

# Customer Segmentation for Marketing Analysis — Project

## Introduction

At my workplace, I've done extensive customer segmentation to better understand and satisfy our customers. Businesses thrive on customer satisfaction, retention, and loyalty, making it essential to know our customers deeply. Customer segmentation, particularly using machine learning, allows us to group customers based on common characteristics for tailored marketing strategies.

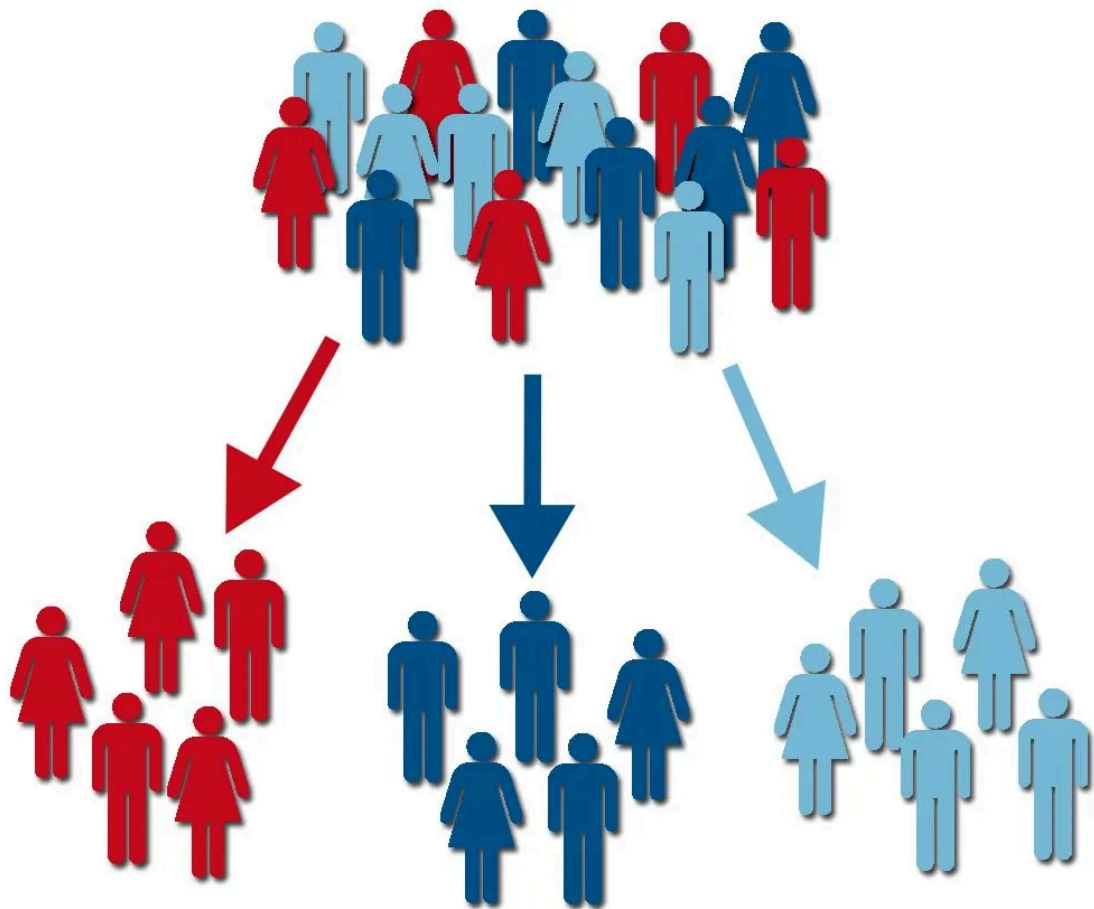
In this article, I'll share my experience in segmenting customers using a machine-learning-based model

# Segmentation client pour l'analyse marketing — Projet

## Introduction

Sur mon lieu de travail, j'ai procédé à une segmentation approfondie de la clientèle pour mieux comprendre et satisfaire nos clients. Les entreprises prospèrent grâce à la satisfaction, à la fidélisation et à la fidélité de leurs clients, ce qui rend essentiel une connaissance approfondie de nos clients. La segmentation des clients, notamment grâce au machine learning, nous permet de regrouper les clients en fonction de caractéristiques communes pour des stratégies marketing sur mesure.

Dans cet article, je partagerai mon expérience dans la segmentation des clients à l'aide d'un modèle basé sur l'apprentissage automatique



## Dataset Overview

Source: Kaggle

Description: This dataset includes demographic and behavioral information about customers, useful for identifying distinct segments within the customer base.

### Key Features:

id: Unique identifier for each customer. age: Age of the customer. gender: Gender of the customer (Male, Female, Other). income: Annual income of the customer (in USD). spending\_score: Spending score (1-100), indicating the customer's spending behavior and loyalty. membership\_years: Number of years the customer has been a member. purchase\_frequency: Number of purchases made by the customer in the last year. preferred\_category: Preferred shopping category (Electronics, Clothing, Groceries, Home & Garden, Sports). last\_purchase\_amount: Amount spent by the customer on their last purchase (in USD). Understanding the dataset is crucial as it forms the foundation of our segmentation model.

Potential Uses:

- Customer Segmentation: Identify different customer segments based on their demographic and behavioral characteristics.
- Targeted Marketing: Develop targeted marketing strategies for different customer segments.

## Aperçu de l'ensemble de données

Source : Kaggle

Description : cet ensemble de données comprend des informations démographiques et comportementales sur les clients, utiles pour identifier des segments distincts au sein de la clientèle.

### Principales caractéristiques :

id : Identifiant unique de chaque client. age : Âge du client. sexe : sexe du client (homme, femme, autre). revenu : Revenu annuel du client (en USD). saving\_score : score de dépenses (1 à 100), indiquant le comportement d'achat et la fidélité du client. member\_years : nombre d'années pendant lesquelles le client est membre. Purchase\_Frequency : nombre d'achats effectués par le client au cours de la dernière année. Preferred\_category : catégorie d'achats préférée (électronique, vêtements, épicerie, maison et jardin, sports). last\_purchase\_amount : Montant dépensé par le client lors de son dernier achat (en USD). Comprendre l'ensemble de données est crucial car il constitue la base de notre modèle de segmentation.

Utilisations potentielles :

- Segmentation client : Identifiez différents segments de clientèle en fonction de leurs caractéristiques démographiques et comportementales.
- Marketing ciblé : Développer des stratégies marketing ciblées pour différents segments de clientèle.
- Programmes de fidélisation des clients : concevoir des programmes de fidélité basés sur le comportement et les préférences des clients en matière d'achat.

## Data Collection and Exploration

The dataset for this project is readily available on Kaggle. We begin by downloading it and importing it into the Jupyter Notebook where we'll build our segmentation model.

## Step 1: Import Libraries and Load Data

First, import the necessary libraries and load the dataset. This dataset is in CSV format, so we'll load it directly into our notebook.

## Step 2: Data Exploration

Exploring the data helps us understand its structure and content. Here's how we proceed:

- Understand the Data Structure: Use the `shape()` function.
- Summary Statistics: Use the `describe()` function to understand central tendency, dispersion, and distribution shape.
- Missing Values: Check for completeness.
- Data Distribution: Visualize key features to understand their distributions.

## Step 3: Data Preprocessing

Preprocessing is essential for accurate analysis and modeling. It includes:

Handling Missing Values: We have no missing values in our dataset. Removing Duplicates: Ensure no duplicate entries. Scaling Numerical Data: Improves model performance, especially for algorithms sensitive to feature scales. Encoding Categorical Data: One-Hot Encoding: Converts categorical variables into binary vectors. Label Encoding: Converts categorical labels into numeric values.

## Step 4: Model Building and Customer Segmentation

Using K-means clustering, we segment customers. Steps include:

Determine Optimal Number of Clusters: Use the Elbow Method. Apply K-means Clustering: Initialize and fit the model with the optimal number of clusters.

## Step 5: Evaluation and Interpretation

After building the model, evaluate and interpret the clusters to understand the characteristics of each segment and derive actionable insights.

Here are the characteristics of each cluster based on the profiles and visualizations, followed by some actionable insights and recommendations.

## Aperçu de l'ensemble de données

Source : Kaggle

Description : cet ensemble de données comprend des informations démographiques et comportementales sur les clients, utiles pour identifier des segments distincts au sein de la clientèle.

## Principales caractéristiques :

id : Identifiant unique de chaque client. age : Âge du client. sexe : sexe du client (homme, femme, autre). revenue : Revenu annuel du client (en USD). saving\_score : score de dépenses (1 à 100), indiquant le comportement d'achat et la fidélité du client. member\_years : nombre d'années pendant lesquelles le client est membre. Purchase\_Frequency : nombre d'achats effectués par le client au cours de la dernière année. Preferred\_category : catégorie d'achats préférés (électronique, vêtements, épicerie, maison et jardin, sports). last\_purchase\_amount : Montant dépensé par le client lors de son dernier achat (en USD). Comprendre l'ensemble de données est crucial car il constitue la base de notre modèle de segmentation.

Utilisations potentielles :

- Segmentation client : Identifiez différents segments de clientèle en fonction de leurs caractéristiques démographiques et comportementales.
- Marketing ciblé : Développer des stratégies marketing ciblées pour différents segments de clientèle.
- Programmes de fidélisation des clients : concevoir des programmes de fidélité basés sur le comportement et les préférences des clients en matière d'achat.
- 

### Segment 1:

- Average Age: 32
- Average Income: \$95,749
- Gender: Male
- Spending Score: 42
- Purchase Frequency: 36
- Preferred Category: Electronics Recommendations: Focus on electronics promotions tailored to this segment, highlight new products, and implement loyalty programs.

### Segment 2:

- Average Age: 52
- Average Income: \$75,836
- Gender: Other
- Spending Score: 52

- Purchase Frequency: 28
- Preferred Category: Groceries Recommendations: Diversify grocery offerings, increase engagement with personalized recommendations, and initiate community programs.

### Segment 3:

- Average Age: 46
- Average Income: \$96,057
- Gender: Female
- Spending Score: 56
- Purchase Frequency: 14
- Preferred Category: Sports Recommendations: Emphasize sports and fitness products, promote health and wellness events, and introduce subscription services. Customer segmentation is a powerful tool that helps businesses understand and serve their customers better. By leveraging machine learning, we can uncover insights that would be difficult to obtain through traditional methods. This project demonstrates the importance of data-driven decision-making in crafting targeted marketing strategies and enhancing customer satisfaction.

As we continue to collect and analyze customer data, our segmentation models will evolve, providing even deeper insights. Remember, the goal is not just to segment customers but to use these segments to deliver personalized experiences that meet their unique needs and preferences.

I hope this project inspires you to explore customer segmentation in your work and harness the power of data to drive business success.

Thank you for following along, and I look forward to sharing more insights and projects in the future.

### Segment 1 :

- Âge moyen : 32 ans
- Revenu moyen : 95 749 \$
- Sexe : Homme
- Score de dépenses : 42
- Fréquence d'achat : 36
- Catégorie préférée : Électronique Recommandations : privilégier les promotions électroniques adaptées à ce segment, mettre en avant les nouveaux produits et mettre en place des programmes de fidélité.

### Segment 2 :

- Âge moyen : 52 ans

- Revenu moyen : 75 836 \$
- Sexe : Autre
- Score de dépenses : 52
- Fréquence d'achat : 28
- Catégorie préférée : Épicerie Recommandations : diversifiez l'offre d'épicerie, augmentez l'engagement avec des recommandations personnalisées et lancez des programmes communautaires.

### Segment 3 :

- Âge moyen : 46 ans
- Revenu moyen : 96 057 \$
- Sexe : Femme
- Score de dépenses : 56
- Fréquence d'achat : 14
- Catégorie préférée : Sports Recommandations : mettre l'accent sur les produits de sport et de fitness, promouvoir les événements de santé et de bien-être et introduire des services d'abonnement. La segmentation client est un outil puissant qui aide les entreprises à mieux comprendre et servir leurs clients. En tirant parti de l'apprentissage automatique, nous pouvons découvrir des informations difficiles à obtenir par les méthodes traditionnelles. Ce projet démontre l'importance de la prise de décision basée sur les données pour élaborer des stratégies de marketing ciblées et améliorer la satisfaction des clients.

À mesure que nous continuons à collecter et à analyser les données clients, nos modèles de segmentation évolueront, fournissant des informations encore plus approfondies. N'oubliez pas que l'objectif n'est pas seulement de segmenter les clients, mais d'utiliser ces segments pour offrir des expériences personnalisées qui répondent à leurs besoins et préférences uniques.

J'espère que ce projet vous incitera à explorer la segmentation des clients dans votre travail et à exploiter la puissance des données pour favoriser la réussite de votre entreprise.

Merci d'avoir suivi et j'ai hâte de partager plus d'idées et de projets à l'avenir.

In [ ]: