



# CUSTOMER CHURN IN THE BANKING REPORT

Written By :  
**Nguyen Huong Giang**

## Table of Contents

<b>A. Introduction.....</b>	<b>4</b>
<b>B. Data Collection.....</b>	<b>4</b>
<b>C. Data Exploration.....</b>	<b>4</b>
1. <i>Data Loading.....</i>	<i>4</i>
2. <i>Dataset Overview.....</i>	<i>4</i>
3. <i>Cleaning Data.....</i>	<i>5</i>
<b>D. Exploratory Analysis.....</b>	<b>6</b>
1. <i>Churn Distribution.....</i>	<i>6</i>
2. <i>Churn and Gender.....</i>	<i>7</i>
3. <i>Churn and Age.....</i>	<i>8</i>
4. <i>Churn and Credit Score.....</i>	<i>9</i>
5. <i>Churn and Account Balance.....</i>	<i>10</i>
<b>E. Recommendations.....</b>	<b>11</b>
1. <i>Strategies Focused on Female Clients.....</i>	<i>11</i>
2. <i>Engaging in Age-Specific Activities.....</i>	<i>12</i>
3. <i>Tracking Credit Scores.....</i>	<i>12</i>
4. <i>Handling Account with No Balance.....</i>	<i>12</i>
5. <i>Improving the Customer Experience.....</i>	<i>13</i>
<b>F. Conclusion.....</b>	<b>13</b>
<b>G. References.....</b>	<b>14</b>

## Table of Figures

Figure 1: Data loading .....	4
Figure 2: Dataset Overview .....	4
Figure 3: Check for duplicated row .....	5
Figure 4: Check for missing value row .....	5
Figure 5: Churn Distribution - code .....	6
Figure 6: Churn Distribution - bar chart .....	6
Figure 7: Churn and Gender - code .....	7
Figure 8: Gender Distribution of Churn - bar chart .....	7
Figure 9: Churn and Age - code .....	8
Figure 10: Age Distribution of Churned vs. Non-Churned Customers - box plot .....	8
Figure 11: Churn and Credit Score - code .....	9
Figure 12: Credit Score and Churn - scatter plot .....	9
Figure 13: Churn and Account Balance - code .....	10
Figure 14: Account Balance and Churn - histogram plot .....	10

## A. Introduction

In the banking sector, customer attrition is a substantial challenge, as it has a direct influence on customer loyalty and revenue. The objective of this project is to predict customer attrition by the analyzing of many customer attributes, including credit scores, estimated salary, and engagement trends. By recognizing at-risk customers early on, banks may reduce the financial impact of lost consumers by implementing focused retention initiatives.

The findings and actionable recommendations derived from the analysis are detailed in the report, which also includes the data collection, cleaning, and analysis methodologies employed.

## B. Data Collection

This [Google Drive link](#) provided the dataset used in this report. The data contains a number of consumer attributes that are important for predicting turnover.

## C. Data Exploration

### 1. Data Loading

To begin our initial review and comprehend the data's structure, the dataset is loaded by using Pandas.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

df = pd.read_csv('/Users/apple/Documents/IT/TERM 6/COMP1844/Bank Customer Churn Prediction.csv')
✓ 0.0s
```

Figure 1: Data loading

Because the dataset is Excel file, the `read_csv()` method of Pandas is used to read data in file.

### 2. Dataset Overview

To get an initial understanding of the dataset, we looked at the number of rows and columns, data types, and basic statistics. Using `info()` method of Pandas is suitable to get general information of the dataset.

```
print(df.info())
✓ 0.0s

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 12 columns):
#   Column                Non-Null Count  Dtype  
---  -
0   customer_id           10000 non-null  int64  
1   credit_score           10000 non-null  int64  
2   country                10000 non-null  object  
3   gender                 10000 non-null  object  
4   age                    10000 non-null  int64  
5   tenure                 10000 non-null  int64  
6   balance                10000 non-null  float64 
7   products_number        10000 non-null  int64  
8   credit_card            10000 non-null  int64  
9   active_member          10000 non-null  int64  
10  estimated_salary        10000 non-null  float64 
11  churn                   10000 non-null  int64  
dtypes: float64(2), int64(8), object(2)
memory usage: 937.6+ KB
None
```

Figure 2: Dataset Overview

The dataset have 10000 rows (each row have its index from 0 to 9999) with 12 columns belong with its data type. Data frame have 2 columns with float64 data type are balance and estimated\_salary, 2 columns with object data type are country and gender, the remaining 8 columns are int64 data type: customer\_id, credit\_score, age, tenure, products\_number, credit\_card, active\_number, churn.

### 3. Cleaning Data

We checked for infeasible, duplicate, missing data and determined the best ways to handle them.

```
duplicate_df = df[df.duplicated()]
duplicate_df.shape[0]
```

✓ 0.0s

0

Figure 3: Check for duplicated row

Using duplicated() method to find any duplicated rows. The shape() method of Pandas return the dimensionality of the data frame. The output of below code is zero, it mean that the dataset does not contain any duplicated row.

```
null_df = df.isnull().sum() # total number of missing data
null_df
```

✓ 0.0s

customer_id	0
credit_score	0
country	0
gender	0
age	0
tenure	0
balance	0
products_number	0
credit_card	0
active_member	0
estimated_salary	0
churn	0
dtype: int64	

Figure 4: Check for missing value row

Using isnull() method to find any missing values and sum() method to calculate the summary of missing values. The output of below code is all zeros, it means that the dataset does not contain any missing values in any column.

The data set does not have any duplicated row and does not have any missing values in any column, so that we do not need to drop any rows.

## D. Exploratory Analysis

We created a number of visualizations using Matplotlib in order to comprehend the distribution and relations between churn and customer attributes.

### 1. Churn Distribution

To compare the number of customers who left with those who stayed, using bar chart is suitable.

```
# churn distribution
churn = df['churn'].value_counts()

# bar chart
churn.plot(kind='bar')
plt.title('Churn Distribution')
plt.xlabel('Churn')
plt.ylabel('Count')
plt.xticks(rotation=0)
plt.show()
```

Figure 5: Churn Distribution - code

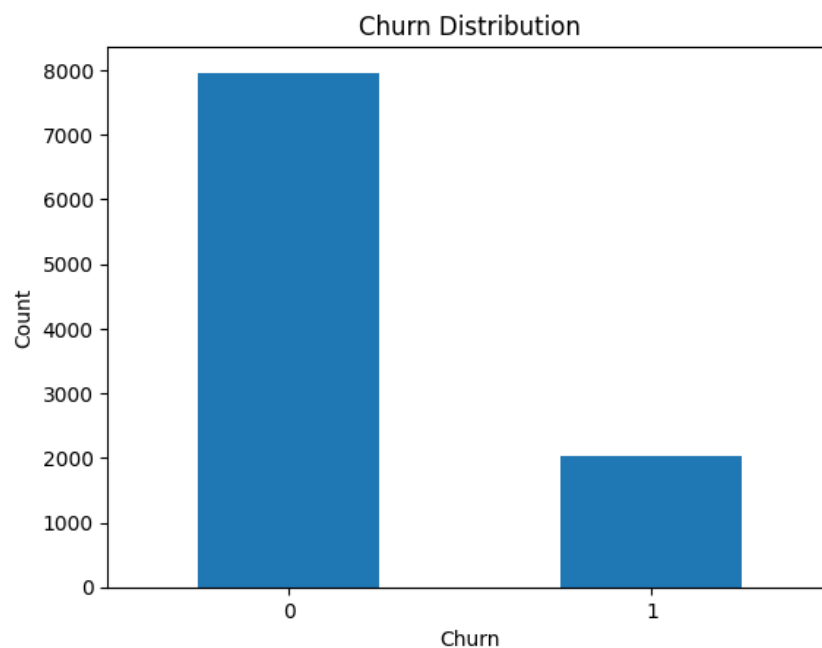


Figure 6: Churn Distribution - bar chart

The "Churn Distribution" bar chart shows the number of clients who have left the bank, in comparison to those who have not. The x-axis is called "Churn," and it has two categories: 0 (which represents customers who haven't churned) and 1 (which represents customers who have churned). Labeled "Count," the y-axis spans from 0 to 8000. Given that there are around 8000 clients in the non-churned group (0) and 2000 in the churned category (1), it is clear from the chart that the majority of consumers have not left.

The retention rate and churn rate can be easily calculated over a specific period.

$$\text{Churn rate} = \frac{\text{Number of Churned Customers}}{\text{Total Number of Customers}} = \frac{2000}{10000} = 20\%$$

$$\text{Retention rate} = \frac{\text{Number of Non-Churned Customers}}{\text{Total Number of Customers}} = \frac{8000}{10000} = 80\%$$

With an 80% retention rate, the bank is effectively keeping the vast majority of its clients. A churn rate of 20%, it mean that 1 in every 5 customers leave the bank. 20% is higher than the average churn rate for a bank (Grigovera, 2023). This comparatively high percentage of customer attrition indicates that the bank is having trouble keeping clients, which might have serious financial repercussions.

The high churn rate can be due to many different customers attributes such as gender, credit scores, age, ... Let's analyze these attributes to a clearly analytical view.

## 2. Churn and Gender

First of all, a bar chart identify the distribution of churn across genders.

```
churn_gender = df.groupby(['gender', 'churn']).size()

# bar chart
churn_gender.plot(kind='bar', color=['green', 'red'])
plt.title('Gender Distribution of Churn')
plt.xlabel('Churn')
plt.ylabel('Count')
plt.xticks(rotation=0)
plt.show()
```

✓ 0.1s

Figure 7: Churn and Gender - code

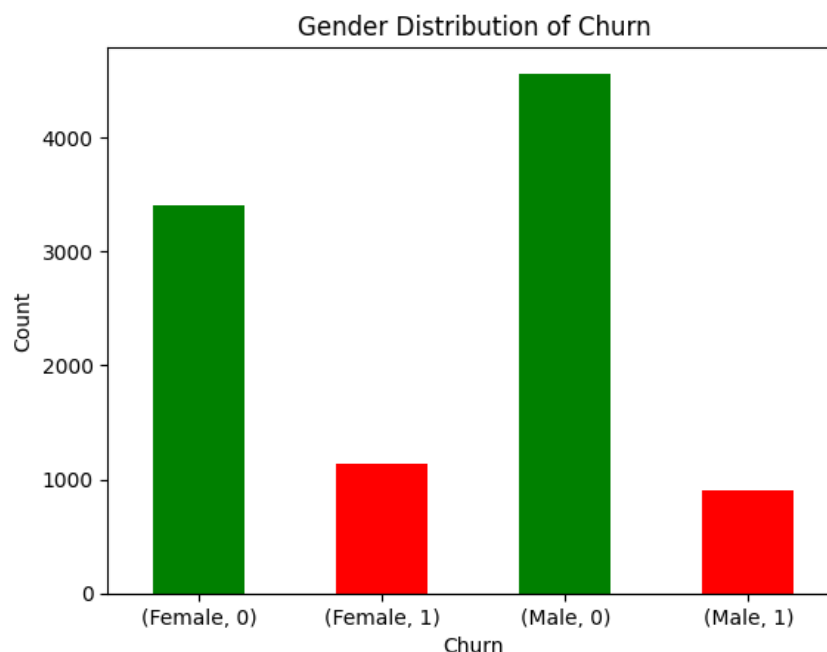


Figure 8: Gender Distribution of Churn - bar chart

The "Gender Distribution of Churn" bar chart shows the proportion of male and female customers who have left compared to those who have stayed. There are two bars for each gender group, one for churned customers (red) and one for non-churned customers (green).

### Churn Rate by Gender:

#### ❖ Female Churn Rate:

- Non-Churned: 3500
- Churned: 1000
- Female Churn Rate =  $\frac{1000}{3500+1000} \approx 22\%$

#### ❖ Male Churn Rate:

- Non-Churned: 4500
- Churned: 500
- Male Churn Rate =  $\frac{500}{4500+500} \approx 10\%$

When compared to male consumers (10%), the attrition rate for female customers is significantly higher (22%). According to this discrepancy, female consumers are more likely than male customers to leave the bank. Improving overall retention rates may depend on identifying and resolving the underlying causes of this discrepancy. In other studies, woman left at a high rate than men (Kuo, 2024).

### 3. Churn and Age

In customer churn, age play an important role. A box plot is used to visualize the age distribution of churn vs. non-churned customers.

```
df.boxplot(column='age', by='churn')
plt.suptitle('')
plt.title('Age Distribution of Churned vs. Non-Churned Customers')
plt.xlabel('Churn')
plt.ylabel('Age')
plt.xticks([1, 2], ['Non-Churned', 'Churned'])
plt.show()
```

0.2s

Figure 9: Churn and Age - code

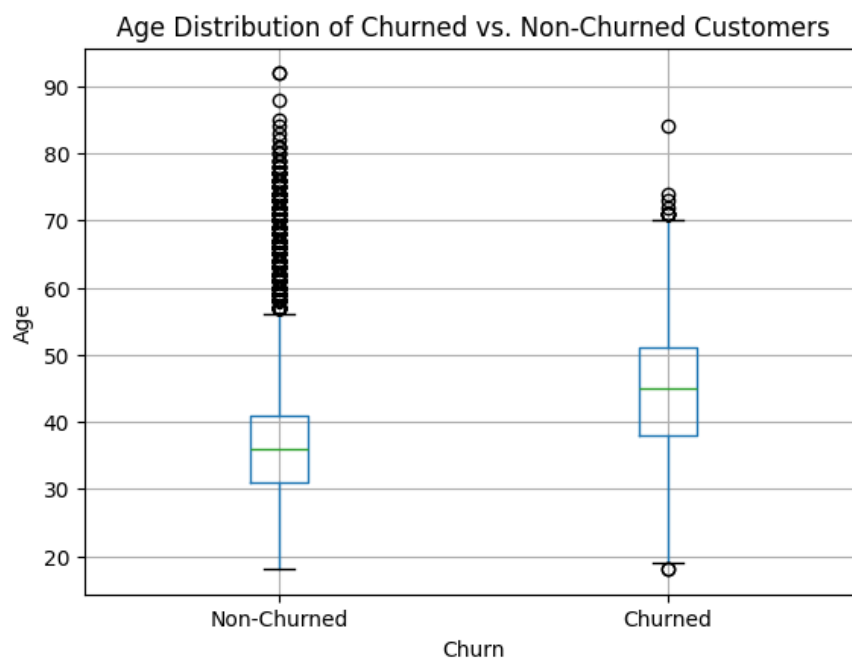


Figure 10: Age Distribution of Churned vs. Non-Churned Customers - box plot



#### ❖ Non-Churned Customers:

Non-churned consumers (labeled as "Non-Churned") had a median age of around 40 and an interquartile range (IQR) of roughly 30 to 50, according to the box plot. When outliers are excluded, whiskers expand to display the whole age range.

#### ❖ Churned Customers:

A median age of about 40 is also shown in the box plot for churned customers (labeled as "Churned"), with an IQR of roughly 30 to 50. A number of outliers fall below 20 and over 60.

The age distributions of churned and non-churned consumers differ, as the box plot illustrates. The dispersion (IQR) may vary even when the median age of both groups is comparable, indicating that age may be a factor in customer attrition. According to this information, some age groups may be more likely to churn, and by comprehending these trends, retention tactics may be more successfully customized.

#### 4. Churn and Credit Score

Credit score also play a role in churn rate of bank market (Ryan Flanigan, 2024). Are clients with lower scores more likely to churn? Let's using a scatter plot to determine the relationship between credit scores and churn.

```
df.plot(kind='scatter',x='churn',y='credit_score')
plt.title('Scatter Plot: Credit Score vs Churn')
plt.xlabel('Credit Score')
plt.ylabel('Churn')
plt.show()
```

Figure 11: Churn and Credit Score - code

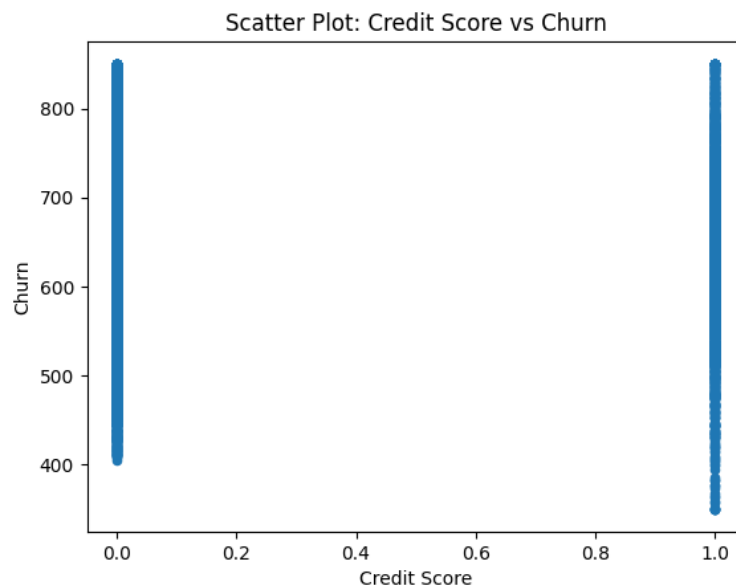


Figure 12: Credit Score and Churn - scatter plot

The "Credit Score vs. Churn" scatter plot illustrates the connection between churn status and credit scores. The y-axis displays the churn status, with 0 denoting customers who have not churned and 1 denoting customers who have. The x-axis depicts credit ratings, which range roughly from 400 to 850.

- ❖ **Lower Credit Scores and Churn:** Customers with lower credit ratings have a higher concentration of churned customers ( $y=1$ ), according to the scatter plot. This suggests that a lower credit score increases the likelihood of client attrition.
- ❖ **Higher Credit Scores and Non-Churn:** The existence of points with a  $y$ -value of 0 (non-churned) in this range, on the other hand, indicates that consumers with higher credit ratings (for example, above 700) are less likely to churn.
- ❖ **Wide Range of Credit Scores:** The vast variety of credit scores for both churned and non-churned clients suggests that while churn can happen at any credit score level, the probability varies with score.

Customers who have lower credit ratings are more likely to churn than those with higher credit scores, according to the scatter plot. Customers with lower credit ratings may be more likely to leave the bank, indicating that credit score plays a major role in customer turnover. By comprehending this connection, the bank may create focused retention plans to assist clients with lower credit ratings and minimize attrition rates.

## 5. Churn and Account Balance

Account balance are one of many factors cause churned customers. To visualize the distribution of account balances among churned and non-churned customers, we created a histogram with 2 groups (churned vs. non-churned). By doing this, comparing the distributions side by side in a single plot can be easier.

```
# histogram for non-churned customers
plt.hist(df[df['churn'] == 1]['balance'], color='green', alpha=0.5, label='Non-churned')
# histogram for churned customers
plt.hist(df[df['churn'] == 0]['balance'], color='pink', alpha=0.5, label='Churned')

plt.xlabel('Account Balance')
plt.ylabel('Frequency')
plt.title('Histogram Plot: Account Balance vs Churn')
plt.legend(loc='upper right')
plt.grid()
plt.show()
```

Figure 13: Churn and Account Balance - code

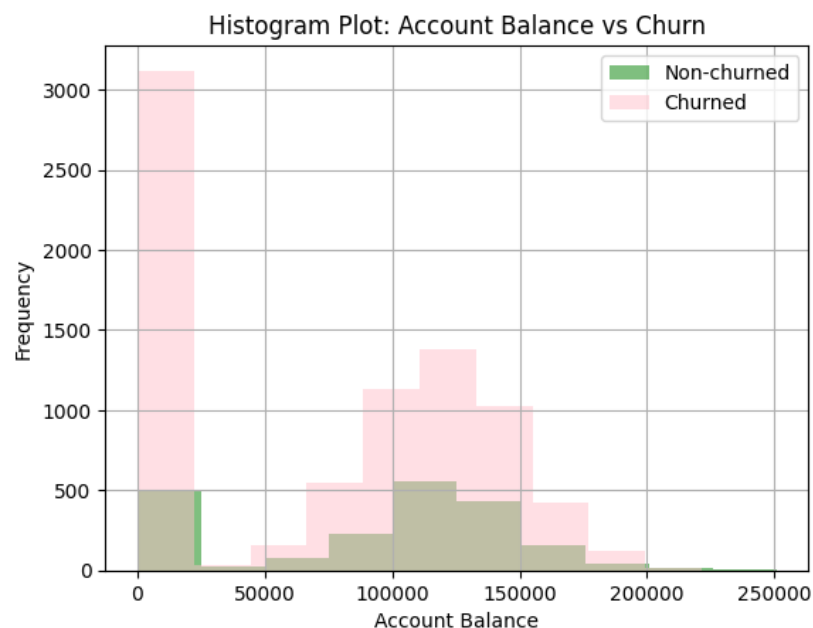


Figure 14: Account Balance and Churn - histogram plot

The histogram “Account Balance vs Churn” shows how account balances are distributed between customers who have churned and those who have not. An account balance between 0 and 250,000 is shown by the x-axis, while the frequency of clients is represented by the y-axis. To distinguish between consumers who have churned and those who have not, the histogram employs two colors: pink for churned customers and green for non-churned ones.

The high frequency bar at the very beginning of the x-axis of the histogram indicates that a sizable portion of churned customers have an account balance of 0. On the other hand, there aren't many non-churned consumers with a zero account balance. The distribution of both churned and non-churned clients is more dispersed for account balances larger than 0, peaking between 100,000 and 150,000.

- ❖ **High Churn Rate at Zero Balance:** The most startling finding is how frequently customers with account balances of 0 are churned. This implies that clients who lack equilibrium are more prone to leave.
- ❖ **Distribution Spread:** Although churned customers still seem to be more common in the lower value ranges, the distribution of churned and non-churned customers is more even for account balances greater than 0.
- ❖ **Overlap in Higher Balances:** Higher account balance ranges (100,000 to 150,000) show an overlap in the distribution of churned and non-churned clients, suggesting that churn is not always prevented by larger balances.

Account balance and customer attrition are clearly correlated, according to the histogram. Compared to customers with larger amounts, those with a zero account balance are far more likely to churn. The probability of churn, however, seems to be more uniformly distributed for account balances over zero, with some overlap between customers who have left and those who have not.

## E. Recommendations

We may provide focused suggestions to assist the bank in lowering the churn rate and enhancing client retention based on the analysis and insights obtained from the five charts: Churn Distribution, Churn and Gender, Churn and Age, Churn and Credit Scores, and Churn and Account Balance.

### 1. Strategies Focused on Female Clients

According to the gender-based data, female consumers had a higher churn rate (22%) than male customers (10%).

Recommendations:

- **Customized Offers:** Create promotions and offers that are tailored to the unique requirements and tastes of female clients. Discounts on goods and services that are well-liked by women may fall under this category.
- **Better Customer Service:** Improve customer service by educating agents on how to properly and sympathetically handle the particular issues faced by female consumers (PwC, 2017).
- **Financial Products Customized:** Create financial services and products that are suited to the requirements of female clients, such as loans for female entrepreneurs, investment plans, and flexible savings accounts (Iqbal, 2017).

## 2. Engaging in Age-Specific Activities

According to the age distribution box plot, some age groups could be more likely to experience churn.

Recommendations:

- **Youth Programs:** Create engagement initiatives tailored to younger clients, such as workshops on financial literacy, student savings accounts, and investment plans for young people. These initiatives can foster enduring loyalty and financial knowledge (Annamaria Lusardi, Olivia S. Mitchell, 2014).
- **Retirement Planning:** Provide senior clients with expert retirement planning services. To meet their specific financial needs, offer tailored guidance on asset management, pension plans, and retirement savings (Alders, n.d).
- **Middle-Aged Clients:** Pay particular attention to family financial planning services, mortgage alternatives, and investment possibilities that fit the objectives and life stage of middle-aged clients.

## 3. Tracking Credit Scores

Customers with lower credit ratings are more likely to leave, according to the scatter plot.

Recommendations:

- **Credit Score Monitoring:** Put in place a mechanism to keep an eye on credit ratings on a regular basis and spot clients whose scores are dropping. This can assist the bank in proactively contacting clients who are at danger.
- **Financial Counseling:** To assist clients in understanding the elements influencing their credit ratings and offering solutions to raise them, give financial counseling and credit score development programs (Agarwal, Sumit, and Bhashkar Mazumder, 2013).
- **Customized Loan Products:** Provide clients with lower credit ratings with loan products that have adjustable terms and conditions. To promote responsible financial conduct, provide rewards for on-time repayments (Anon., 2001).

## 4. Handling Account with No Balance

There is a significant concentration of churned customers with an account balance of 0 in the histogram.

Recommendations:

- **Proactive Outreach:** Put in place a system to contact clients who have accounts with no balances. Get in touch with them via customized phone calls, emails, or SMS to find out why they keep their balance at zero.
- **Incentives:** Provide rewards for keeping a minimum balance, such as fee exemptions, interest rate advantages, or modest cash bonuses. Customers may be encouraged to keep a positive balance by these incentives (Frederick F. Reichheld, W. Earl Sasser, Jr. , 1990).
- **Financial Counseling:** To assist clients in improving their money management and preserving equilibrium, offer individualized financial counseling sessions. Inform them of the advantages of maintaining a balanced, active account (V. Kumar, Denish Shah, 2004).

## 5. Improving the Customer Experience

The overall distribution of attrition indicates that all segments need to increase customer happiness and loyalty.

Recommendations:

- **Outstanding Customer Service:** Make an investment in educating customer support agents to deliver outstanding customer service (Christian Homburg, Nicole Koschate, Wayne D. Hoyer, 2005). Make sure that consumer questions and concerns are addressed in a timely and efficient manner.
- **User-Friendly Digital Experience:** By offering an intuitive and user-friendly interface, you may improve the digital banking experience. Provide tools like real-time alerts, customized dashboards, and simple account administration.
- **Feedback Mechanisms:** To continually collect and examine consumer input, put feedback mechanisms in place. To improve the overall customer experience, use this input to pinpoint areas that need work and implement the required adjustments (Christian Homburg, Nicole Koschate, Wayne D. Hoyer, 2005).

## F. Conclusion

This report has shed important light on customer attrition in the banking industry. We've found important trends that help guide focused retention tactics by looking at a variety of variables, including gender, age, credit ratings, and account balances. The bank can lower churn, improve client happiness, and cultivate enduring loyalty by addressing these elements with customized interventions. In order to adjust and improve these tactics to satisfy changing client requirements, it will be essential to continuously monitor and analyze customer data.

## G. References

- Agarwal, Sumit, and Bhashkar Mazumder, 2013. Cognitive abilities and household financial decision making. *American Economic Journal: Applied Economics* 5 (1): 193–207.
- Alders, P., n.d. *Human Capital and Retirement*. [Online]  
Available at: <https://papers.tinbergen.nl/99056.pdf>  
[Accessed 2025].
- Annamaria Lusardi, Olivia S. Mitchell, 2014. The Economic Importance of Financial Literacy: Theory and Evidence†. *Journal of Economic Literature* 2014, 52(1), 5–44.
- Anon., 2001. *Do liquidity constraints and interest rates matter for consumer behavior? Evidence from credit card data*, Cambridge: Nber Working Paper Series.
- Christian Homburg, Nicole Koschate, Wayne D. Hoyer, 2005. Do Satisfied Customers Really Pay More? A Study of the Relationship Between Customer Satisfaction and Willingness to Pay. *Journal of Marketing*, 69(2), pp. 84-96.
- Frederick F. Reichheld, W. Earl Sasser, Jr. , 1990. *Zero Defections: Quality Comes to Services*. [Online]  
Available at: <https://hbr.org/1990/09/zero-defections-quality-comes-to-services>  
[Accessed 2 2025].
- Grigovera, S., 2023. *How to approach customer churn measurement in banking*. [Online]  
Available at: <https://uxpressia.com/blog/how-to-approach-customer-churn-measurement-in-banking#:~:text=The%20average%20churn%20rate%20for,10%25%20is%20acceptable%20for%20banks.>  
[Accessed 2025].
- Iqbal, S., 2017. *Women, Business and the Law 2016: Getting to Equal*. [Online]  
Available at: <http://documents.worldbank.org/curated/en/619981501477827438>  
[Accessed 10 2 2025].
- Kuo, H.-L., 2024. *Customer Churn Rate Analysis*. [Online]  
Available at: <https://www.linkedin.com/pulse/customer-churn-rate-analysis-hsiu-ling-kuo-xd7hc/>  
[Accessed 10 2 2025].
- PwC, 2017. *Winning the battle for female talent in banking*. [Online]  
Available at: <https://www.pwc.com/gx/en/about/diversity/iwd/iwd-female-talent-report-web.pdf>  
[Accessed 10 2 2025].
- Ryan Flanigan, C., 2024. *What is credit card churning?*. [Online]  
Available at: <https://www.bankrate.com/credit-cards/advice/what-is-credit-card-churning-and-how-does-it-affect-your-credit-score/>  
[Accessed 10 2 2025].
- V. Kumar, Denish Shah, 2004. Building and sustaining profitable customer loyalty for the 21st century. *Journal of Retailing*, 80(4), pp. 317-329.