CREDIT CARD CUSTOMER CHURN ANALYSIS

BY: AKANCHHA CHOUDHARY

Abstract

Customer Churn Dataset

For this project I am using credit card customer churn prediction/behaviour. I have collected csv dataset from kaggle which can be downloaded with this link.

“

https://www.kaggle.com/code/kmalit/bank-customer-churn-prediction/input

“ It includes 14 features for three countries.

Exploring various behavious of customer churn.

I have tried to understand reason for retaining or churning of customer of a creditcard company.I wants to understand how the variables or features given in dataset related with each other and what features has impact on customer churning or retaining.

Motivation

Customer Churn is one of the important concern in business,banking,marketing and sales sector. This issue is important because it costs more to acquire new customers than it does to retain existing customers. This gives some hard truth to the business owner and they always look for best way to find solution for this problem. This project explores the underlying issues that contribute to churn, predict how churn will impact your business, and find ways to combat it within a given dataset. This will give a great intuitive to business owner to retain more customers and keep their company healthy. If not given required

Dataset(s)

For this project I am using credit card customer churn prediction/behaviour. I have collected csv dataset from kaggle which can be downloaded with this link.

“https://www.kaggle.com/code/kmalit/bank-customer-churn-prediction/input

It includes 14 features for three countries and has around 12000 datasample. The dataset contains 14 columns namely “customerId” , “surname” , “creditScore”, “ Geography” , “Gender” , “Age” , “tenure” , “Balance” , “NumberOfProducts” , “HasCard” , “IsActiveMember” , “EstimatedSalary” , and Exited. It has three data types : int , object and float.

For the analysis I have choosen “Exited” column as target variable and find how other independent features impact it.

Data Preparation and Cleaning

To analyse data I need to collect data and load into jupyter notebook. To remove the column or features that I think is not needed for the analysis. To remove or replace any null values. Its important to do proper check if data types are correctly placed. Wrong data types could lead to incorrect conclusion. Also it very important to see if data distribution of the column is balanced or not. An unbalanced distribution could lead to biased result. To perform correlation I need to scale the dataset with string colum into categorical value.

Research Question(s)

What and how do independent features of dataset causes in churning of the customer of credit card holder and what solution can be thought of to retain the customer.

Methods

For this project I have used EDA(Exploratory data analysis) method because I needed to know the relationship between independent variables and how do they impact the target variable which in this case is the column ‘Exited’ of dataset. I used statistical and visualisation tool to get insight of the relation .

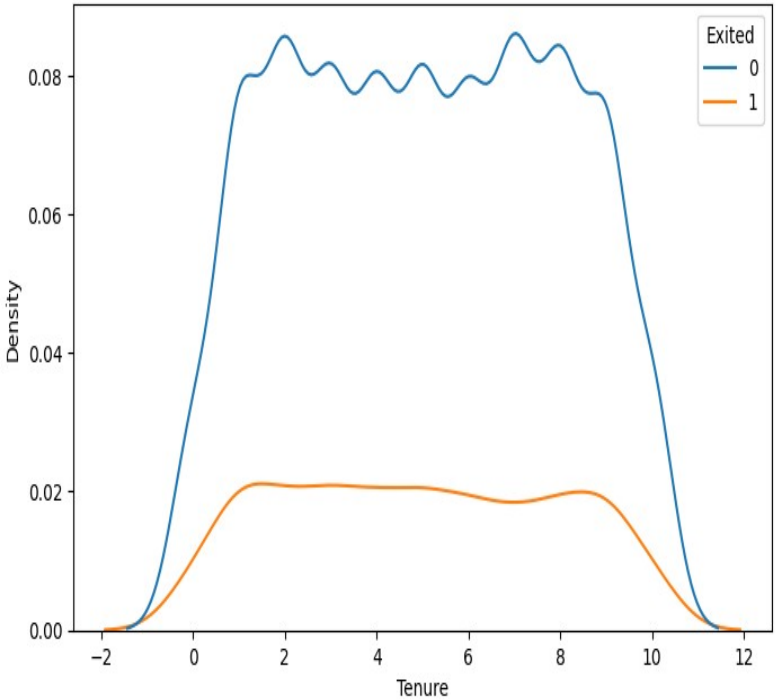
FINDINGS: I list my finding below for this dataset

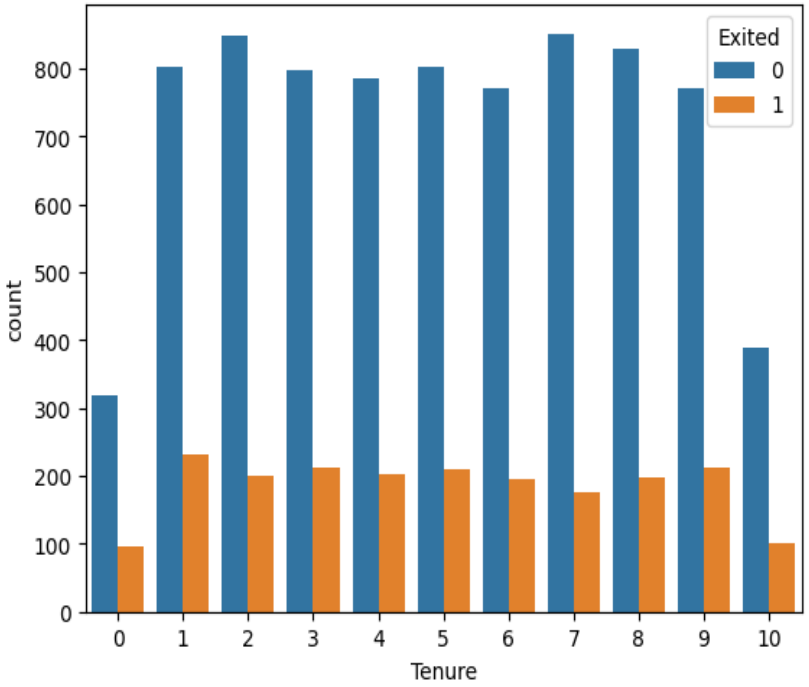
- Tenure" and "Exited" are negatively correlated means less the tenure more the chances of Exit. So company need to increase tenure time to retain customer
- customer from France are less likely to churn and customer from Germany are most likely to churn so bank should focus on undrstanding the customer issues in Germany specifically
- Gender has impact on churn. Female are most likely to churn.
- the customer who are more active are less likely to churn which means companies need to take measure that create more engagement for customer to retain them, for example giving more offers
- customer with more number of products are more likely to be more active hence less likely to churn
- the customer with more credit score are less likely to churn
- the most regular customer are from France are less likely to churn and customer who are from Germany are more likely to churn
- Customer above the age of 40 to 50 are more likely to churn.

I have added some visualisation of some of the above finding in subsequent slides for audience to get better insight.

Findings : Visualisation.

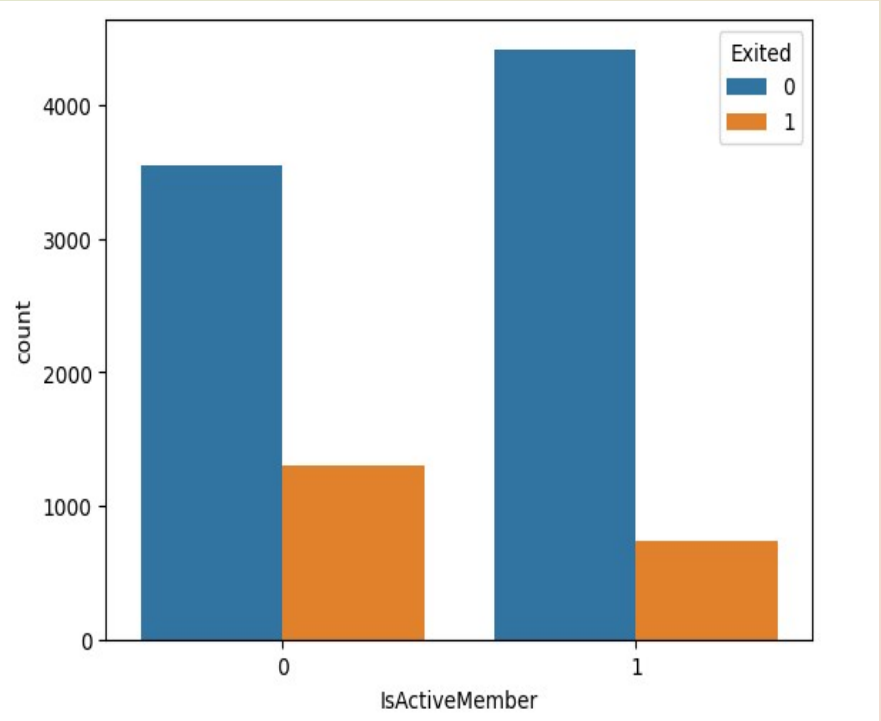
Tenure of the contract are positively correlated with churining of customer

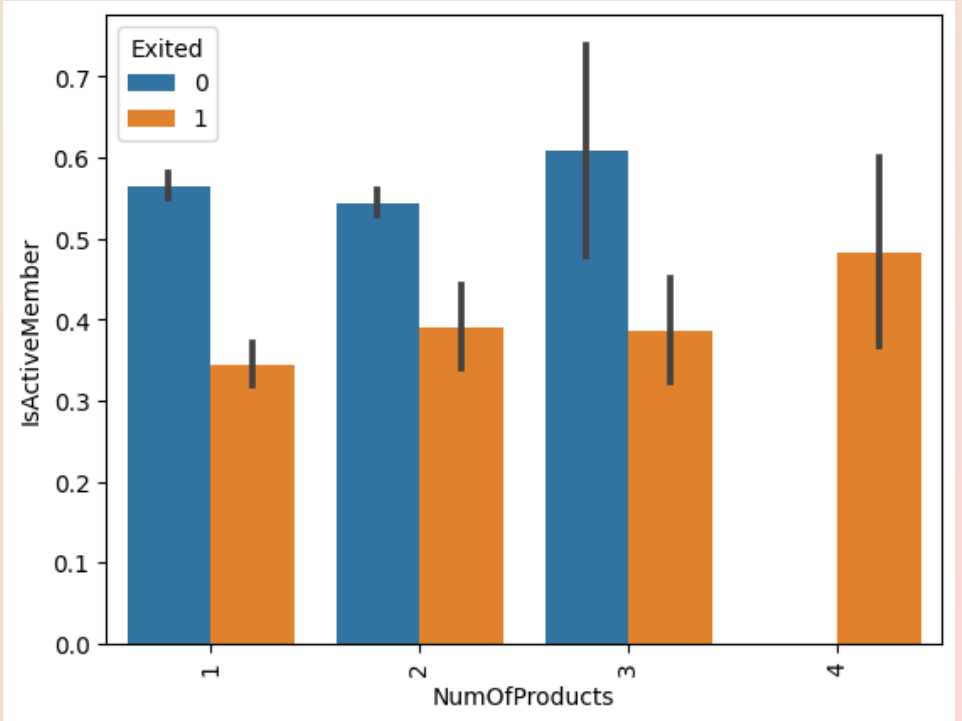




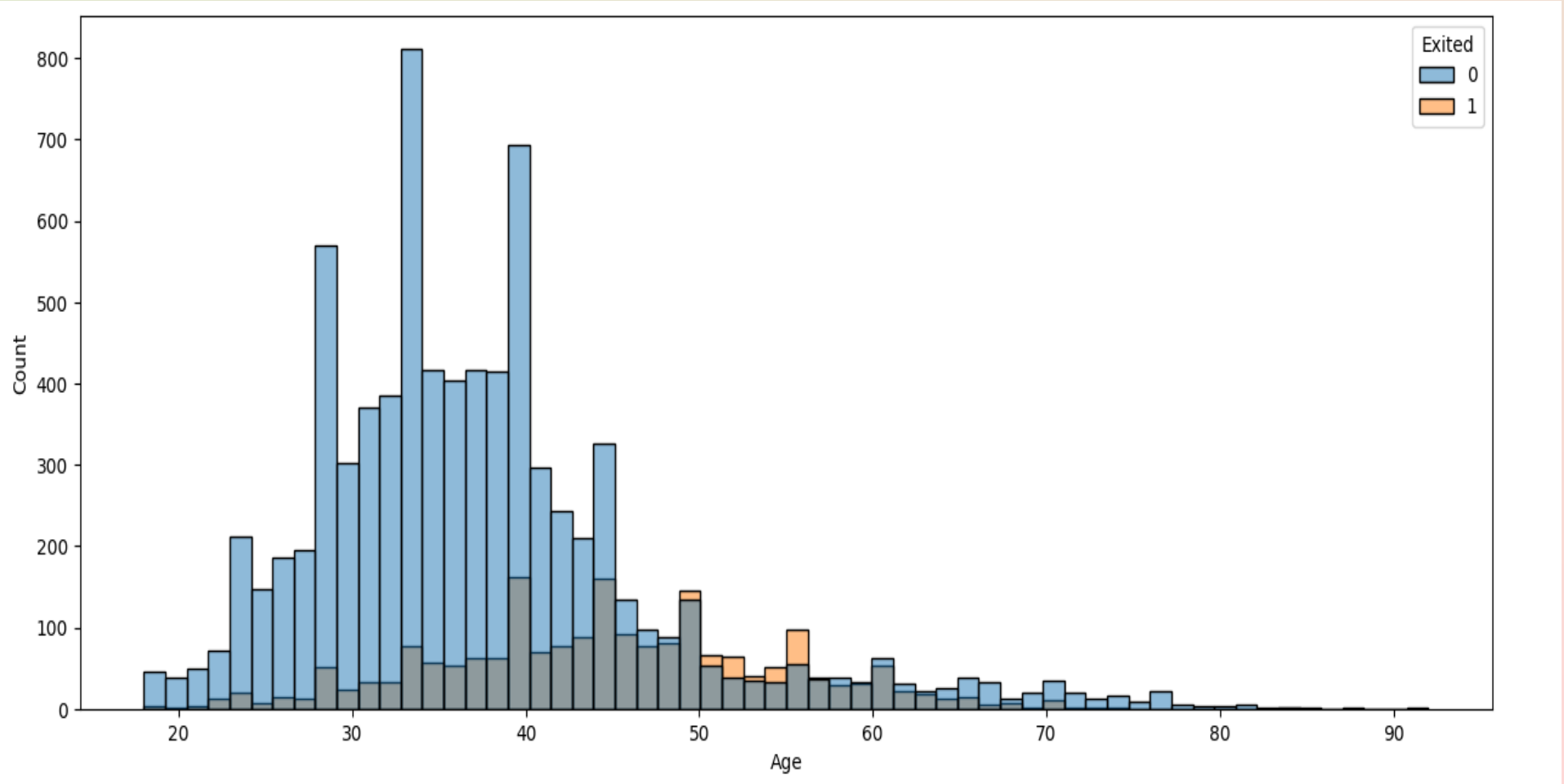
The 1st figure is a line grpah that shows with more tenure period the customer is less likely to churn and 2nd figure give same information in detailed views. From 2nd figure we can say that customer have already churned within 1st tenure.

Another interesting finding in this analysis is that independent variable also shows some correlation. For example I found that if a customer is active member then they are less likely to churn and if a customer has more number of products of a company then are more likely to be a more active member. See the grpah below/.





Age have also shown correlation with customer churn. It seems customer with younger age are less likely to churn. See the graph below.



Limitations

The analysis that I performed here was for 10000 samples. The data distribution of target variable were not even. Only 20 percent of customer have churned according to data but I believe if there had enough sample and more balanced dustribution then there could have been possibility of getting more accurate insight. However I assume that the findings presented in this project is yet very impactful and effective .It can help company in understanding the customer behaviour pattern and make marketing startegies according to findings to reatain the existing customer.

Conclusions

- Tenure" and "Exited" are negatively correlated means less the tenure more the chances of Exit. So company need to increase tenure time to retain customer
- customer from France are less likely to churn and customer from Germany are most likely to churn so bank should focus on undrstanding the customer issues in Germany specifically
- Gender has impact on churn. Female are most likely to churn.
- the customer who are more active are less likely to churn which means companies need to take measure that create more engagement for customer to retain them, for example giving more offers
- customer with more number of products are more likely to be more active hence less likely to churn
- the customer with more credit score are less likely to churn
- the most regular customer are from France are less likely to churn and customer who are from Germany are more likely to churn
- Customer above the age of 40 to 50 are more likely to churn.

Acknowledgements

I have done this work independently but have looked outside for more information to make my project look like professional. I have taken feedback from my friends as well and have improved where there was possibility to do better.

References

I have not read or cite any paper for this project. But i did read on internet to know more about customer churning and how it impacts the industries.