XS

Xabier Sagarzazu
VIRTUAL ASSISTANT

DATA ANALYSIS

MARKET CLUSTERING

CONTEXT

A company has created a big database with information related to its customers.

The goal will be to to Analyze this database to be able to identify different groups.

By using "Principal Component Analysis" technique we will be able to label the customers in different groups.

Once the the customers are labeled in different groups Machine Learning algorithms could be used for prediction models (not done in this project).

DATABASE

The database of the customer contains the following information:

- CUSTOMER ID.
- BALANCE
- BALANCE_FREQUENCY
- PURCHASES
- ONEOFF_PURCHASES
- INSTALLMENTS_PURCHASES
- CASH_ADVANCE
- PURCHASES_FREQUENCY
- ONEOFF_PURCHASES_FREQUENCY

- PURCHASES_INSTALLMENTS_FREQUENCY
- CASH_ADVANCE_FREQUENCY
- CASH_ADVANCE_TRX
- PURCHASES_TRX
- CREDIT_LIMIT
- PAYMENTS
- MINIMUM_PAYMENTS
- PRC_FULL_PAYMENT
- TENURE

RESULTS

After using PCA technique and "elbow" method the data was labeled in 4 different groups.

Now the model is trained to be able to label future customers .