Exploratory Data Analysis

Customer Segmentation – Case study

11/14/2021

Agenda

Executive Summary

Problem Statement

EDA

EDA Summary

Problem Statement

- XYZ bank wants to roll out Christmas with personalized offers to their customers. The group up process needs to be automated and can’t exceed 5 groups in total as a result.

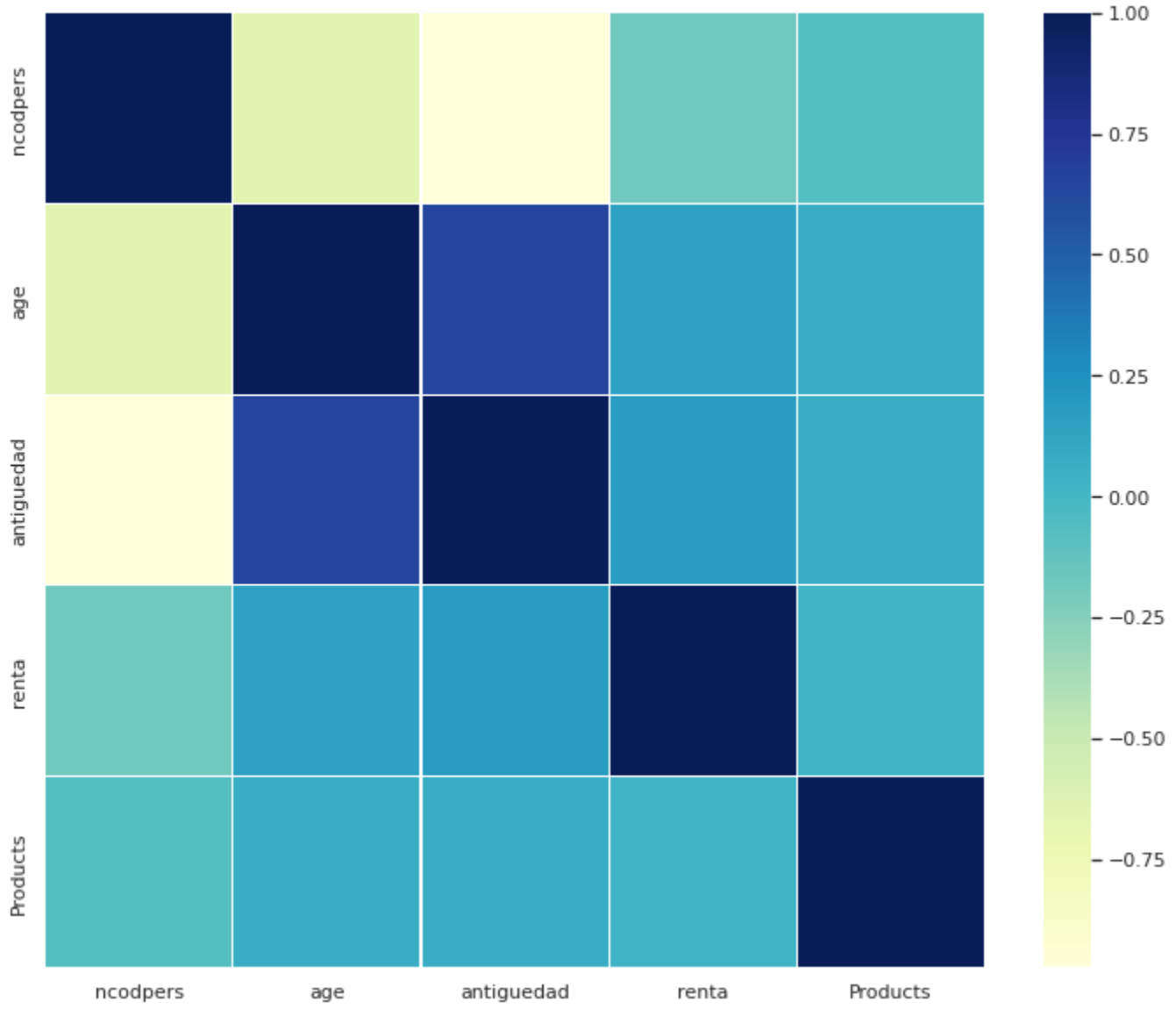
EDA

Source Data

- **Customer_set.csv –** this file includes details of transactions from different clients, each entry has taken when a client acquire a new financial product.

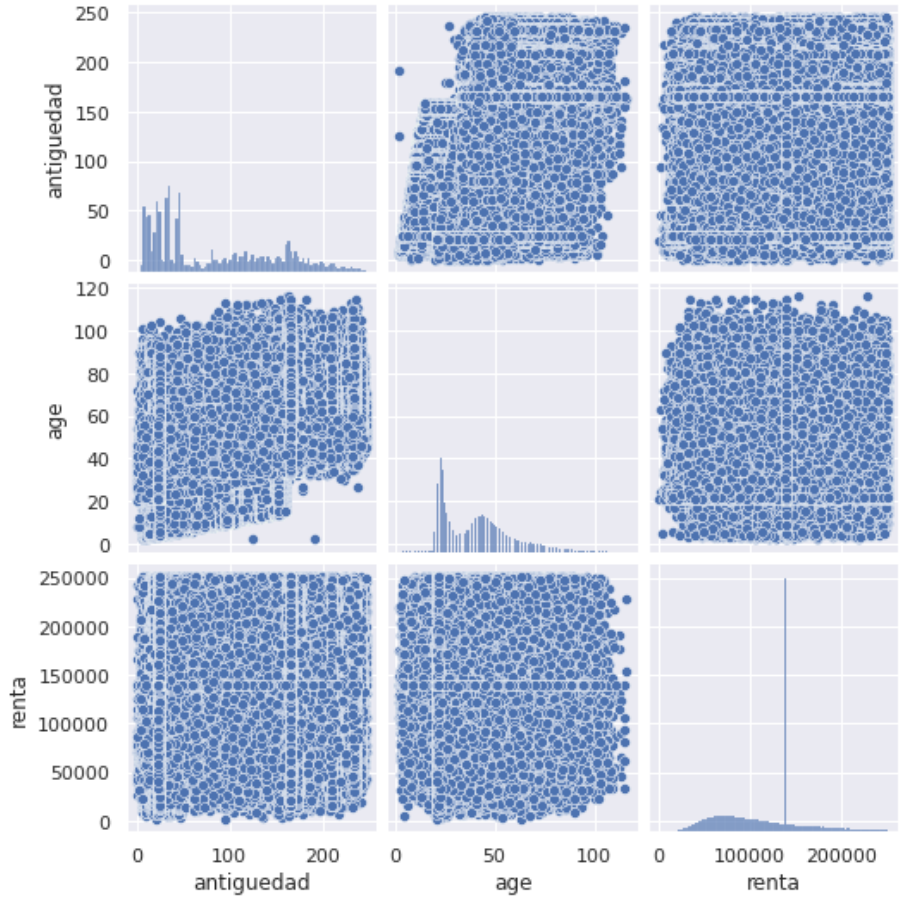
Selection of features

- In the graph we show a strong correlation from “antiguedad” and “age”, so we can only take one of this features
- We create a new feature called “products” to express the total number of products that each client has.



Spread from numerical features

- We detect an strong dispertion from the numerical values, this can affect the model selected to get the classification of clients



Model Technique Proposed

Clustering – non supervised classification

- We suggest to utilize a classification algorithm as K-Means, because from high dispertion of data, DBSCAN and OPTICS can be wrong with the classification results. (the centroids can be too near one to the other)

Thank You