Impact of Customer Complaints on Churn in Banking : Statistical Analysis Report

Alexander Alboukhari

ABSTRACT

The following statistical report investigates customer behavior within the banking sector, with a specific focus on exits, complaints, and satisfaction.This report uses statistical analysis and different visualizations to investigate the relationship between complaints and customer exit, the demographic and financial characteristics of customers who are likely to be complaining, and the predictive value of satisfaction ratings. Data shows that there is a positive correlation between customer complaints and customer exits, offering detailed information on the behavioral characteristics of unsatisfied consumers.

1 INTRODUCTION

Banking institutions face the ongoing challenge of customer churn the first and foremost factor is that understanding the behavior and retention of customer mechanisms in the banking industry is more imperative than ever. Such a scenario when customer expectations are high in this sector due to a dynamic environment can be compelling. This report provides more information about the dataset that contains details about how a client interacts and his saving behavior, as well as the level of satisfaction that the client has with banking. The report uses statistics and multiple data displays, including pie charts, scatter plots, heat maps, and histograms, to directly show the causes of churns and customer complaints and highlights a few factors like demography and economics that identify customers with a higher tendency to complain and exit. In regards to this, the report essentially tries to give a clear and detailed explanation of the variables behind customer loyalty in an attempt to highlight the complexity of the space among these variables, which affects the consumer's choice to either stay or change the bank they are with through the data and visuals. This approach offers a chance for banks to display working strategies that can help them to increase customer satisfaction, neutralize customer attrition, and, in the end, promote customer devotion in the competitive market.

2 DATASET DESCRIPTION

The dataset includes a variety of factors that offer a complete picture of consumer profiles and behaviors in the banking setting. Furthermore, it contains personal data, financial activities, and product as interaction parameters including age, gender, account balance, and credit score. It also includes customer satisfaction indicators such as complaints and leave status allowing for a more in-depth investigation of the issues influencing customer retention and churn. However, this dataset serves as the basis for investigating the client loyalty in the banking industry.

3 DATA ANALYSIS

3.1 The proportion of the customers that are still using the banking services compared to those that have left in the period.

An analysis was undertaken to ascertain the proportion of bank clients who have chosen to stay rather than those who have departed. Upon reviewing the bank dataset to differentiate between the clients who stayed and those who left. Identified A pie chart was employed to depict the clientele. Based on the pie chart 80% of the customers which is equivalent to 7,949 clients Stay and 2,030 customers have left, accounting for 20% of the total. While most customers remain loyal a minority has opted to go. An in-depth analysis should be conducted to ascertain the underlying factors behind the departure of 20% of customers since comprehending their motivations can be essential for performance improvement. Moreover, some smart predictions might be employed to ascertain their probable future departures and ensure their pleasure in retaining them.

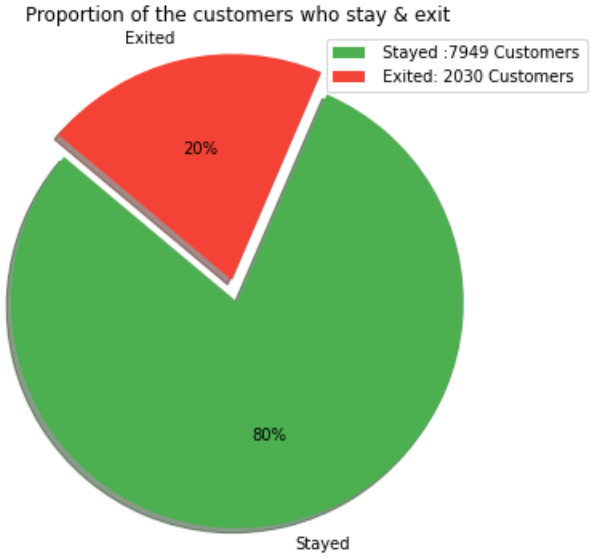

Figure 1: Pie chart showing Proportion of the customers who stay & exit

3.2 The relationship between the number of complaints received by the bank authorities and the number of exited customers.

The analysis was based on the assumption that there is a strong correlation between the number of customer complaints received by bank authorities and the subsequent leaving of customers. Increased

complaints could result in a higher rate of customers leaving their banking services.

The Pearson correlation coefficient was chosen as the appropriate statistical measure to evaluate the relationship between the number of complaints (*Complain*) and the number of customers who exited (*Exited*). [1] The coefficient measures how strong the linear relationship between two variables is giving values ranging from -1 to 1. A value of 1 signifies a perfect positive correlation, 0 represents no linear correlation, and -1 signifies a perfect negative correlation.The Pearson Correlation Coefficient follows the formula below:

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$$

where n is the number of observations, x and y are the variables being analyzed, and $\sum$ denotes the summation.

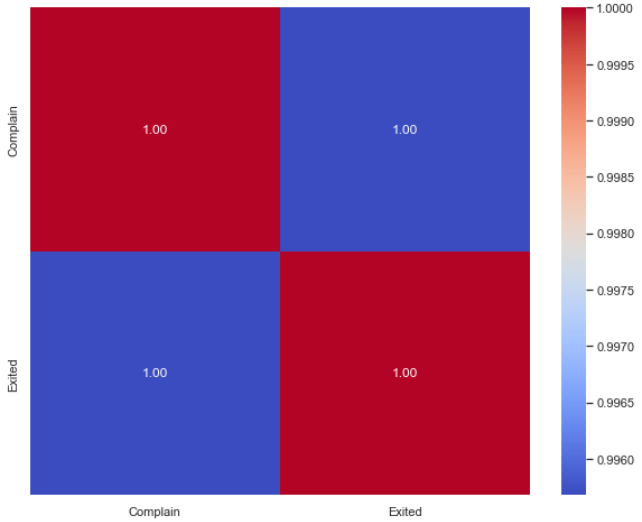

Figure 2: Relationship between Customer complaints and customer exits

The correlation matrix showed a coefficient of about 0.996 for the variables *Complain* and *Exited*, showing a solid positive linear relationship and implying that the rate at which customers complain is an essential sign of customer churn.

3.3 The characteristics and statistics (in terms of gender, age groups, and tenure etc,) of the customers that are more likely to complain.

The analysis assumes that there is a notable correlation between the number of customer complaints received by bank authorities and the subsequent customer departures. It was suggested that complaints could result in a higher rate of clients discontinuing their banking services. The correlation matrix provides a comprehensive overview of the relationships and offers significant insights into factors influencing customer behavior and retention.Through the correlation matrix below, it is concluded that there is a positive relationship between age and complaint (+)0.28, in addition to a positive relationship between balance and complaint(+)0.12.

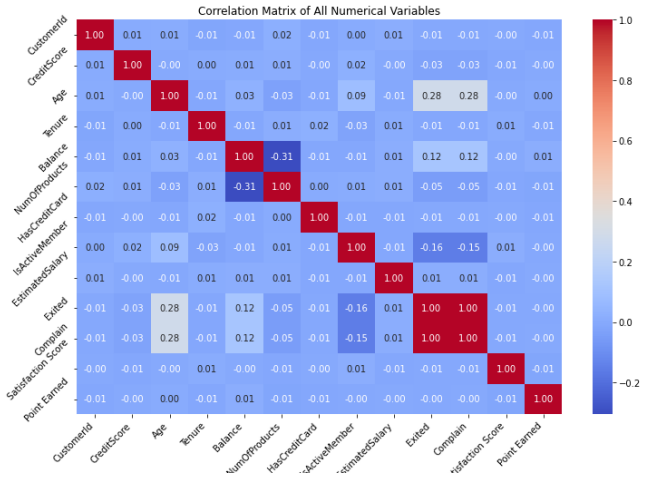

Figure 3: Heatmap of (Persons) correlation matrix of the customers that are more likely to complain

[1] The sample mean $\bar{x}$ is calculated as the sum of all sample values divided by the number of samples $\bar{n}$, which can be expressed as:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

The formula for calculating the standard deviation of a sample is:

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Metric	Mean	Standard Deviation
Age	44.78	9.78
Tenure	4.93	2.94
Balance	91203.84	58347.19
Estimated Salary	101449.58	57852.89

Table 1: Means (x), Standard deviations of groups among those who have complained, which will be used in the formula

First, from the the average customers tenure who file complaints is approximately 4.93 years, with a standard variation of 2.94 years. The duration spans from 0 to 10 years, indicating that the customer who have filed complaints include both recently acquired clients and those who have been with us for an extended period.on another hand ,The most significant complaints came from customers in the 40-50 age group and The average age of those who filed complaints was 44.78 years, with a standard deviation of 9.78 years, indicating a moderate age range among the complainants. The age distribution of the dissatisfied customersspanned from 18 to 84 years.

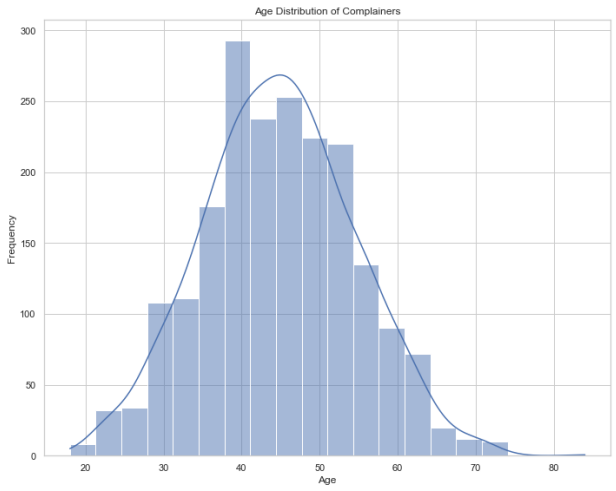

Figure 4: The distribution of age groups among those who have complained

Second, the mean account balance was 91,203.84 $ and had an average salary of 101,449.58$ contradicting the idea that financial difficulties cause most complaints. According to the data the clients with more products and the 'Diamond' card type are associated with a higher rate of complaints. So customer retention data indicates that financial considerations may have an impact on the choice to file a complaint or end banking services.

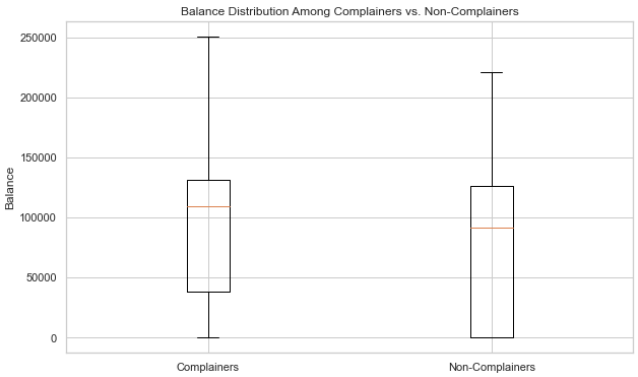

Figure 5: The balance distribution among customers who have complained and non-Complainers

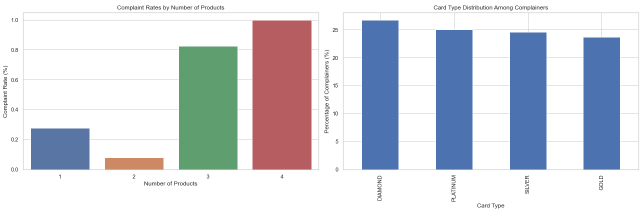

Figure 6: 2 bar chart comparing The distribution Card Type and Number of Products Among Complainers

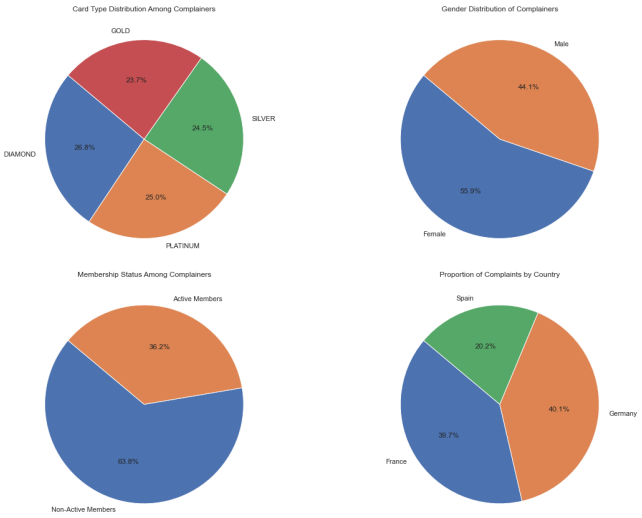

Figure 7: Active Membership Status, Location ,Gender and Card type Distribution Among Complainers

Third, As shown in Figure 7 the data indicates Gender Distribution approximately 55.87% of the complaints are female, while 44.13% are male. These findings suggest that females are significantly more likely to complain than males Also, a geographic distribution analysis of the sources of complaints revealed that Germany and France accounted for almost equal proportions, with Germany representing 40.11% and France representing 39.72%. Spain contributed 20.18% of the total complaints, indicating potential variations in regional service experiences or disparities in customer base sizes.A notable aspect of the analysis is the proportion of complainers who were active members versus non-active members. Specifically, 63.77% of the persons who filed complaints were non-active members, while 36.23% were active. This indicates a greater probability of receiving complaints from less involved or interested customers.

At the end, the analysis identifies a client profile more inclined to complain as generally middle-aged and possibly more frequently female. Typically, most complaints come from customers with 1 to 3 years of tenure. Economically, these individuals usually possess more significant account balances and higher incomes indicating heightened service expectations and the importance of trust and satisfaction in the initial banking relationship these customers generally have higher bank balances and salaries, which may lead to higher service expectations. Geographical trends in complaints also emerged in specific locations showing a higher propensity for complaints. The higher average salary and balance among complainers suggest a correlation between financial expectations and the likelihood of complaints.

3.4 The satisfaction scores on complain resolution provide indication of the customers.

This plot offers views on the satisfaction score distribution between customers and their correlation with customer exit, and

the count of customers in each satisfaction score category enables us to determine significant changes in satisfaction levels between customers who leave the bank and those who stay.However, no striking differences existed between the numbers of Exited and not Exited customers across the satisfaction scores. As captured by these scores, it is similar between those who stay and those who leave, suggesting that satisfaction scores are not a strong indicator of exit likelihood.

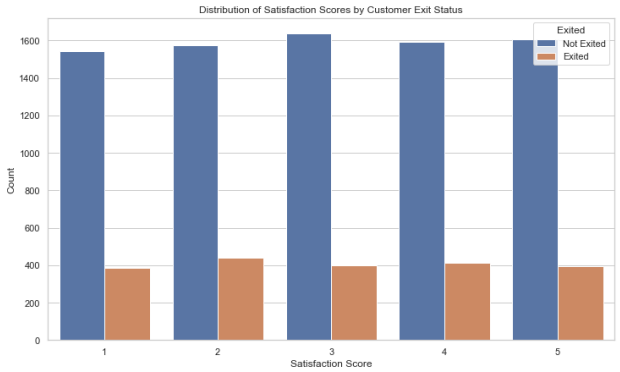

Figure 8: Distribution of Satisfaction Scores by Customer Exit Status

4 LIMITATIONS

The report admits a number of limitations. The results might not apply to different financial institutions or geographic areas, and the dataset might not fairly represent all demographic groups. Furthermore, the research is restricted to the data that is accessible, and other unknown variables that impact client retention but are not included in the dataset could exist.

5 CONCLUSION

The analysis within this report has provided significant insights into customer behavior and retention in the banking sector also highlighting the strong correlation between customer complaints and exist. However, financial stability does not guarantee customer loyalty which highlights the importance of banks focusing on customer service and satisfaction in their efforts to retain customers. Future studies should conduct long-term research to validate these findings and investigate the effects of specific customer service improvements over an extended period of time.

REFERENCES

[1] D. Lane, D. Scott, M. Hebl, R. Guerra, D. Osherson, and H. Zimmer. *Introduction to Statistics.* David Lane, 2003.