Thera Bank

Credit Card Review

August 7th, 2021

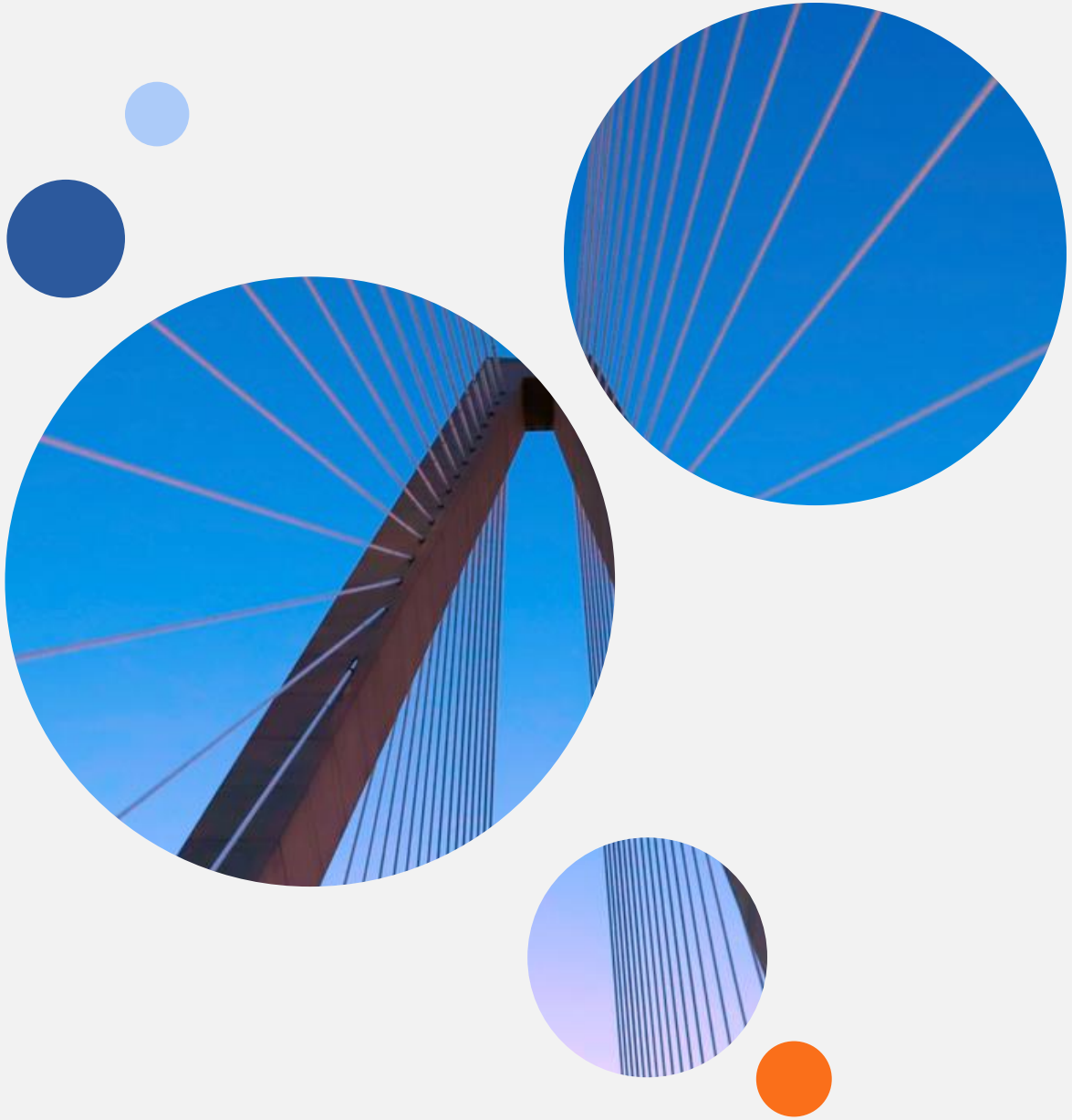
Data Science Team

Contact

Joe Balog

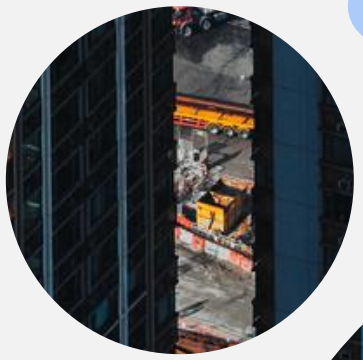
josephmbalog@gmail.com

www.linkedin.com/in/joebalog/



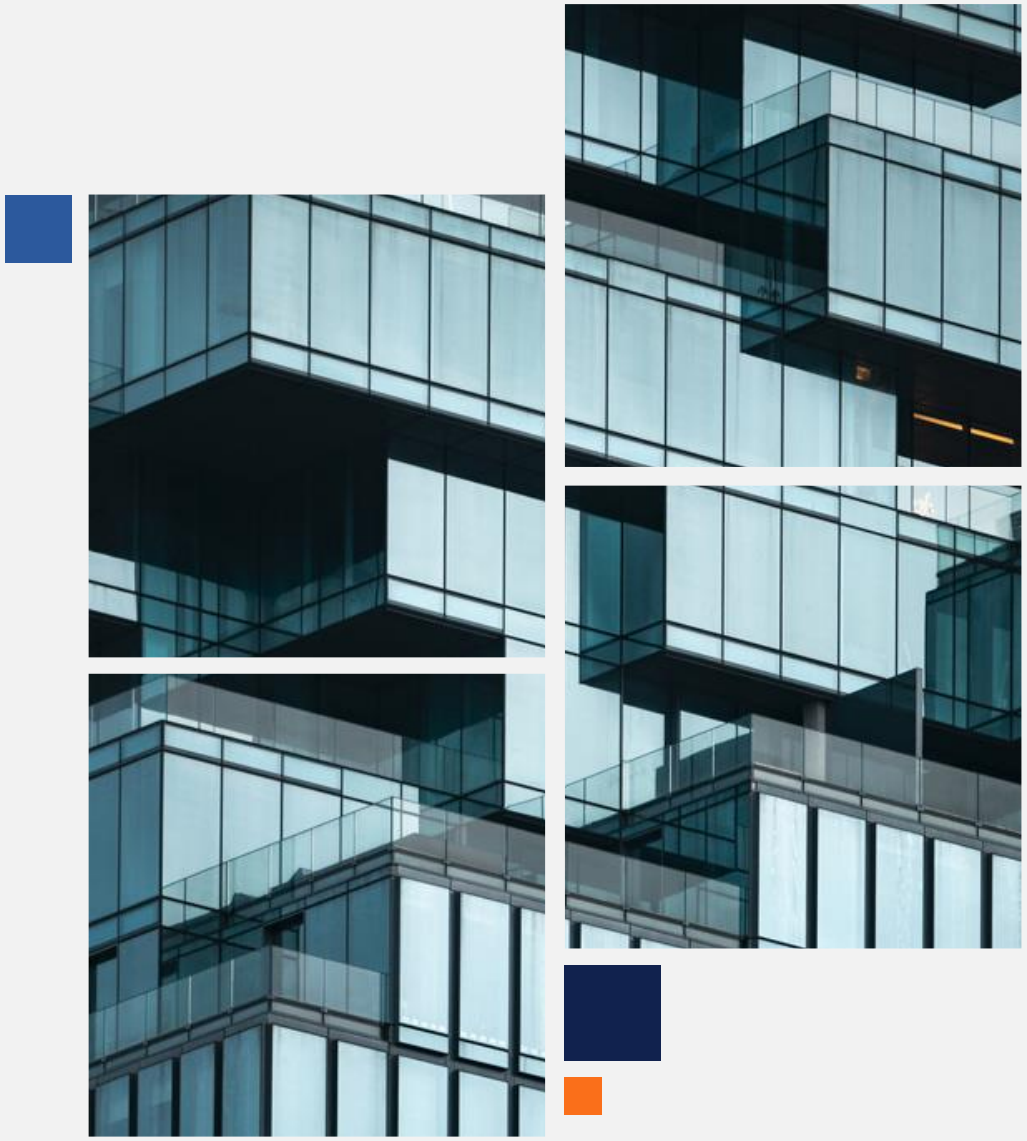
Agenda

- Business Overview of the Problem and Solution Approach
- Key Findings and Insights Which Can Drive Business Decisions
- Model Overview and Performance Summary
- Business Recommendations



Introduction

The Thera bank recently saw a steep decline in the number of users of their credit card, credit cards are a good source of income for banks because of different kinds of fees charged by the banks like annual fees, balance transfer fees, and cash advance fees, late payment fees, foreign transaction fees, and others. Some fees are charged to every user irrespective of usage, while others are charged under specified circumstances.



Business Overview of the Problem and Solution Approach

Let’s dive in

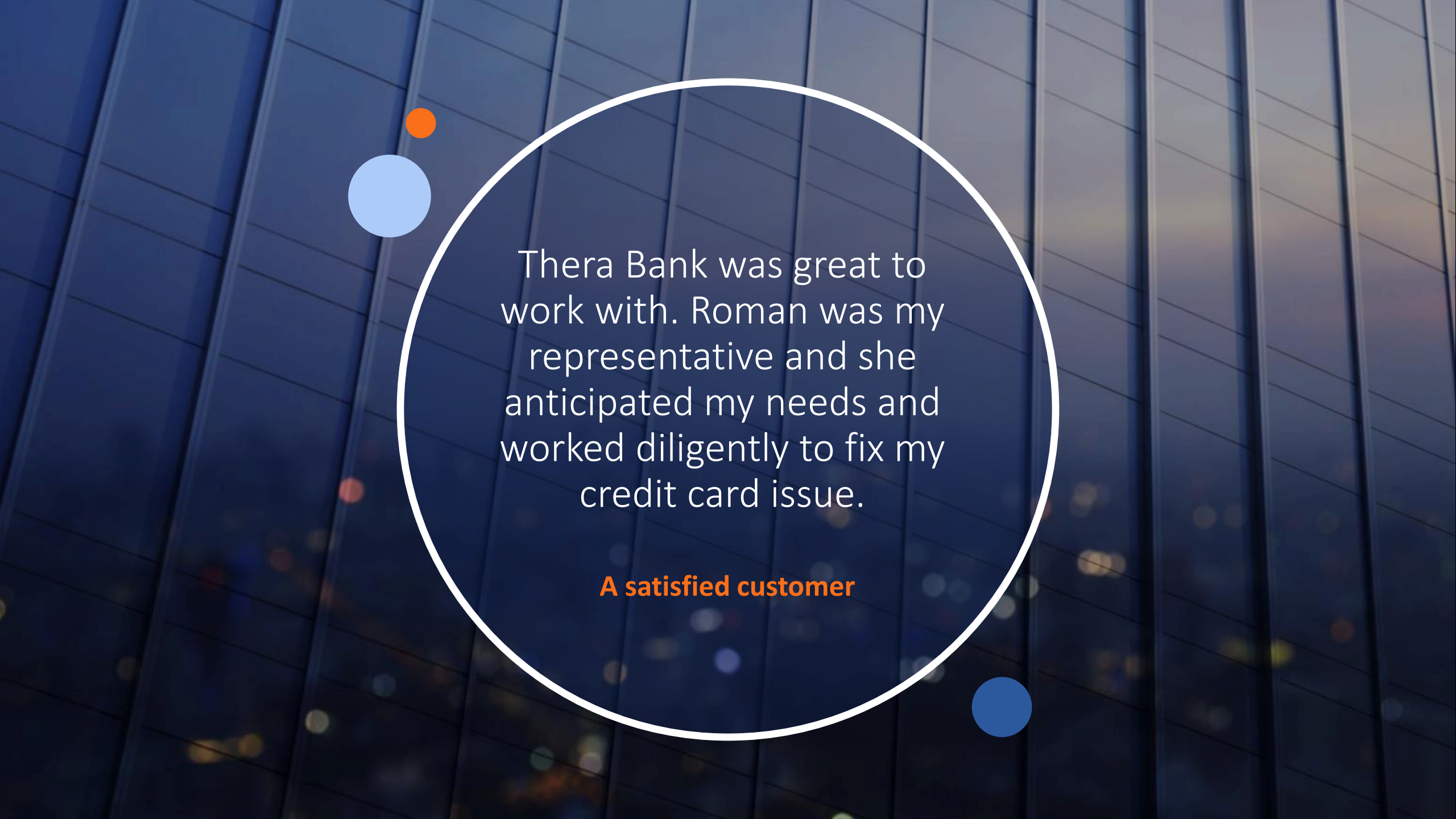
Problem

Customers’ leaving credit cards services would lead bank to loss, so the bank wants to analyze the data of customers and identify the customers who will leave their credit card services and reason for same – so that bank could improve upon those areas

Solution

You as a Data scientist at Thera bank need to create a classification model that will help the bank improve its services so that customers do not renounce their credit cards

You need to identify the best possible model that will give the required performance


Thera Bank was great to work with. Roman was my representative and she anticipated my needs and worked diligently to fix my credit card issue.
A satisfied customer

Team

Ana

Data Engineer

Roman

Credit Sales Representative

Joe

Data Scientist

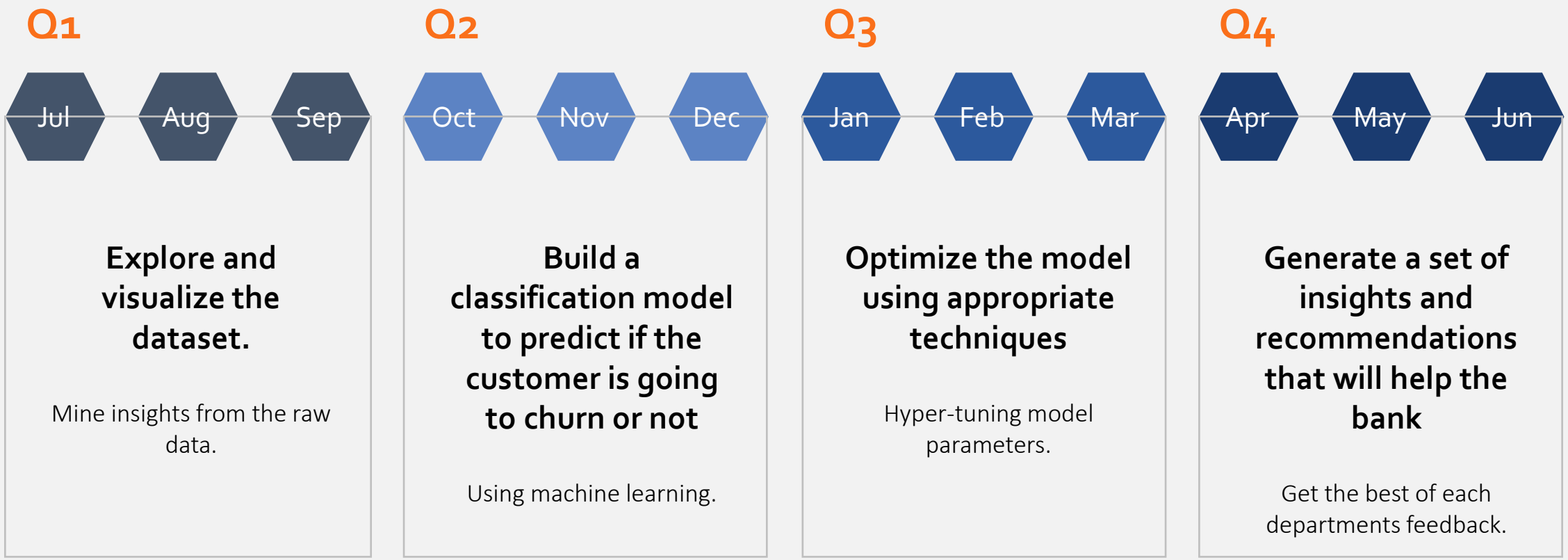
Jim

Credit Services Representative

Larissa

Marketing Manager

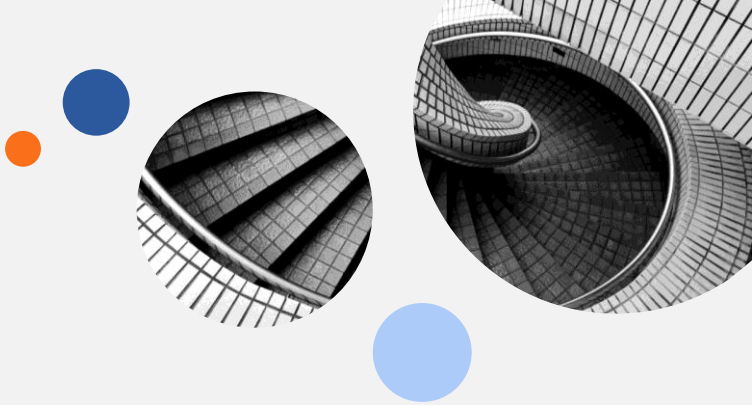
Project Timeline



Data Catalog

Feature	Description
CLIENTNUM	Client number. Unique identifier for the customer holding the account
Attrition_Flag	Internal event (customer activity) variable - if the account is closed then "Attrited Customer" else "Existing Customer"
Customer_Age	Age in Years
Gender	Gender of the account holder
Dependent_count	Number of dependents
Education_Level	Educational Qualification of the account holder - Graduate, High School, Unknown, Uneducated, College(refers to a college student), Post-Graduate, Doctorate.
Marital_Status	Marital Status of the account holder
Income_Category	Annual Income Category of the account holder
Card_Category	Type of Card
Months_on_book	Period of relationship with the bank
Total_Relationship_Count	Total no. of products held by the customer
Months_Inactive_12_mon	No. of months inactive in the last 12 months
Contacts_Count_12_mon	No. of Contacts between the customer and bank in the last 12 months
Credit_Limit	Credit Limit on the Credit Card
Total_Revolving_Bal	The balance that carries over from one month to the next is the revolving balance
Avg_Open_To_Buy	Open to Buy refers to the amount left on the credit card to use (Average of last 12 months)
Total_Trans_Amt	Total Transaction Amount (Last 12 months)
Total_Trans_Ct	Total Transaction Count (Last 12 months)
Total_Ct_Chng_Q4_Q1	Ratio of the total transaction count in 4th quarter and the total transaction count in 1st quarter
Total_Amt_Chng_Q4_Q1	Ratio of the total transaction amount in 4th quarter and the total transaction amount in 1st quarter
Avg_Utilization_Ratio	Represents how much of the available credit the customer spent

Key Findings and Insights



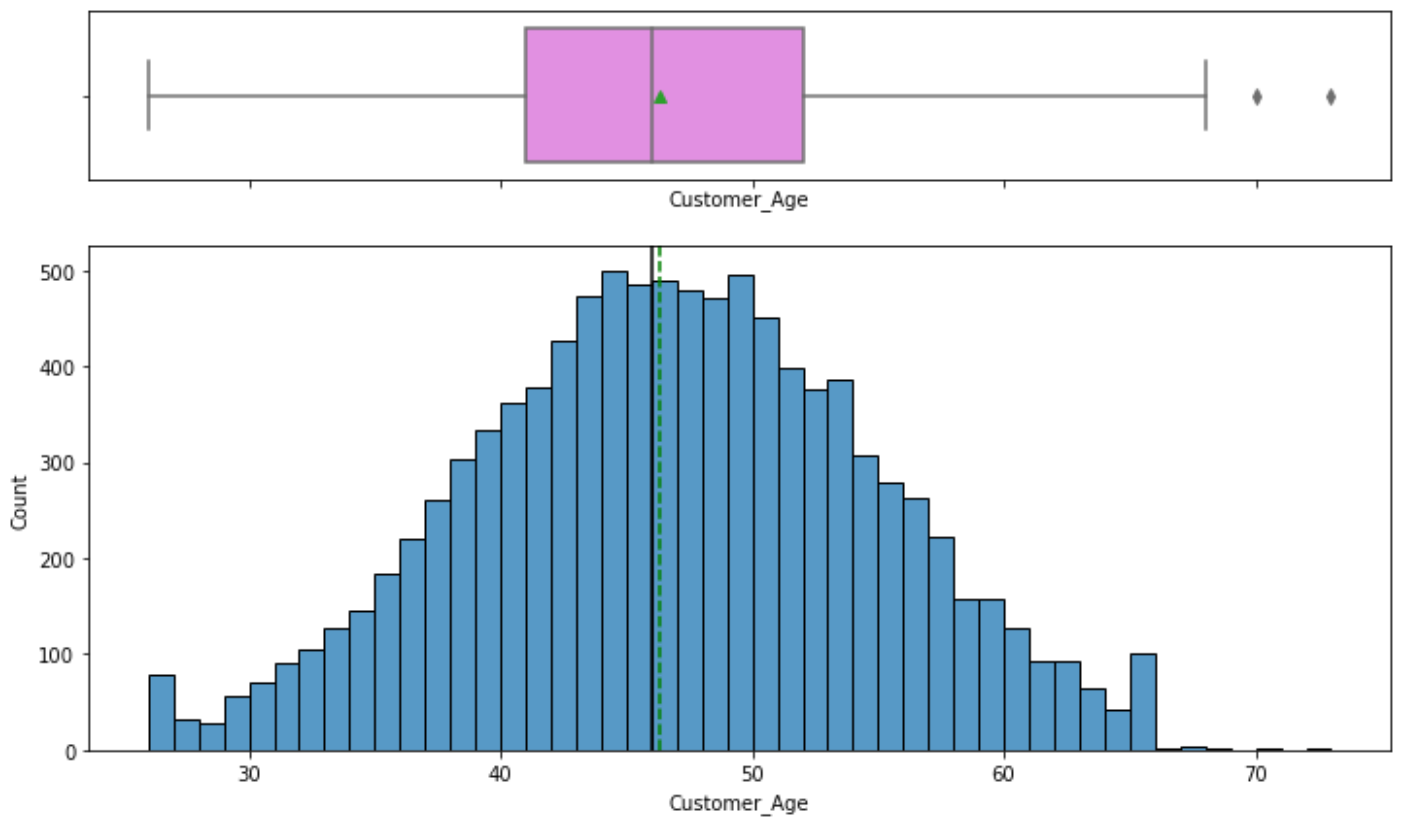
Exploring the data

- Most customers are female
- Most have graduate degrees
- Most make less than $40k a year.

Attrition Identifiers

- There is little difference in likelihood of attrition based on age, gender, or marriage status
- Doctorate degree holders have the highest ratio.
- 6 contacts in the last 12 months is a historical predictor for attrition.

Customer Age



The average age is 46.

The curve follows a normal distribution that coincides with the general population (following the Law of Big Numbers).

The median and mean are very close.

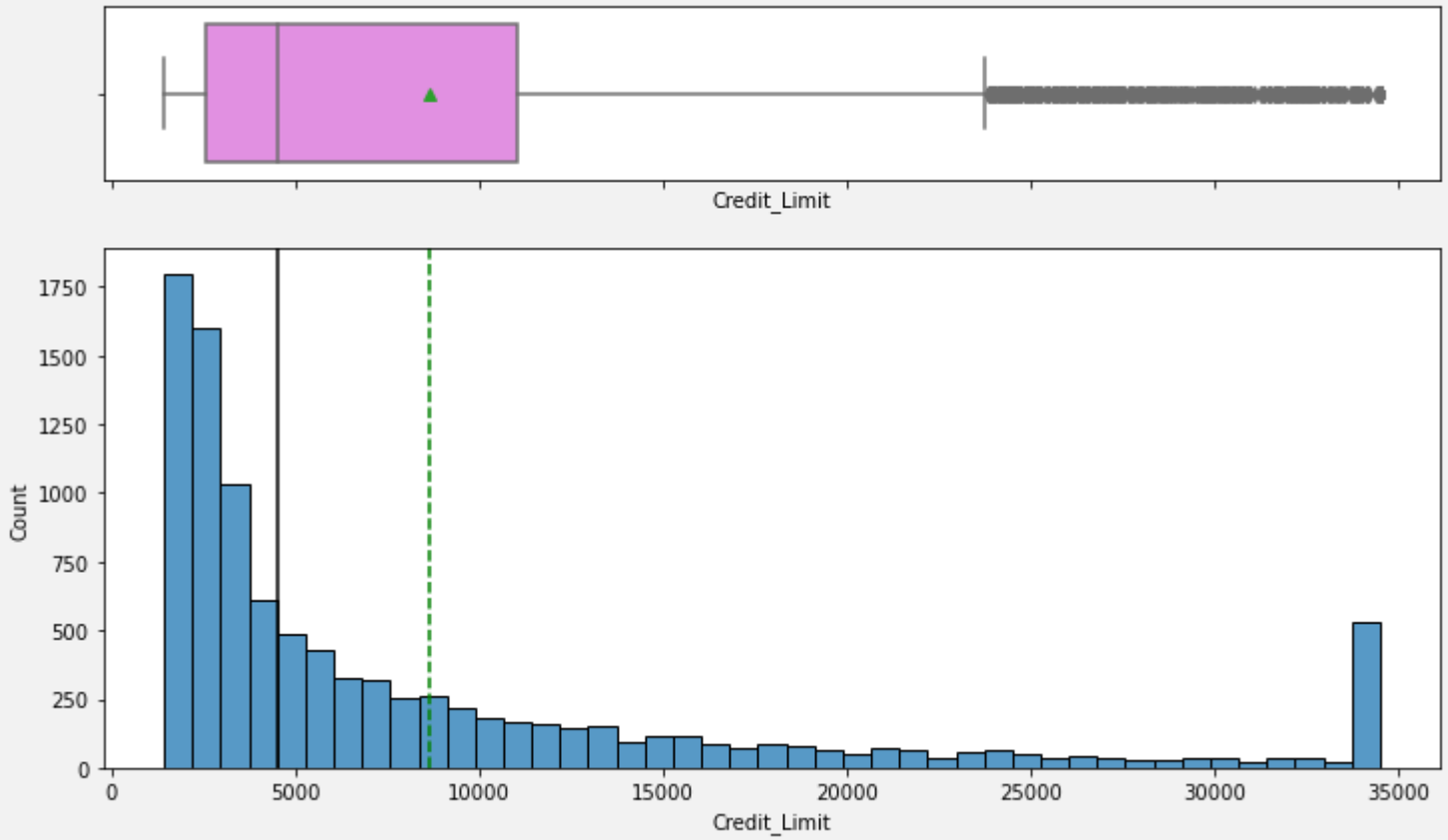
There are some outliers above 65 years of age.

Credit Limit

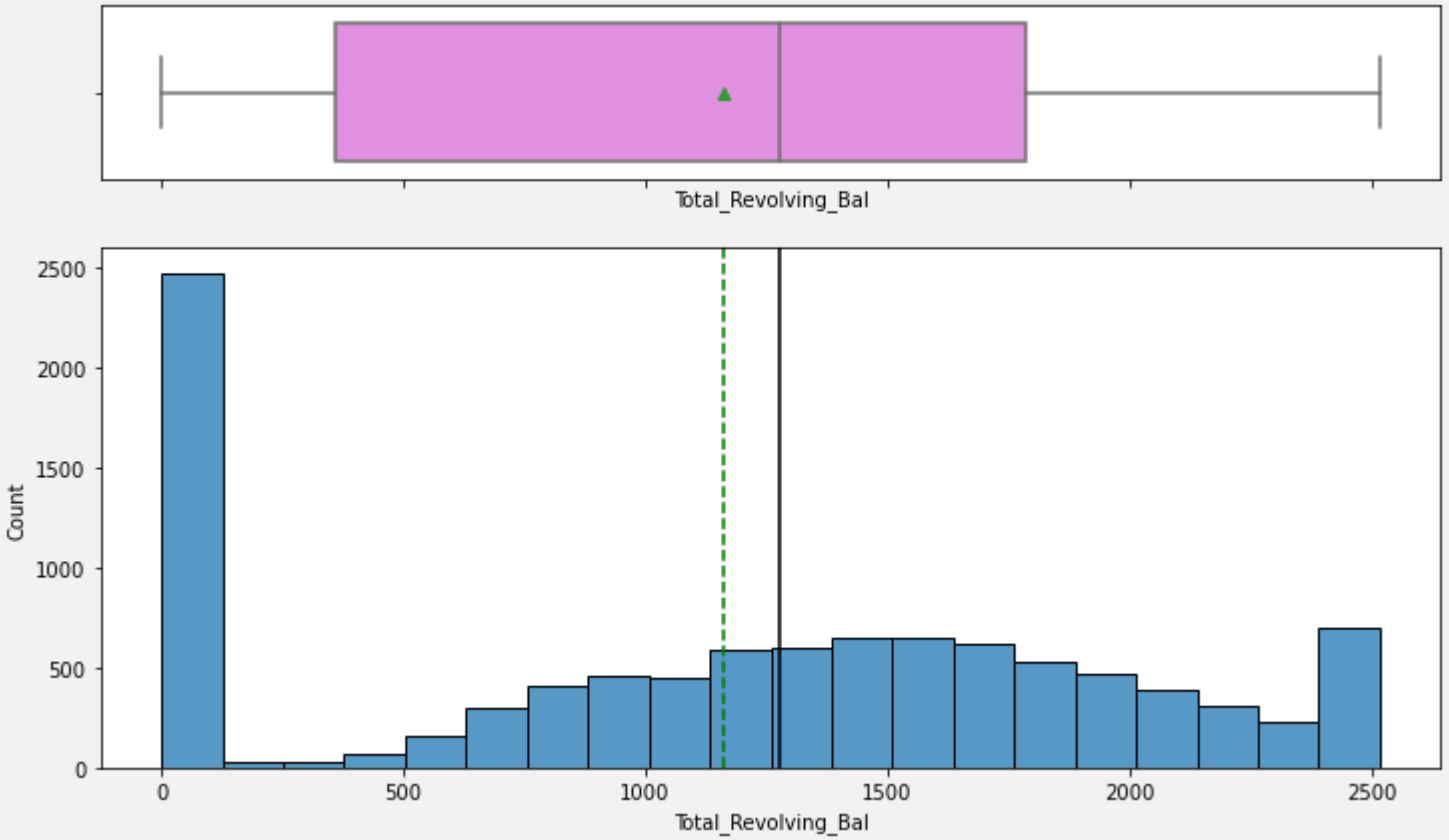
The credit card limit is right skewed and leads data to the right.

There is a set of outliers right below $35,000.

The majority of customers have a credit limit less than $10,000.



Month-To-Month Balance

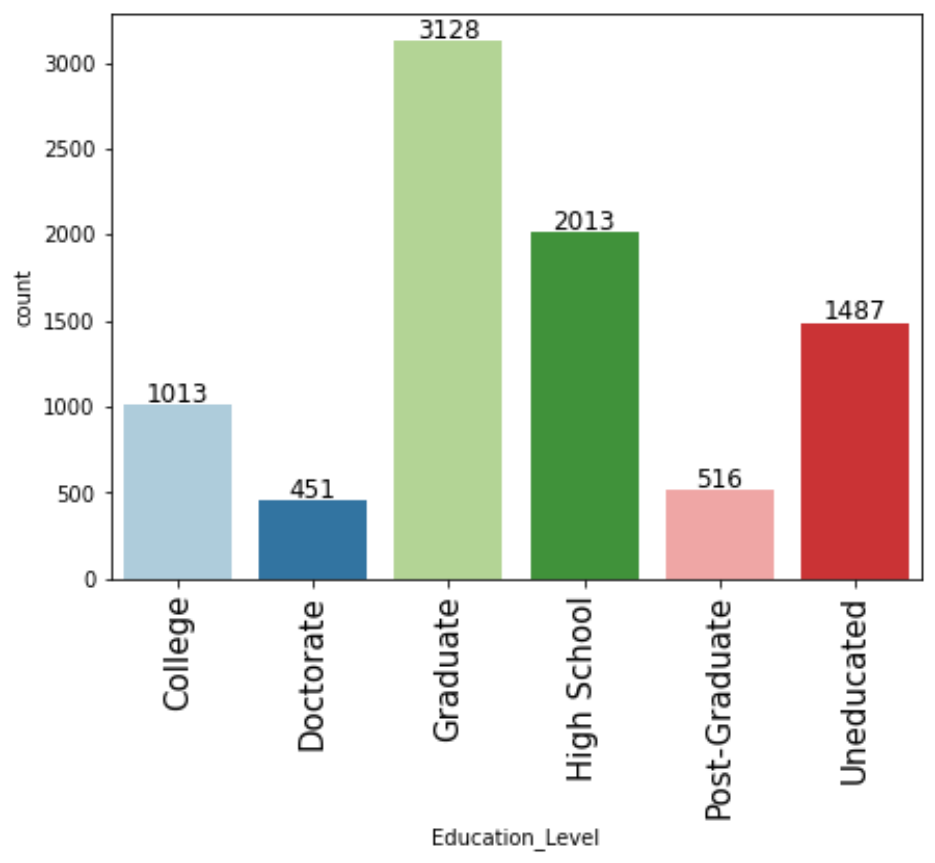


The monthly balance would be normally distributed if it weren't for a large constituency of customers that keep balances at or near zero every month.

50% of customers keep balances below $1,276.

The highest balance is $2,517.

Types of Customers



Customers tend to hold graduate degrees or high school diplomas.

There is a large amount of uneducated customers using credit cards at TheraBank.

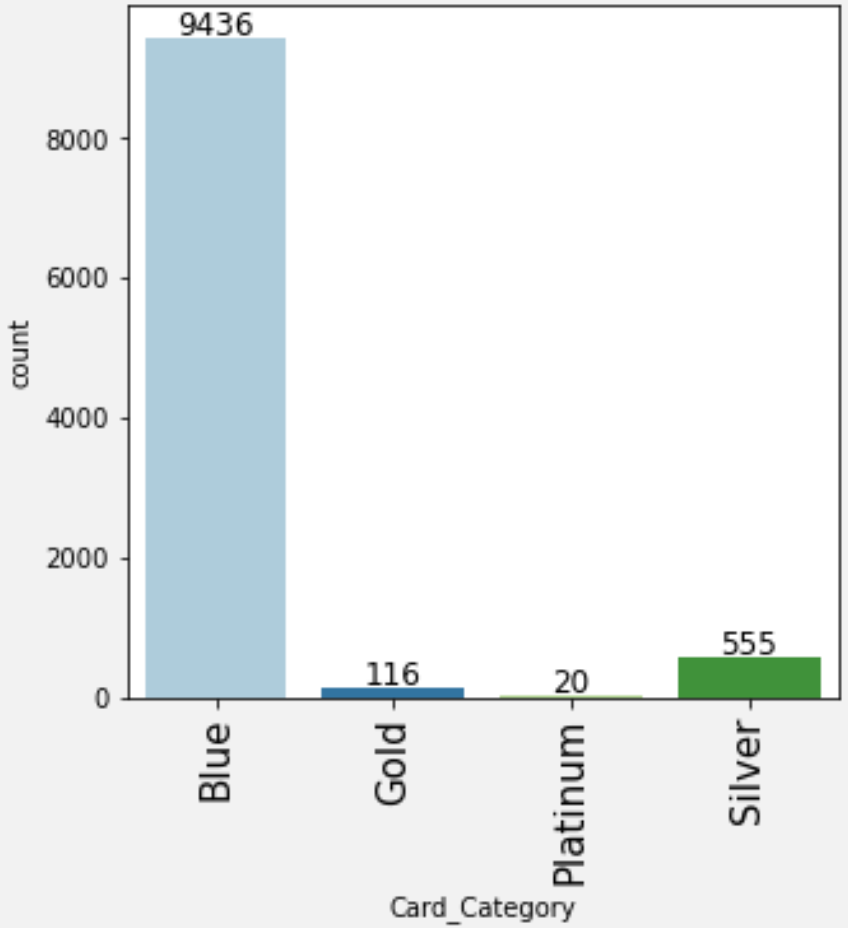
There is a skew in graduate degree holders that seems out of balance with the fact that college is less represented.

Types of Customers

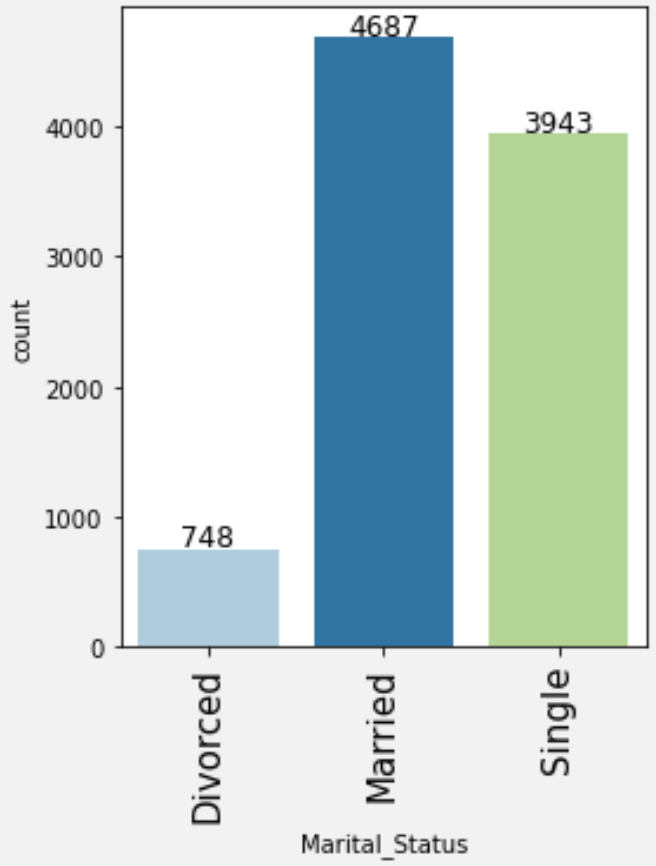
Most of these customers use the Blue credit card.

The least used credit card type is Platinum.

The reasons for people going from Blue to Silver or Gold must be compelling as there are very few compared to Blue.



Types of Customers



Customers tend to be married or single.

Most customers are married and might be using more than one credit card.

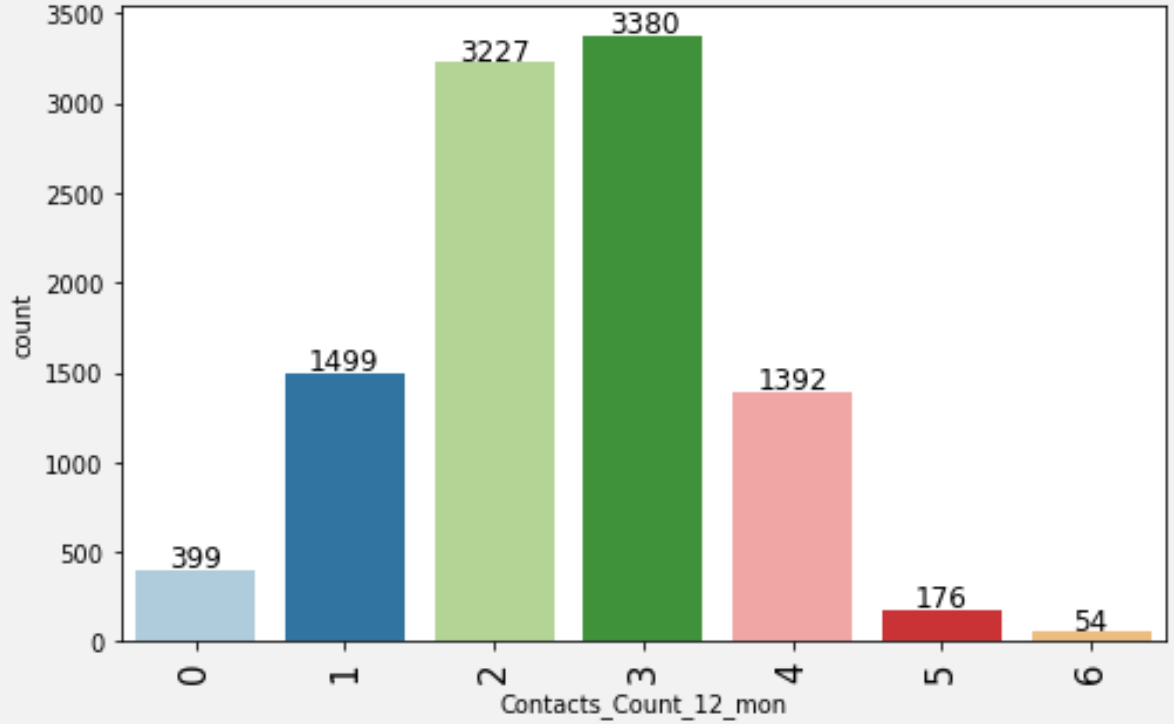
The differentiation between a divorced person and single person needs to be investigated further to further specify these categories.

Types of Customers

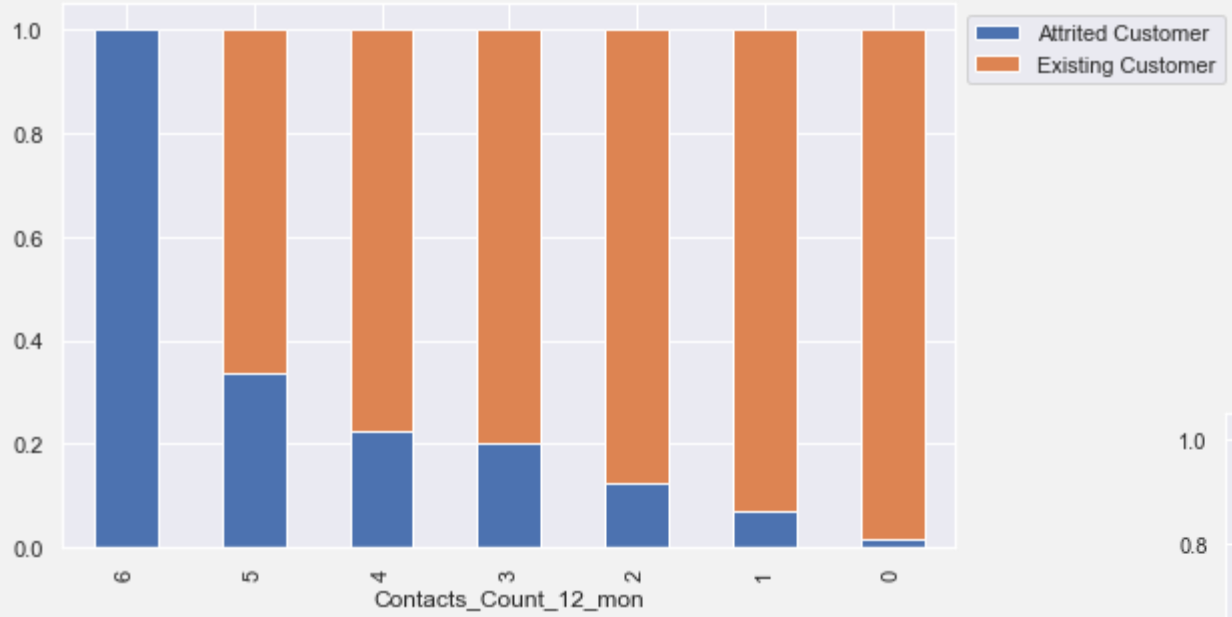
Most customers have had at least 2 contacts from a bank representative in the last 12 months.

There is a target of ~400 customers who have had no contact.

The drop-off from 3 to 4 calls is dramatic and represents the shift from 25% to 33% in contact rate over 12 months.

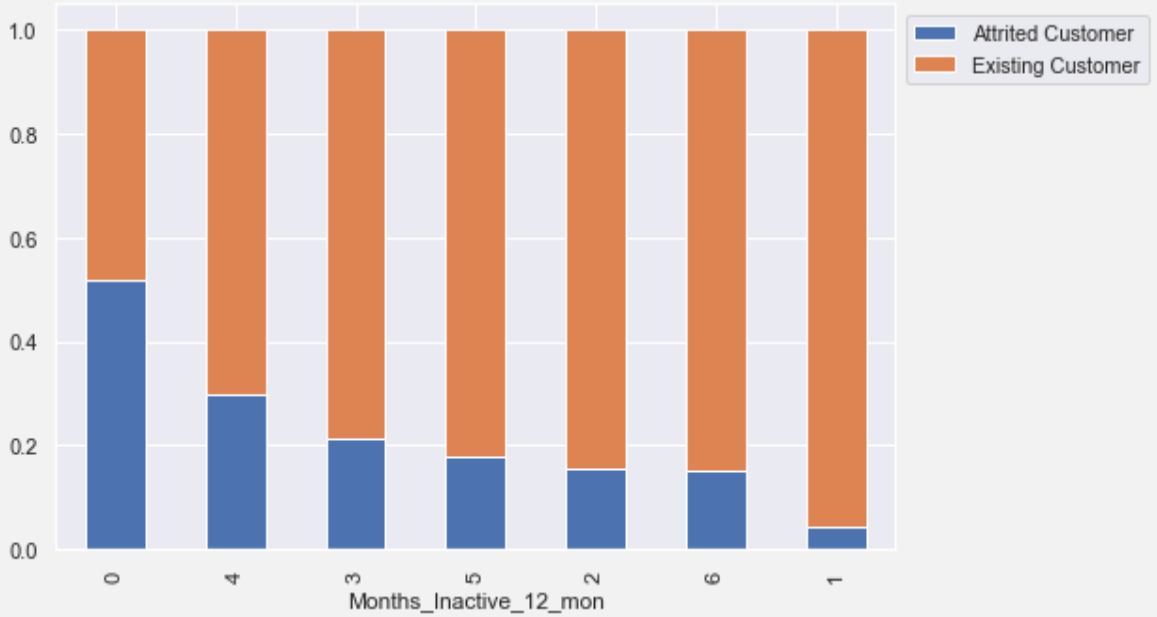


Attrition Indicators

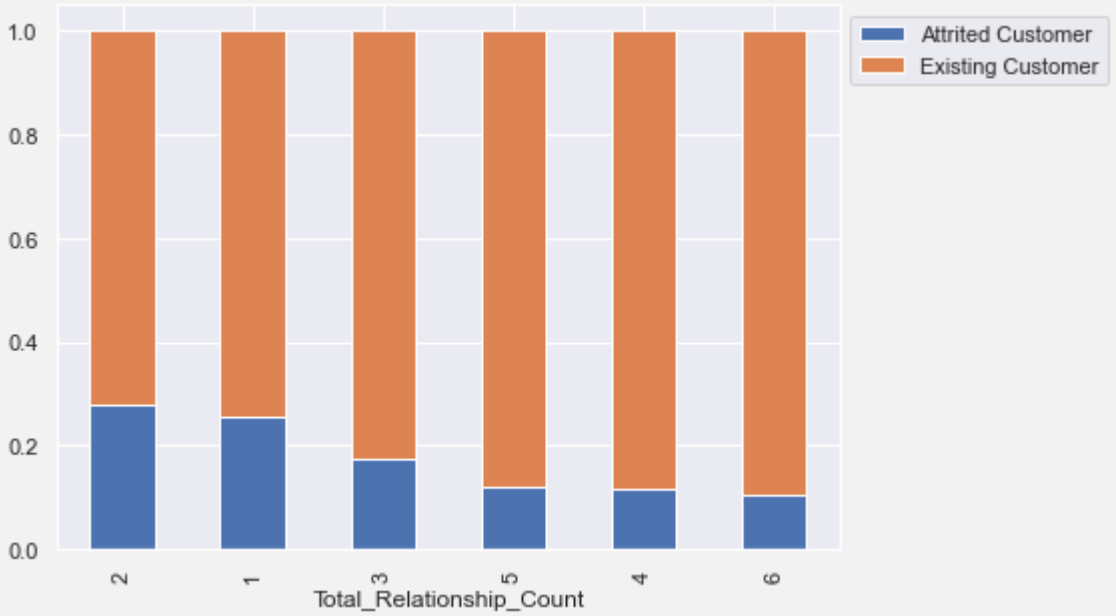


The people who get called the most tend to drop their credit cards.

Watch out for customers who have not used their card in 4 months but don't count on it as the majority of attritees use their card regularly.

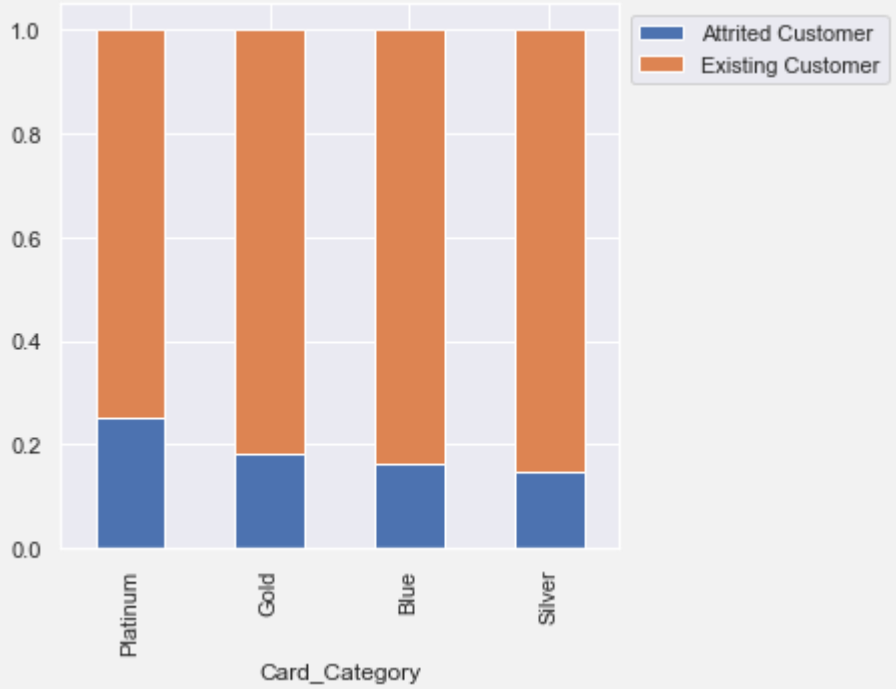


Attrition Indicators



The more products a customer's hold doesn't necessarily correlate to lower attrition events.

The very few Platinum card holders have attired which points to a problem with that card's plan.



Attrition Indicators

The major factors in historical attrition by data variable.

43 - 48 years old	Females	3 dependents	Graduate	Married	Less than $40k
Blue & Platinum Card holders	36-39 month on the books or more	2 - 3 products	3 months of inactivity	3 Contacts in the last 12 moths.	~$1,438.00 credit limit
~$0.00 monthly carry-over balance.	~ 1,438.00 average amount left on the credit card in the last 12 months	0.6% of the Q4 amount left in the account in Q1.	$2,108 - $2,312 Last 12 months of transaction total	40 - 44 Total transactions in the last 12 months	0.5 - 0.6% of Q4 transactions found in Q1
		0.0% Of amount of credit the customer spent			

Attrition: High Likelihood

Customers around these features and values need to be focused on.

- Doctorate degrees
- Inactivity in the last 12 months: 0 months
- No. of Contacts between the customer and bank in the last 12 months: 6
- The balance that carries over from one month to the next: $0
- Credit Limit: ~$3,261.00
- Ratio of the total transaction amount in 4th quarter and the total transaction amount in 1st quarter: ~0.7
- Ratio of how much of the available credit the customer spent: ~0.3

Model Overview and Performance

Models Tested

- Logistic Regression
- DecisionTree Classifier
- GradientBoosting Classifier
- AdaBoost Classifier
- XGBoost Classifier
- Bagging Classifier

Metrics for Success

- Recall to focus on positivity for attrition.
- Target of reducing false negative error for prediction of closing an account when they have not in the past.

Scores Used

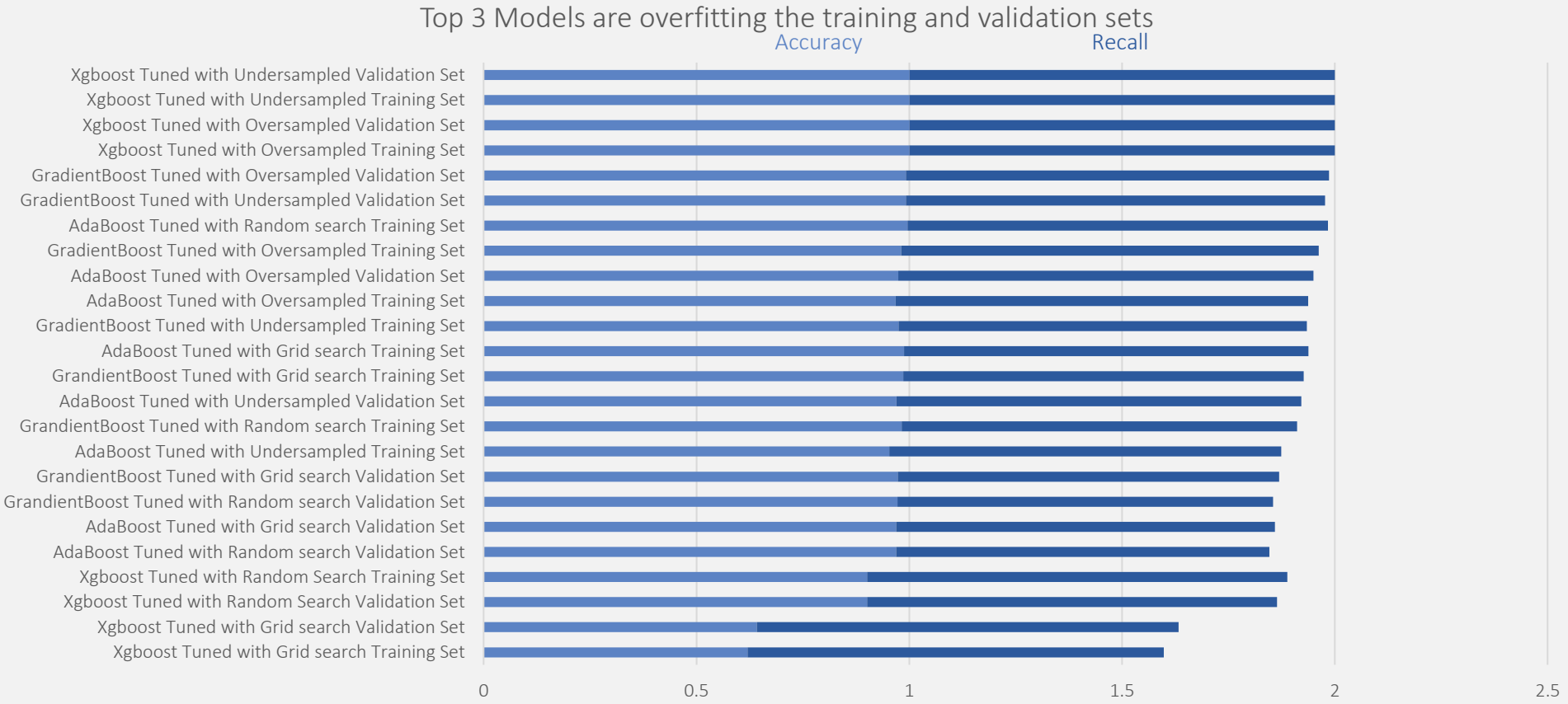
- Models were hyper tuned for best performance on metrics.
- Data was under and over-sampled to account for a class imbalance.
- Best model was chosen from over 20 models.

Model Overview and Performance

Sample of modeling scores

Metric	GradientBoost GS	AdaBoost GS	XGBoost GS
Accuracy	Train: 0.98 Validation: 0.97 Test: (oversampled / undersampled)	Train: 0.98 Validation: 0.96 Test: (oversampled)	Train: 0.62 Validation: 0.64 Test: (undersampled)
Recall	Train: 0.94 Validation: 0.89 Test: (oversampled / undersampled)	Train: 0.95 Validation: 0.89 Test: (oversampled)	Train: 0.97 Validation: 0.99 Test: (undersampled)
Precision	Train: 0.97 Validation: 0.93 Test: (oversampled / undersampled)	Train: 0.97 Validation: 0.91 Test: (oversampled)	Train:0.29 Validation: 0.30 Test: (undersampled)

Model Overview and Performance



Model Overview and Performance

Final Model

XGBoost Classifier with Oversampled Validation Data

- Accuracy > 70 and Recall & Precision > 0.95 on test set

Metric	Accuracy	Recall	Precision	F1
Score	0.96	0.96	0.83	0.89

- False Negatives <1%

- Time to return model was ~15 minutes.

Model Overview and Performance

Important Features for Customer Segmentation and Targeting

Transaction count, product count, and account month-to-month balance.

Flag customers who fall around thresholds as likely to attrite.

- Total transactions below 100.
- Monthly balance carry-over below $500.
- Total product holdings between 2 - 3.

From the historical data we can see that these are likely to fall into credit card attrition and could be offered different incentives for staying with their cards.

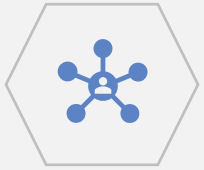
Business Recommendations

Get a Cost-Matrix

Use the same data but an alternative approach to modeling the prediction.

Get environmental data

Gather data about customers from external sources to include in analysis.



Target market attritors

Make sure to test each campaign with a small test group with different features.

Target the historical attrition customers

These customers should be targeted more by the company and offer them added benefits.

Measure what's important

Track the transaction count, product count, and revolving balance closely to identify potential attritees within ranges.

Use the production model

The model provides a way to create a new table with predicted attritees. Identify and strategize how to prevent them from attriting.

Conclusion

Thanks to your commitment and strong work ethic. We know next year will be even better than the last.

We look forward to working together.

Thank You!