Customer Retention Project

Pig E. Bank

1. Checking the quality of the data, by checking for missing values, errors, and inconsistencies:

  1.1. Assessing Missing Values and Inconsistencies:

- Missing Values, NULL, or Inconsistencies:

  - Missing Last_Name (null or Y?an): 0.30%

  - Missing Credit Score: 0.30%

  - Missing Gender: 0.10%

  - Age is 2 or NULL: 1.21%

  - Balance = $0: 35.22%

  - Missing Estimated Salary: 0.20%

The balance column, where 35.22% of the rows have a balance of $0 will stay. The other missing values represent a relatively small percentage (less than 5%) of the total data. Thus, these rows (except Balance = $0) will be deleted. Besides, most of these clients are inactive or have already exited from the bank.

1.2. Fixing Inconsistent Formats:

- Inconsistent Formats:

  - Gender and Country columns use abbreviation spelling (e.g., 'M' instead of 'Male', FR instead of France) on every value in these two columns.

Country values were standardized by replacing the abbreviations with their corresponding full names.

Based on the descriptive statistics for both the "Exited" and "No-Exited" groups, here are some insights into the factors that contribute to client loss:

1. Age:

- **Observation:**

  o The average age of clients who have exited is **45.29**, while the average age of clients who stayed is **39.16**.

  o The data suggests that older clients are more likely to leave Pig E. Bank.

- **Implication:**

  o Clients around middle age (45+) might be at higher risk of leaving, perhaps because they seek better services or products as their financial needs change.

2. Balance:

- **Observation:**

  o The average balance of clients who have left is **$90,495.71**, which is higher than those who stayed (**$78,153.57**).

  o Interestingly, the mode (most frequent value) for balance in both groups is **$0.00**, indicating that many clients (in both groups) might have no balance at all.

- **Implication:**

  o Higher average balances among those who left suggest that customers with higher balances might be dissatisfied with the bank's services, despite having more invested in Pig E. Bank.

  o It may also indicate that a lack of balance (e.g., having $0.00) is not a clear indicator of leaving, as many clients with $0.00 remain.

3. Tenure:

- **Observation:**

  o The average tenure for those who left is **4.71** years, while those who stayed have an average tenure of **5.09** years.

  o However, the mode (most common value) is quite different: for those who left, the mode is **1 year**, while for those who stayed, the mode is **9 years**.

- **Implication:**

  o Shorter tenure (especially around 1 year) seems to be a significant indicator of client loss, suggesting that newer clients are at higher risk of leaving.

  o Customers who stay beyond a certain point (like 9 years) are likely to remain loyal.

4. Number of Products:

- **Observation:**

  - The average number of products used by clients who left is **1.46**, while for those who stayed, it's slightly higher at **1.52**.

- **Implication:**

  - Clients who use fewer products are more likely to leave. This makes sense because clients with more products are generally more engaged with Pig E. Bank.

  - Banks should focus on increasing product engagement to improve retention.

5. IsActiveMember:

- **Observation:**

  - A much smaller percentage of clients who left were active members (**30%**) compared to those who stayed (**51%**).

- **Implication:**

  - Being an active member (e.g., using banking services regularly, making transactions, etc.) seems to be a key factor in retaining clients. Banks should focus on increasing customer engagement.

6. Has Credit Card (HasCrCard?):

- **Observation:**

  - There's a very small difference in the percentage of clients who have credit cards between those who left (70%) and those who stayed (71%).

- **Implication:**

  - Having a credit card doesn't appear to be a major factor in client retention. Other factors like age, tenure, and product engagement seem more influential.

Leading Risk Factors for Client Loss:

1. **Age:** Older clients (around 45+) are more likely to leave.

2. **Tenure:** Shorter tenures, especially around 1 year, are associated with higher client loss.

3. **Product Engagement:** Clients with fewer products are more likely to leave.

4. **Active Membership:** A lack of active engagement is a significant indicator of client loss.

Next Steps for Pig E. Bank:

- Focus on increasing product engagement and ensuring new clients (within the first few years) feel valued to improve retention.

- Develop strategies to re-engage older clients and offer services that better meet their changing needs.

- Investigate why higher balance clients are leaving despite their substantial investment in the bank and address their concerns.

Decision Tree

Predicting the probability of customers leaving the bank will be based it on the top risk factors identified from the data: Age, Tenure, Product Engagement, and Active Membership

Tenure > 1 year

Age >= 45 years & number of products = 1

Age < 45 years

Medium-High Probability of Leaving if only 1 product

More than 1 Product - Low Probability

Low Probability of Leaving

IsActiveMember = 0
Then
High probability of
leaving

IsActiveMember = 1
Then
Medium probability
of leaving

Tenure ≤ 1 year.

Tenure