



Data Glacier

Your Deep Learning Partner

XYZ Credit Union

(Banking industry)

Nicolette Peterkin

Virtual Internship

14-Aug-2021

Background

XYZ credit union in Latin America is performing very well in selling banking products like; Credit card, deposit account, retirement account, safe deposit box etc) but their existing customer is not buying more than 1 product which means bank is not performing good in cross selling. The bank is not able to sell their other offerings to existing customers and would like that to change.

Business Problem

Changing to a new selling technique requires a lot of research and data analyzing to ensure that the right cross selling strategy is used to re-enter the market. The issue XYZ credit union is having currently is that they do not know what actions to take in order to increase cross selling in the credit union.

SOLUTION

Our goal is to create an exploratory data analysis for XYZ credit unions' dataset and create a cross-selling plan solution. The EDA report will be specific and meaningful so there is no room for assumptions. As the data analyst assigned to this project my job is to analyze the given data with much scrutiny in order to provide XYZ credit union with vital insights on cross selling. Once the exploratory data analyze report is completed XYZ credit union will have actionable insights which can then be used to inform their cross-selling campaigns.

Layout of Case Study

The analysis has been divided into four parts:

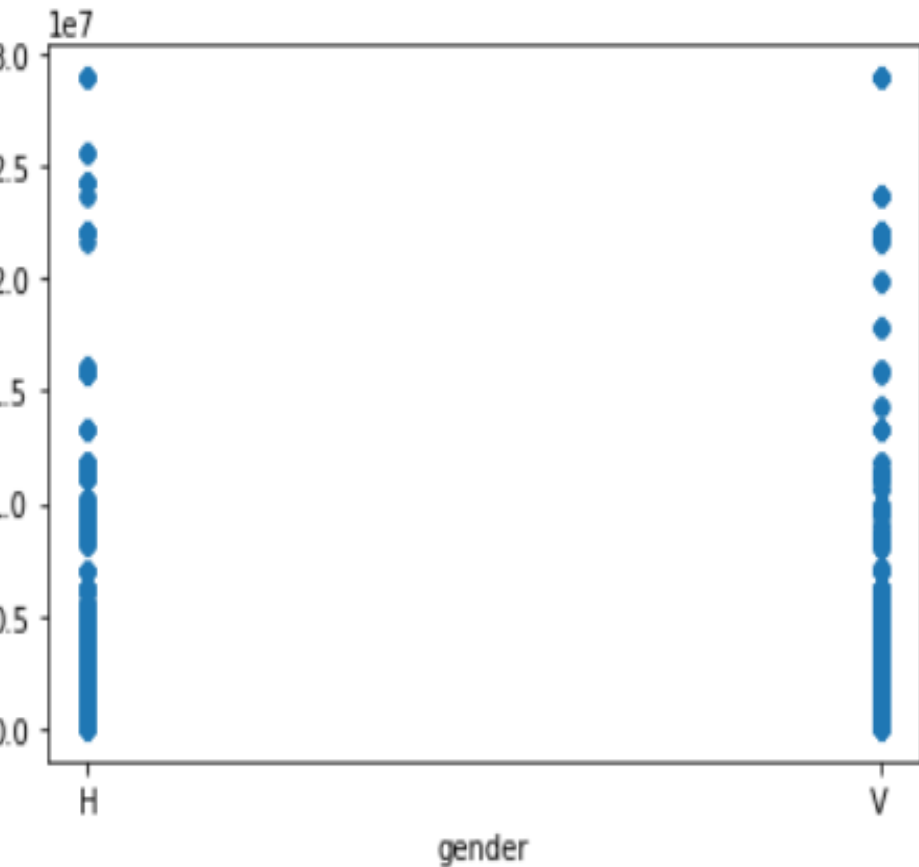
- Data Understanding which includes data cleaning and transformation
- Data exploring – checking correlations of dataset
- Recommendations Model for cross-selling

Data Understanding

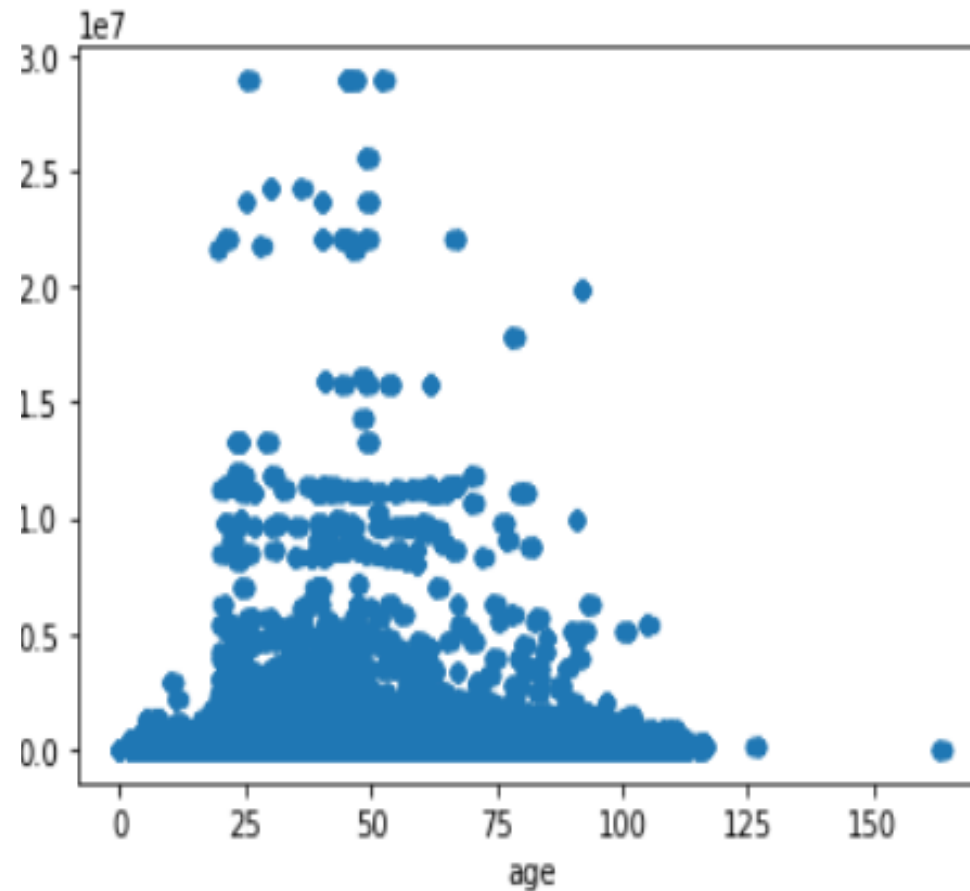
- Timeframe of the data: July 2021 to August 2021
- Total data points : 13,647,309
- In order to overcome some of the NA values we will use the mean and mode values to fill in the missing information and for the ones that's we are unable to fill we will remove them.
- For duplicate values we will remove them as they are not needed in the dataset.
- Also we intend on removing the outliers in the dataset as they will not change the actionable insights from the dataset.¶

Univariate Plots Analysis

Customer Analysis based on Gender



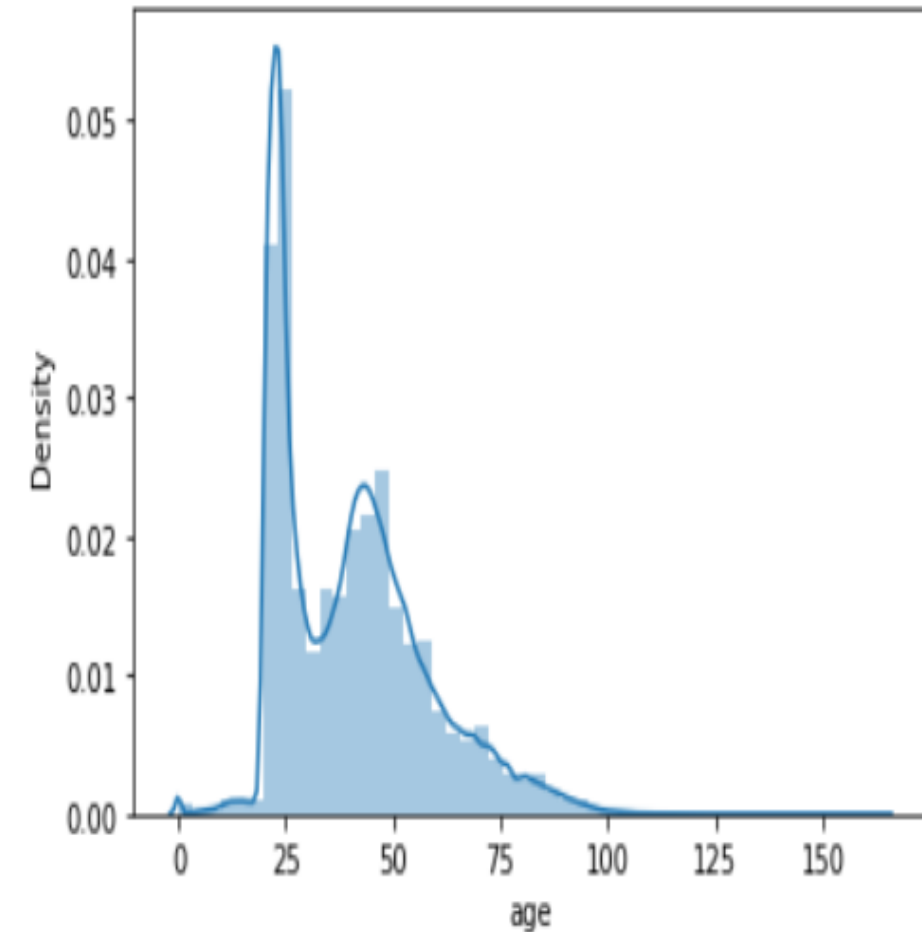
Customer Analysis based on Age and Gross Income



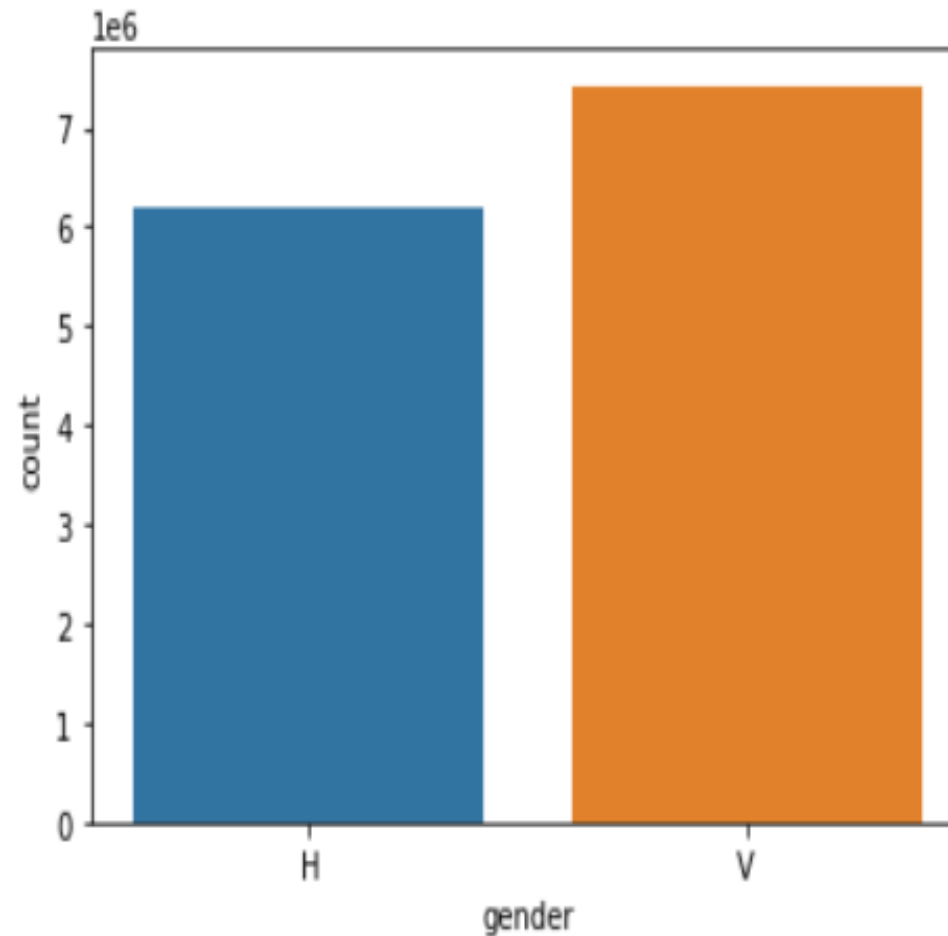
- The majority of customers earn below \$1.5 Million yearly.

Univariate Plots Analysis

Customer Analysis based on Age



Customer Analysis based on Gender

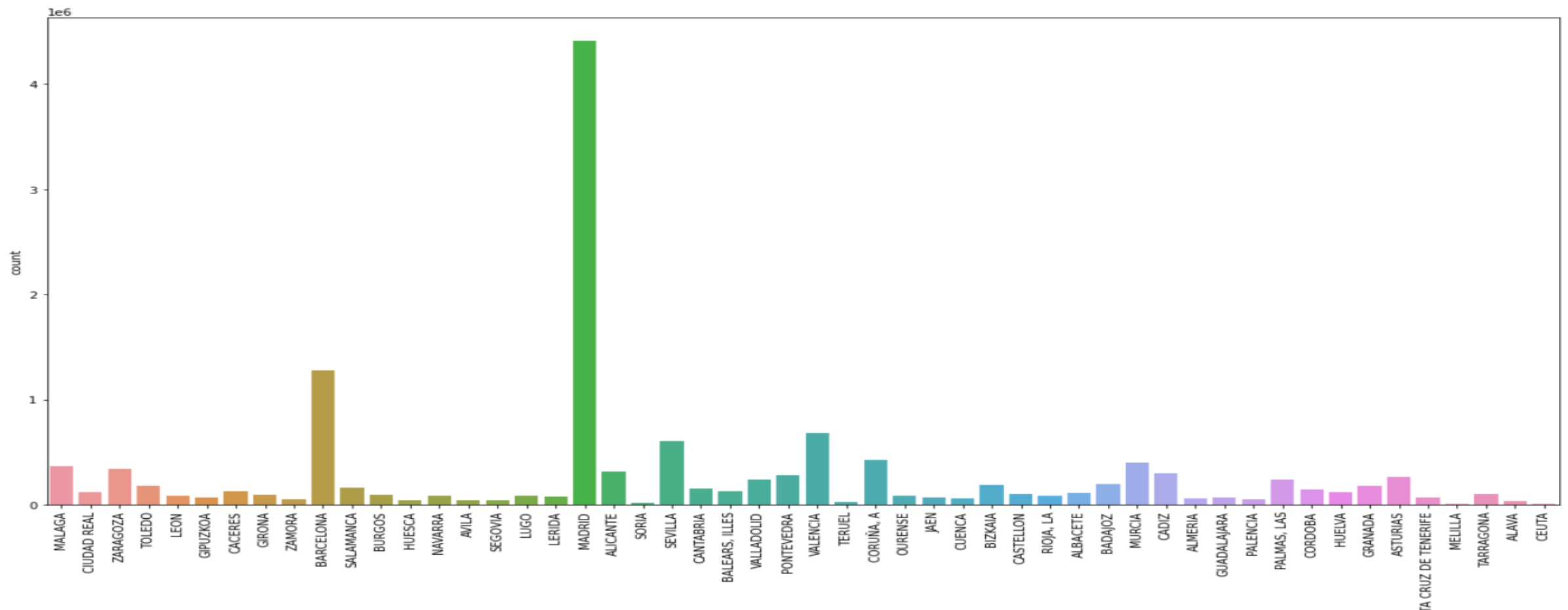


- Majority of the customers are from their 20s to late 40s.
- There are more female customers than male.

Univariate Plots Analysis

Customer Analysis based on Province.

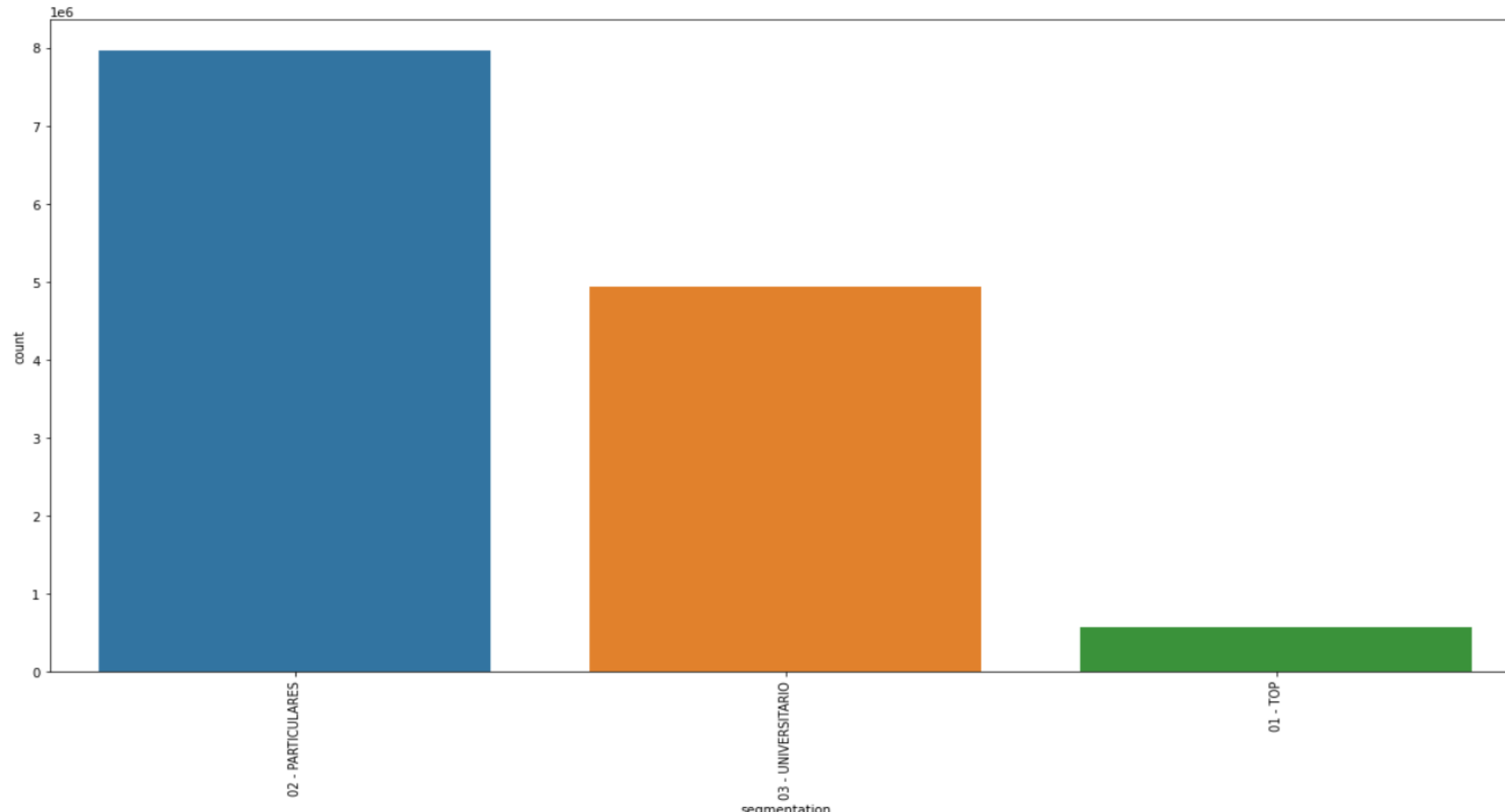
- It is evident that Madrid has the highest amount of customers at XYZ Credit Union.



Univariate Plots Analysis

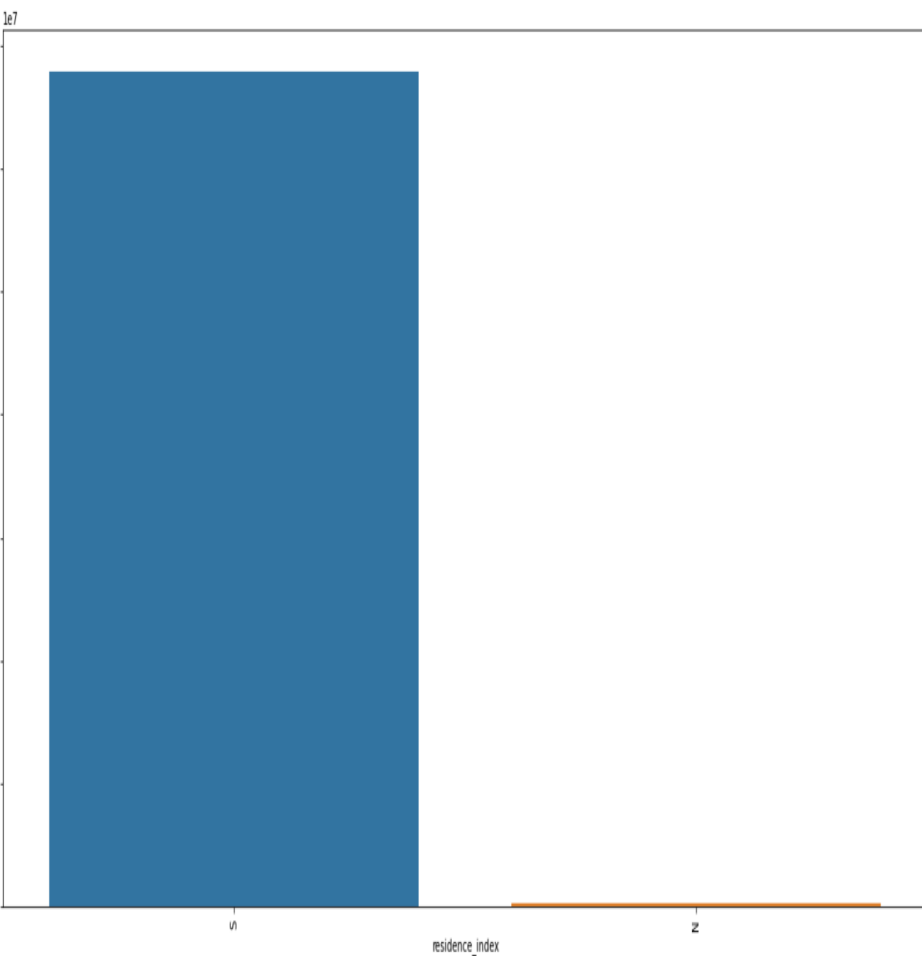
Customer Analysis based on Segmentation.

➤ The Individual segmentation has the most customers.

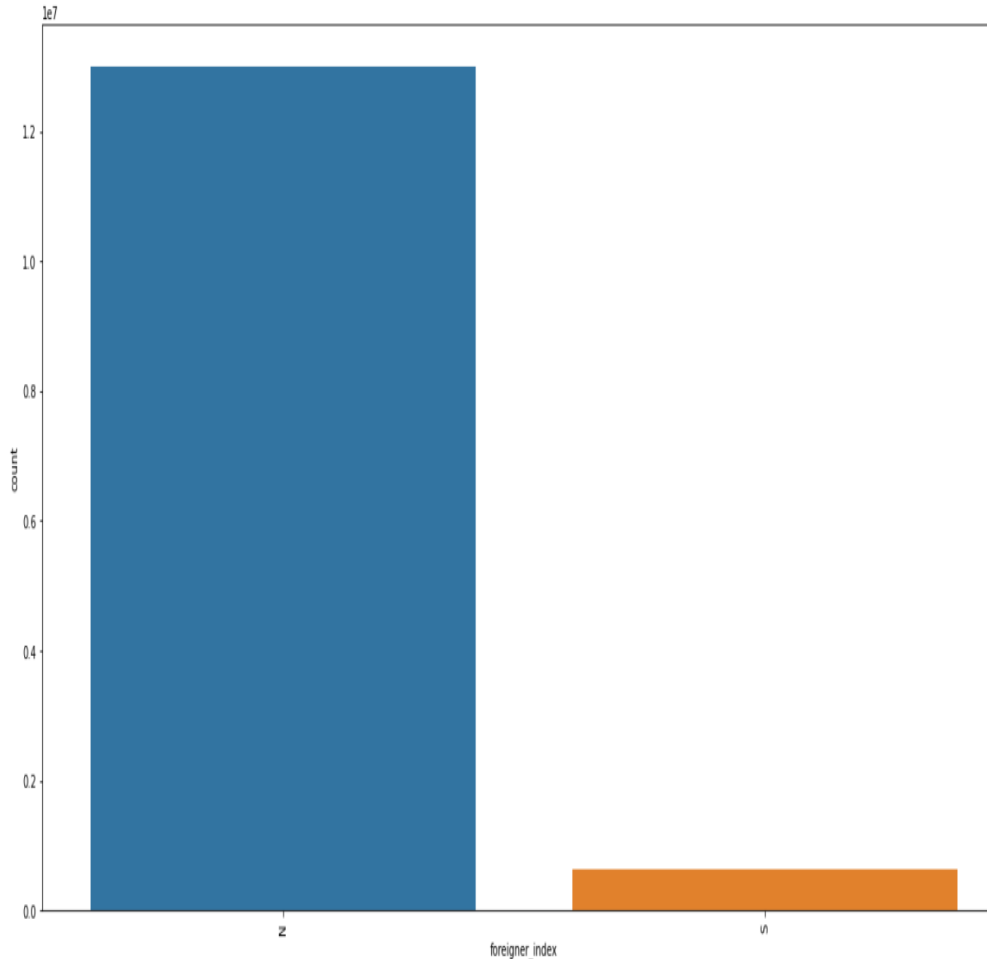


Univariate Plots Analysis

Customer Residence Index



Customer Foreign Index

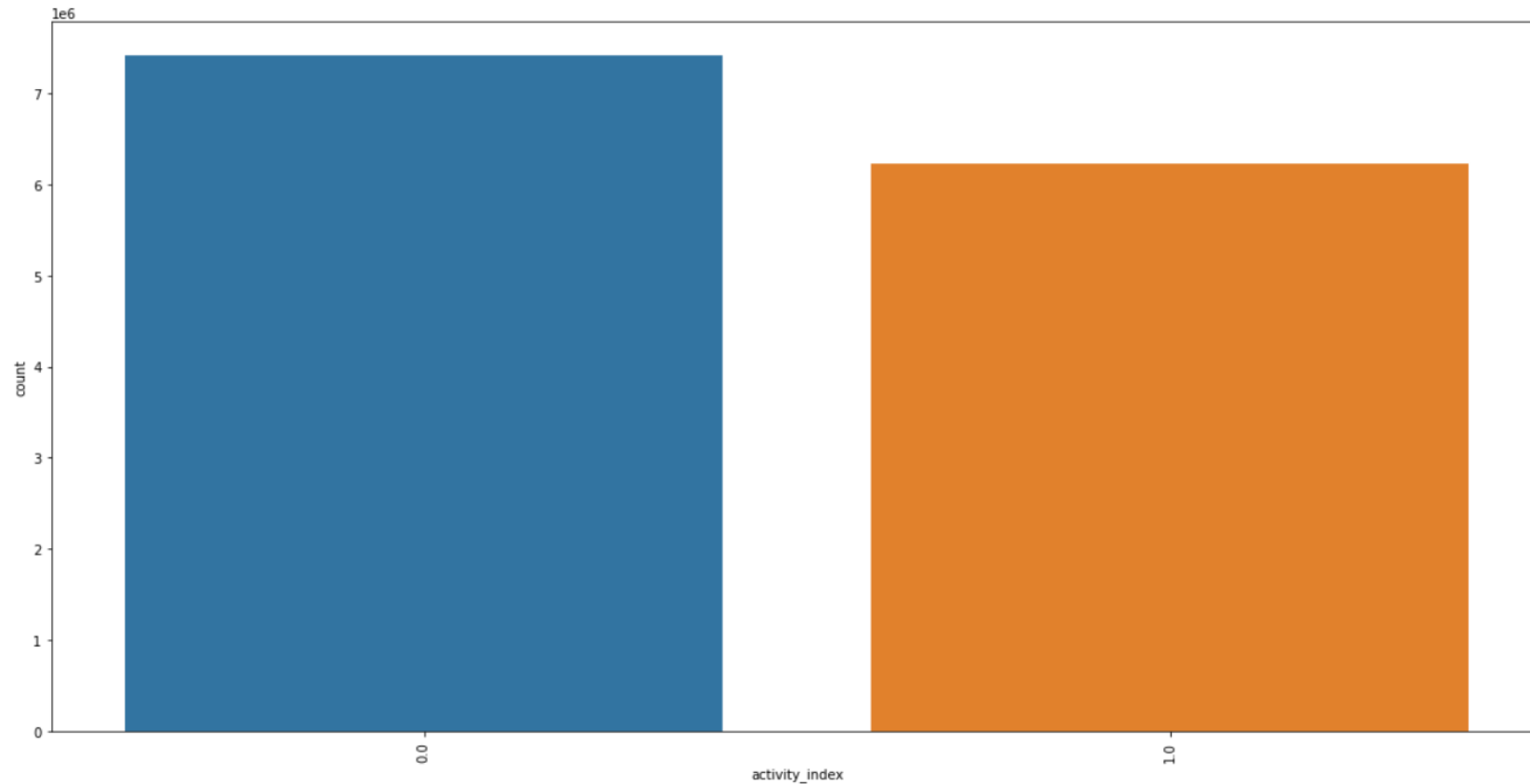


- The majority of the customers(residence) live in the same country as the bank.
- Majority of the foreigner have banks outside of the country..

Univariate Plots Analysis

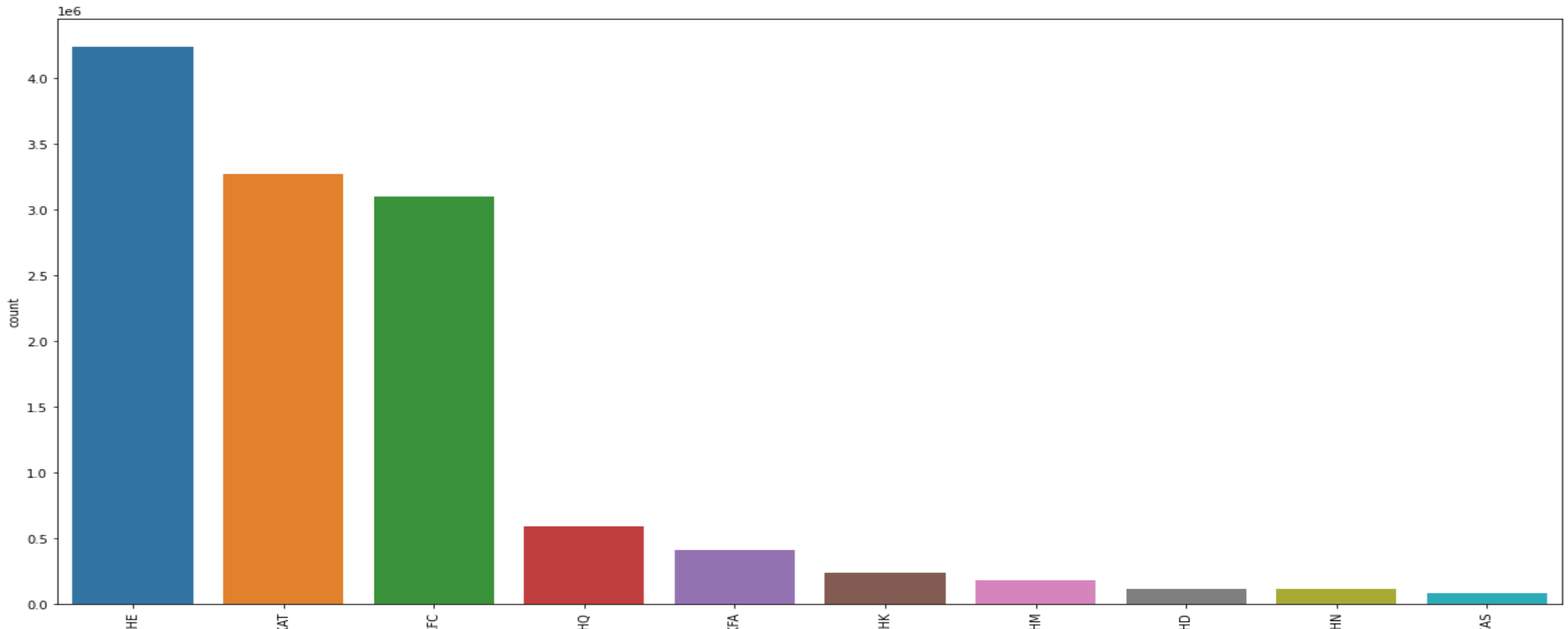
Activity Index (None-Active or Active Customers)

- There are more inactive customers than there are active customers.



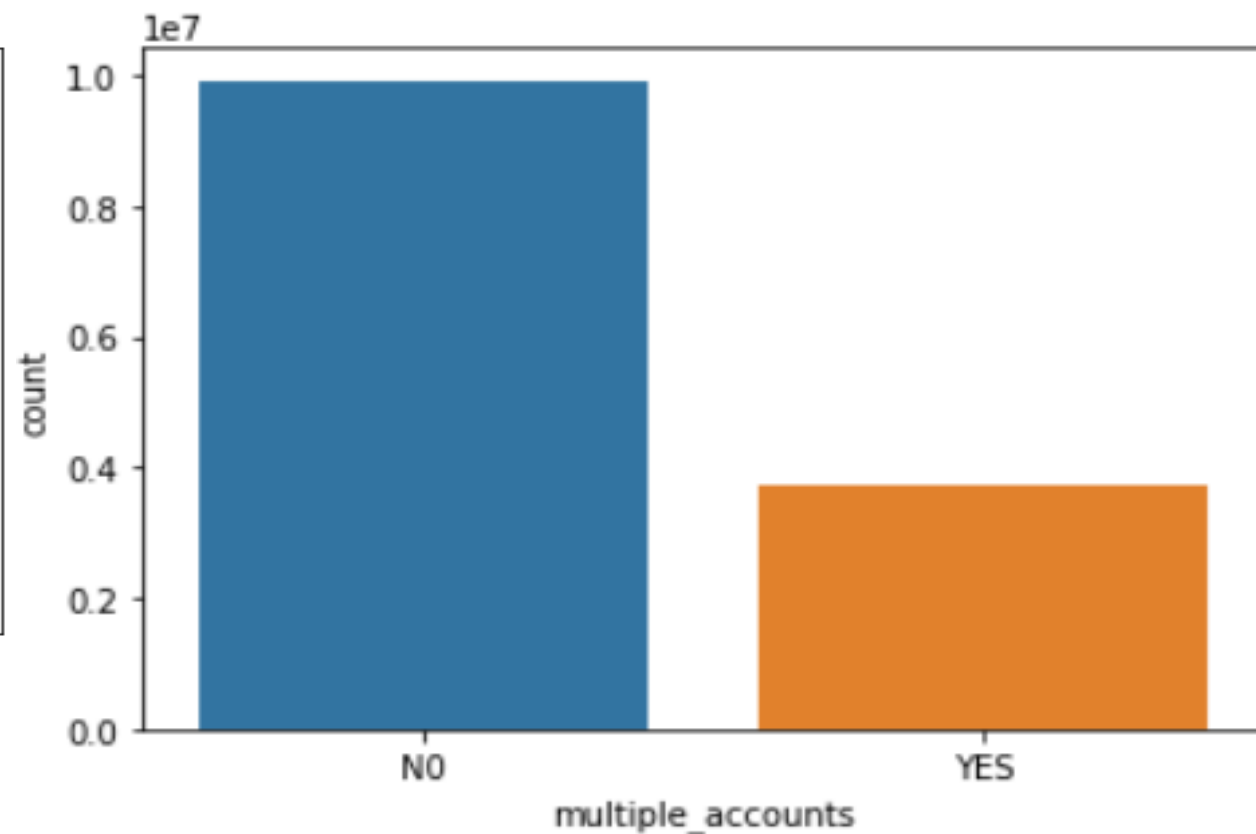
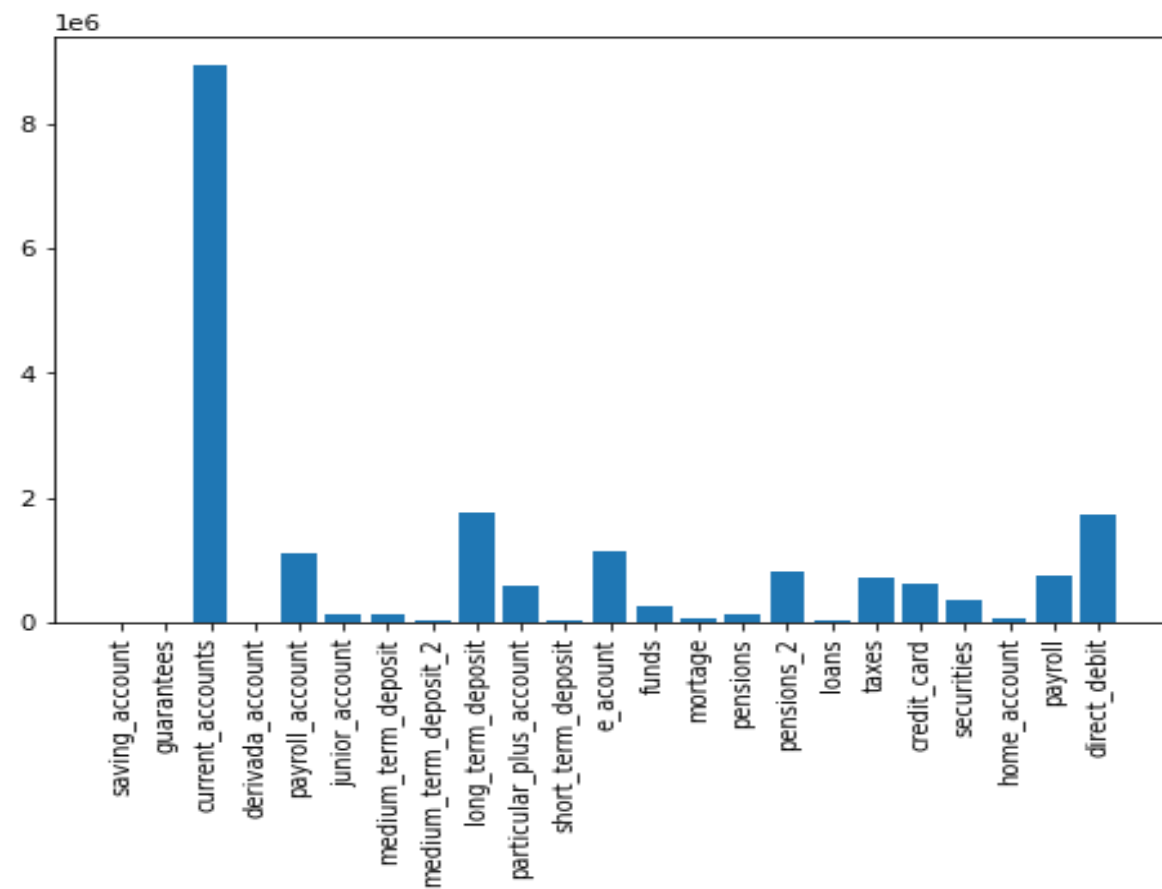
Univariate Plots Analysis

Top 10 Methods customers used to join



Univariate Plots Analysis

Customers that has more than one Account type.



Recommendation For Cross-Selling at XYZ Credit Union

Based on the exploratory data analysis ran, below are some recommendations that the analyst team at ABC analytics came up with to best help XYZ credit union use cross selling with existing customers.

- Use the mediums that the majority of customer sign up with to retarget them. Do ads that properly explain the services that XYZ credit union offers and how they will benefit an old client.
- Customer who currently only have one account – we recommend figuring out their needs based on age, income and other criteria then calling them for an appointment to suggest a personalized financial plan that is specifically suited for them. This will let XYZ credit union customers know that they are paying unique attention to their needs as clients.
- A complementary service or as we know it 'Bundle sales' can be provided for customers who add a new type of account. This helps the customers who are already a client to understand that things are changing and as such new customer are currently buying "x" and "y" together.

Recommendation For Cross-Selling at XYZ Credit Union

- With the exploratory data analysis completed XYZ credit union can also use a data driven suggestion to cross-sell. Like, for example; if customers are always coming into the bank trying to make payments or do transaction, it can be suggested that they download the app for ease of going back and forth. Also, if a 65-year-old customer updated her pension to say she has a new grand baby; we can suggest that she opens a saving account for the new baby.

Recommended Models for XYZ Credit Union Customer Dataset

Based on the Exploratory Data Analysis performed we recommend the below models be used on the dataset:

1. Segmentation and Logistics Regression Model – which will divide the broad dataset of customers into subsets of entities with common characteristics and homogeneous groups — then designing and implementing strategies specific to these segments makes easier decision making.
2. Response Model – Since the primary goal is to get existing customer to open other account types this model will be effective as a black box analytics. We will also used it as a regression model to predict what type of accounts a customer might open.

Five Models for XYZ Credit Union Customer Dataset

- The base models chosen were all classifiers and Regression.
1. Ada-Boosting Classifier - It combines multiple classifiers to increase the accuracy of classifiers. AdaBoost is an iterative ensemble method.
 2. GaussianNB- It's specifically used when the features have continuous values. It's also assumed that all the features are following a gaussian distribution.
 3. Random Forest Classifier - The algorithm uses randomness to build each individual tree to promote uncorrelated forests, which then uses the forest's predictive powers to make accurate decisions.

Five Models for XYZ Credit Union Customer Dataset

1. Decision Tree Classifier - On each step or node of a decision tree, used for classification, we try to form a condition on the features to separate all the labels or classes contained in the dataset to the fullest purity.
1. Logit Regression - Logistic regression analysis is used to examine the association of (categorical or continuous) independent variable(s) with one dichotomous dependent variable.

Best Prediction Models for XYZ Credit Union Customer Dataset

- After running all five models it was concluded that the Ada-Boosting model is the best one for predicting what type of accounts to recommend to users based on age and gross income. The Ada-Boosting Model provided the highest level of accuracy for the data frame.

Thank You 😊



Data Glacier

Your Deep Learning Partner