Customer Segmentation Using Data Science

# Problem Statement ;

In today's highly competitive business landscape, understanding customers and tailoring your marketing strategies to their needs is crucial for success. Our Project aims to improve customer engagement, increase customer satisfaction, and boost profitability by creating meaningful customer segments. By segmenting customer base, we can deliver more personalized experiences and targeted marketing campaigns.

# Key Objectives:

1. Enhance Marketing Effectiveness: Target customers with tailored marketing campaigns to improve conversion rates and ROI.
2. Improve Customer Retention: Identify and engage with high-value customers to reduce churn.
3. Optimize Product Offerings: Understand customer preferences to refine product and service offerings.
4. Increase Revenue: Maximize the lifetime value of customers by providing them with what they need.

# Design Thinking Process Empathize

In this phase, we will gather data and insights to understand our customers better. This includes analyzing historical customer data, conducting surveys, and collecting feedback through various channels. The goal is to empathize with the customer's perspective.

# Define

Based on the insights gained, we will define specific customer segments that align with our business goals. These segments could be based on demographics, behavior, purchase history, or any other relevant criteria.

# Ideate

In this phase, we will brainstorm and ideate potential solutions for segmenting customers. We will explore various data-driven approaches, including clustering algorithms, machine learning models, and statistical techniques.

# Prototype

We will create prototype customer segments and test them against historical data. This will involve selecting and implementing the most appropriate data science tools and techniques for segmentation.

# Phases of Development

**Phase 1: Data Collection and Preprocessing**

Gather historical customer data, including demographics, purchase history, website interactions, and customer support interactions. Clean and preprocess the data to remove duplicates, handle missing values, and ensure data quality.

# Phase 2: Exploratory Data Analysis (EDA)

Conduct EDA to identify patterns, trends, and outliers in the data. Visualize customer characteristics and behaviors to gain insights.

# Phase 3: Customer Segmentation

Apply clustering algorithms (e.g., K-Means) and machine learning models (e.g., decision trees, random forests) to segment the customer base. Evaluate and compare different segmentation approaches.

# Phase 4: Segment Validation and Refinement

Validate the segments by measuring their performance in real-world marketing campaigns.Refine the segments based on the feedback and results.

# Dataset Description:

The dataset provided contains information about customers, including CustomerID, Gender, Age, Annual Income (in thousands of dollars), and Spending Score (ranging from 1 to 100). Each row represents a different customer. Here are the first few rows of the dataset

| CustomerID | Gender | Age | Annual Income (k$) | Spending Score (1-100) |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| | | | | | | | | | | | |  |
| | 0001 | | Male | | 19 | | 15 |  | |39 | | |
| | 0002 | | Male | | 21 | | 15 |  | |81 | | |
| | 0003 | | Female | | 20 | | 16 |  | |6 | | |
| | 0004 | | Female | | 23 | | 16 |  | |77 | | |
| | ... | | ... | | ... | | ... |  | | ... | | |

# Data Preprocessing:

Before performing customer segmentation, data preprocessing is necessary to ensure the data is clean and ready for analysis. The following data preprocessing steps were applied:

1. Handling Missing Data: Check for and handle missing values in the dataset, if any.
2. Data Cleaning: Remove any duplicate records to maintain data integrity.
3. Data Transformation: Depending on the analysis technique chosen, data transformation might include one-hot encoding for the "Gender" variable or normalizing numeric features like "Age," "Annual Income," and "Spending Score."

# Analysis Techniques:

The goal of the customer segmentation project is to group customers with similar characteristics together. Several analysis techniques can be applied, including but not limited to:

# K-Means Clustering:

This technique groups customers into K clusters based on their similarity. The "K" value needs to be determined based on the data.

We applied K-Means clustering to segment customers based on their characteristics. The optimal number of clusters (K) was determined using the elbow method, and we identified the following clusters:

# Annual Income vs Spending score :

* Cluster 1: Low Income, High Spending Score
* Cluster 2: Low Income, Low Spending Score
* Cluster 3: High Income, Low Spending Score
* Cluster 4: High Income, High Spending Score
* Cluster 5: Average Income ,Average Spending Score

# Key Findings:

**Cluster 1: Low Income, High Spending Score**

These customers have potential for upselling and cross-selling, as they exhibit a willingness to spend despite limited incomes.

# Cluster 2: Low Income, Low Spending Score

These customers are price-sensitive and may respond well to cost-effective offerings and discounts.

# Cluster 3: High Income, Low Spending Score

They may have unmet needs or dissatisfaction with current offerings, which need to be addressed to increase their spending. Personalized and premium offerings might resonate with this group

# Cluster 4: High Income, High Spending Score

These customers are high-value and should be retained through excellent customer service and exclusive offerings.Implementing loyalty programs and providing personalized recommendations can further enhance their loyalty and spending.

# Cluster 5: Average Income ,Average Spending Score

Customers in this cluster exhibit a balanced behavior in terms of income and spending.They are neither high spenders nor extremely thrifty, indicating they spend within their means.

# Age vs Spending Scores:

-Cluster 1: Young, High Spending Customers

* Cluster 2: Middle-Aged, Moderate Spending Customers
* Cluster 3- Older,Moderate Spending Customers

-Cluster 4- Middle-Aged,Lower Spending Score

Certainly, here are key findings for each cluster based on the analysis of age vs. spending scores:

**Cluster 1: Young, High Spending Customers**: This cluster represents a group of young customers who exhibit high spending behavior.

* + They have the potential to be brand-loyal and may respond well to trendy and youth-centric marketing strategies.
  + Strategies to retain and upsell this group may lead to increased profits.

**Cluster 2: Middle-Aged, Moderate Spending Customers** :This segment includes middle-aged customers with a moderate spending score.They are likely to be more financially stable and considerate in their spending.Tailored marketing and promotions may encourage them to spend more without overspending.

**Cluster 3: Older, Moderate Spending Customers:**This cluster consists of older customers who exhibit a moderate spending score.They are financially responsible and may value quality and reliability in their purchases.Offering loyalty programs or products tailored to their preferences may increase their spending.

**Cluster 4: Middle-Aged, Lower Spending Score**: This segment represents middle-aged customers who tend to spend less.They may be price-conscious and seek value-for-money products.Marketing strategies should focus on providing cost-effective options and addressing potential barriers to spending.

# Insights:

1. **Cross-Selling Opportunities**- The "Low Income, High Spending Score" segment presents cross-selling opportunities, as these customers have a high propensity to spend.
2. **Price Optimization**- The "Low Income, Low Spending Score" segment is price-sensitive, indicating the need for price optimization strategies.
3. **Product and Service Enhancement**- The "High Income, Low Spending Score" segment may require product or service enhancements to meet their expectations and encourage higher spending.
4. **Customer Loyalty**- The "High Income, High Spending Score" segment represents valuable customers. Building and maintaining customer loyalty programs is essential to keep them engaged.

# Recommendations:

1. **Personalized Marketing:** Implement personalized marketing campaigns tailored to each customer segment's preferences and behaviors.
2. **Pricing Strategies**: Experiment with pricing strategies, including discounts for price-sensitive segments and premium offerings for high-income customers.
3. **Product Development**: Continuously improve and expand products and services to better cater to the needs of each segment.
4. **Customer Engagement**: Engage with high-value customers through loyalty programs, exclusive offers, and exceptional customer service.
5. **Monitoring and Feedback:** Regularly monitor the effectiveness of strategies and collect customer feedback to refine and adapt your approach.