



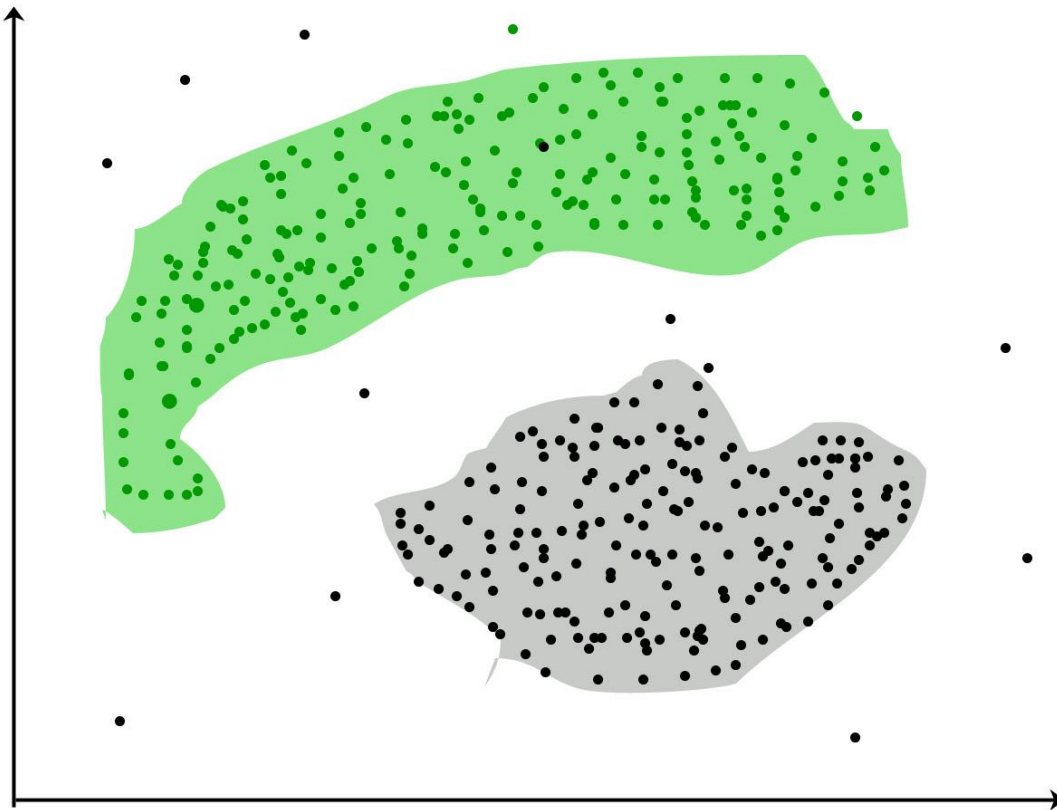
## The 3<sup>rd</sup> DeepDive Challenge

### Credit Card Clustering

Khordad 1402



Clustering is the task of dividing the data points into a number of groups such that data points in the same groups are more similar to other data points in the same group and dissimilar to the data points in other groups.



“Credit Card Clustering” is a dataset of customer credit card details that summarizes the usage behavior of about 9000 active credit card holders during 6 months. The file is at a customer level with 18 behavioral variables. It requires to develop a customer segmentation to define marketing strategy.

Clustering credit card details can help banks and other financial institutions understand their customers' spending habits and preferences. This information can be used to segment customers into different groups based on their spending behavior, allowing companies to tailor their marketing efforts and product offerings to the needs of each group.

The objectives of this challenge are to perform exploratory data analysis (EDA) on the dataset using clustering methods that will involve analyzing the distribution and correlations between the various features in the dataset, as well as identifying any outliers or missing values. The clustering phase will involve grouping the data into different clusters based on their similarities using methods such as k-means, hierarchical, or density-based clustering.

The "Credit Card Clustering" dataset includes the following features:

- **CUST\_ID**: Identification of Credit Card holder (Categorical)
- **BALANCE**: Balance amount left in their account to make purchases
- **BALANCE\_FREQUENCY**: How frequently the balance is updated, score between 0 and 1 (1 = frequently updated, 0 = not frequently updated)
- **PURCHASES**: Amount of purchases made from account
- **ONEOFF\_PURCHASES**: Maximum purchase amount done in one-go
- **INSTALLMENTS\_PURCHASES**: Amount of purchase done in installment
- **CASH\_ADVANCE**: Cash in advance given by the user
- **PURCHASES\_FREQUENCY**: How frequently the purchases are being made, score between 0 and 1 (1 = frequently purchased, 0 = not frequently purchased)
- **ONEOFF\_PURCHASES\_FREQUENCY**: How frequently purchases are happening in one-go (1 = frequently purchased, 0 = not frequently purchased)
- **PURCHASES\_INSTALLMENTS\_FREQUENCY**: How frequently purchases in installments are being done (1 = frequently done, 0 = not frequently done)
- **CASH\_ADVANCE\_FREQUENCY**: How frequently the cash in advance being paid
- **CASH\_ADVANCE\_TRX**: Number of transactions made with "Cash in Advanced"
- **PURCHASES\_TRX**: Number of purchase transactions made

- **CREDIT\_LIMIT**: Limit of credit card for user
- **PAYMENTS**: Amount of payment done by user
- **MINIMUM\_PAYMENTS**: Minimum amount of payments made by user
- **PRC\_FULL\_PAYMENT**: Percent of full payment paid by user
- **TENURE**: Tenure of credit card service for user