

Project Team 3: #Green

EDA Phase

Srushti Kankotiya

The background of the slide features a close-up, slightly blurred image of several credit cards stacked on top of each other. The cards are in various colors, including blue, brown, and grey. Overlaid on the right side of the image are several overlapping, semi-transparent green geometric shapes, primarily triangles and polygons, creating a modern, abstract design. The text is centered on a white, angular background that cuts across the image.

Problem Statement

To find out why Customers are leaving a bank's credit card services.

Goal of the Exploratory Data Analysis

To provide an initial set of predictor variable and possible machine learning model or models.

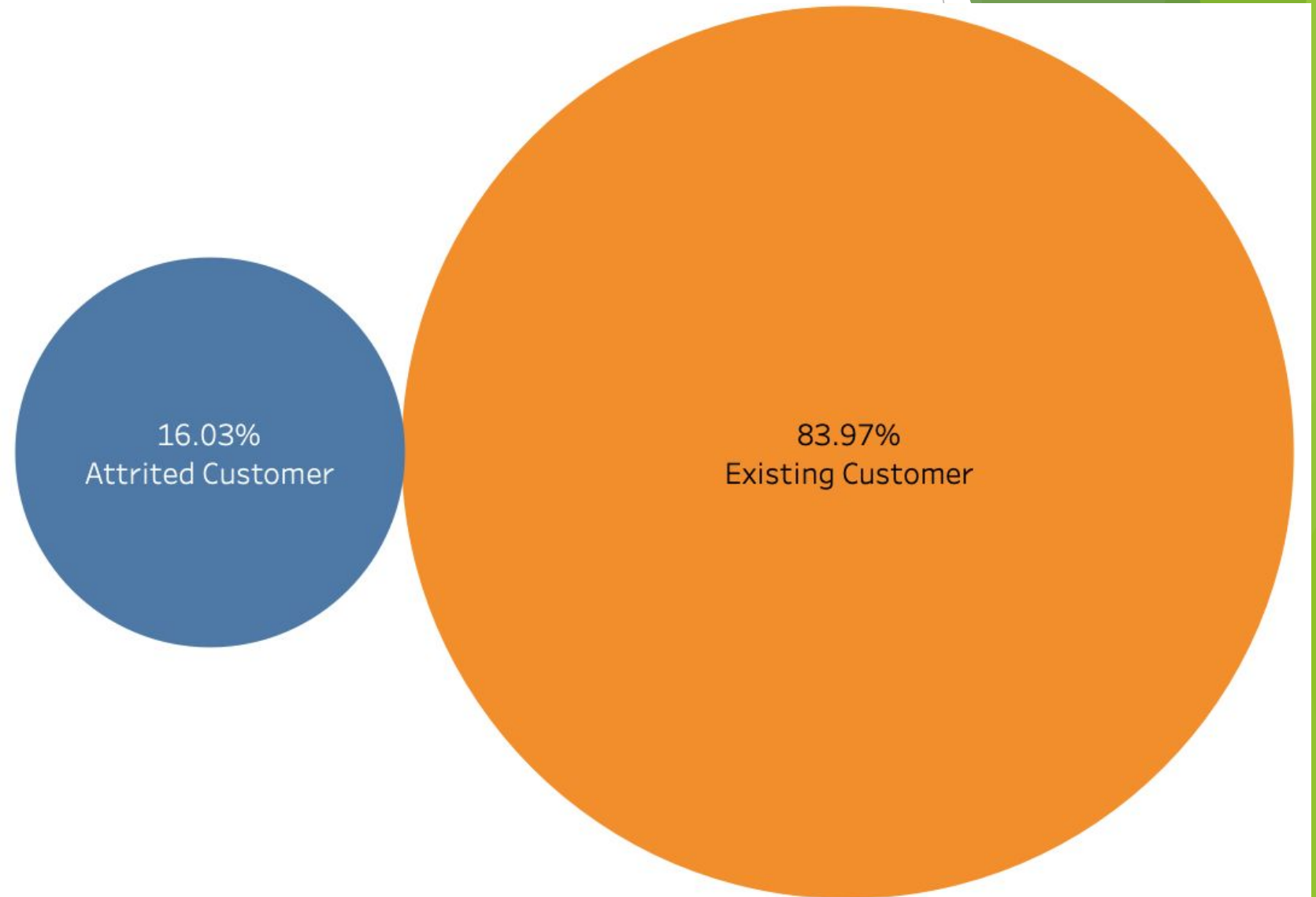
Independent variables chosen for analysis by each member

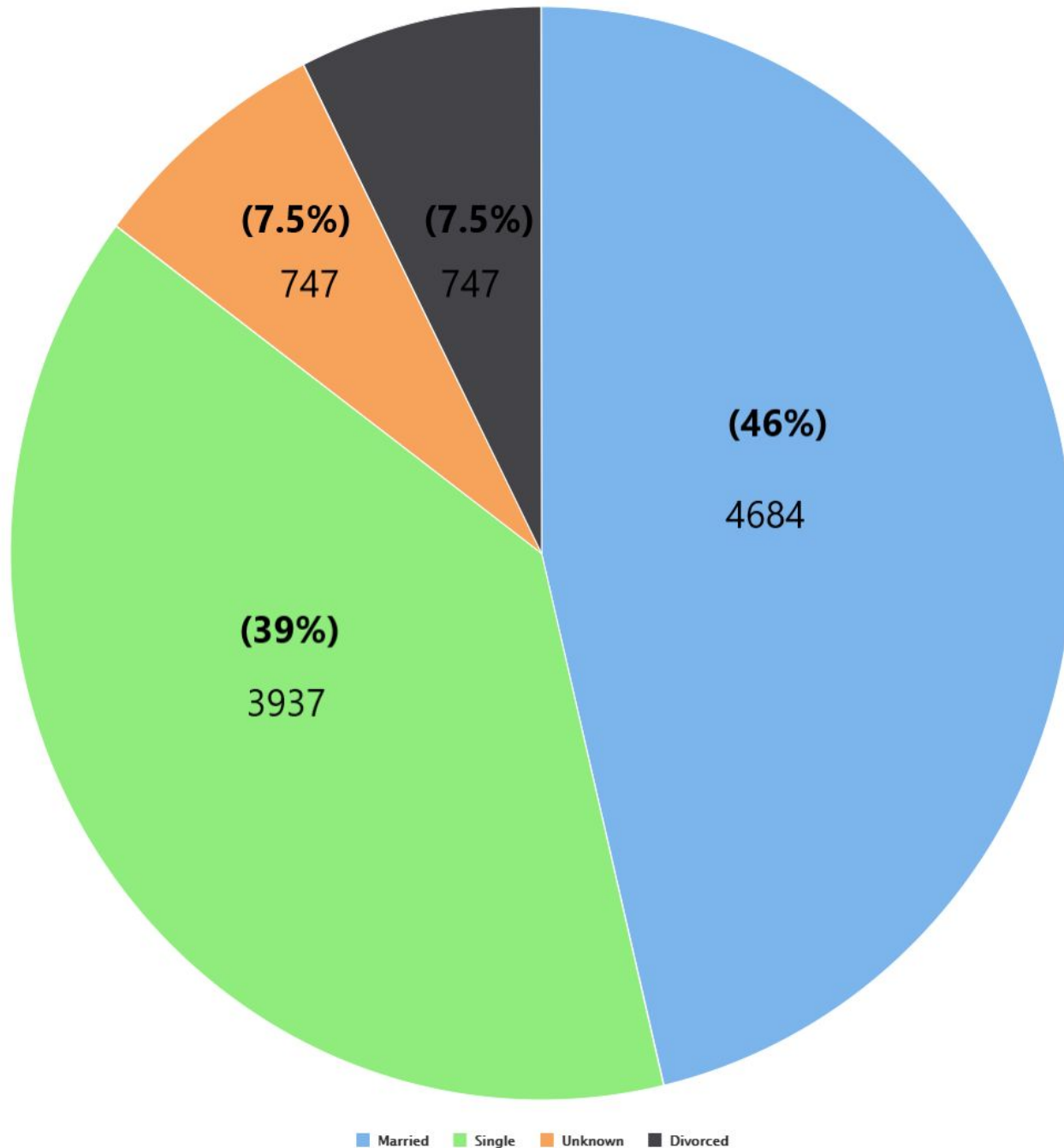
1. Gender
2. Total Revolving Balance
3. Marital status
4. Income Category
5. Card Category

Total of number of customers 10115

Existing :- 8494

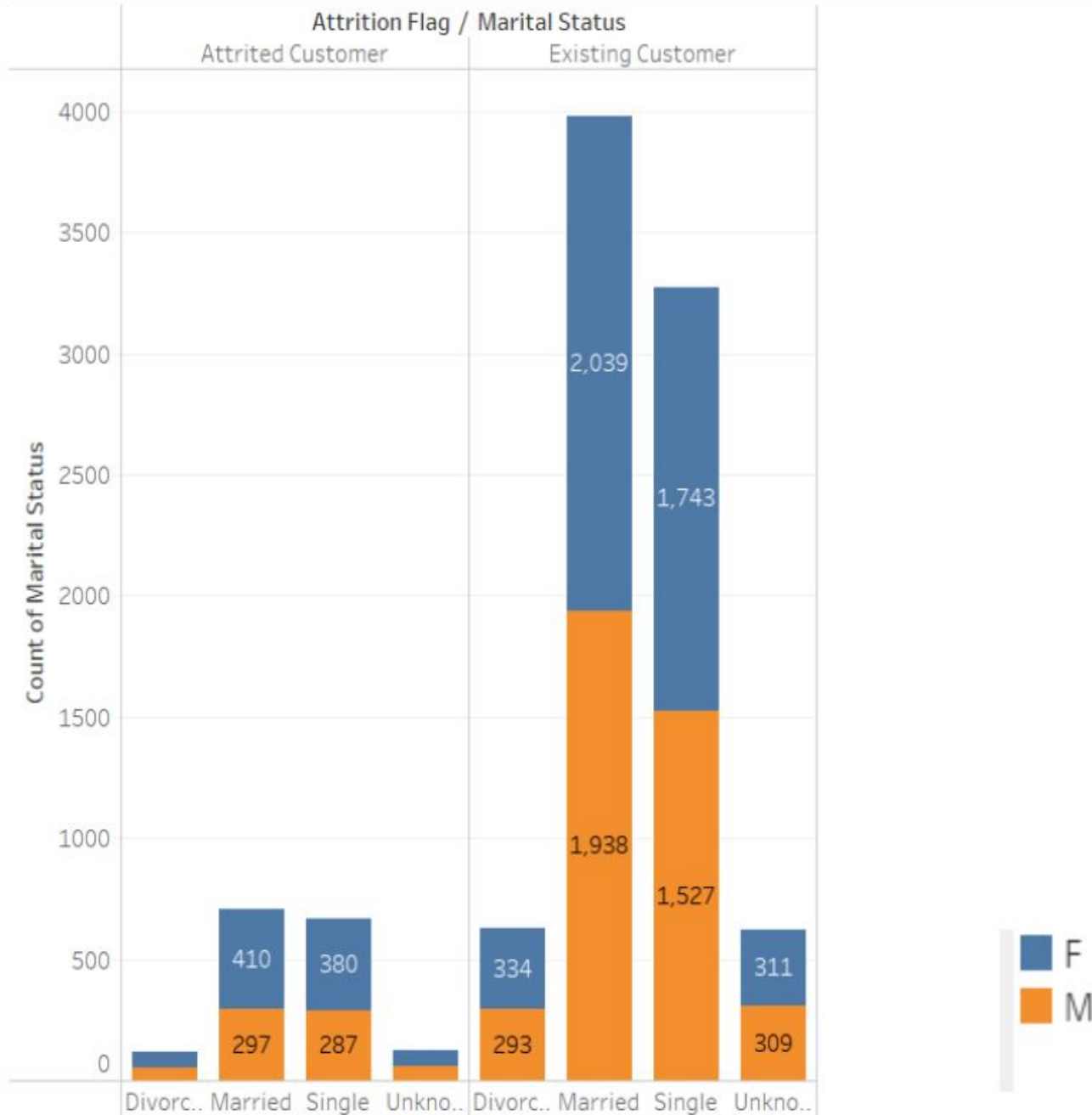
Attrited :- 1621





Marital status

- ▶ Variable: Total marital status
- ▶ Target variable: Attrition Flag
- ▶ The EDA is performed through Rapid Miner and Tableau
- ▶ Pie charts outlines following things
- ▶ The total number of clients is 10115
- ▶ Out of which 4684 are married which contributes around 46% of aggregate
- ▶ 3937 are single that contribute around 39% of aggregate
- ▶ Rest divorced and unknowns are 747 in numbers which both contribute 7.5% rate



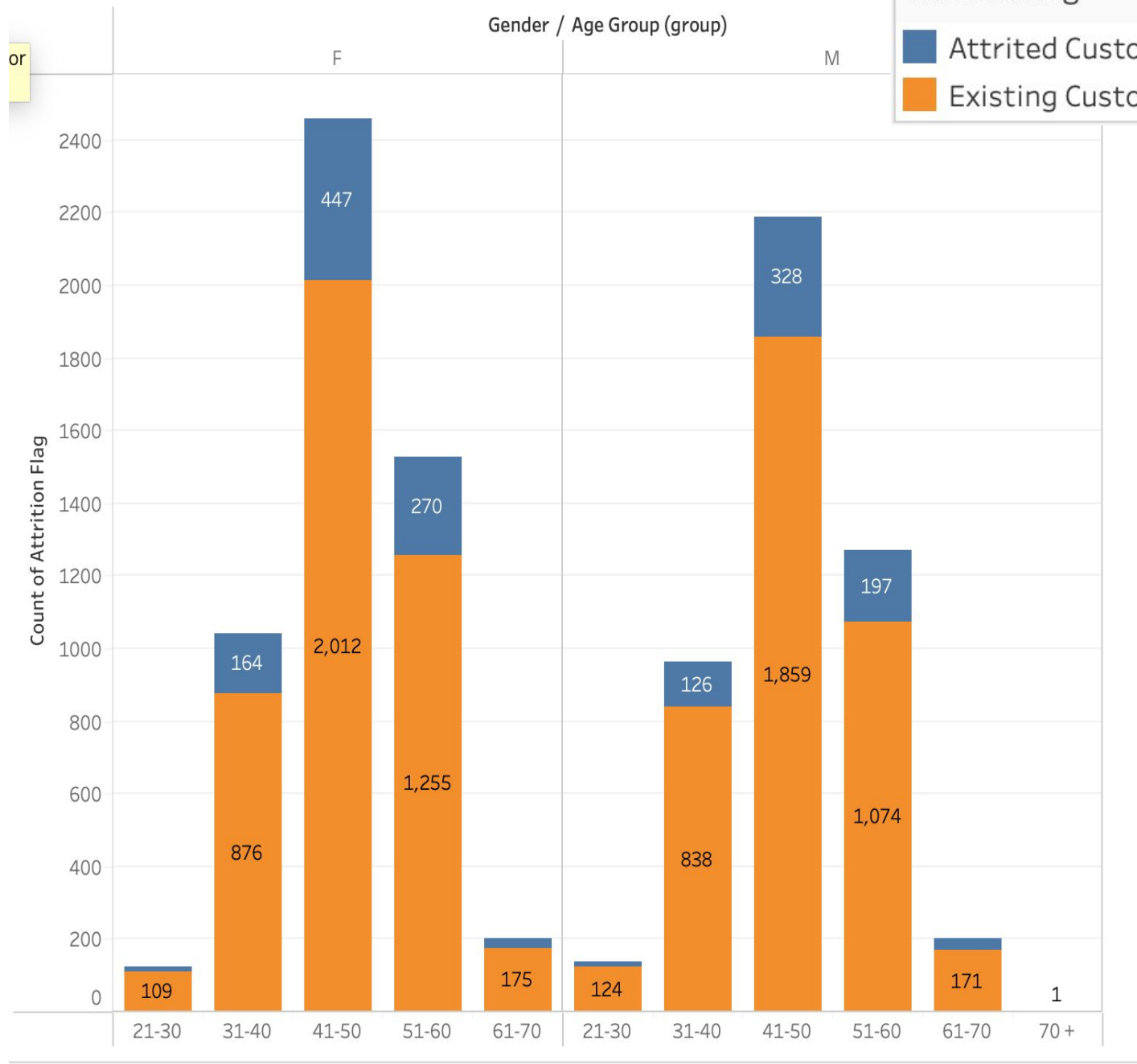
Marital status

- The number of males and females who are attrited and our current clients are being estimated on the foundations of marital status.
- As the diagram shows the largest number of clients who are married and single will generally be a regular customer
- Then again, on the other hand, clients who are divorced and are obscure will generally have less pace of churn rate than existing clients.
- After cautious perception, it very well may be observed that the extent of females will quite often have a larger number of credit card users than men.

Marital status

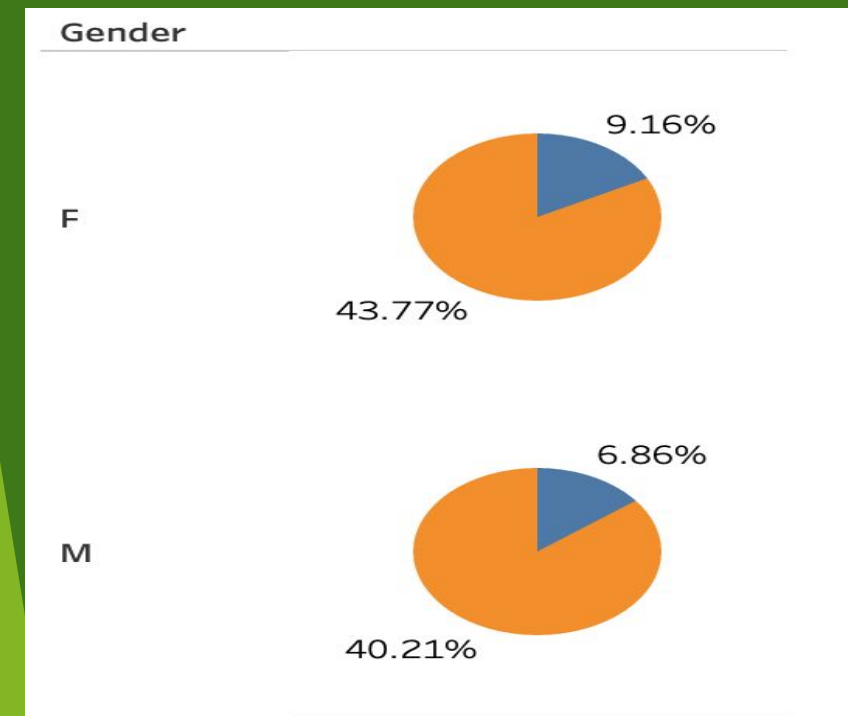
Observation

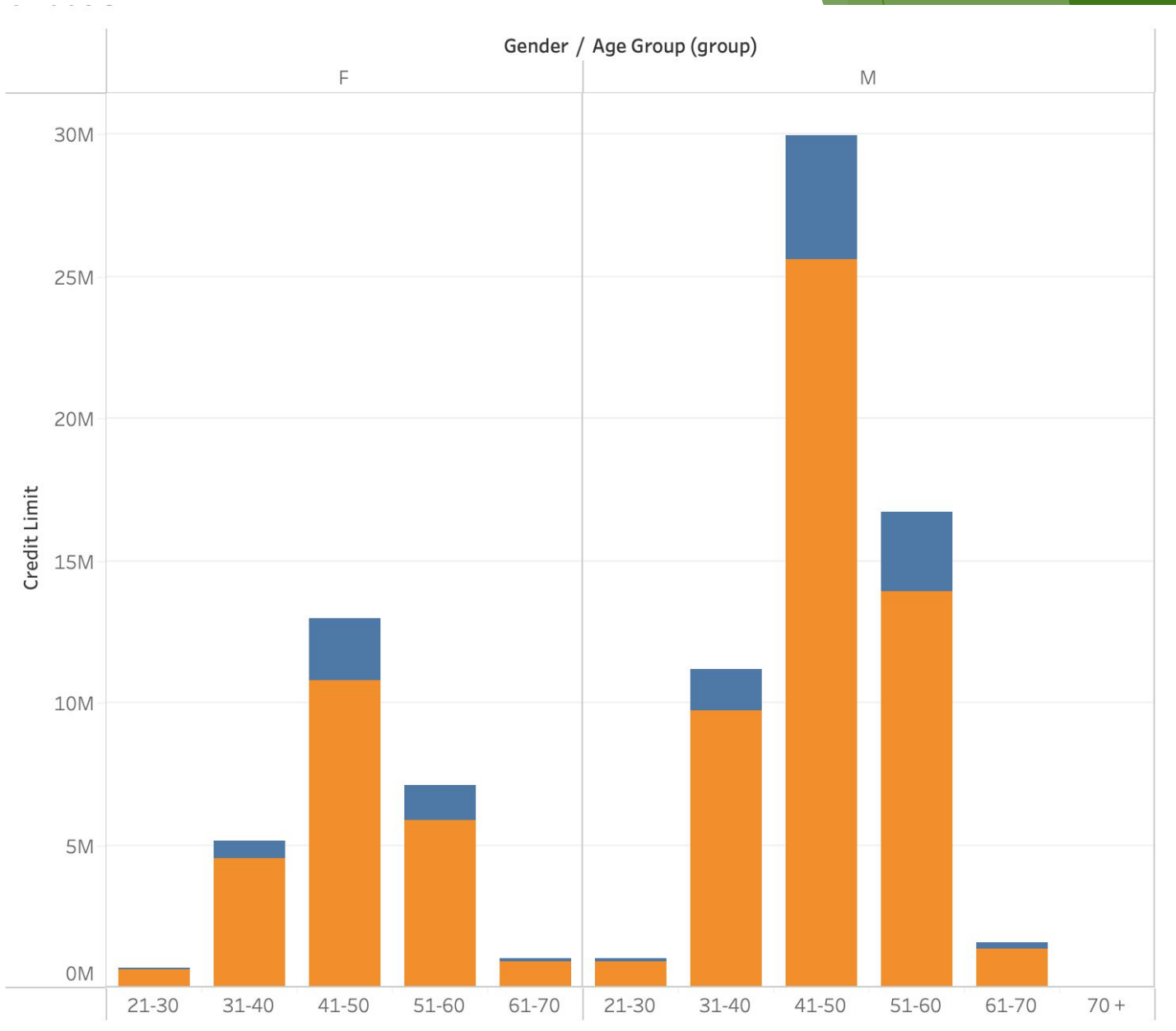
- After careful observation, it tends to be observed that the attrited individuals are for the most part the married and single clients.
- There could part of possibility married people tend to go for separate bank accounts and might have left the usage of credit cards.
- Married individuals will generally convey more debt than that of singles.
- Married customers carry more private loans, credit card debt.
- As per the Good Credit Often Means High Debt and On-Time Payments.
- The manager should look for alternative perspectives so that the customer, By giving arrangements such as lowering the interest rate or annual percentage rate (APR), providing cashback for combined choice category/grocery store/wholesale club purchases, Joining bonuses or rewards, and so on.



Gender

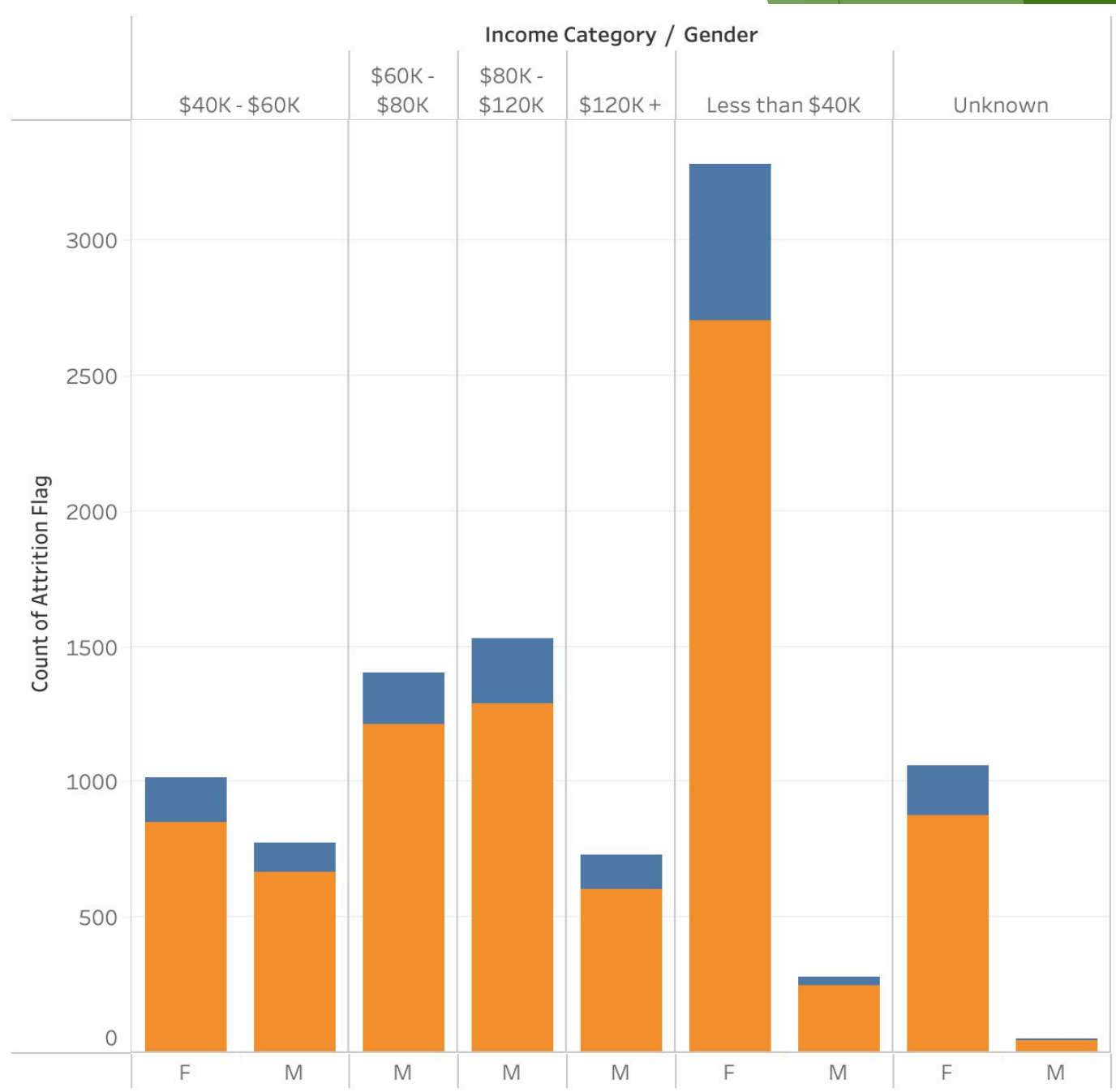
- Variable: Gender
- Target variable: Attrition Flag
- The data analysis is done by Tableau data visualization





Gender/ Age group Vs Credit limit

- In this graph we can see that the credit limit provided to men are much higher than credit limit provided to women.
- We can see that men in age group of 41-50 have the highest credit limit.
- Also, Credit limit for men in age group 51-60 is higher than that of women.
- We can see that there is big difference between the credit limit provided to men Vs Credit limit provided to women.



Gender Vs Income Category

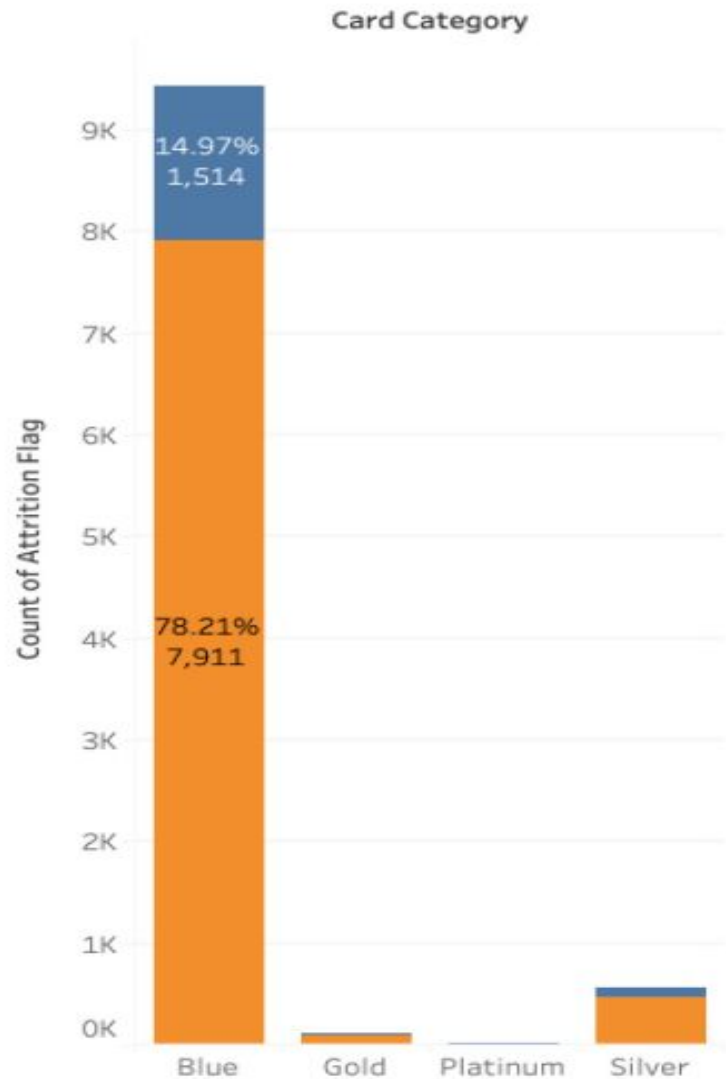
- In graph we can see that most of the women belong to income category which is less than \$40k.
- There are very few men in income category which is less than \$40k.
- Also, there are more number of women in income category of \$40k - \$60k.
- But most of the men belong to income category \$60k - \$120k and \$120k+.
- We can see that there are no women in income category of \$60k - \$120k.
- This can be one of the reason why there is big difference between the credit limit provided to men Vs credit limit provided to women.

Gender Observations

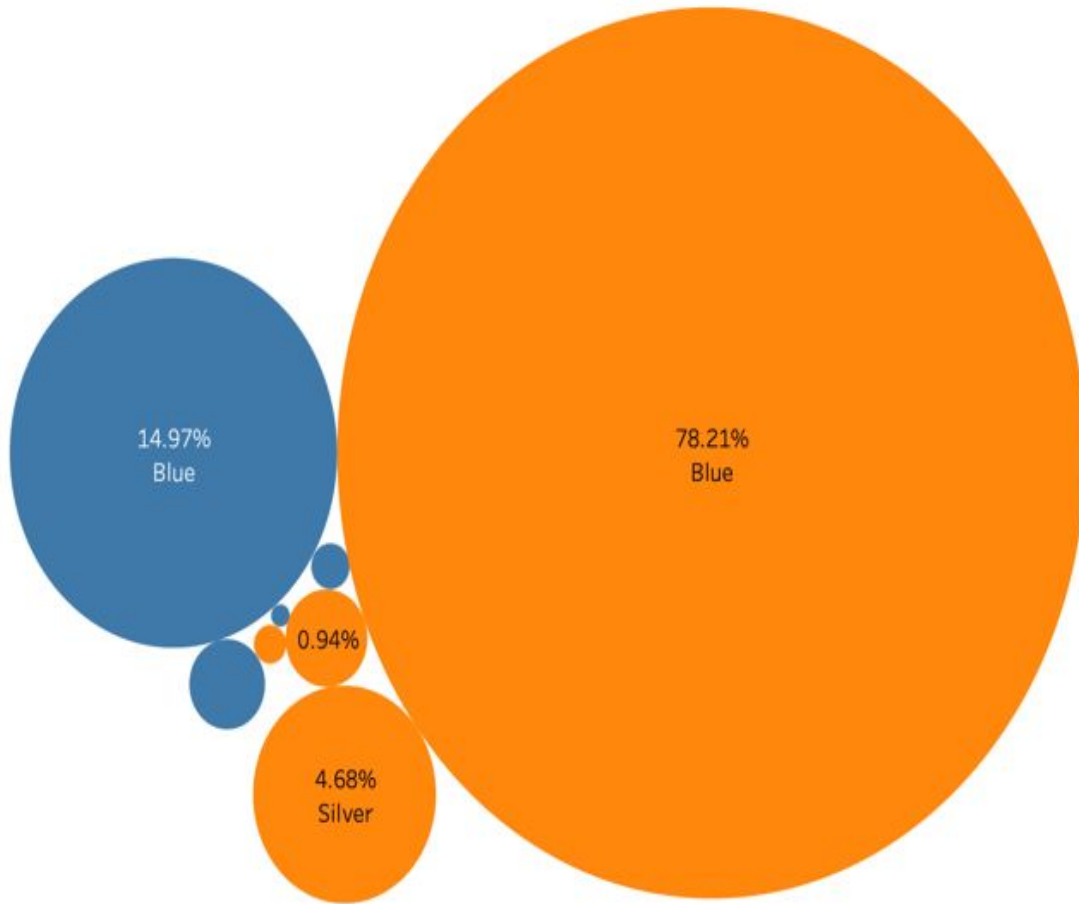
- There are more women customers leaving the bank credit card services. Most of the women leaving bank credit card services belong to the age group of 41-50 and 51-60.
- There is big difference between the credit limit provided to men Vs credit limit provided to women. The men in age group of 41-50 have the highest credit limit.
- Most of the women fall under income category of less than \$40k. This can be one of the reason why women have low credit limit.
- There are very few men in income category less than 40k. Most of the men fall under the income category of \$60k - \$120k. And this can be one of the reason why men have high credit limit.
- According to the analysis, Women with low credit limit is the area where one should focus.

Card Category

- Variable : Card Category
- Target Variable : Attrition Flag
- Data Analysis is by Tableau Data Visualization
- The initial number of Customers before the attrition trickle was 10,115
- Of the four cards, the Blue card is by far the customers card of choice garnering 9425 (93.2%) cumulated customers
- As a result, the biggest number of Attrited customers falls under that same category (14.97%)
- The other three cards (Gold, Platinum and Silver) take up the remaining 6.8% customers



Card Category



➤ By the time of the report, Existing customers were 8494 (84%) while Attrited customers were 1621 (16%)

➤ The bubbles depicts customers, both existing and attrited based on their card category, thus:

- Blue - Existing 7911, Attrited 1514
- Gold - Existing 95, Attrited 21
- Platinum - Existing 15, Attrited 5
- Silver - Existing 473, Attrited 81

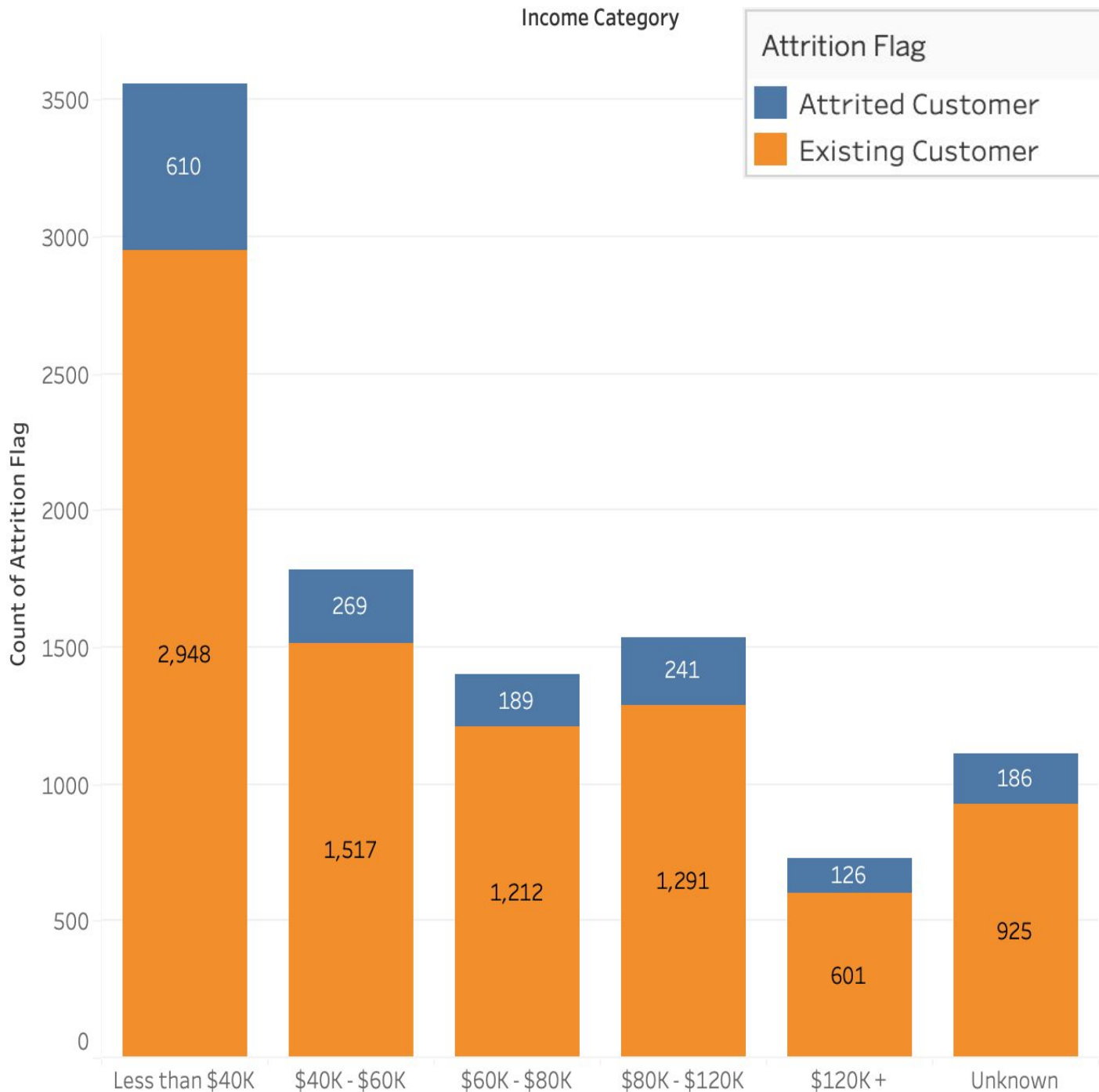
Card Category



- ▶ Based on our deduction from the previous analysis, over 93.2% of cumulated customers subscribe to the Blue Card Category
- ▶ With this in mind, we looked at the income category in relation to the blue card category.
- ▶ It is evident that the income range of less than \$40k has the biggest clientele
- ▶ And when we added gender to the two variables, it is the Females on that income bracket that make the majority of Blue card customers.

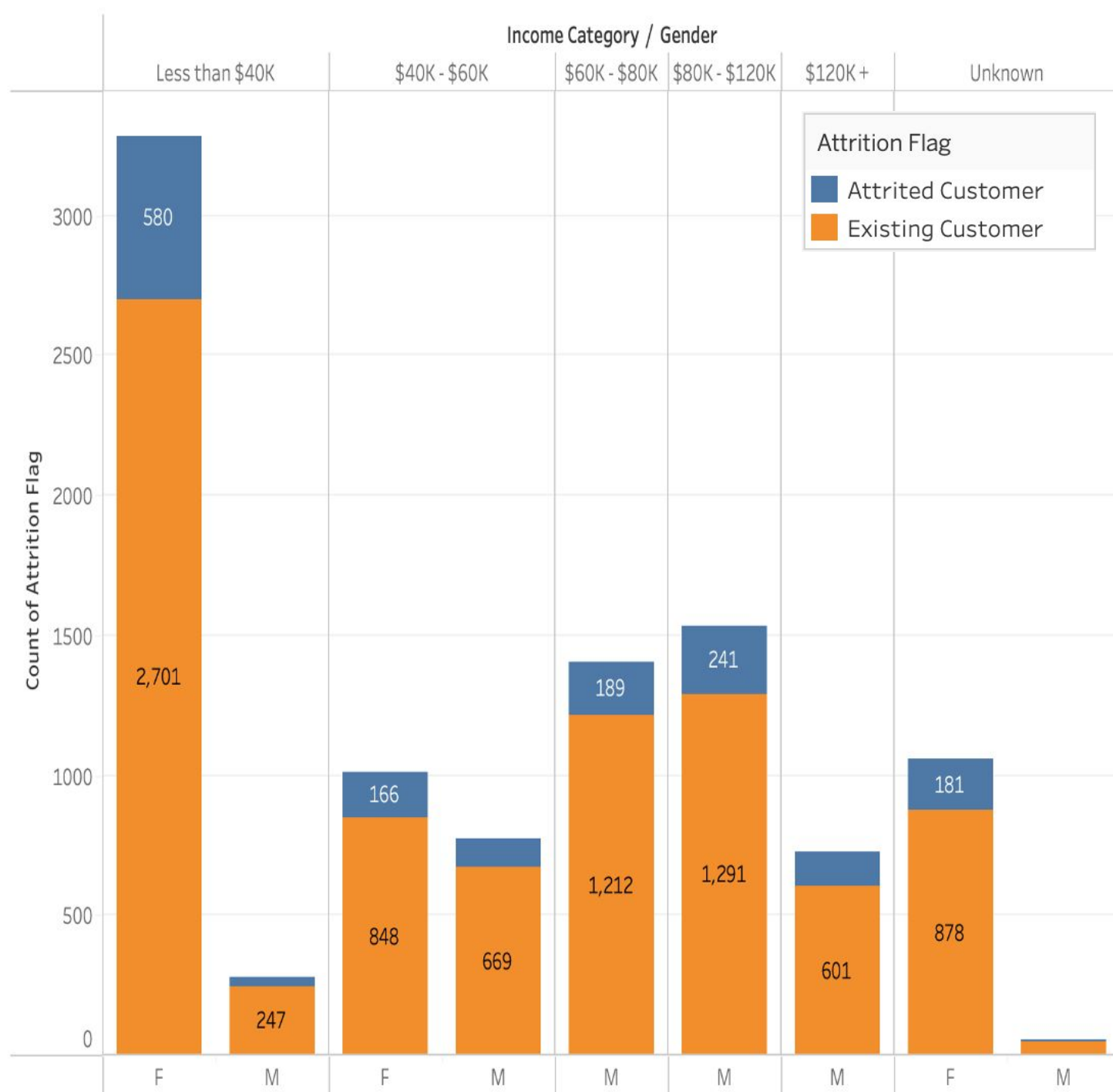
Card Category Observation

- ▶ The Blue card category is so dominating of other cards that data on cards like Platinum almost seems inconsequential (93.2% vs 1.1% vs 0.2% vs 5.5%)
- ▶ Data analysis establishes a trend inferring that, the Blue Card Category is the customers favorite and the majority of the clientele is Female
- ▶ There is an enormous concentration of customers of the Blue Card Category on the income bracket of less than \$40k over and above other income brackets
- ▶ This establishes a clear relationship between the Blue card category, Income bracket of less than \$44k and the Female gender



Income Category

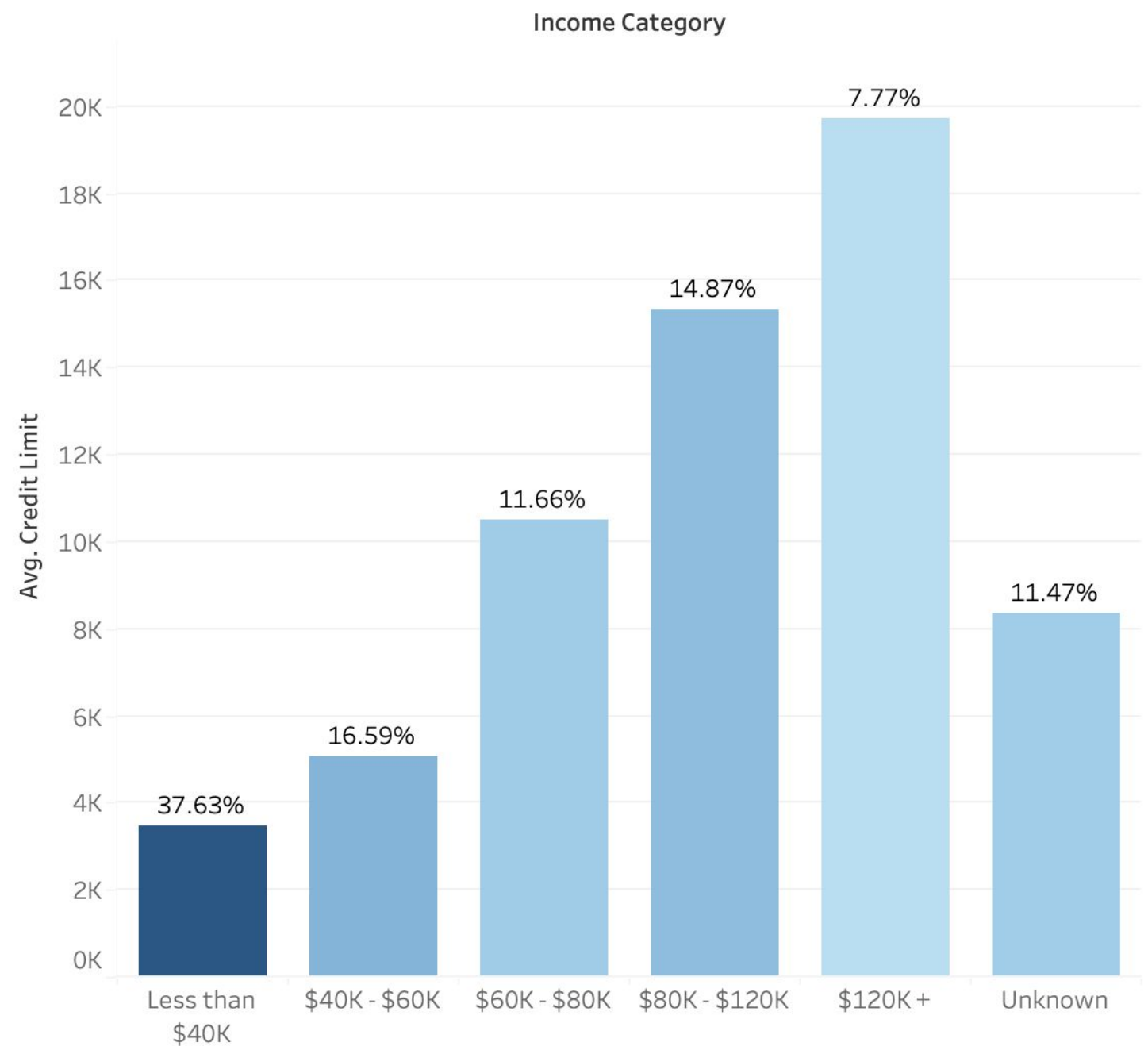
- ▶ Variable: Income Category
- ▶ Target variable: Attrition Flag
- ▶ The data analysis is done by Tableau data visualization
- ▶ As we can see from the bar chart that the most customers fall under income category “Less than \$40k” in which 610 customers churns out of 3558
- ▶ Second highest churners from the category \$40k-\$60k from which 1517 is existing customers and 269 churned customers.
- ▶ 241 customers churned from the category \$80k-\$120k



Income Category Gender

- ▶ As per the information from the graph we can see that the most females fall into the less than \$40k category in which 580 females left the credit card service.
- ▶ In terms of males, they mostly fall under the category \$80k-\$120k and \$60k-\$80k in which 241 and 189 customers respectively left the credit card service.
- ▶ Interestingly, in the less than \$40k category the male proportion is the lowest compared to the female portion. From total 277 male customers 30 customers left.
- ▶ Surprisingly, there is no female under the category of \$60k-80k, \$80k-120k and \$120k + income.

Income category with Credit Limit



Income Category Credit Limit

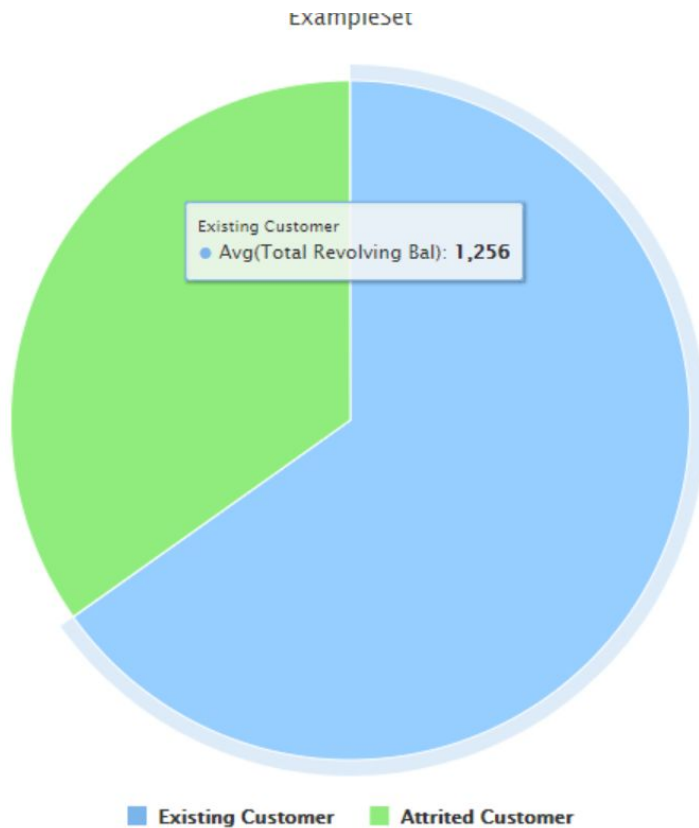
- ▶ The lowest credit limit is given to the income category “less than 40k” and highest credit limit given to the \$120k
- ▶ Average credit limit for customers
 - Less than \$40k :- 3469
 - \$80k-\$120k :- 5052
 - \$60K-\$80k :- 10492
 - \$80k-\$120k :- 15319
 - \$120k + :- 19734
 - Unknown :- 8339



Income Category Observation

- ▶ The analysis of attrition flag with the income category shows that the customers are more under the category of less than \$40k, \$80k-\$120k and \$60k-\$80k because of that the churners are more in this income brackets.
- ▶ The proportion of the females is greater compared the males proportion for the income brackets less than \$60k and no females found under the category of greater than \$60k.
- ▶ The highest churners found under the category of less than 40k for the female customers and \$80k - \$120k for the males.
- ▶ After analysis with the credit limit it says that the customer whose credit limit is less they are most likely to leave the credit card service.
- ▶ As we have seen the result the customers under the category less than \$40k and \$40k-\$60k have lower credit limit might possible for this reason the attrition percentage is higher compared to the other category.

Total Revolving Balance



- ▶ Variable: Total Revolving Balance
- ▶ Target variable: Attrition Flag
- ▶ The EDA is performed through Rapid Miner.
- ▶ The pie chart depicts :
The Customers who Churn has average total Revolving balance of 671
Whereas, the average total revolving balance of Customers who stayed are 1256

Total Revolving Balance

Different perspective to look at variable chosen (Revolving Balance) for EDA to gain more insights

Analysing through Age factor helps in determining below points:

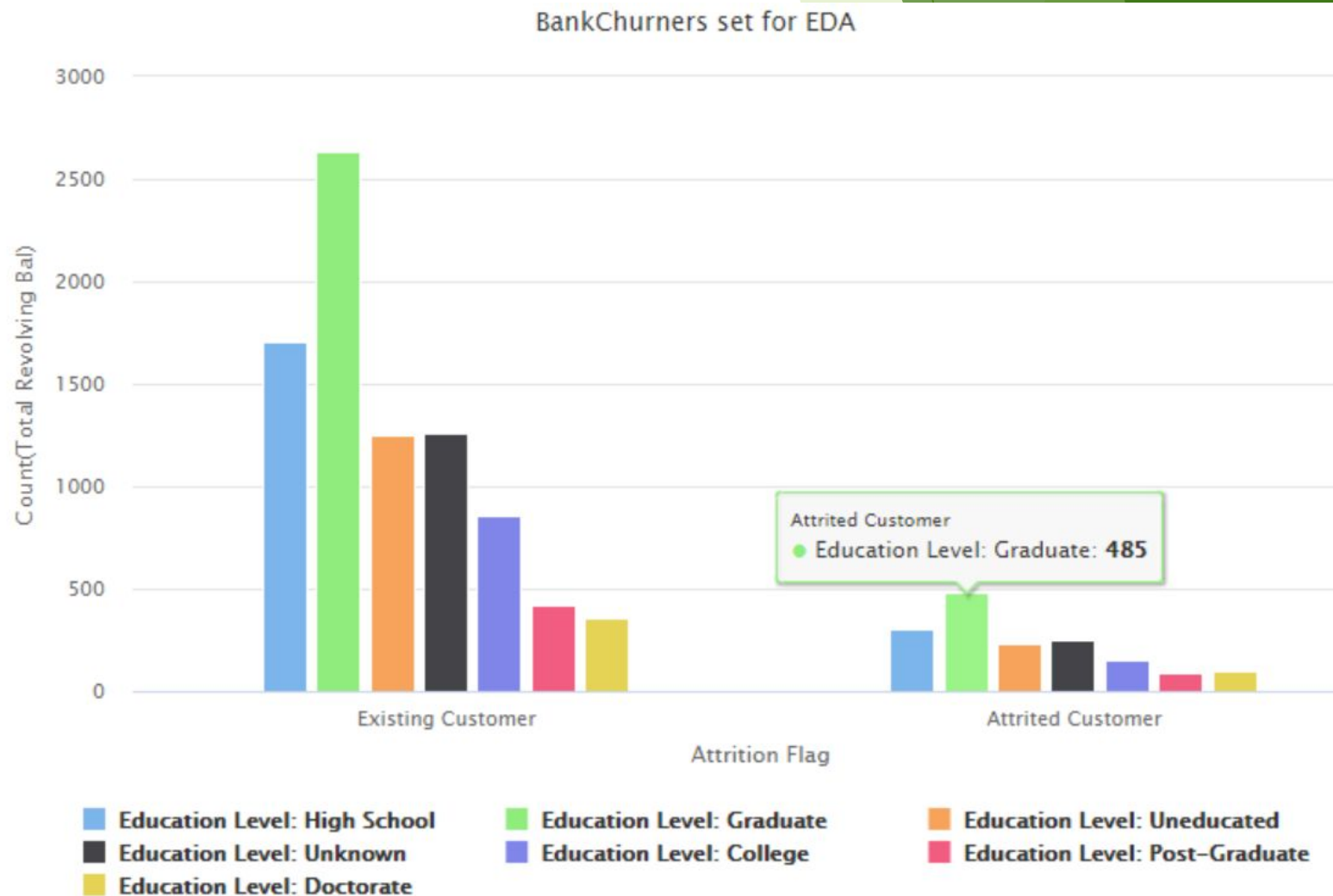
- ▶ The age group that falls under the range 26-30 has on an average churn rate higher as their revolving balance above 1000.
- ▶ Whereas the Customers who stays loyal with the bank has on an average quite similar total revolving balance but are more elder and wiser
- ▶ You might also wonder the sudden spike for existing customer age group 66-68, it could be a potential outlier.



Total Revolving Balance

To support the prior analysis:

Analysing through through Education levels factor:



- ▶ The customers whose attrition rate as per the total revolving balance is high falls under Graduate level students(485)
- ▶ Whereas next to that was High School level students who count turns out to be 305, which entails that this age groups should look into.

Total Revolving Balance Observation

- ▶ The average churning w.r.t Total Revolving Balance seems to be considerably high as per the depiction on Pie Chart
- ▶ It is a factor that should be taken into consideration.
- ▶ Bu performing further analysis, we found another key insight which states that age group 26-30 is to look after where the churn rate w.r.t Total revolving Balance is on an average high and even exceed in few cases.
- ▶ The manager in such cases need to figure a way out in having a strategy to minimize the losses by introducing different scheme to lure them back.

Group

Summary & Conclusions

- ▶ The analysis of attrition of customers with the income category shows the affection when compared with some attributes so as results of infographics need to pay attention to the less credit limit given to the customers who fall under the lower income brackets. Thus, the increment in the credit limit of this customers might help to prevent customers from leaving the credit card service of the bank.
- ▶ The EDA performed on Total Revolving Balance w.r.t Attrition rate seems to minimal in some cases but need to look after as the age group which ranges from 26-30 tends to have more revolving balance than what they can spend.
- ▶ As well as the Customer age group 64-66 shows similar concerns but still need to look for by the manager as it might be due to various factors that might not even be considered as insurance claims didn't come through, uncalled events like injuries, etc, It can be an outlier as well.
- ▶ The analysis of this data establishes a trend inferring that the Blue Card Category is the customers favorite, majority of whom are Female within the Income bracket of less than \$40k

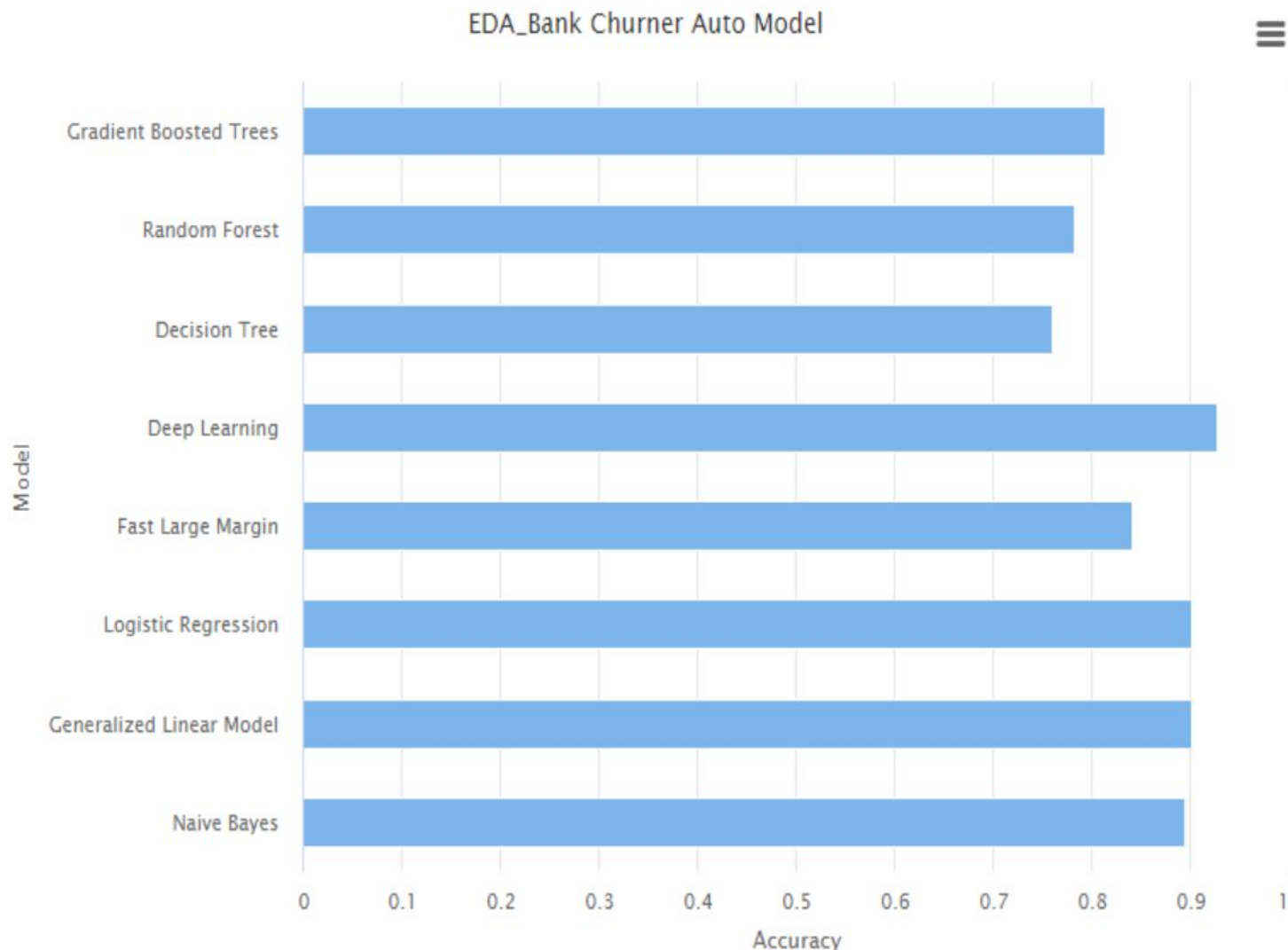
Team Green

Models to choose from

As per the analysis as well as through test conducted on Auto model, We came to conclusion that below models might provide us the accurate analysis as the target attribute presented which is in our case: **Attrition flag**

Models that can be used for Prediction based on their Accuracy:

- ◆ **Deep Learning (92%)**
- ◆ **Logistic Regression / Generalised Linear Model (90%)**
- ◆ **Naive Bayes (89%)**



A close-up, shallow depth-of-field photograph of several credit cards stacked on top of each other. The top card is dark grey or black, showing a gold-colored chip and embossed numbers including '4711' and '08/1'. The card below it is a lighter color, possibly blue or silver. The background is a white triangular shape pointing towards the right, bordered by green geometric patterns on the left and right sides.

Thank you

