the model.
- Because of this stepwise splitting procedure, the classification and regression tree approach are also referred to as recursive partitioning.

STEP 8 - SELECTING THE SEGMENTS:

SOWMIYA V

Selecting the right target market is crucial. This process involves refining potential customers through segmentation and then choosing the most promising segment(s). Here's a concise overview:

- Refine your options: Ensure remaining segments are good fits after an initial screening.
- Consider two sides: How attractive is the segment to you (size, growth), and how likely are they to choose you (brand fit, etc.)?
- Visualize with a matrix: Plot these two factors for each segment. Bubble size can show additional info like profit potential.
- Data drives decisions: Predefined criteria determine segment attractiveness, while competitor analysis helps with competitiveness.
- Choose wisely: Favor segments that are both attractive and receptive. Reconsider those that seem profitable but lack appeal.
- Team decision: Discuss and decide as a team, considering factors like profit and ease of catering to each segment.

STEP 9 - CUSTOMIZING THE MARKETING MIX :

SOWMIYA V

The marketing mix, traditionally represented by the 4Ps (Product, Price, Place, and Promotion), offers a framework for key factors you can adjust to influence your target market. Here's how to customize each P for targeted impact:

- Product:
 - Tailor existing offerings to better suit segment needs (e.g., a cultural heritage destination might modify tours for history buffs).
 - Develop entirely new products that cater directly to a segment (e.g., a "Family Fun Package" for a segment with young children).
- Price:
 - Adjust pricing strategies based on price sensitivity or willingness to pay a premium (e.g., premium pricing for a segment with higher vacation expenditures).

- Offer discounts or promotions targeted towards specific segments (e.g., student discounts for an education-focused segment).
- Place:
 - Choose distribution channels frequented by the segment (e.g., online booking for a tech-savvy segment, physical stores for those who prefer in-person interaction).
 - Make your product accessible through channels convenient for the segment (e.g., selling hiking gear near popular trails).
- Promotion:
 - Tailor communication messages and channels to resonate with segment interests and media consumption habits (e.g., social media marketing for a younger segment, travel magazines for a segment planning luxury vacations).

CODE CONVERSION TASK :

Market Segmentation Analysis

SOWMIYA V

Link - https://github.com/sowmi234/FeynnLabs-Internship_Task/tree/main/Team%20Project1

ADITYA SANGOLE

Link - <https://github.com/AdityaSangole/Feynn-Labs-Internship-Task/tree/main/Task%2002>

ADWAIT GORE

Link - https://github.com/Adwaitgore2520/Feynn_Labs_Internship

