

CROSS SELLING RECOMMENDATIONS

Exploratory Data Analysis

DATA GLACIER VIRTUAL INTERNSHIP

GROUP NAME: HEGY

B. Harika

Yusuf Yuhan

Ebaghae Imhanlahimi

Gladys Kalas



INTRODUCTION:

- PROBLEM STATEMENT
- PROBLEM DESCRIPTION
- BUSINESS UNDERSTANDING

Problem Statement:

XYZ credit union in Latin America is performing very well in selling the Banking products (e.g.: 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 (Bank is not able to sell their other offerings to existing customer). XYZ Credit Union decided to approach ABC analytics to solve their problem

PROBLEM DESCRIPTION:

- XYZ Credit Union is a financial institution based in Latin America that offers a variety of banking products to its customers, including credit cards, deposit accounts, retirement accounts, and safe deposit boxes. While the credit union has been successful in selling these products individually, it has not been as successful in cross-selling its products to existing customers.
- Cross-selling is the process of selling additional products to existing customers who have already purchased a product or service from the same organization. It is an important strategy for businesses because it can increase revenue, improve customer loyalty, and reduce marketing costs.
- The lack of success in cross-selling suggests that there may be several barriers preventing XYZ Credit Union from selling additional products to its existing customers. These barriers could include a lack of awareness of the products, low perceived value, inadequate marketing efforts, or a lack of trust in the credit union.
- To address this problem, XYZ Credit Union has decided to work with ABC Analytics, a data analytics consulting firm, to identify the barriers to cross-selling and develop strategies to overcome them. ABC Analytics will work with the credit union to analyze customer data, identify patterns and trends, and develop targeted marketing campaigns that are designed to increase awareness and interest in the credit union's other banking products.

Business Understanding:

- The banking industry is highly competitive, and it is crucial for banks to maintain their existing customers while attracting new ones. Cross selling is one of the strategies that banks use to increase their revenue by offering additional products to their existing customers. The benefits of cross-selling are two-fold: firstly, it increases the revenue per customer, and secondly, it strengthens the relationship between the bank and the customer.
- However, cross-selling is not an easy task, and banks need to carefully plan and execute their cross-selling strategies. Banks need to identify the right products to offer to the right customers at the right time. This requires a deep understanding of the customer's financial needs, behavior, and preferences.
- Moreover, banks need to be mindful of their customer's privacy and should not be seen as intrusive. If customers feel that the bank is pushing products aggressively, they may end up leaving the bank. Therefore, it is essential to strike a balance between cross-selling and customer satisfaction.
- In summary, the key to successful cross-selling is to have a deep understanding of customer needs, offer the right products at the right 4 Cross-Selling Recommendation time, and maintain a balance between cross-selling and customer satisfaction.

Data Overview:

File name	Train.csv
No. of observations	13393927
No. features	48
File Size	3.005 GB
File type	CSV file
No. of files	1

Data Quality Highlights:

The data for the EDA is a cleaned data set which is:

- Free from duplicates
- Checked for missing values and handled the missing values
- Detected outliers and treated the outliers with appropriate methods
- Checked the data types and appropriate conversions in places of incorrect types
- Organized feature headings (from Spanish language to English)

DATA CLEANING HIGHLIGHTS:

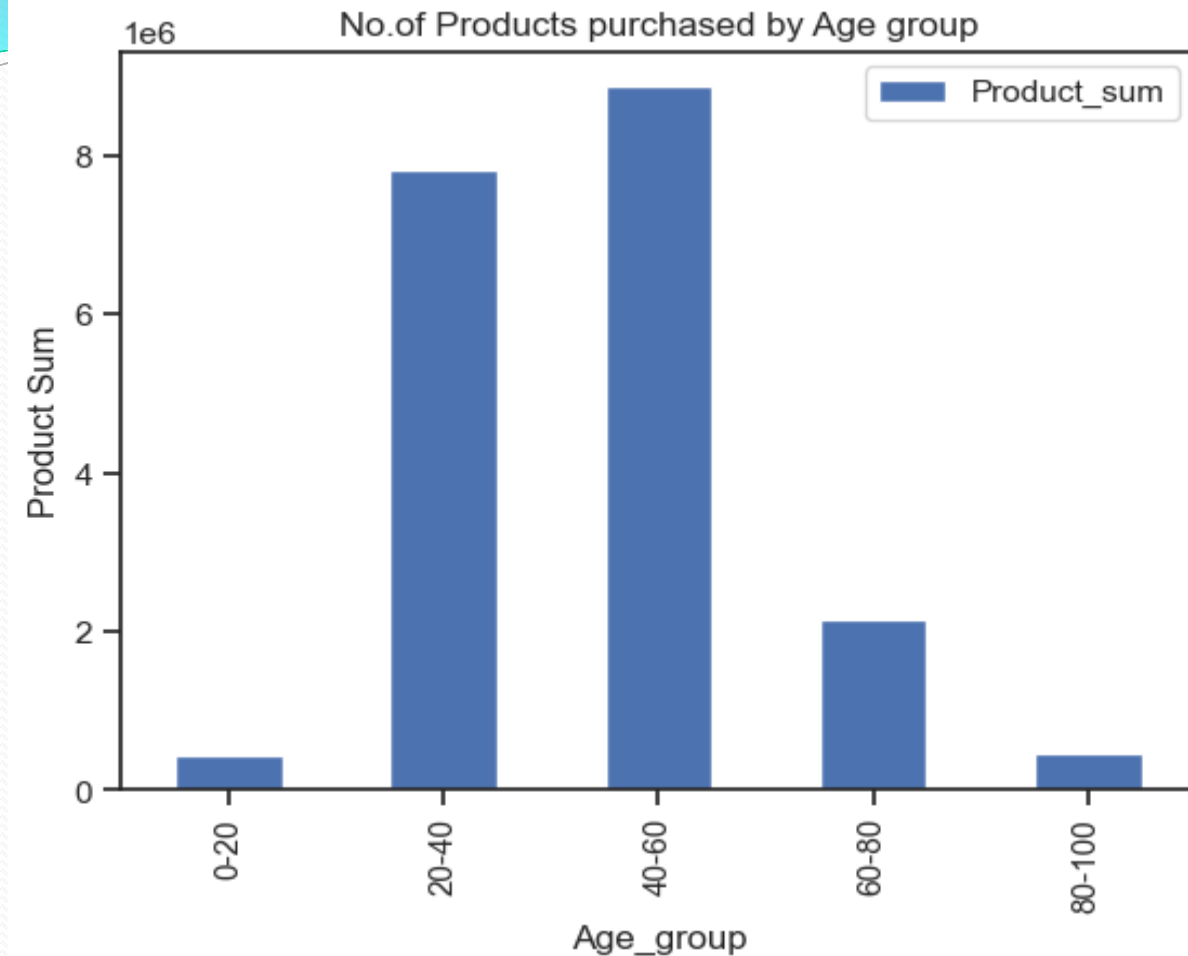
The data received is in its raw format with language in Spanish literacy, erroneous inputs, incompatible data types which needed to be addressed to obtain a clean organized data for further analysis

- Changed the feature names from Spanish literacy to English for easy understanding.
- Dropped some null values and duplicates
- Checked for null values in the data set Train.csv and imputed some features' missing /null values with the Mean, Median, Mode, ffill and bfill techniques for the respective columns.
- Replaced some features with their equivalent category values interpretations.

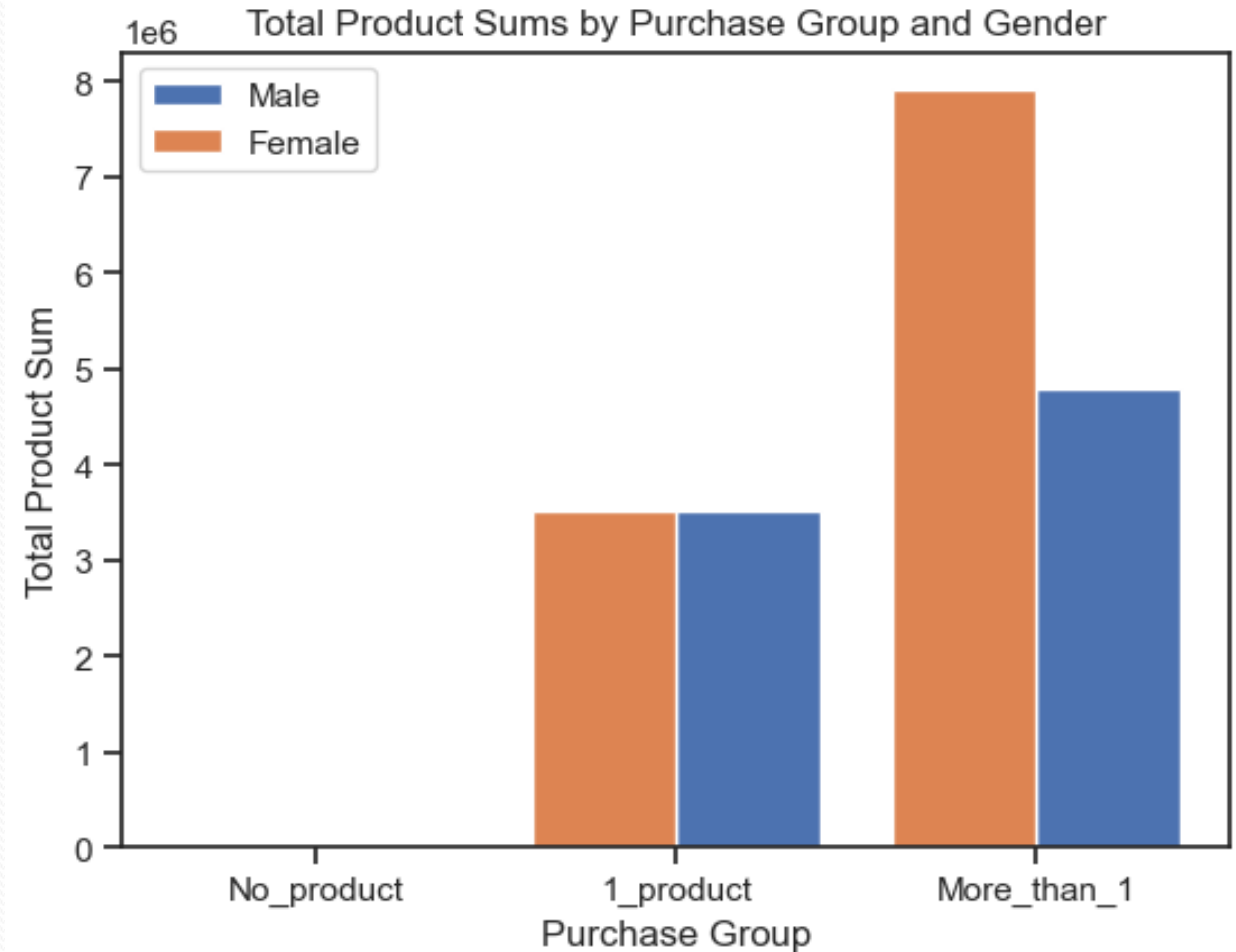
Example- ['H'-Male, 'V'- 'Female']

['S' to 'Yes' and 'N' to 'No']

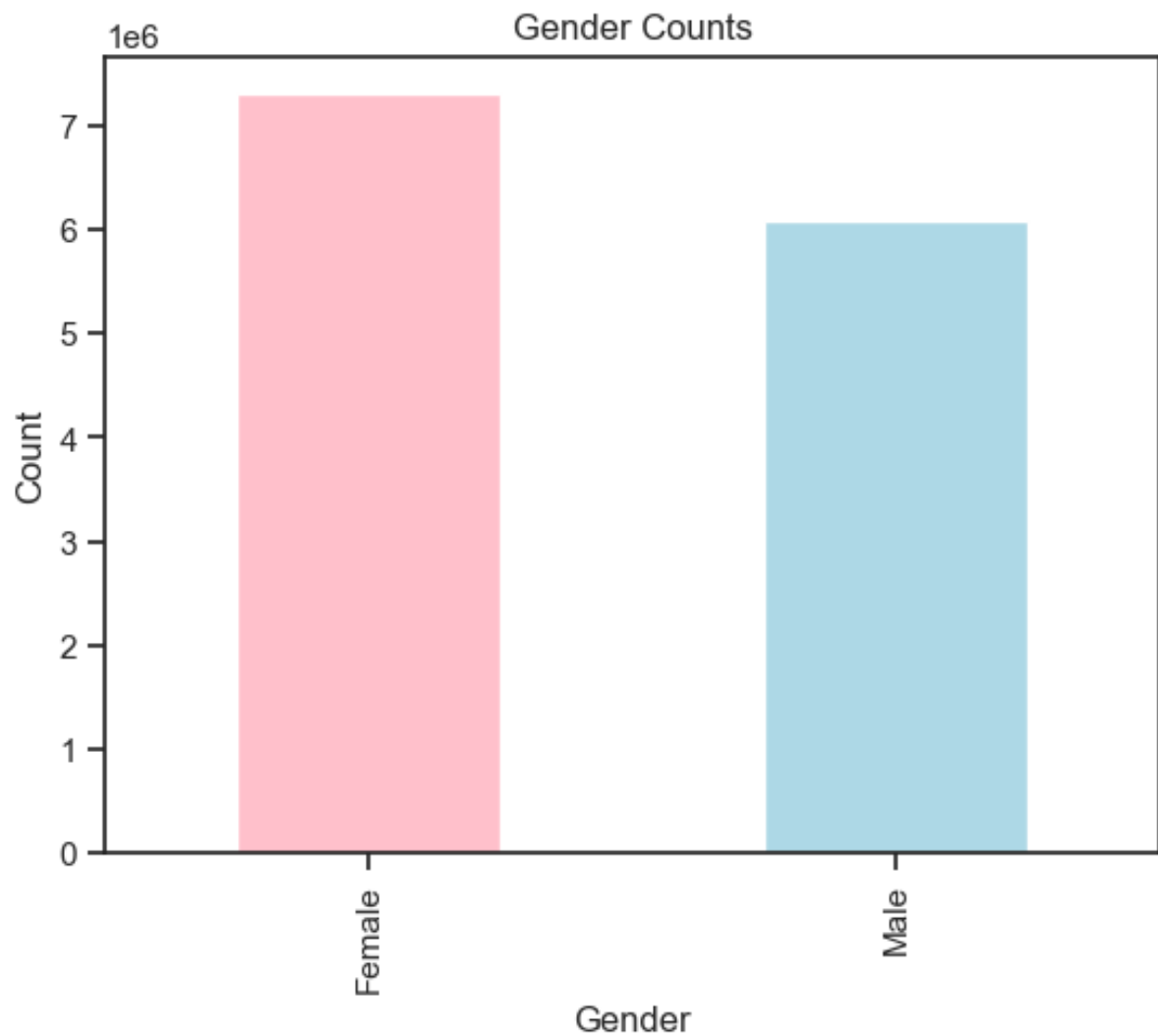
- Detection of outliers and treating Outliers in the data using Standard deviation, IQR, Percentile methods and treating outliers by capping, clipping, trimming outliers.



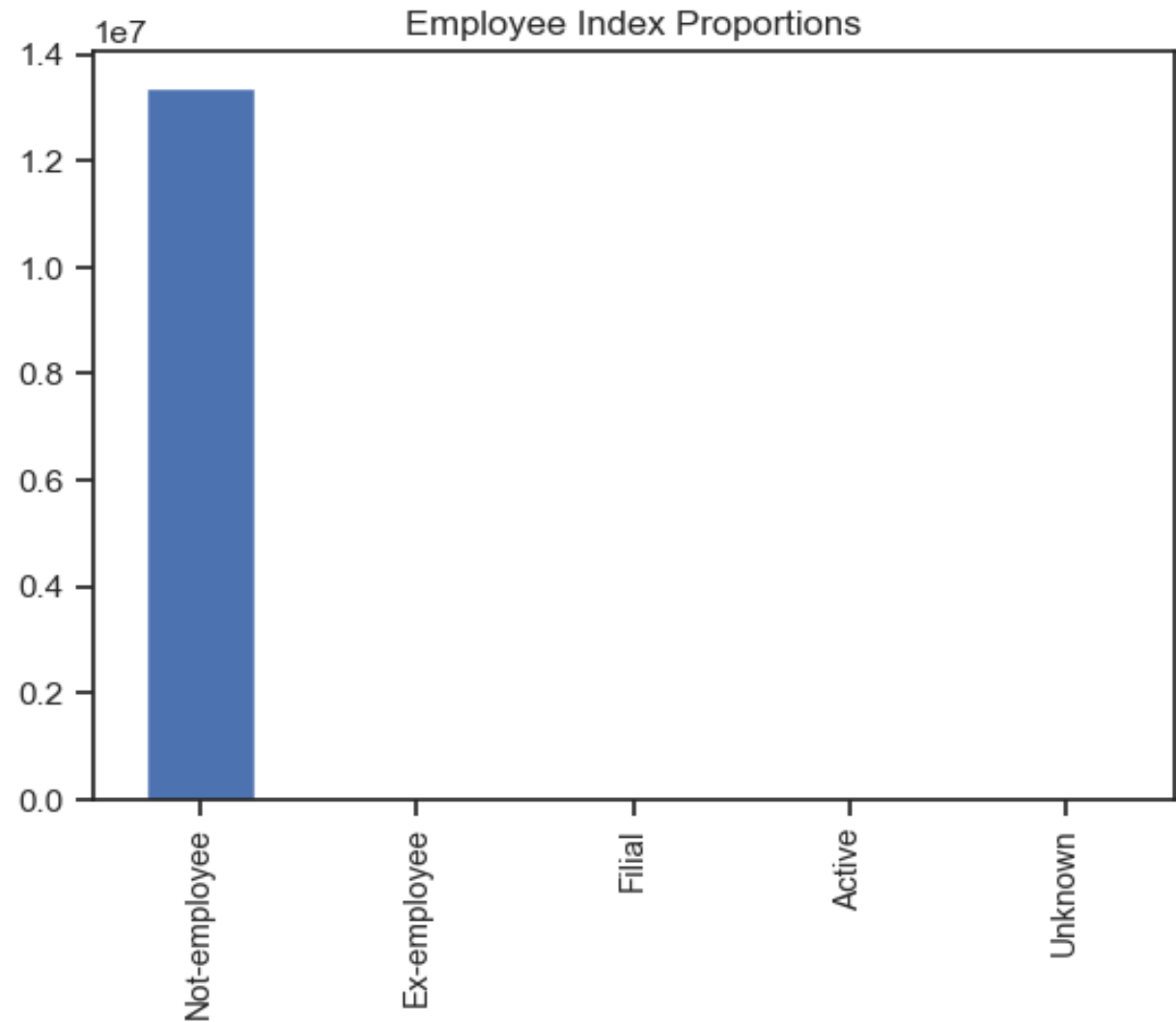
- From the bar graph, it was observed that Age Group of 40-60 has purchased more products than others.
- Age group of 20-40 has second highest purchases than other age groups.



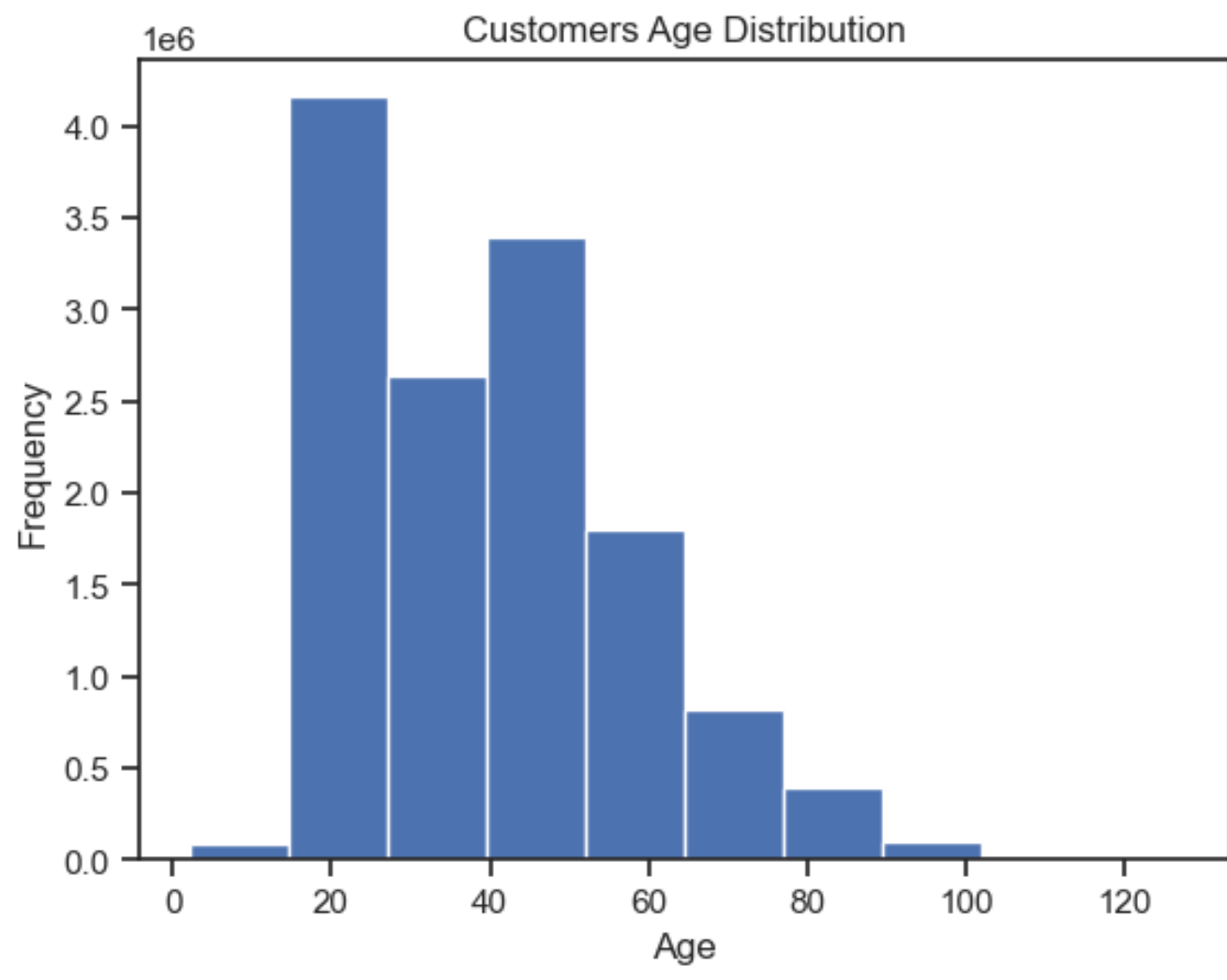
- From the bar graph, both Male and Female has purchased maximum products of “More_than_1”, where Male customers has higher count of purchases than Female customers.
- 1_product has equal no. of purchases among both genders.



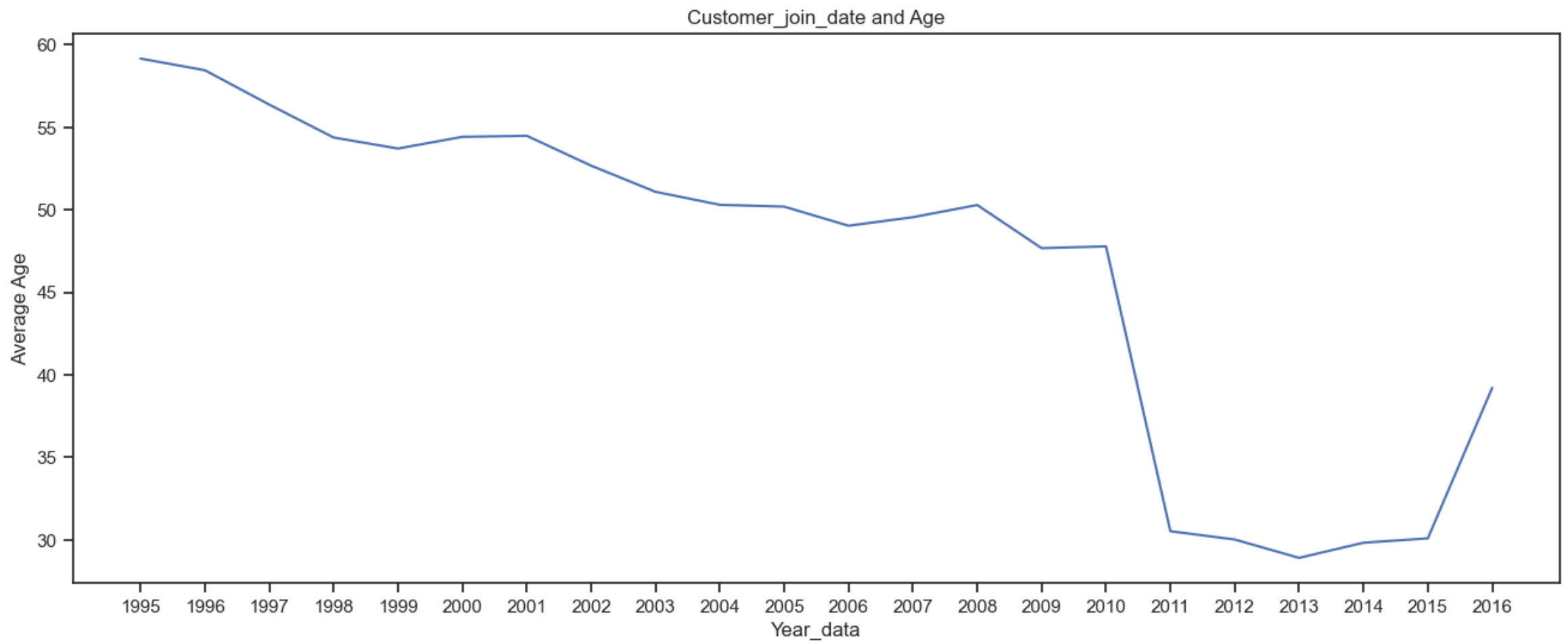
- From graph, Females customers are higher in the company than Male customers.



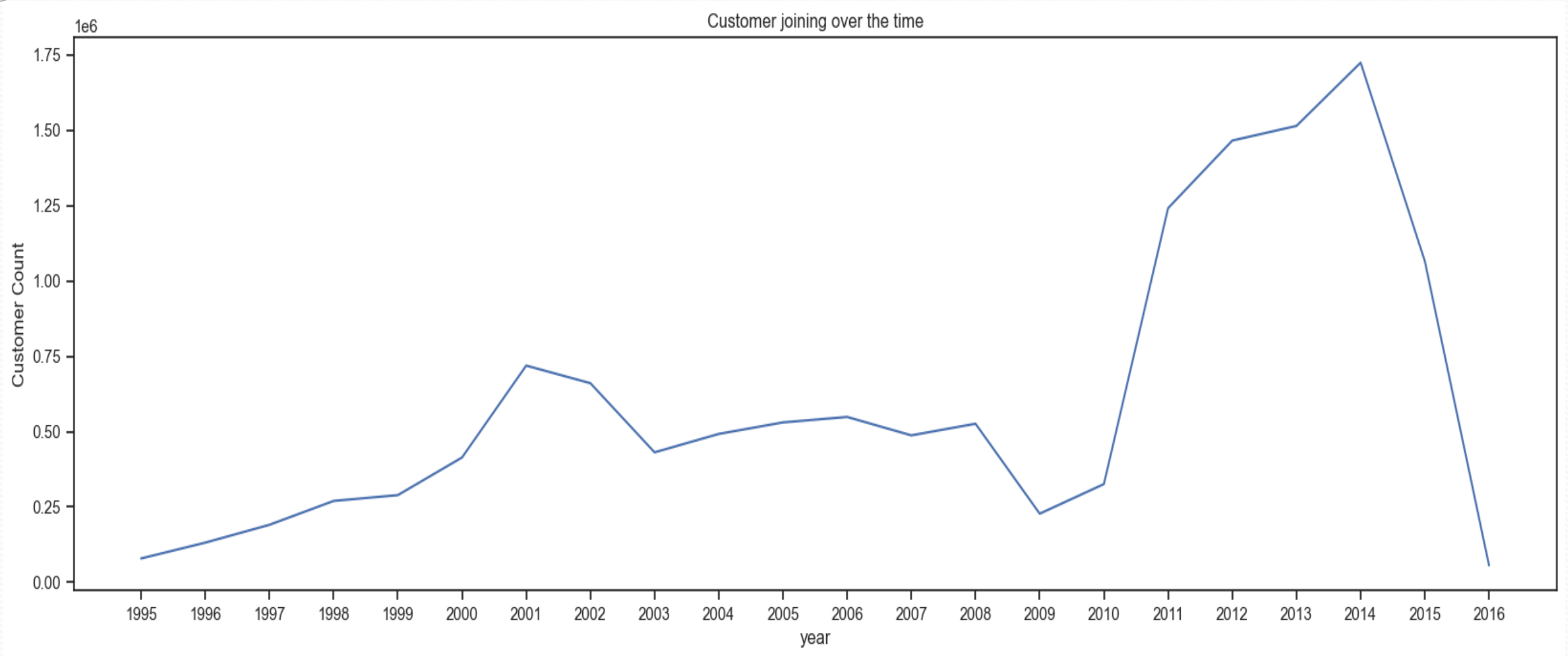
- From graph, we can conclude that Majority of customers are Not-employee.



- The graph shows that the majority of the customers in the bank are in their 20's and second majority of customers are in 50's.

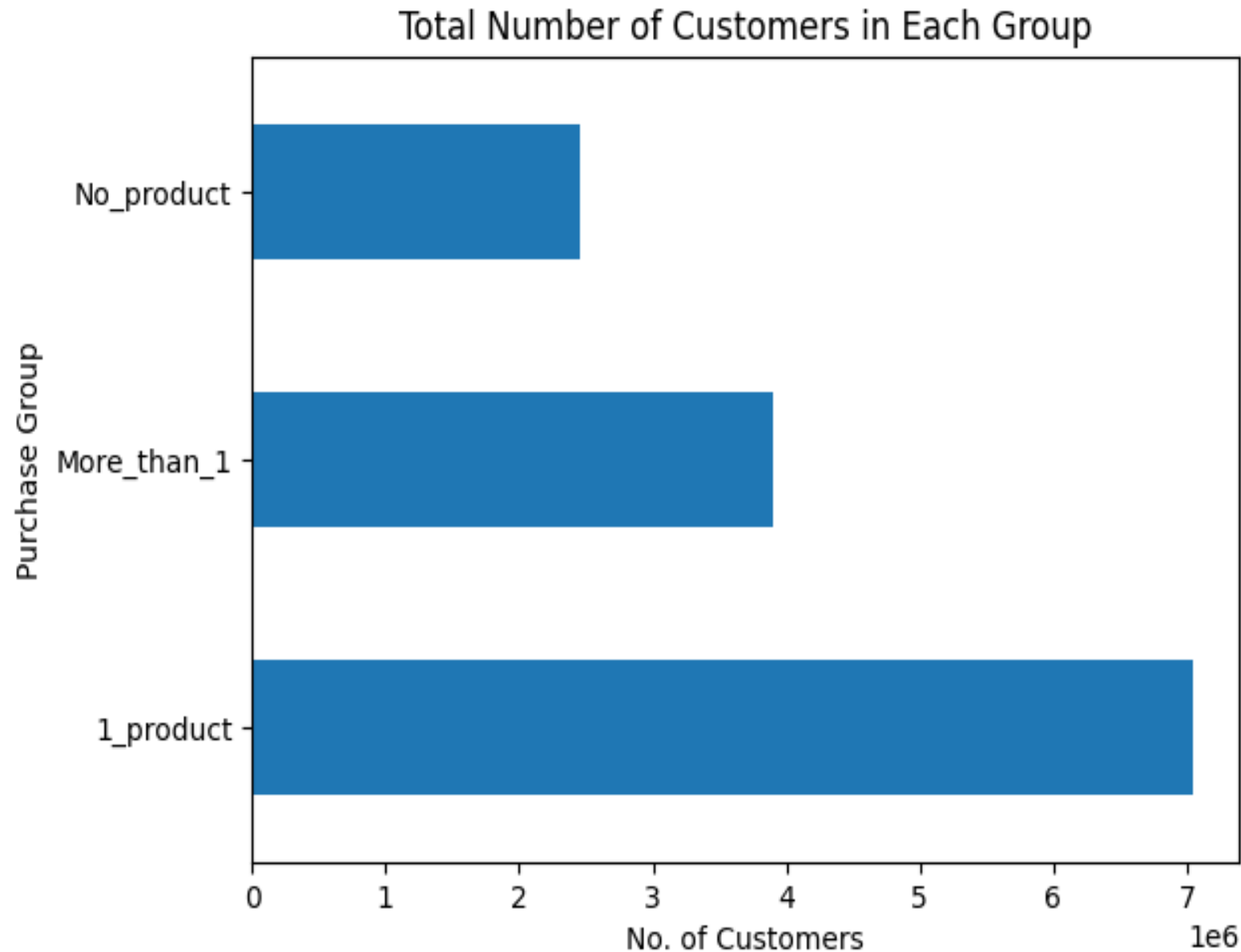


- From graph of Average Age and Customer Join Date, it can be observed that the Average Age of customers who have joined the bank was highest in the year 1995 and gradually the customers joining the bank is decreased. The sudden fall of Average Age of customers is seen in between year 2010 – 2011, and bank started seeing the increase of Customers age in year 2015.



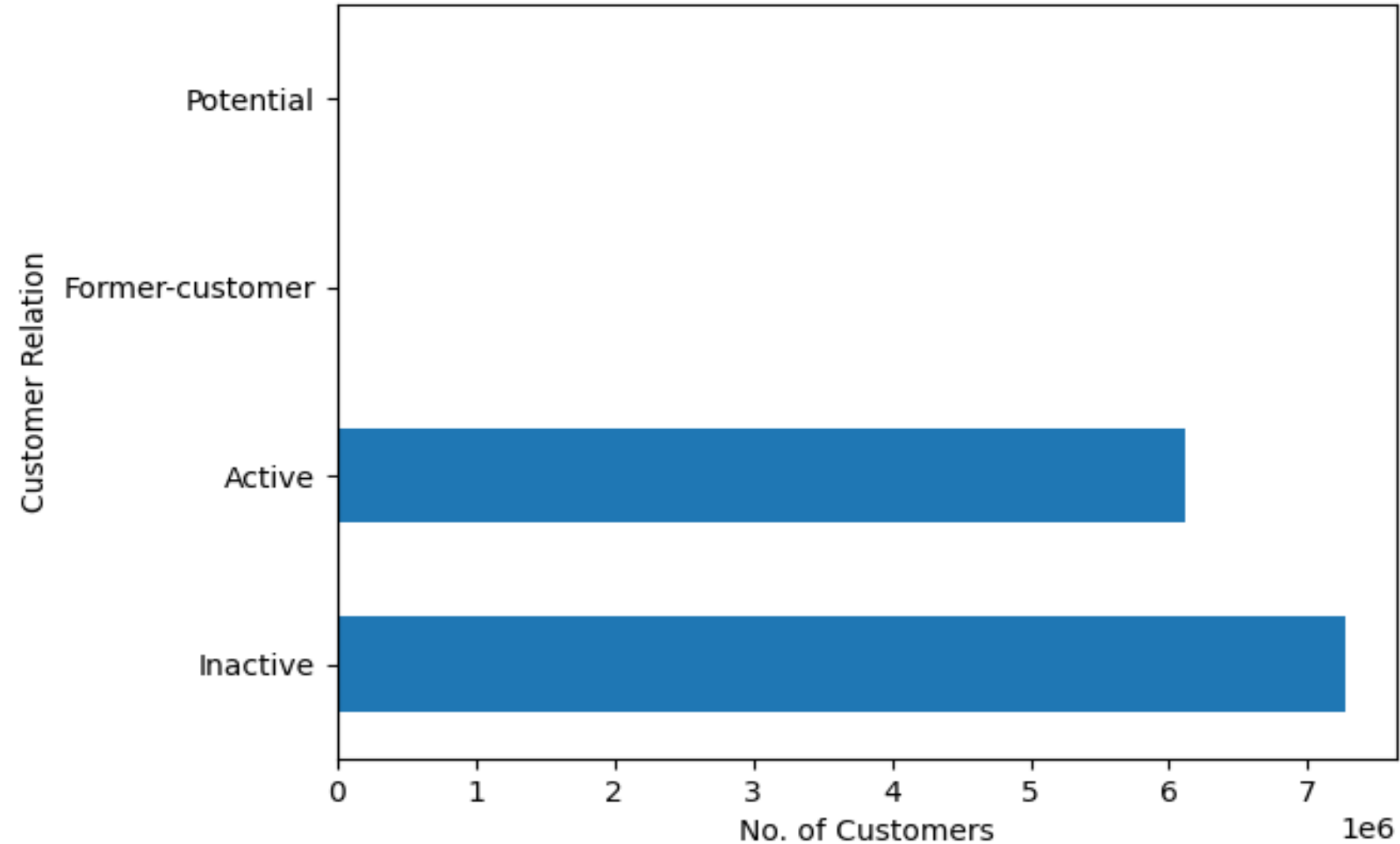
- Based on graph of Customer Join Date and Customer Count, In starting year 1995 the no. of customers joining is very less, Customers joining the bank has been gradually increased after 1995. In year 2011 there was sudden increase in customers joining the bank and sudden drop started from year 2015-2016.

- The bar chart shows the total number of customers grouped according to the amount of products they buy.
- About 7 million customers purchase only one product, about 4 million customers purchase more than one product and about 2 million customers purchase no product.

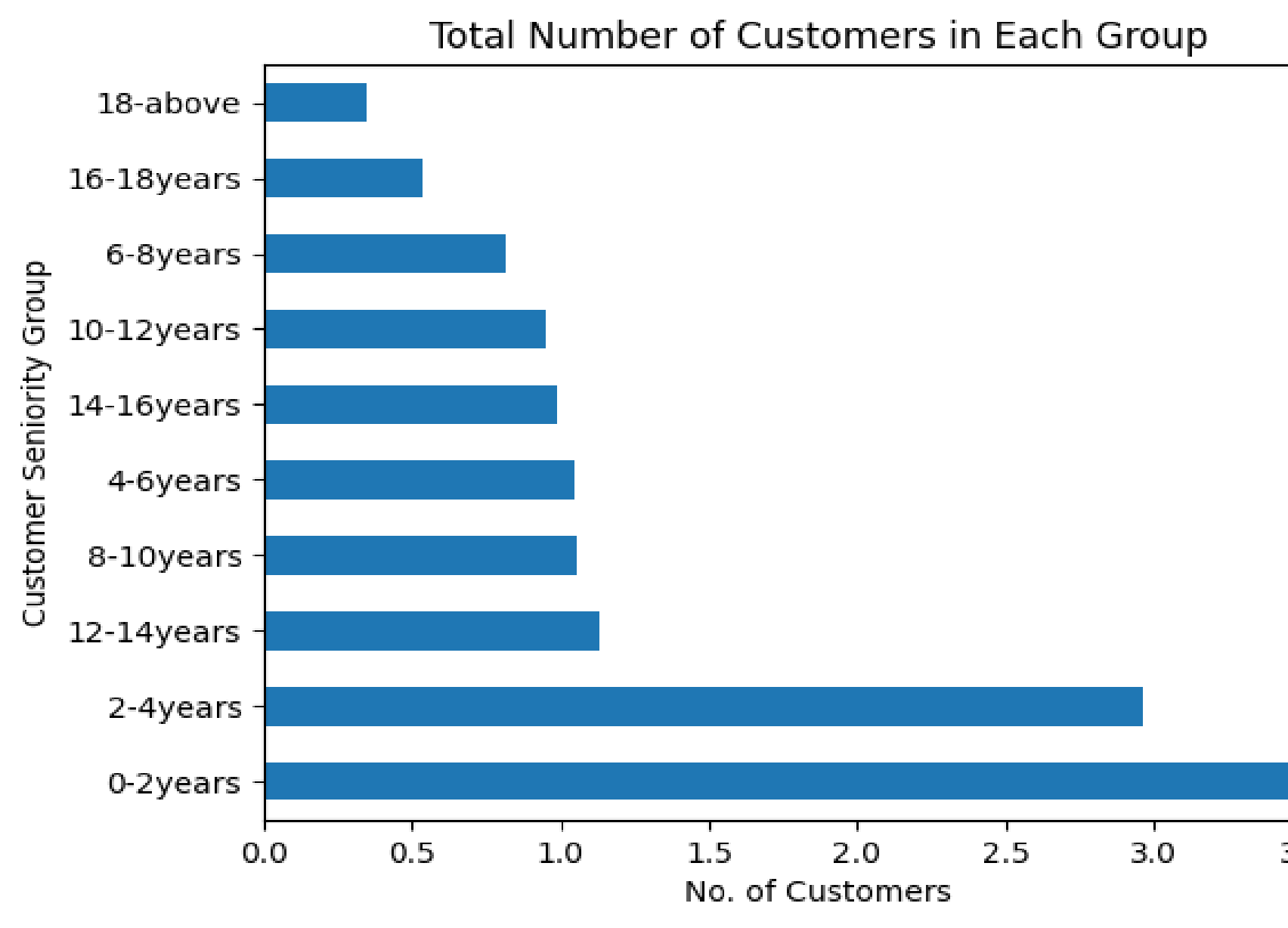




Customers in Customer Relation Classification



- The bar chart shows the total number of customers in each customer relation category.
- It shows that majority of customers are inactive.



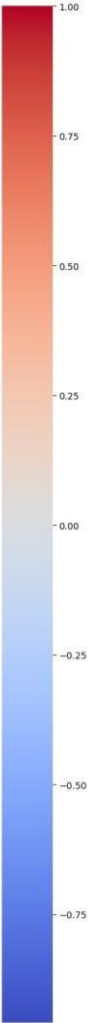
- The bar chart shows the total number of customers grouped according to the period of time they have spent with the bank.
- It shows that majority of customers have been in the bank for less than 4years.
- Also a small amount of customers have been in the bank for at least 18years.

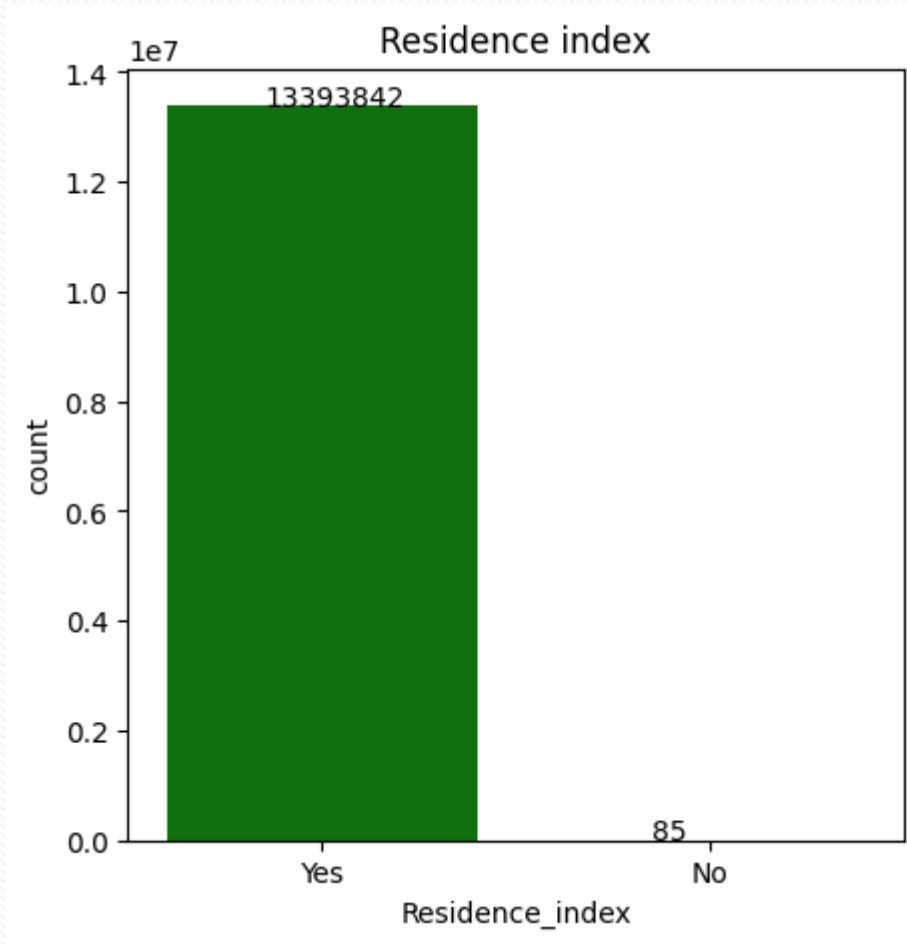
Correlation between Products

Unnamed: 0	1	0.094	-0.02	-0.028	0.0018
Customer_code	0.094	1	-0.61	-0.96	0.016
Age	-0.02	-0.51	1	0.6	0.00071
Customer_seniority	-0.028	-0.96	0.6	1	-0.018
Primary_customer	-0.0018	0.016	0.00071	-0.018	1
Primary_address					
Customer_address	-0.0039	-0.041	0.052	0.042	0.00023
Gross_income	0.035	0.011	0.016	0.002	0.019
Saving_account	-0.0021	-0.013	0.005	0.014	-0.00038
Guarantees	-0.0012	-0.0055	0.00086	0.0058	-0.00018
Current_accounts	-0.12	0.17	-0.17	-0.17	-0.38
Derivative_account	-0.0029	-0.019	0.013	0.019	-0.00074
Payroll_account	-0.027	-0.13	0.068	0.13	-0.098
Junior_account	-0.016	-0.016	-0.15	0.016	-0.0018
More_private_account	-0.013	0.023	0.036	-0.045	0.00013
Private_account	-0.067	-0.45	0.32	0.47	-0.012
Private_plus_account	-0.04	-0.16	0.12	0.15	-0.0067
Short_term_deposits	-0.025	0.02	0.019	-0.023	0.0084
Medium_term_deposits	-0.012	-0.035	0.028	0.035	-0.00083
Long_term_deposits	-0.039	-0.14	0.18	0.12	-0.0068
E_account	-0.026	-0.15	0.14	0.13	-0.0078
Funds	-0.02	-0.14	0.12	0.15	-0.0045
Mortgage	-0.015	-0.098	0.038	0.1	-0.0027
Pensions	-0.015	-0.12	0.066	0.13	-0.0035
Loans	-0.0089	-0.036	0.024	0.036	-0.0016
Taxes	-0.021	-0.19	0.096	0.19	-0.0042
Credit_card	-0.036	-0.19	0.1	0.19	-0.0073
Securities	-0.021	-0.17	0.11	0.18	-0.0054
Home_account	-0.012	-0.075	0.017	0.08	-0.0023
Payroll	-0.026	-0.12	0.043	0.12	-0.0079
Pensions_2	-0.027	-0.13	0.071	0.13	-0.0082
Direct_debit	-0.036	-0.14	0.088	0.14	-0.0093

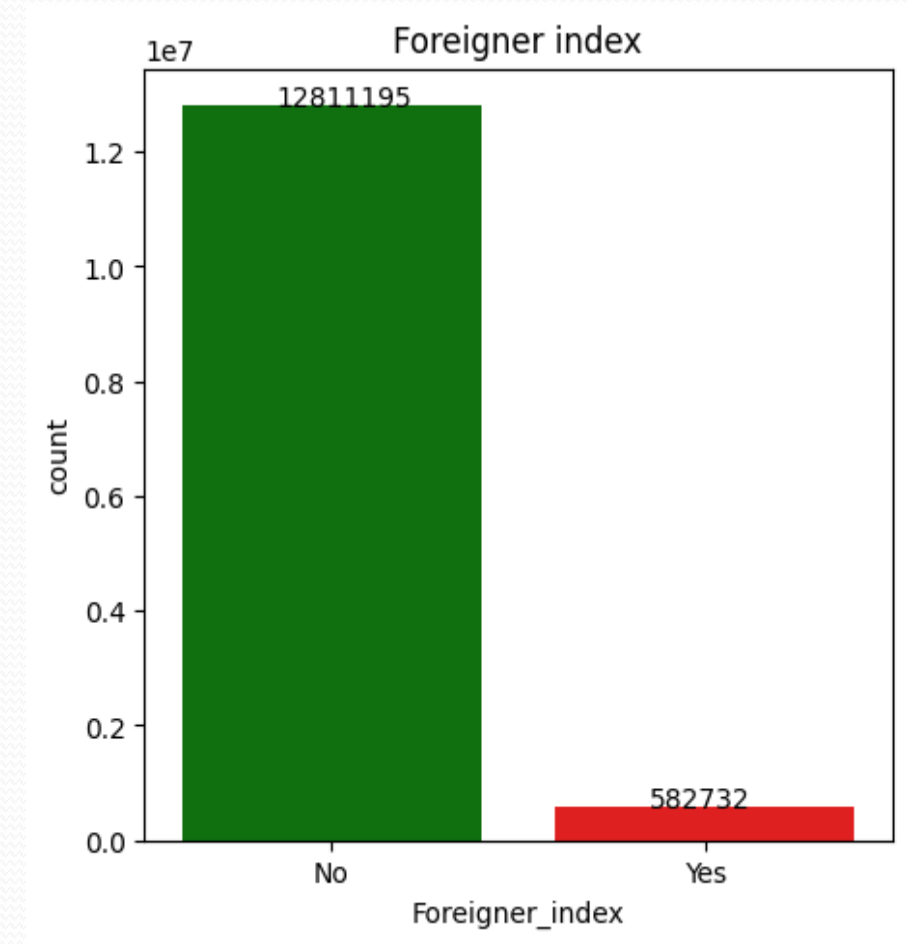
Correlation Matrix for Banking Products																									
-0.0039	0.035	-0.0021	-0.0012	-0.12	-0.0029	-0.027	-0.016	-0.013	-0.067	-0.04	-0.025	-0.012	-0.039	-0.026	-0.02	-0.015	-0.015	-0.0089	-0.021	-0.036	-0.021	-0.012	-0.026	-0.027	-0.036
-0.041	0.011	-0.013	-0.0055	0.17	-0.019	-0.13	-0.016	0.023	-0.45	-0.16	0.02	-0.035	-0.14	-0.15	-0.14	-0.098	-0.12	-0.036	-0.19	-0.19	-0.17	-0.075	-0.12	-0.13	-0.14
0.052	0.016	0.005	0.00086	-0.17	0.013	0.068	-0.15	0.036	0.32	0.12	0.019	0.028	0.18	0.14	0.12	0.038	0.066	0.024	0.096	0.1	0.11	0.017	0.043	0.071	0.088
0.042	0.002	0.014	0.0058	-0.17	0.019	0.13	0.016	-0.045	0.47	0.15	-0.023	0.035	0.12	0.13	0.15	0.1	0.13	0.036	0.19	0.19	0.18	0.08	0.12	0.13	0.14
0.00023	0.019	-0.00038	0.00018	-0.038	-0.00074	-0.0098	-0.0018	0.00013	-0.012	-0.0067	0.0084	-0.00083	-0.0068	-0.0078	-0.0045	-0.0027	-0.0035	-0.0016	-0.0042	-0.0073	-0.0054	-0.0023	-0.0079	-0.0082	-0.0093

1	-0.0039	0.0014	0.00053	-0.0057	0.00088	0.012	0.0093	0.0071	0.042	0.015	0.0013	0.0043	0.016	0.019	0.007	-0.0035	0.0034	-0.0084	0.0071	0.009	0.013	0.0041	0.01	0.011	0.015
-0.0039	1	0.0008	0.00032	-0.015	0.00037	-0.011	-0.00061	0.0065	0.0064	0.0021	0.0014	-0.0011	-0.0016	-0.0038	0.0092	-0.0048	0.0069	-0.0053	0.0054	0.0029	0.012	0.0016	-0.01	-0.0094	-0.011
0.0014	0.0008	1	4.8e-05	-0.0011	-0.0002	0.0011	-0.001	-0.00071	0.0096	0.0042	-0.00042	0.00042	0.002	0.0015	0.0013	0.0024	0.0056	-0.00052	0.0077	0.0014	0.0063	0.014	0.0002	5.3e-05	0.0032
0.00053	0.00032	4.8e-05	1	-0.0021	-9.4e-05	0.011	-0.00047	0.0018	0.002	0.009	-0.0002	0.00058	0.0017	0.012	0.0041	-0.00037	0.00046	0.00024	0.011	0.01	0.012	-0.0003	0.012	0.012	0.012
-0.0057	-0.015	-0.0011	-0.0021	1	0.0032	-0.35	-0.14	-0.025	0.0011	-0.053	-0.028	0.012	0.009	-0.08	0.0091	-0.048	-0.032	-0.027	-0.09	-0.099	0.012	-0.0059	-0.26	-0.27	-0.073
0.00088	0.00037	-0.0002	-9.4e-05	0.0032	1	0.0083	-0.002	-0.0016	0.012	0.0071	0.0007	0.0011	0.0043	0.01	0.017	0.0015	0.013	0.0015	0.011	0.0094	0.033	0.0048	0.0066	0.0072	0.0091
0.012	-0.011	0.0011	0.011	-0.35	0.0083	1	-0.029	0.025	0.02	0.19	-0.0024	0.015	0.095	0.31	0.086	0.16	0.1	0.0029	0.32	0.39	0.12	0.03	0.76	0.79	0.54
0.0093	-0.0061	-0.001	-0.00047	-0.14	-0.002	-0.029	1	-0.0097	-0.038	-0.021	-0.0041	-0.004	-0.02	-0.03	-0.014	-0.0076	-0.0095	-0.005	-0.023	-0.021	-0.016	-0.0062	-0.024	-0.024	-0.038
0.0071	0.0065	-0.00071	0.0018	-0.025	-0.0016	0.025	-0.0097	1	-0.036	-0.018	0.017	-0.00068	0.008	-0.011	-0.002	-0.0051	-0.0026	-0.0041	0.014	0.021	0.0018	-0.0057	0.022	0.022	0.029
0.042	0.0064	0.0096	0.002	0.0011	0.012	0.02	-0.038	-0.036	1	0.037	0.00098	0.04	0.077	-0.012	0.086	0.019	0.049	-0.0087	0.062	0.063	0.096	0.063	0.027	0.027	0.041
0.015	0.0021	0.0042	0.009	-0.053	0.0071	0.19	-0.021	-0.018	0.037	1	8.5e-05	0.023	0.074	0.067	0.068	0.071	0.063	-0.0017	0.17	0.17	0.093	0.041	0.17	0.18	0.2
0.0013	0.0014	-0.00042	-0.0002	-0.0028	0.0007	-0.0024	-0.0041	0.017	0.00098	8.5e-05	1	0.05	0.015	-0.0031	0.0038	-0.0012	-0.0019	-0.0021	-0.0011	-0.0011	0.0014	-0.0011	-0.00096	-0.0012	-0.0046
0.0043	-0.0011	-0.00042	0.00058	0.012	0.0011	0.015	-0.004	-0.00068	0.04	0.023	0.05	1	0.079	0.017	0.03	0.0046	0.011	-0.0015	0.012	0.016	0.018	0.0072	0.016	0.016	0.015
0.016	-0.0016	0.002	0.0017	0.009	0.0043	0.095	-0.02	0.008	0.077	0.074	0.015	0.079	1	0.22	0.18	0.016	0.061	-0.0095	0.064	0.11	0.082	0.0068	0.078	0.09	0.093
0.019	-0.0038	0.0015	0.012	-0.08	0.01	0.31	-0.03	-0.011	-0.012	0.067	-0.0031	0.017	0.22	1	0.14	0.067	0.091	-0.0079	0.2	0.26	0.15	0.01	0.26	0.28	0.3
0.007	0.0092	0.0013	0.0041	0.0091	0.017	0.086	-0.014	-0.002	0.086	0.068	0.0038	0.03	0.18	0.14	1	0.019	0.13	0.0018	0.07	0.1	0.18	0.018	0.073	0.082	0.093
-0.0035	-0.0048	0.0024	-0.00037	-0.048	0.0015	0.16	-0.0076	-0.0051	0.019	0.071	-0.0012	0.0046	0.016	0.067	0.019	1	0.041	0.011	0.09	0.12	0.045	0.006	0.16	0.16	0.17
0.0034	0.0069	0.0056	-0.00046	-0.032	0.013	0.1	-0.0095	-0.0026	0.049	0.063	-0.0019	0.011	0.061	0.091	0.13	0.041	1	0.006	0.083	0.11	0.12	0.022	0.099	0.1	0.1
-0.0084	-0.0053	-0.00052	-0.00024	-0.027	0.0015	0.0029	-0.005	-0.0041	-0.0087	-0.0017	-0.0021	-0.0015	-0.0095	-0.0079	0.0018	0.011	0.006	1	0.03	0.012	0.0087	-0.0003	0.0051	0.0054	0.0045
0.0071	0.0054	0.0077	0.011	-0.09	0.011	0.32	-0.023	0.014	0.062	0.17	-0.0011	0.012	0.064	0.2	0.07	0.09	0.083	0.03	1	0.27	0.11	0.053	0.28	0.29	0.35
0.009	0.0029	0.0014	0.01	-0.099	0.0094	0.39	-0.021	0.021	0.063	0.17	-0.0011	0.016	0.11	0.26	0.1	0.12	0.11	0.012	0.27	1	0.14	0.034	0.36	0.37	0.37
0.013	0.012	0.0063	0.012	0.012	0.033	0.12	-0.016	0.0018	0.096	0.093	0.0014	0.018	0.082	0.15	0.18	0.045	0.12	0.0087	0.11	0.14	1	0.033	0.11	0.12	0.14
0.0041	0.0016	0.014	-0.0003	-0.0059	0.0048	0.03	-0.0062	-0.0057	0.063	0.041	-0.0011	0.0072	0.0068	0.01	0.018	0.006	0.022	-0.0003	0.053	0.034	0.033	1	0.03	0.029	0.036
0.01	-0.01	0.0002	0.012	-0.26	0.0066	0.76	-0.024	0.022	0.027	0.17	-0.00096	0.016	0.078	0.26	0.073	0.16	0.099	0.0051	0.28	0.36	0.11	0.03	1	0.96	0.47
0.011	-0.0094	5.3e-05	0.012	-0.27	0.0072	0.79	-0.024	0.022	0.027	0.18	-0.0012	0.016	0.09	0.28	0.082	0.16	0.1	0.0054	0.29	0.37	0.12	0.029	0.96	1	0.49
0.015	-0.011	0.0032	0.012	-0.073	0.0091	0.54	-0.038	0.029	0.041	0.2	-0.0046	0.015	0.093	0.3	0.093	0.17	0.1	0.0045	0.35	0.37	0.14	0.036	0.47	0.49	1

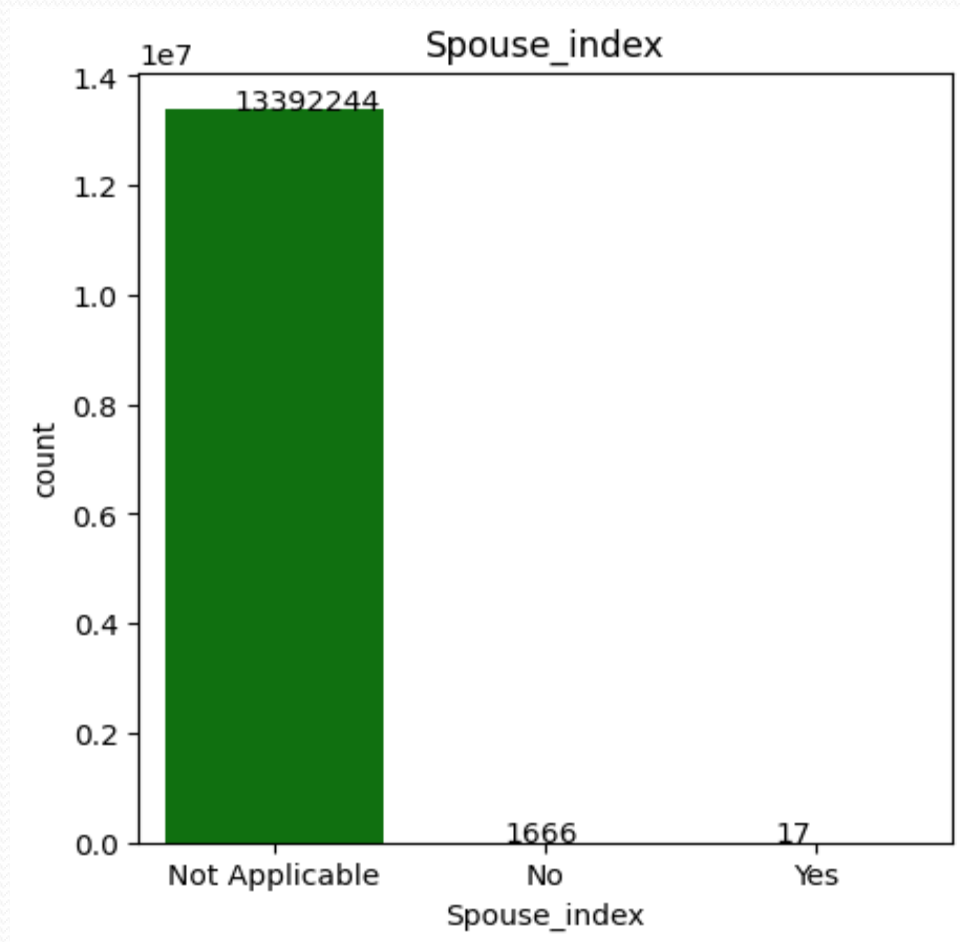




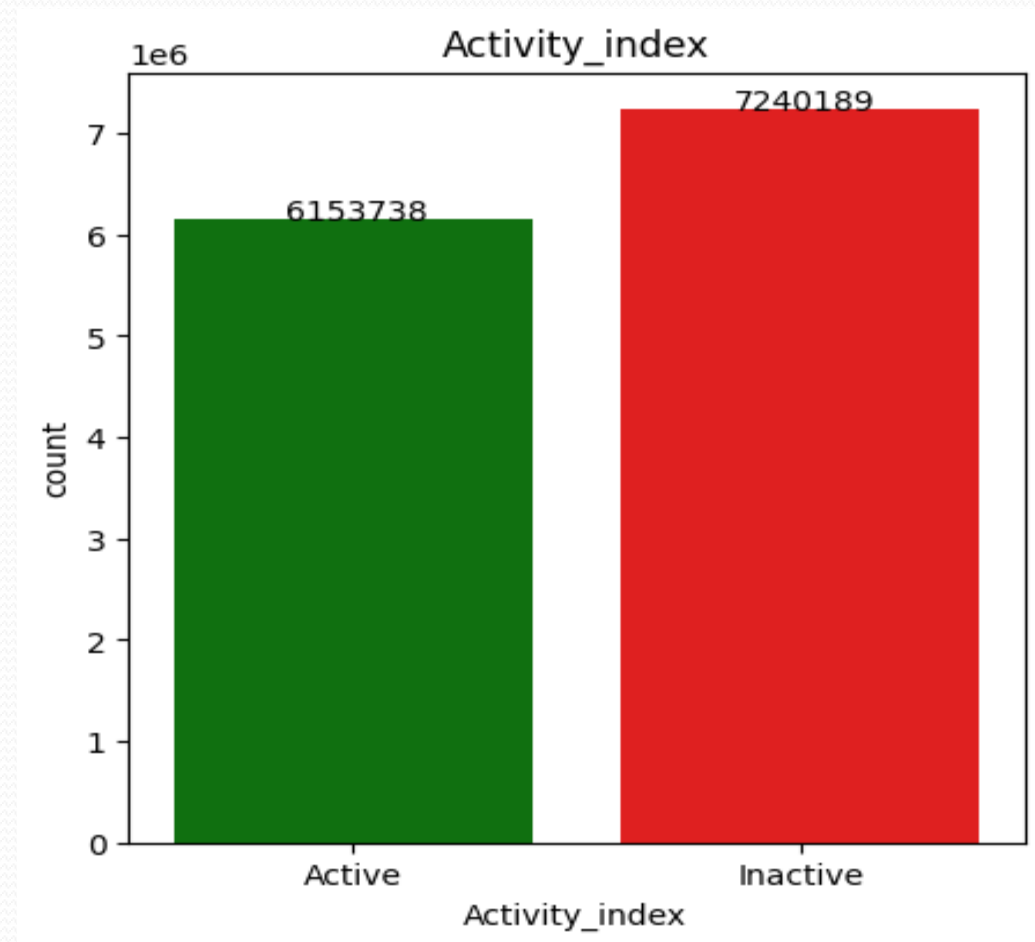
- 99% of the customers reside in the country same as the XYZ credit union bank



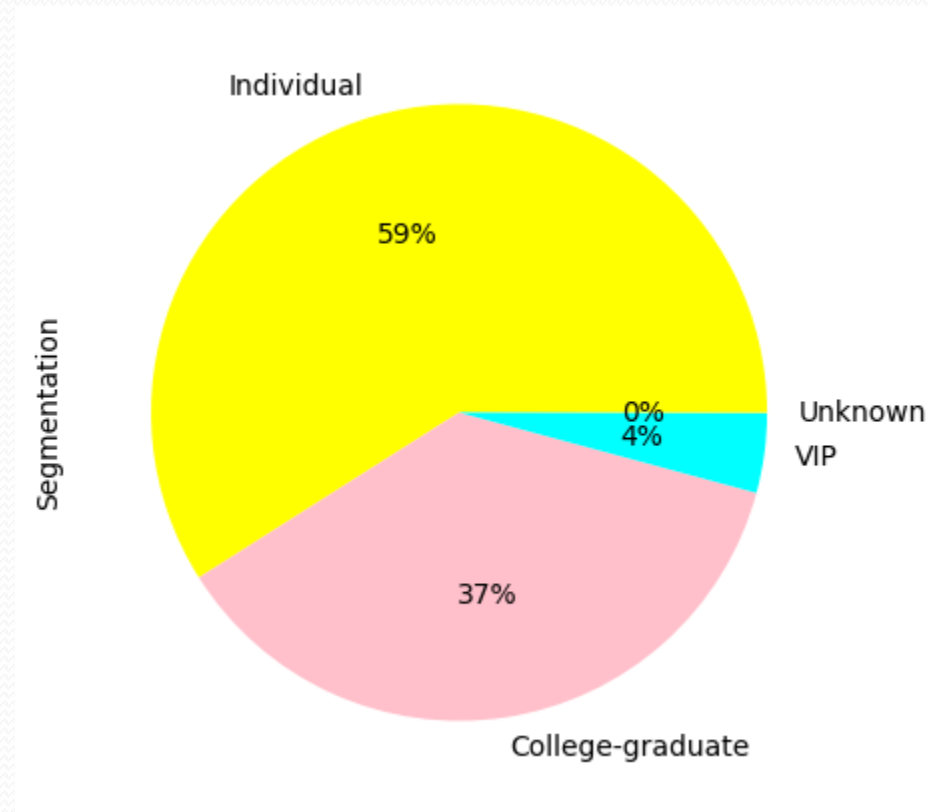
- 95% of the customers are non-foreigners while the 4% population are foreign nationals



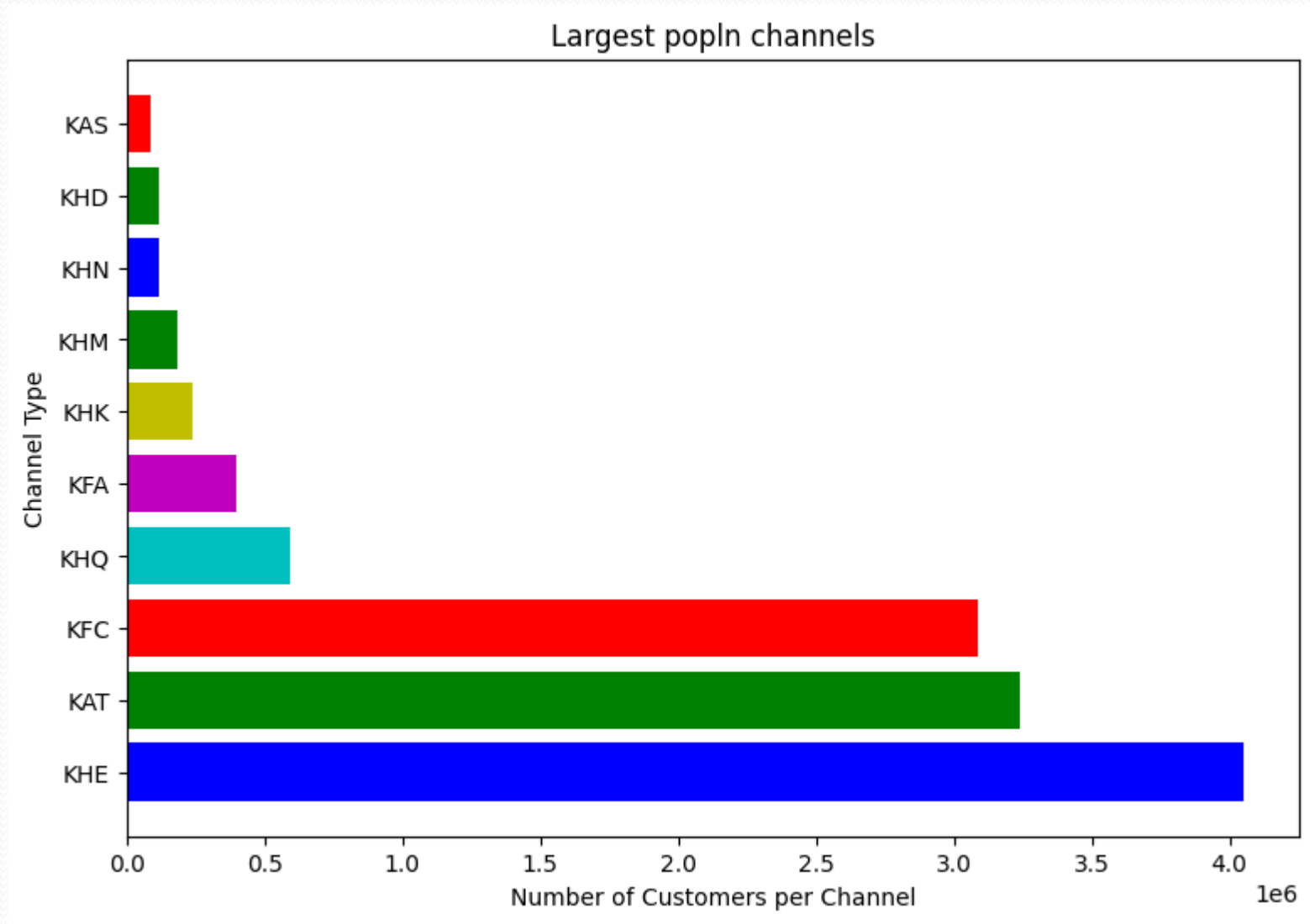
- 95% of the customers are non-foreigners while the 4% population are foreign nationals



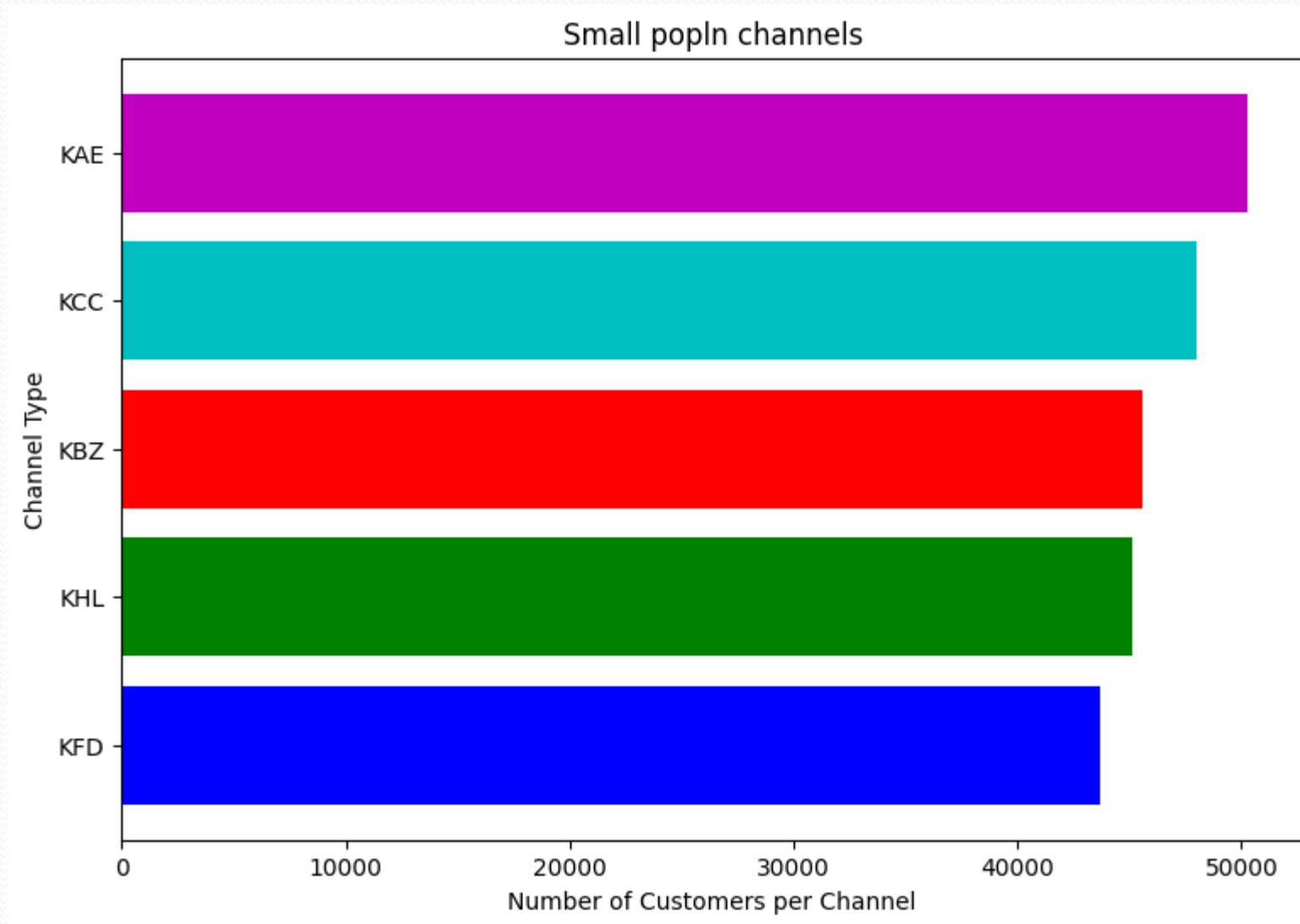
- 45.9% are active account holders in the bank while 54% are inactive.



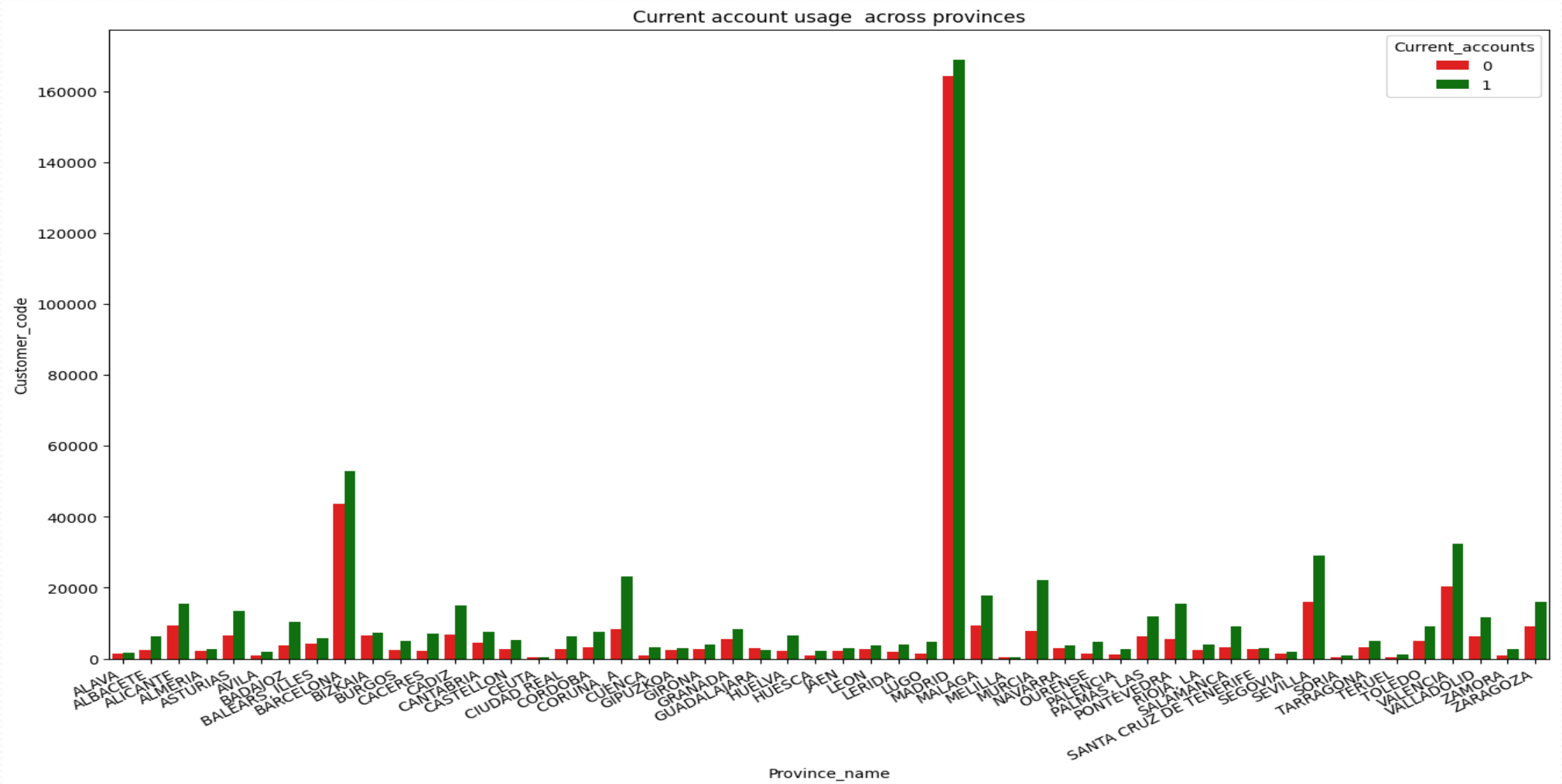
- Individuals form the highest accounts holder population followed by college graduates and VIPs



- Channels KHE, KAT, KFC have the highest number of customers

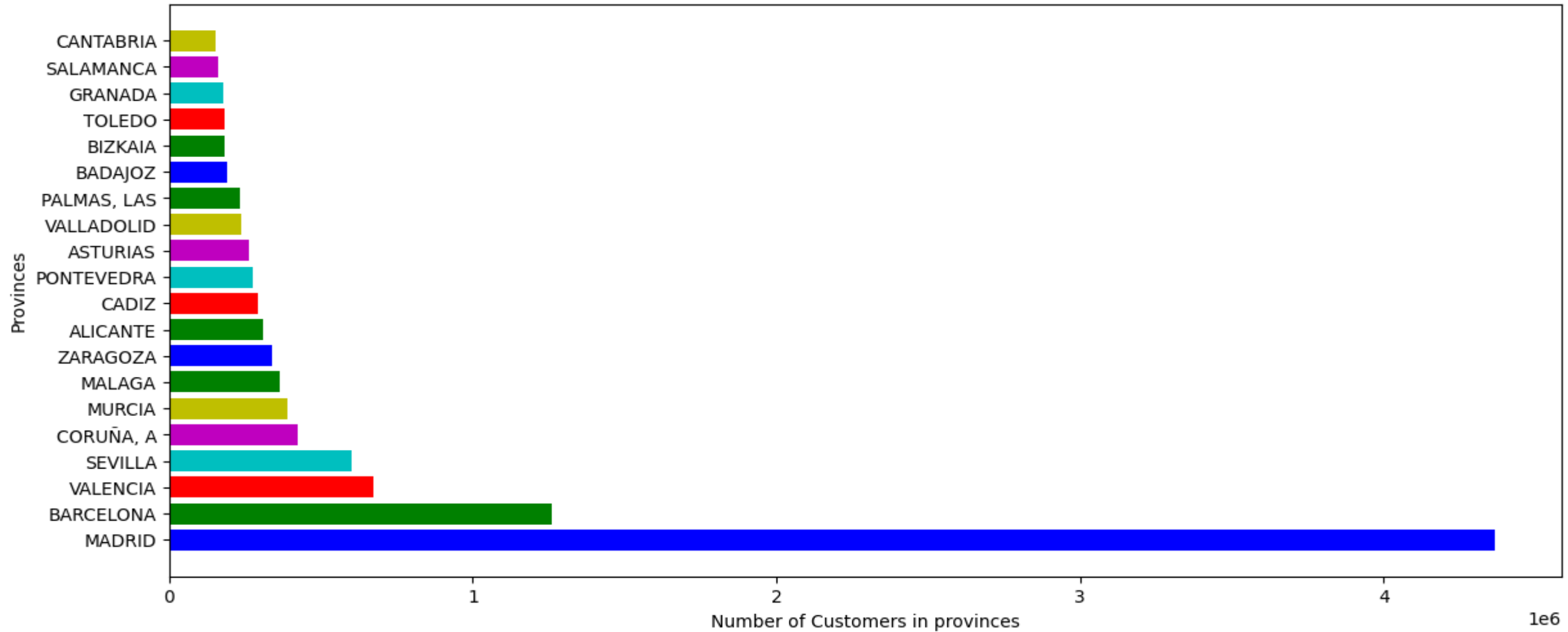


- Channels KFD, KHL, KBZ are among those with least number of customers

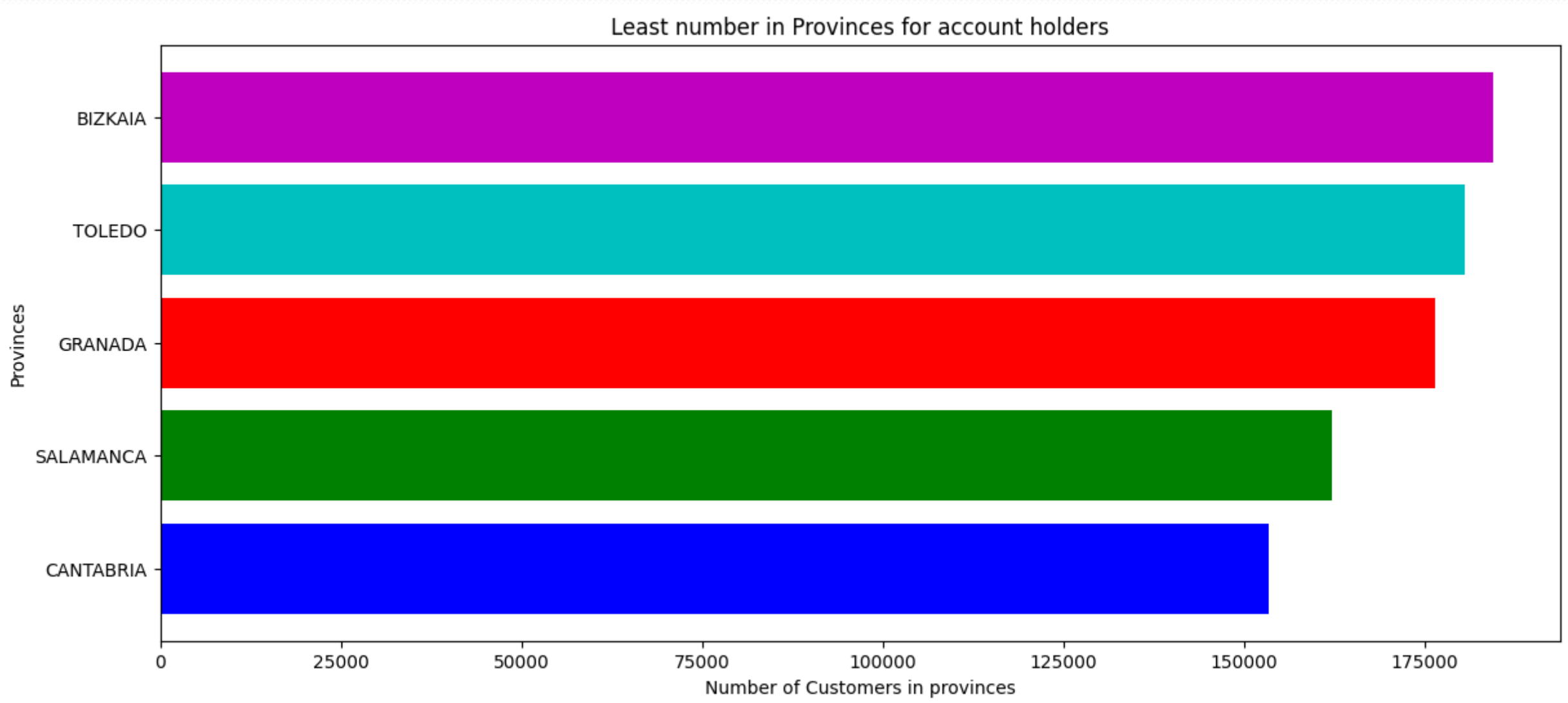


- The above plot depicts Current_accounts across provinces-Madrid has highest current account holders followed by Barcelona and Valladolid.

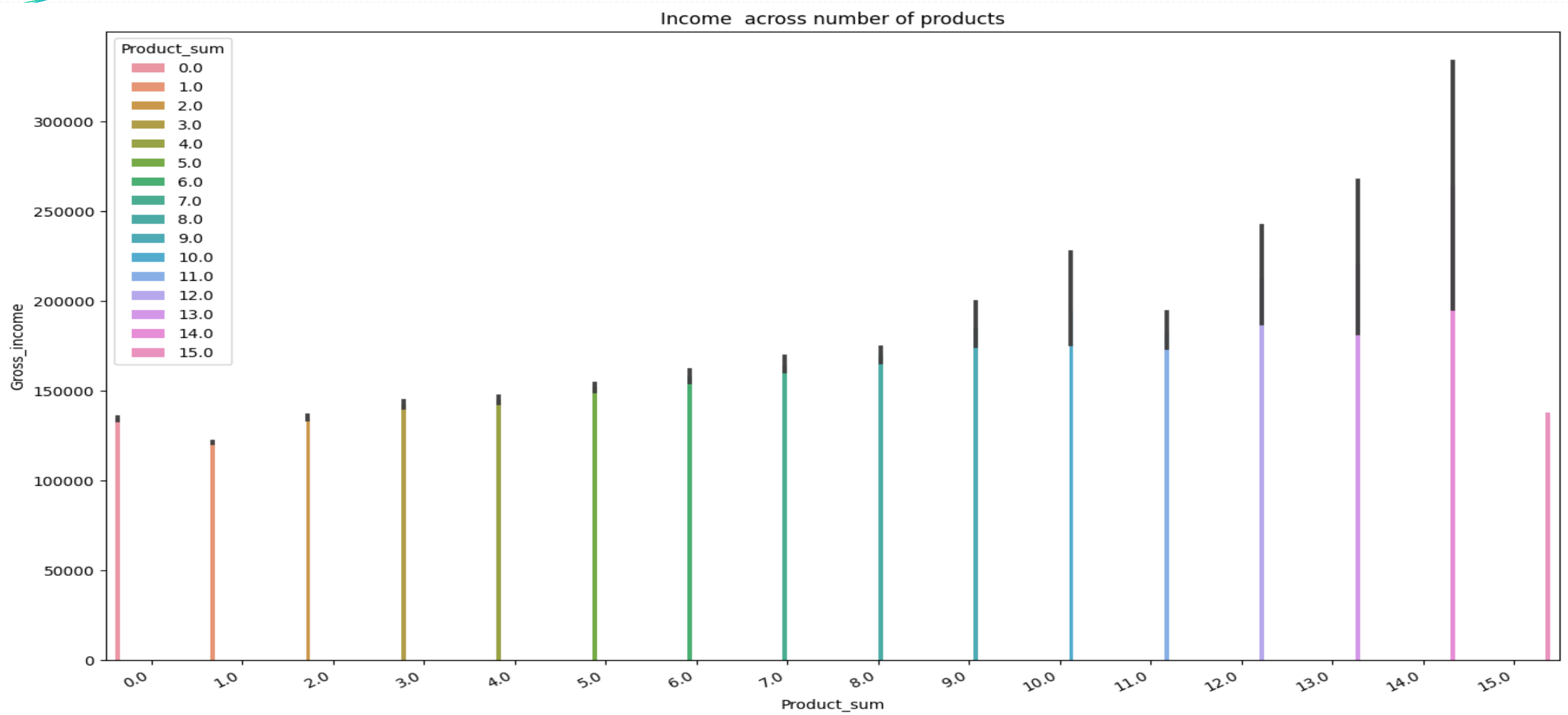
Top 10 Provinces for account holders



- Highest number of customers hail from Province Madrid and the lowest number of customers are from province Cantabria

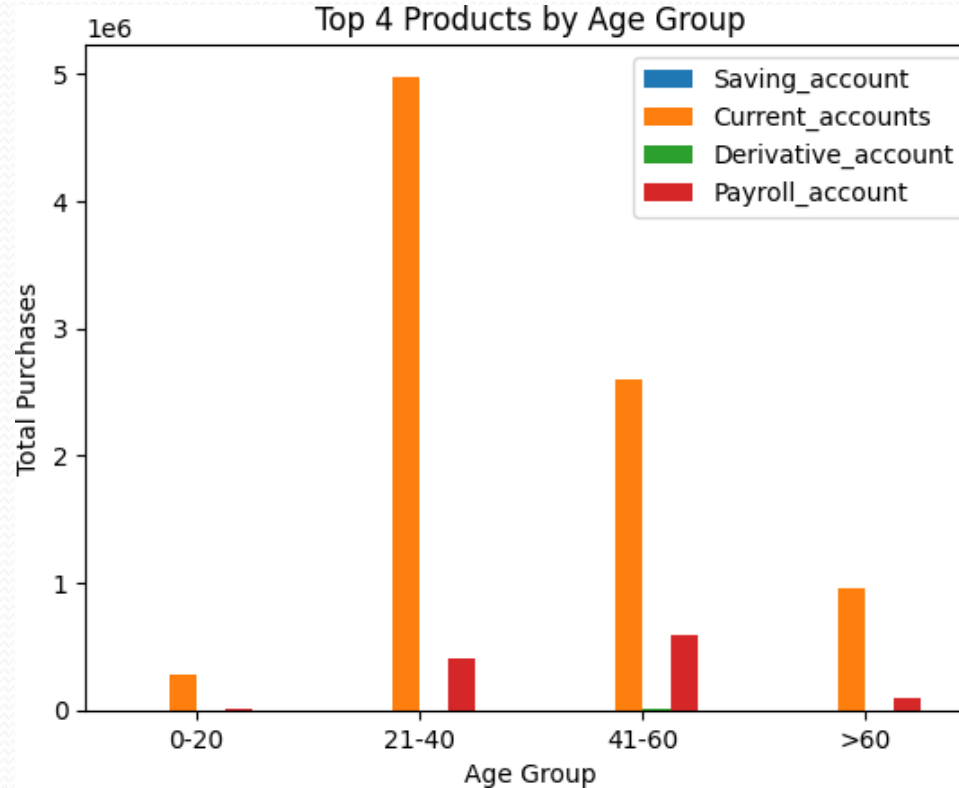


- The above depictions show the population of customers across different provinces in Latin America



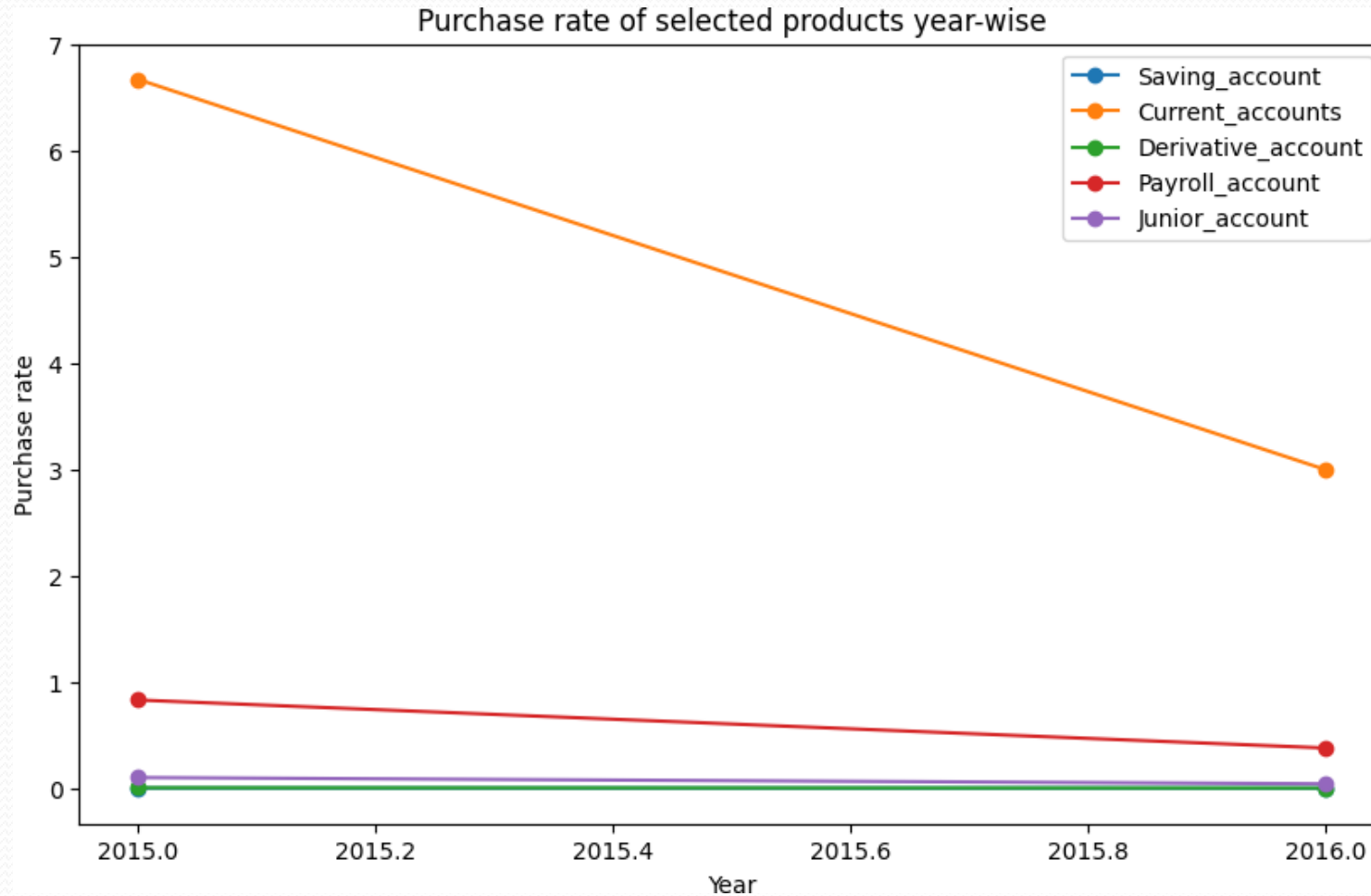
- The plot depicts Gross income is higher among people with multiple account holders

Products by Age group wise



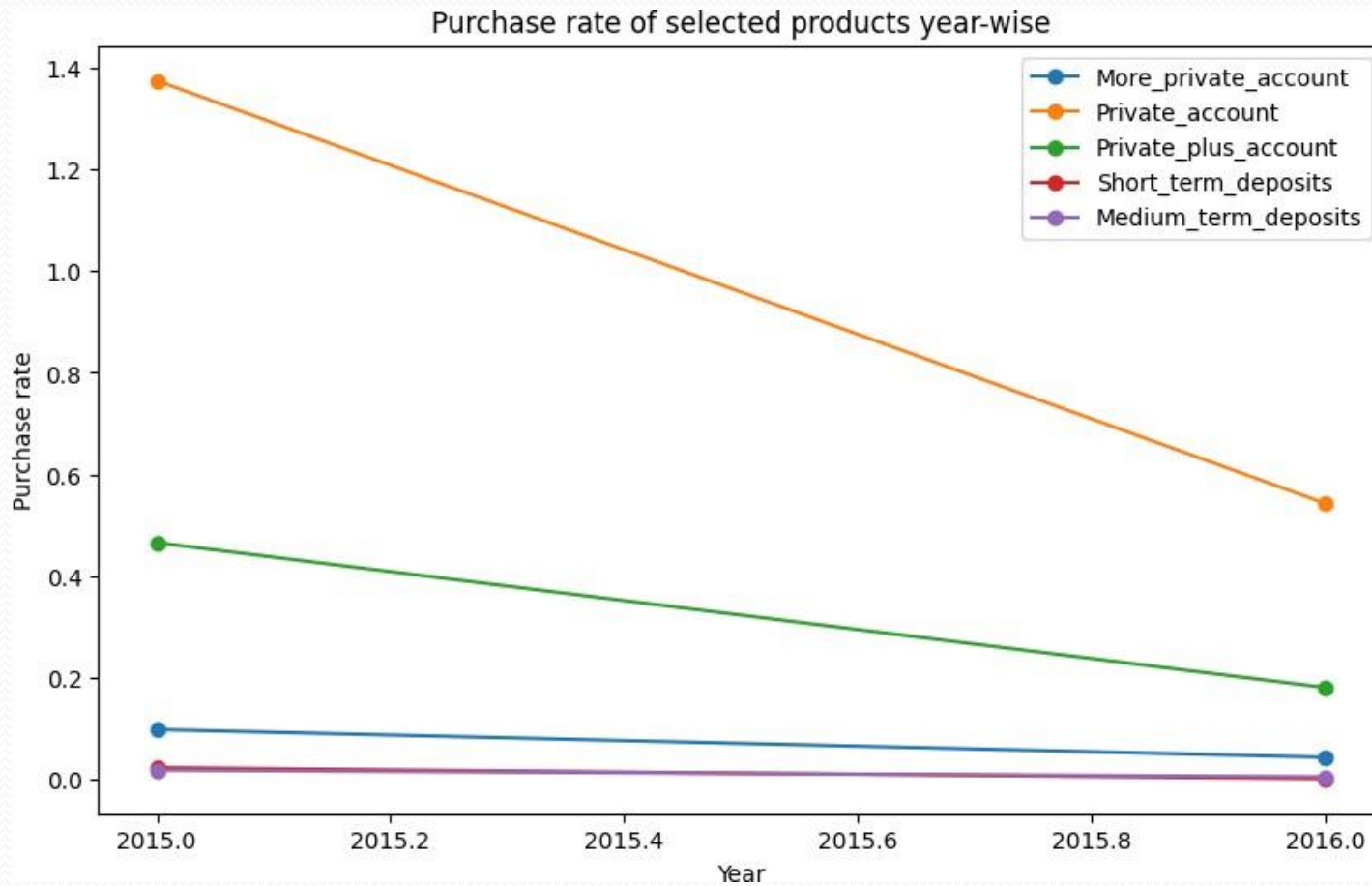
- Younger customers may be more likely to purchase products like savings accounts, payroll accounts, Derivative account and current accounts
- Younger people have more purchase interests in Current_account and Payroll_accounts

Purchase rate of products year wise



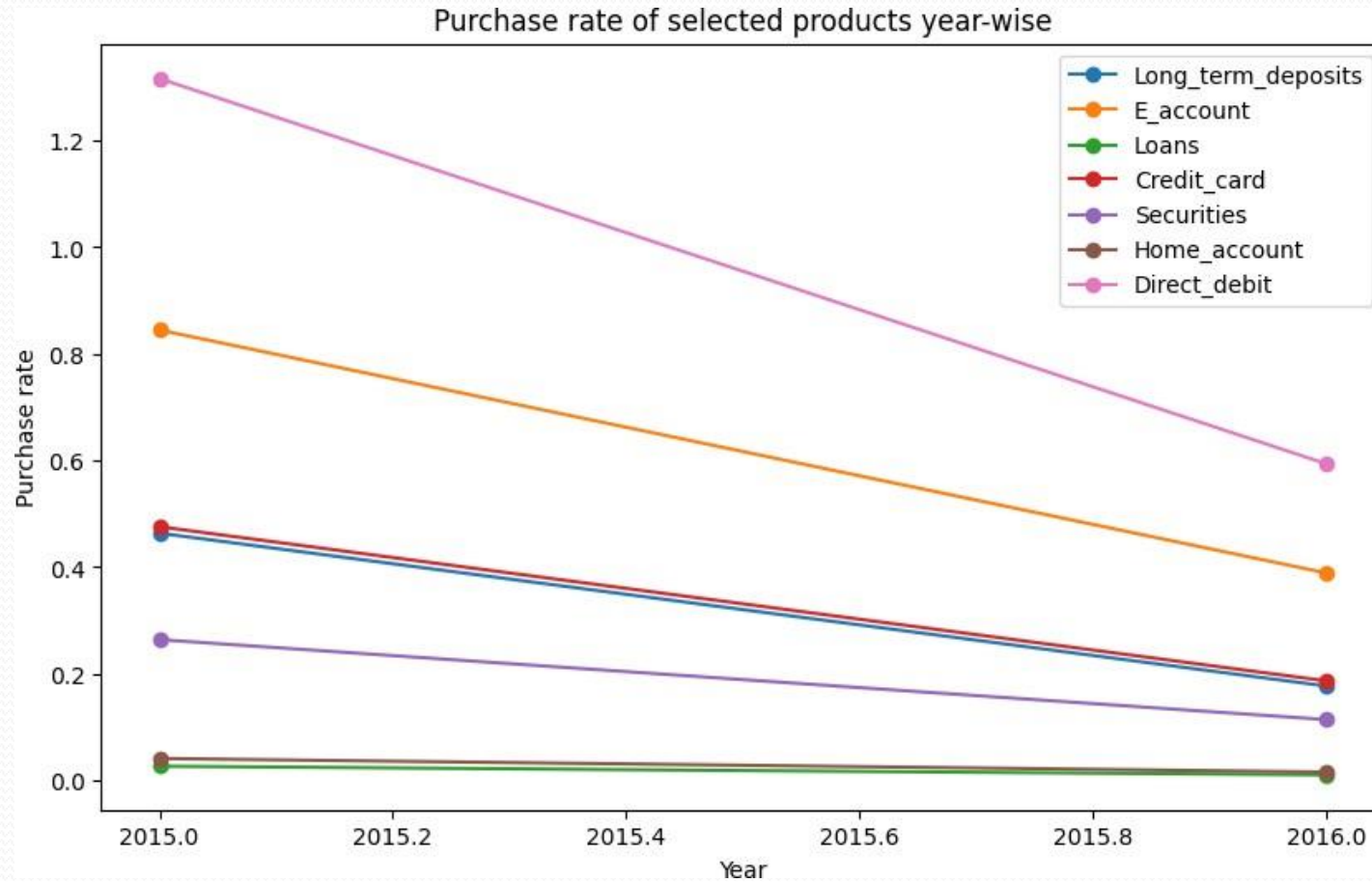
- Product popularity trends over time for 'Saving_account', 'Current_accounts', 'Derivative_account', 'Payroll_account', 'Junior_account'

Purchase rate of products year wise



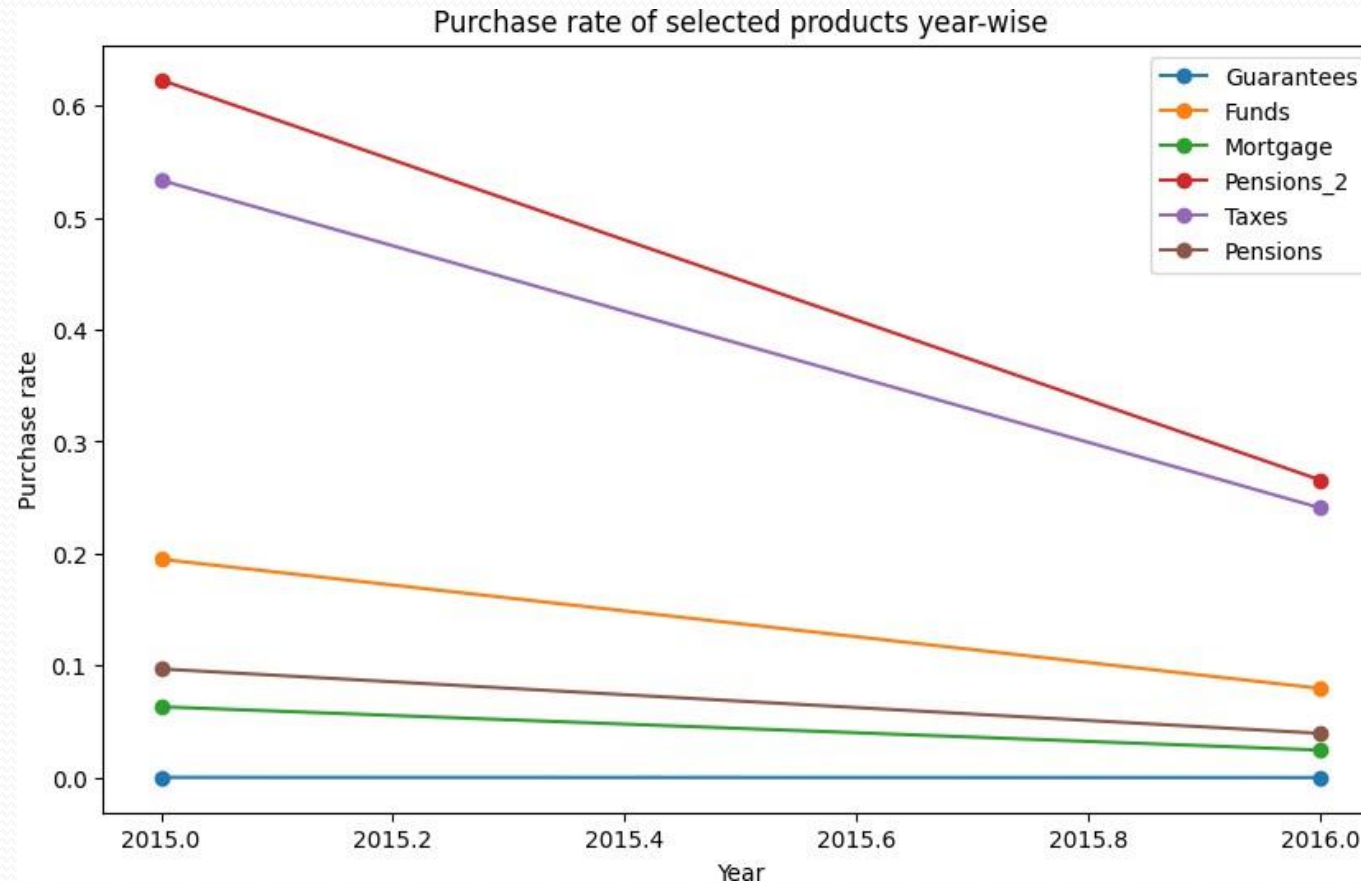
- Product popularity trends over time for 'More_private_account', 'Private_account', 'Private_plus_account', 'Short_term_deposits', 'Medium_term_deposits'

Purchase rate of products year wise



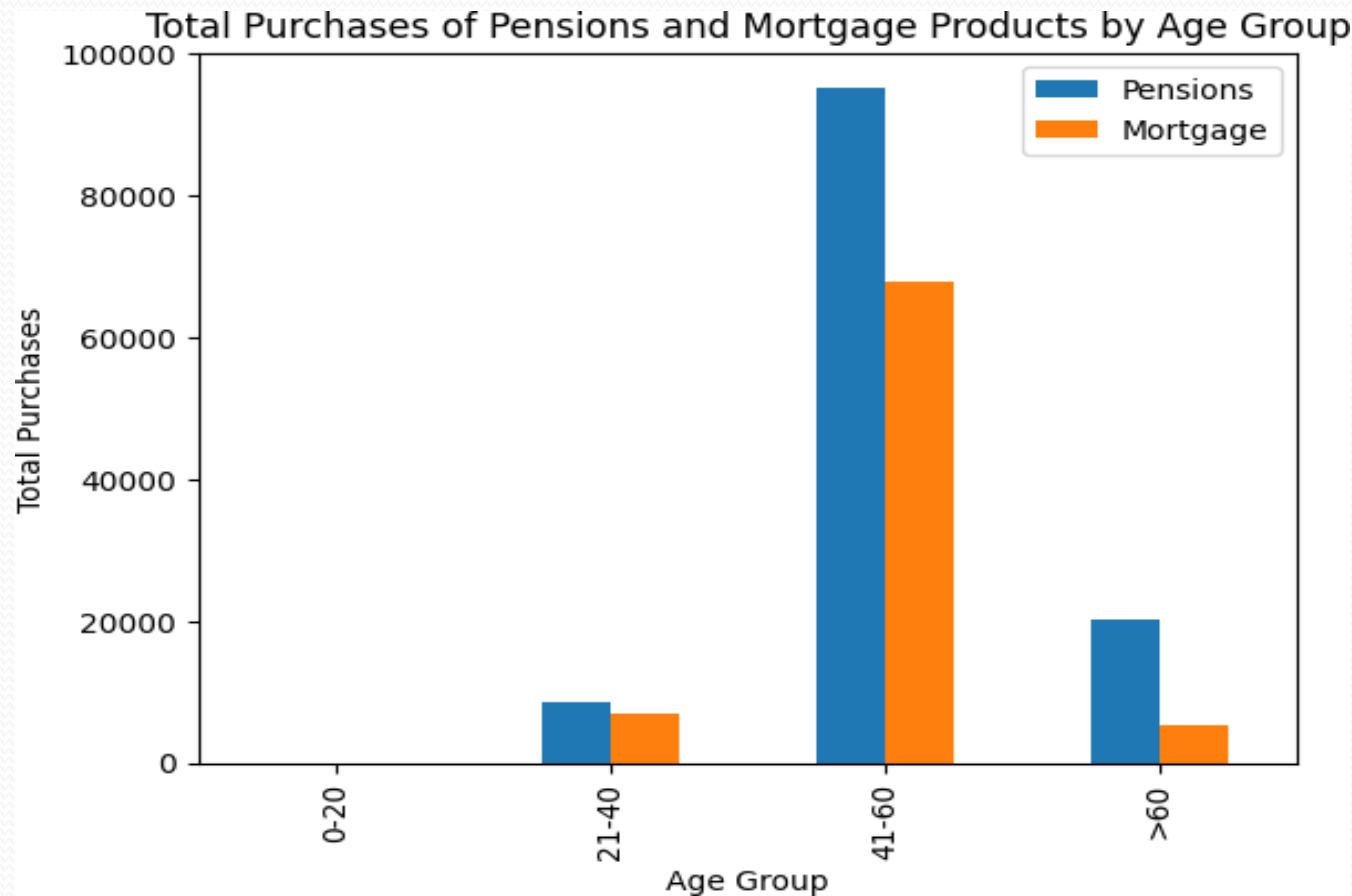
- Product popularity trends over time for 'Long_term_deposits', 'E_Account', 'Loans', 'Credit_card', 'Securities', 'Home_account', 'Direct_debit'

Purchase rate of products year wise



- Purchase rate trend over time for Guarantees, Funds, Mortgage, Pensions2, Taxes, Pensions

Purchase rate of Pensions and Mortgages products



- Bank has less amount of age >60 peoples so when its comes to age older peoples buying more Pensions and Mortgages compare to Younger group

RECOMMENDATIONS:

- From the data, we saw that customers who have been with the bank for at least 10 years are few, but majority of these customers purchase more than 1 product. It is recommended that promotional products should be offered to customers who have been in the bank for at least 10 years as a way of rewarding their loyalty, this may encourage purchase of more than one product by these customers.
- The data also showed that majority of existing customers have been in the bank for less than 10 years, and a good number of these customers purchase only one product. It is recommended that customer loyalty be built by implementing policies that will build customer trust in the bank and its products. This may help to build customers confidence in the bank and encourage purchase of multiple products.
- The data also showed that majority of existing customers have been in the bank for less than 10 years, and a good number of these customers purchase only one product. It is recommended that customer loyalty be built by implementing policies that will build customer trust in the bank and its products. This may help to build customers confidence in the bank and encourage purchase of multiple products.
- From analysis of the data, we find that majority of existing customers who have an active relation with the bank purchased more than 1 product. The results from the chi square test shows that purchase of products is associated with customer relation categories; this implies that customers who have active relation with the bank are more likely to purchase multiple products. XYZ credit union should target active customers when advertising products, this may increase purchase of multiple products by customers.
- Since Payroll Account and Direct Debit are commonly purchased together, it would be beneficial for the bank to create bundle offers for these two products, this may help to encourage purchase of multiple products by customers.
- Saving Account is commonly purchased alone, which could indicate that customers are not aware of the benefits of bundling their products. The bank should consider creating awareness campaigns targeted towards customers with only savings account, to create awareness of the benefits of bundling.

RECOMMENDATIONS:

- As there has been a decline in the purchase of several products year on year, the bank should review and analyze the causes of the decline. This will enable them to identify the issues and take corrective actions to prevent future occurrence.
- The four most popular products are Current Accounts, Private Account, Direct Debit, and E-Account. The bank should focus on improving these products further and continue to promote them to attract more customers.
- The data suggests that older age customers are buying more Pensions and Mortgage products than younger age customers. The bank should consider promoting these products more to the older age group, while also introducing more age-specific products.
- Younger age customers are buying more Current Accounts and Payroll Accounts than older age customers. The bank could create more personalized products to attract more younger age customers. These products should cater to their needs and offer additional benefits to meet their specific requirements.
- From analyses, the trend in customers joining the bank overtime showed that in the year 2014, the highest number of customers joined the bank. After this year, there was a great decline in the number Of customers who joined the bank. It is recommended that the bank implements strategies to attract more customers towards the bank.
- From the analysis, we also see that female customers purchase multiple products more than male customers. Based on this, it is recommended that promotions should be made to encourage increased purchase of products by male customers. Also, since female customers purchase multiple products more, they can serve as a target market when advertising or marketing products for cross selling.
- Establishments of more banking facilities and XYZ branches across provinces with lower customer strength For provinces with higher customer presence like Madrid, Barcelona engaging customers with new or unique programs in mortgages(for businesses), travel/foreigner accounts(leisure activity, travel industry),insurances ,car loans etc. Marketing through a broad spectrum of Channels and introducing engagement programs will increase the customer presence throughout demographics Engaging more with college-graduates with their upcoming transitions may be from short term account to long term, in terms of Education loans, travel insurances, rental loans.

CONCLUSION

This presentation summarized the analysis and results from the analysis carried out on XYZ credit union data. From the data, various visualizations have been created to show; the demographic features of customers, customers categories etc. and how these various features and categories affect or relate to the purchase behavior of customers. Results from the analysis provided great insights from which recommendations have been made in hopes of boosting cross selling in XYZ credit union.