

Group Name: DATA ALCHEMISTS

Team Members:

Name: Yuhong Chen
Email: ryechenn@gmail.com
Country: Canada
College/Company: McGill University

Name: Haoyue Chang
Email: kerrychy1215@gmail.com
Country: USA
College/Company: Northeastern University

Name: Gifty Osei
Email: gifty18osei@gmail.com
Country: USA
College/Company: Montana State University

Specialization: Data Analyst

Problem description

We are determined to help XYZ Bank improve its cross-selling strategies and enhance customer engagement. The bank offers a wide array of financial products and services, including savings accounts, credit cards, mortgages, loans, and investment options. However, we've observed that many of our customers have limited product adoption and aren't fully utilizing the range of services available to them.

To tackle this challenge head-on, we plan to implement customer segmentation techniques to gain deeper insights into our customer base. By dividing our customers into distinct groups based on their demographics, financial behavior, and product usage patterns, we hope to identify specific customer segments that are more likely to use products and services. Armed with this valuable information, we aim to create personalized marketing strategies and tailored cross-selling initiatives to boost customer satisfaction and encourage higher product adoption.

As part of our data analysis team, the objective is to thoroughly analyze the extensive customer dataset provided by XYZ Bank and conduct a comprehensive customer segmentation analysis. The dataset includes detailed information about each customer, such as age, gender, income, transaction history, product holdings, and tenure with our bank.

GitHUB repo Link

<https://github.com/giphitie/Data-Glacier-Internship/tree/main/Week%2010>

Final Recommendation

1. We can see from the Age distribution plot that the majority of customers are between ages 20 and 50, meaning the products are preferred by the working force.
2. In the future, companies will want to make changes to their advertising to be able to retain most of their young customers because they tend to want to explore other options.
3. Cod_prov has an average of 27.0 with a standard deviation of 13 and might be a variable to consider in this cross-selling recommendation system.
4. ncodpers and renta variables might be significant in our model building
5. The x-axis and y-axis display the variable names ind_actividad_cliente, ind_nomina_ult1, ind_nom_pens_ult1, and ind_recibo_ult1.
6. "ind_actividad_cliente" and "ind_nomina_ult1":The correlation coefficient between "ind_actividad_cliente" and "ind_nomina_ult1" is close to zero. This suggests that there is little to no linear relationship between a customer's activity index and whether they received a payroll payment.
7. "ind_actividad_cliente" and "ind_nom_pens_ult1":The correlation between "ind_actividad_cliente" and "ind_nom_pens_ult1" is -1, indicating a negative linear

association between a customer's activity index and whether they received a pension payment.

8. "ind_actividad_cliente" and "ind_recibo_ult1": The correlation between "ind_actividad_cliente" and "ind_recibo_ult1" (direct debit) indicates that there is a strong linear relationship between a customer's activity index and their participation in direct debit transactions.
9. "ind_nomina_ult1" and "ind_nom_pens_ult1": Depending on the sample data, you might observe a correlation coefficient close to zero. This could indicate a possible connection between receiving payroll and receiving pension payments, though further analysis would be needed to establish causation.

In the future, we can use the customer segmentation model to conduct the analysis by using clustering algorithms like K-Means or hierarchical clustering to group customers based on Use clustering algorithms like K-Means or hierarchical clustering to group customers based on their attributes. This could help identify different customer segments with distinct behaviors and preferences. This could help identify different customer segments with distinct behaviors and preferences.

Meanwhile, we can use product recommendation models to build a recommendation system using collaborative filtering or matrix factorization techniques to suggest financial products to customers based on their historical interactions and preferences.

