SmartCart Clustering System

Problem Statement

SmartCart is a growing e-commerce platform serving customers across multiple countries. The company collected extensive customer data consisting of 2240 customer records and 22 attributes including demographics, purchase behaviour, website activity, and response.

Currently, SmartCart uses **generic marketing and engagement strategies** for all customers, without clearly understanding different customer behaviour patterns. This results in **inefficient marketing**, missed opportunities to retain high-value customers, and delayed identification of churn-prone users.

To solve this, SmartCart aims to build an **intelligent customer segmentation system** using **unsupervised machine learning**. The system will analyse historical customer transaction data and **group customers into meaningful clusters** based on purchasing behaviour, engagement levels, and loyalty indicators.

You are hired as an **AI/ML Engineer** to develop this system using **clustering algorithms** to discover hidden patterns in customer behaviour and support **data-driven decision-making** for personalised marketing and customer retention.

Dataset Description

Each row in the dataset represents a **customer** and contains multiple attributes describing their spending and activity on platform.

1. **Customer Demographics**

Feature	Description
ID	Unique customer identifier
Year_Birth	Year of birth of the customer
Education	Highest education level achieved
Marital_Status	Marital status of the customer
Income	Yearly household income
Kidhome	Number of small children in household
Teenhome	Number of teenagers in household
Dt_Customer	Date when customer enrolled with SmartCart

2. **Purchase Behaviour** (Amount Spent)

Feature	Description
MntWines	Amount spent on wine products
MntFruits	Amount spent on fruits
MntMeatProducts	Amount spent on meat products
MntFishProducts	Amount spent on fish products
MntSweetProducts	Amount spent on sweet products
MntGoldProds	Amount spent on gold products

3. **Purchase Behaviour (Frequency)**

Feature	Description
NumDealsPurchases	Purchases made using discounts
NumWebPurchases	Purchases made through website
NumCatalogPurchases	Purchases made through catalog
NumStorePurchases	Purchases made in physical stores
NumWebVisitsMonth	Number of visits to website per month

4. **Customer Feedback & Constants**

Feature	Description
Recency	Number of days since last purchase
Complain	Customer complained in last 2 years (1 = Yes, 0 = No)