Clustering Credit Card Customers for Behavior Analysis and Segmentation

Clustering is a fundamental technique in unsupervised machine learning that groups similar data points into clusters or segments based on shared characteristics. The goal of clustering is to discover inherent patterns or structures in the data without prior knowledge of the target labels. It has numerous applications, including customer segmentation, anomaly detection, and pattern recognition.

In the context of financial data, clustering can be especially useful in segmenting customers based on their spending patterns, payment behavior, and credit usage. By grouping similar customers together, financial institutions can gain valuable insights that help with targeted marketing, personalized offers, and risk management strategies.

This project focuses on clustering credit card customer data, where each customer's behavior is captured through various features, such as account balance, purchase frequency, cash advances, payments, and credit limits. By applying clustering algorithms, we aim to identify distinct groups of customers who exhibit similar financial behaviors. These groups can then be analyzed to understand customer needs better and optimize business strategies.

Problem Statement:

The goal of this project is to apply **clustering techniques** to the provided customer data to segment customers based on their credit card usage patterns. The dataset contains various features, such as:

- **BALANCE**: The current balance of the customer.
- **PURCHASES**: The total value of purchases made by the customer.
- **CASH_ADVANCE**: The total amount of cash advances used.
- **PAYMENTS**: The amount paid towards the balance.
- **CREDIT_LIMIT**: The credit limit available to the customer.
- **TENURE**: The number of months the customer has been using the credit card.

The main objective is to identify distinct groups of customers who have similar financial behaviors. For instance, some customers might use their credit cards for large purchases, while others might use cash advances frequently. By clustering customers into these groups, we can better understand the characteristics of each segment and tailor financial products, marketing strategies, or customer support to their needs.

The clustering process will help us:

- Segment customers based on their spending and payment behaviours.
- Identify high-risk or low-risk customer groups.
- Explore patterns that can lead to better customer retention and targeted offerings.

By performing this clustering analysis, we aim to provide a clear understanding of the customer base, which can then inform decision-making in areas such as marketing, credit management, and service offerings.

Dataset: CC_GENERAL