### Project Objective: Marketing campaign analysis

The goal of this project is to perform **customer segmentation** using clustering techniques to help businesses understand their customer base better. By analyzing **demographic, purchase, and behavioral attributes**, we can group customers into distinct segments, allowing companies to **target their marketing strategies more effectively** and **personalize product recommendations**.

### Key Goals:

🔹 **Identify distinct customer groups** based on spending habits, demographics, and engagement.  
🔹 **Optimize marketing efforts** by focusing on high-value customer segments.  
🔹 **Improve customer retention** by understanding and addressing specific needs.  
🔹 **Enhance product recommendations** by analyzing customer purchase patterns.  
🔹 **Compare different clustering models** like **K-Means, Hierarchical Clustering, and DBSCAN** to determine the best approach.

### Dataset & Attributes

The dataset includes a range of **demographic, behavioral, and transactional attributes**:

📌 **Demographics:**

* **ID:** Unique customer identifier
* **Year\_Birth:** Year of birth
* **Education:** Education level
* **Marital\_Status:** Marital status
* **Income:** Yearly household income
* **Kidhome / Teenhome:** Number of children & teenagers in the household

📌 **Customer Behavior:**

* **Dt\_Customer:** Date of customer enrollment
* **Recency:** Number of days since last purchase
* **Complain:** Whether the customer has complained in the last two years

📌 **Purchase History:**

* **MntWines, MntFruits, MntMeatProducts, MntFishProducts, MntSweetProducts, MntGoldProds:** Amount spent on different product categories in the last two years

📌 **Promotional Engagement:**

* **NumDealsPurchases:** Number of purchases made with a discount
* **AcceptedCmp1-5 & Response:** Whether the customer accepted promotional campaigns

📌 **Shopping Channels & Engagement:**

* **NumWebPurchases:** Online purchases
* **NumCatalogPurchases:** Catalog-based purchases
* **NumStorePurchases:** In-store purchases
* **NumWebVisitsMonth:** Website visits in the last month

### Clustering Approach:

📌 **Feature Engineering** – Handling missing data, normalizing numerical variables, and encoding categorical data.  
📌 **Dimensionality Reduction** – Using **PCA** (Principal Component Analysis) to improve clustering efficiency.  
📌 **Model Selection & Evaluation** – Applying and comparing **K-Means, Hierarchical Clustering, and DBSCAN** to determine optimal segmentation.  
📌 **Cluster Interpretation** – Analyzing cluster characteristics to derive meaningful insights.

### Deployment Plan:

🚀 Build an **interactive dashboard** using **Streamlit/Flask** to visualize customer segments, explore spending behaviors, and provide actionable insights for businesses.