

Data Analytics Capstone Project

Foo Soo Hian



Problem Statement

A manager at a bank is disturbed with more and more customers leaving their credit card services. Using the available data, we aim to **identify the high attrition risk customer groups** so that the bank can provide the customer with better services and reverse churn.



Dataset Information

source: <https://www.kaggle.com/datasets/sakshigoyal7/credit-card-customers>

Step 1: Data Cleaning/Preparation:

- The data is checked for any errors in data values. Null values and unknown values are removed.
- However, for the marital status column, the “Unknown” value is not removed because it would likely not affect our findings. Removing them might result in a smaller dataset, so they are kept.
- Two columns named “Naive_Bayes_...” and “Naive_Bayes_...” are removed as they are not relevant for our analysis.

Step 2: Data Visualization & Analysis

- The data is imported into Tableau.
- And visualized using different charts for analysis.
- We also create a series of Dashboards to group the charts.

Dashboard: Customers

[View on Tableau Public](#)

Total customers

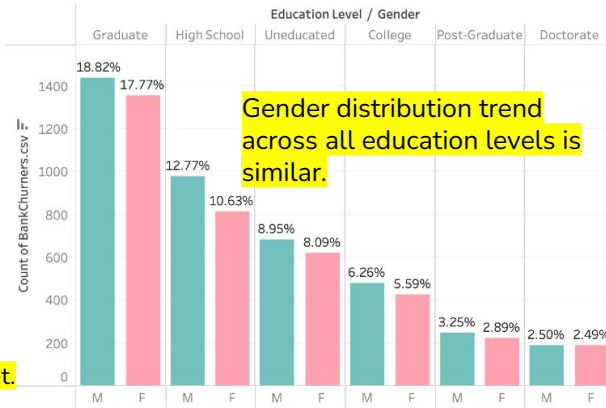
7,641

Customers by gender



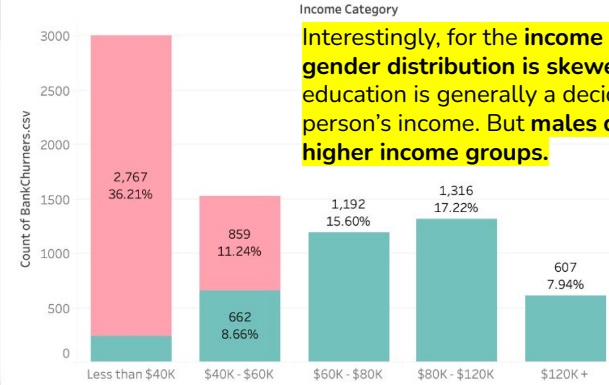
More males than females in the data set.

Customers by education level



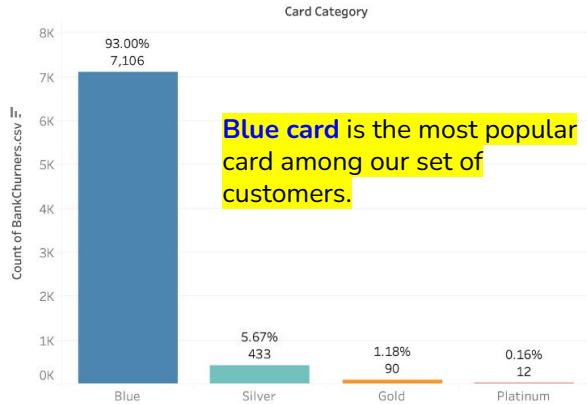
Gender distribution trend across all education levels is similar.

Customers by income categories



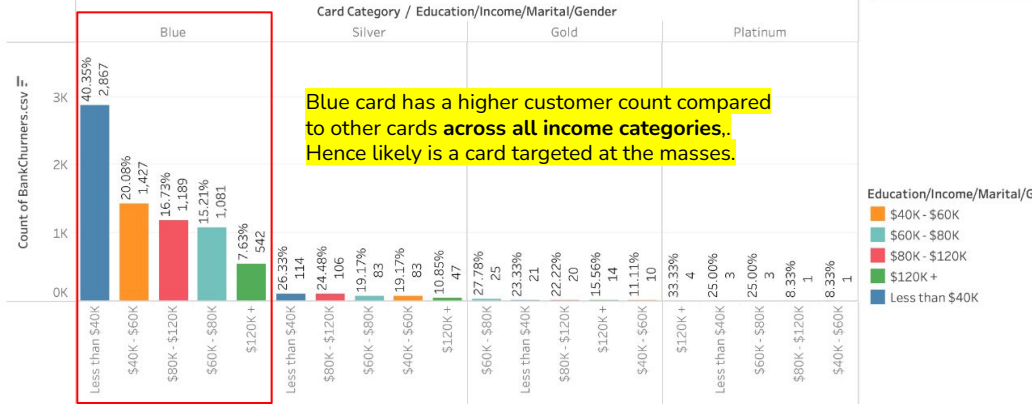
Interestingly, for the income categories, gender distribution is skewed. Even though education is generally a deciding factor for a person's income. But males dominant in the higher income groups.

Customers count for card categories



Blue card is the most popular card among our set of customers.

Customers income by card

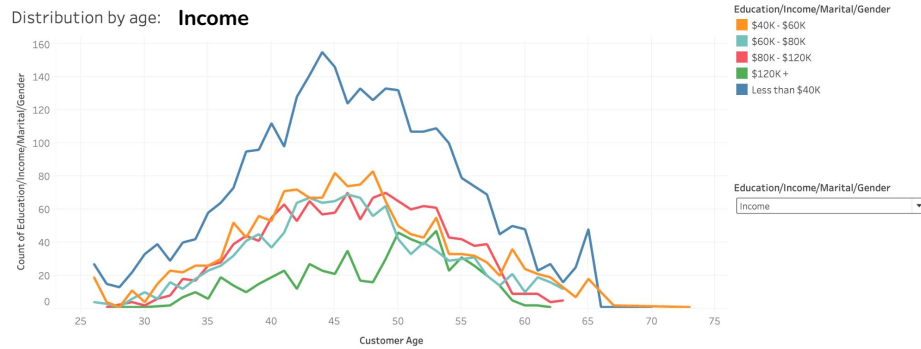


Blue card has a higher customer count compared to other cards across all income categories. Hence likely is a card targeted at the masses.

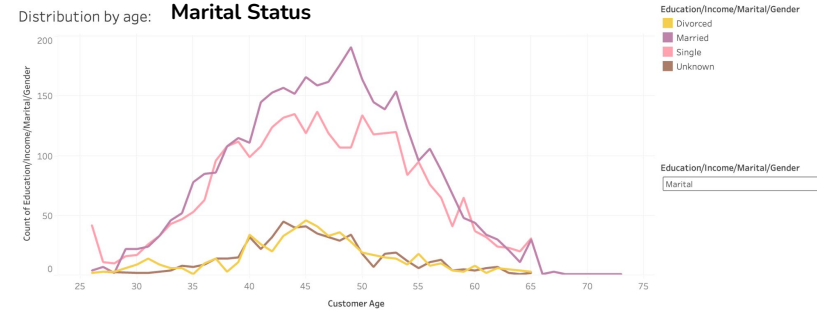
Customers distribution by age

[View on Tableau Public](#)

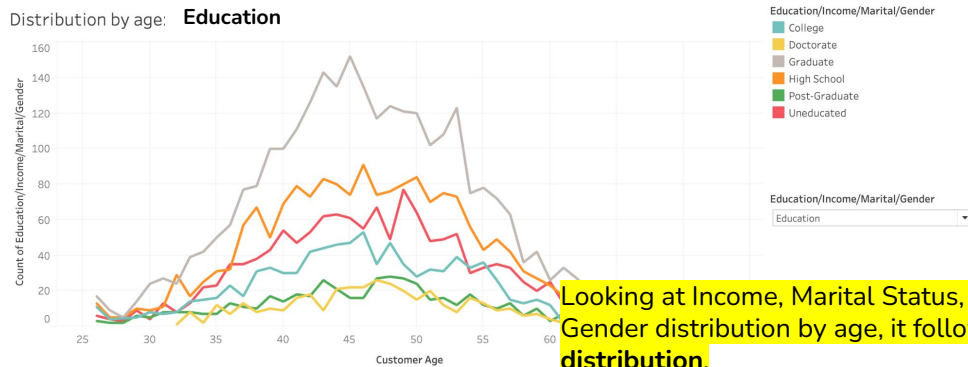
Distribution by age: **Income**



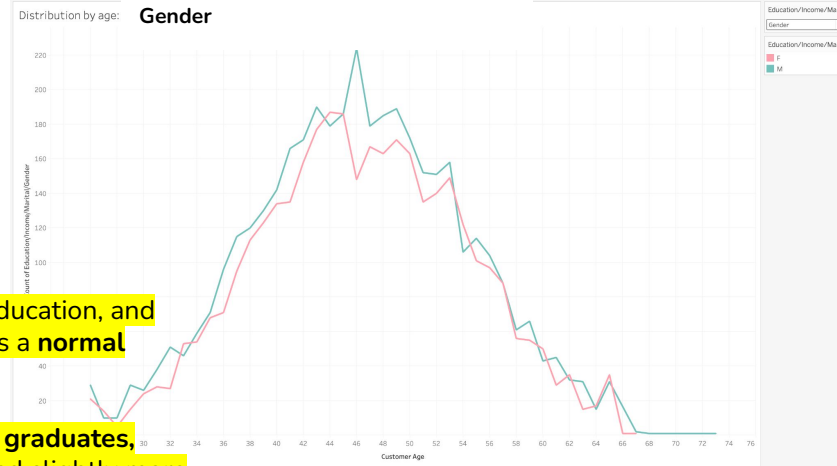
Distribution by age: **Marital Status**



Distribution by age: **Education**



Distribution by age: **Gender**



Looking at Income, Marital Status, Education, and Gender distribution by age, it follows a **normal distribution**.

Higher proportion of customers are **graduates**, earning less than **\$40K**, **married**, and slightly more **males than females**.

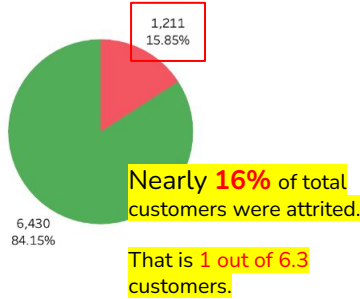
Dashboard: Attrited vs Existing Customers

[View on Tableau Public](#)

Attrited customers

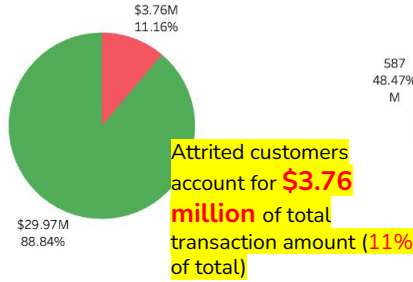
1,211

Attrited vs Existing Customers count/%

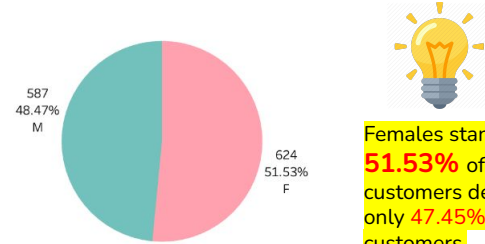


Attrition Flag
■ Attrited Customer
■ Existing Customer

Total transaction amounts: Attrited vs Existing

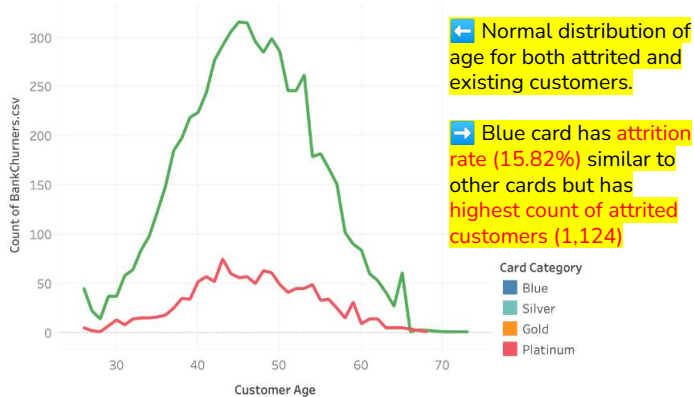


Attrited customers by gender

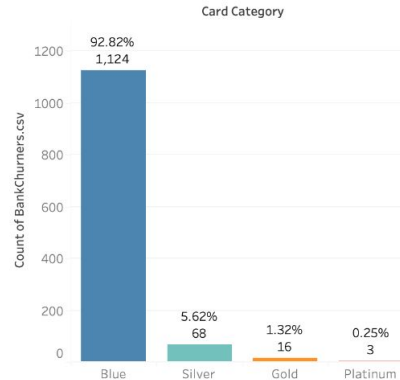


Females stand for **51.53%** of attrited customers despite being only **47.45%** of the total customers.

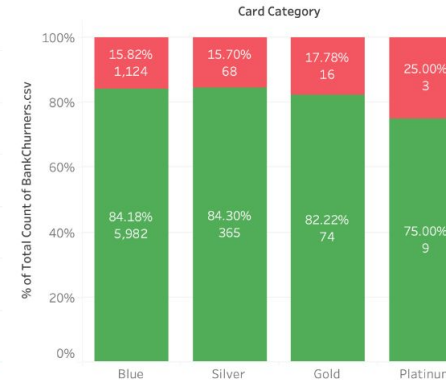
Attrited vs Existing customers: Age Distribution



Attrition count by card category (% across)



Attrition count by card category
 Count and % of total attrited within card category

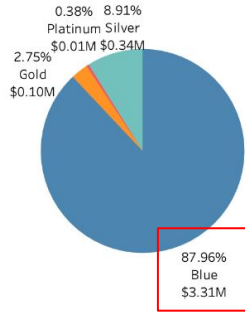


Attrition rate (%) is higher for Gold and Platinum cards but by count, they are quite low numbers, only 16 and 3.

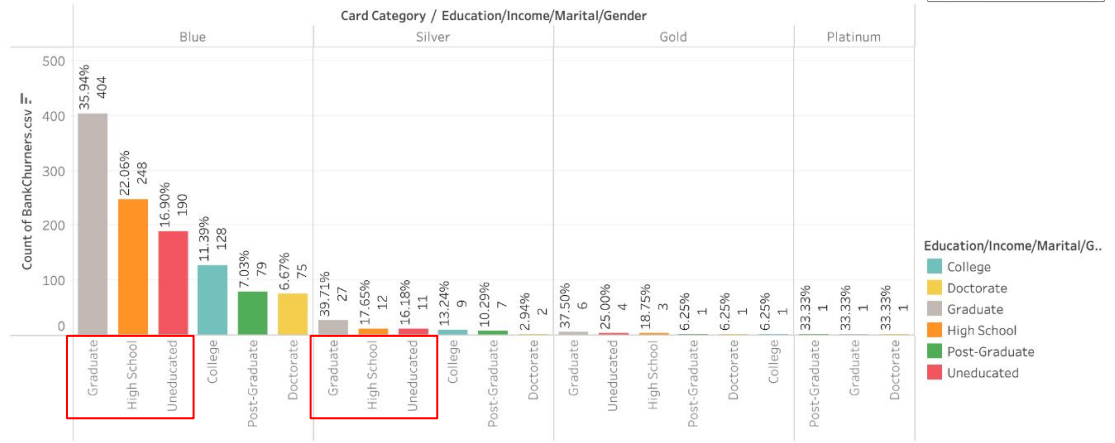
Dashboard: Attrited Customers by Card

[View on Tableau Public](#)

Blue card — highest proportion of attrited total transaction amount at **\$3.31M** (87.96% of total attrited \$3.76M).



Attrited customers by Education



Blue and Silver card customers stand for a collective **96.87%** of the attrited total transaction amount.

Graduates, high schoolers, uneducated account for high proportion of attrited customers for Blue and Silver card.

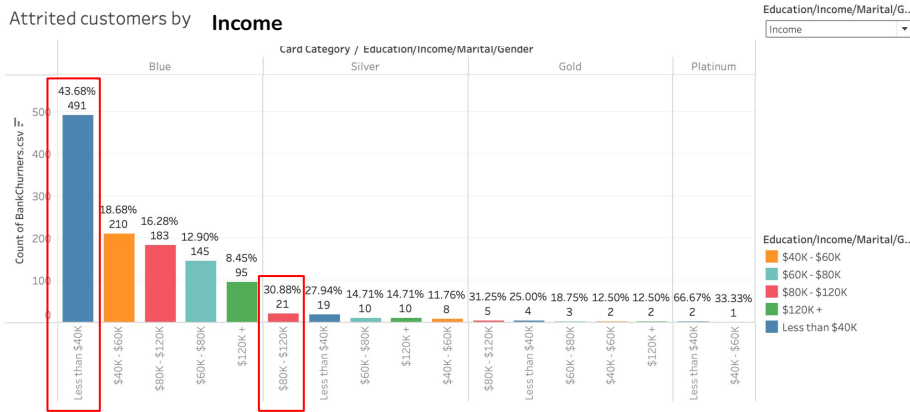
Hence, a **bigger business impact** if our remedial measures are focused on **Blue and Silver** card customers.

Note: Education is usually self-reported by customers, and might not be accurate.

Dashboard: Attrited Customers by Card

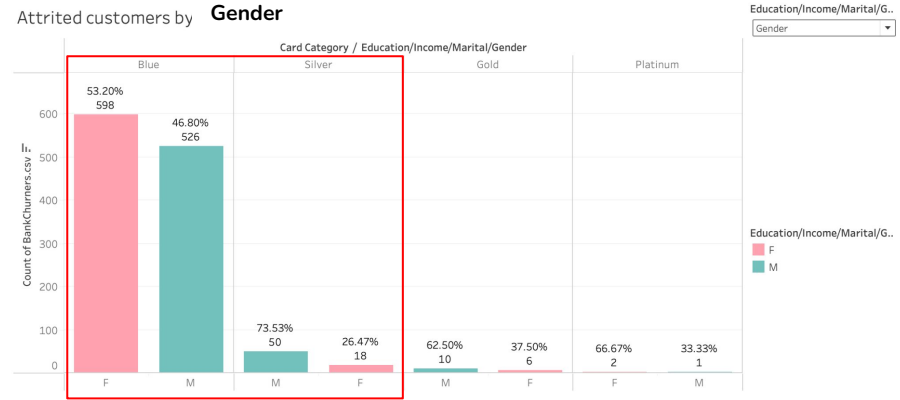
[View on Tableau Public](#)

Attrited customers by **Income**



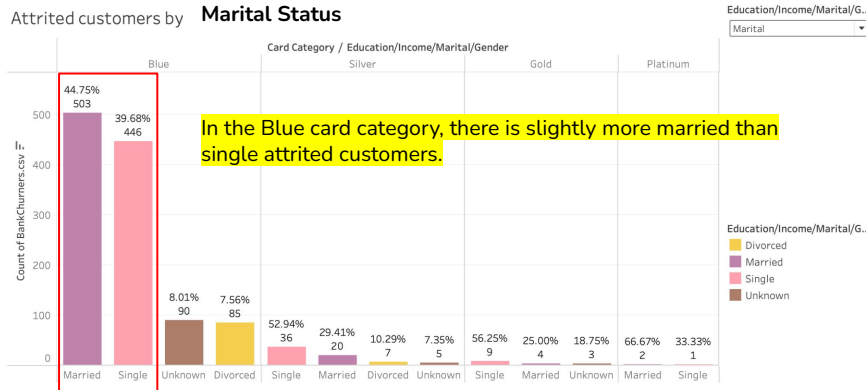
Blue card: Less than \$40K income category stands for large proportion of attrited customers (43.68%) compared to other incomes. For Silver card, \$80K-130K is the largest (30.88%) followed closely by Less than \$40K (27.94%).

Attrited customers by **Gender**



Blue card: Slightly more female customers (53.2%) attrited than males.
Silver card: More male customers attrited (73.53%).

Attrited customers by **Marital Status**

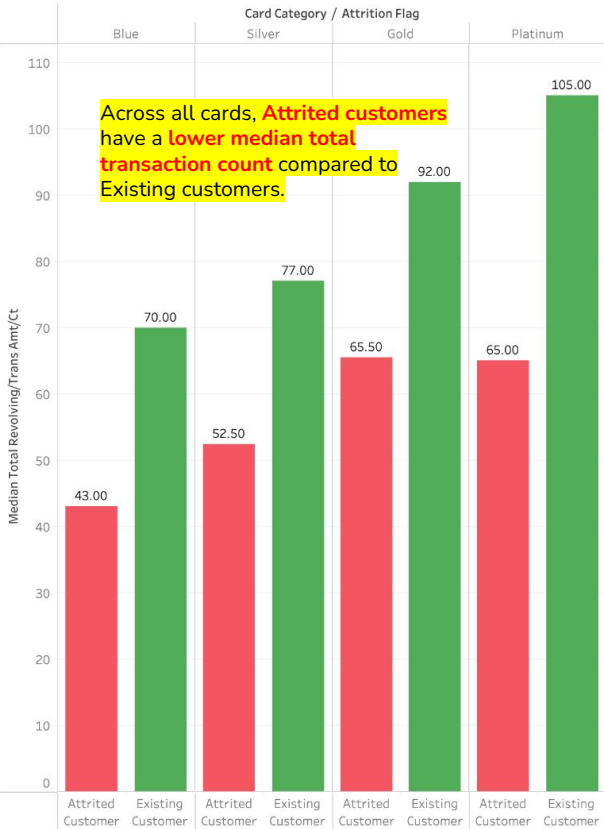


In the Blue card category, there is slightly more married than single attrited customers.

Spending habits of Attrited vs Existing customers

[View on Tableau Public](#)

All Cards: Attrited vs Existing by median
Total Transaction Count



Across all cards, Attrited customers have a lower median total transaction count compared to Existing customers.

- Card Category
- (All)
 - Blue
 - Gold
 - Platinum
 - Silver

Total Revolving/Trans Amt/Trans Ct

Total Trans Ct

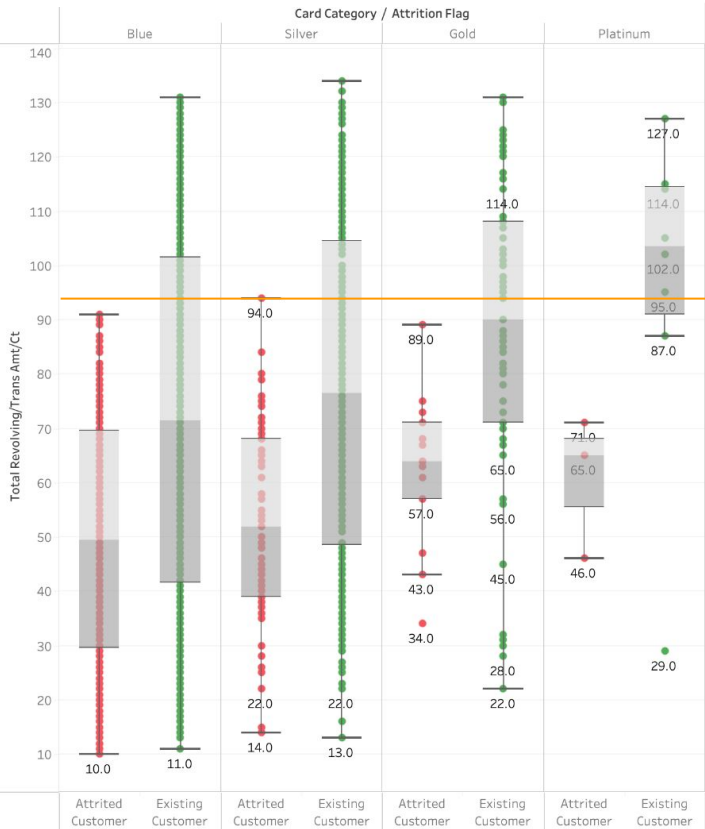
- Attrition Flag
- Attrited Customer
 - Existing Customer

→ The boxplot gives us the distribution of transaction counts for attrited vs existing.

For all cards, there is no attrited customers with transaction counts above 90-94.

Perhaps customers near and below 90 transactions (i.e. 70-80) can be encouraged via promotions to use their cards more often.

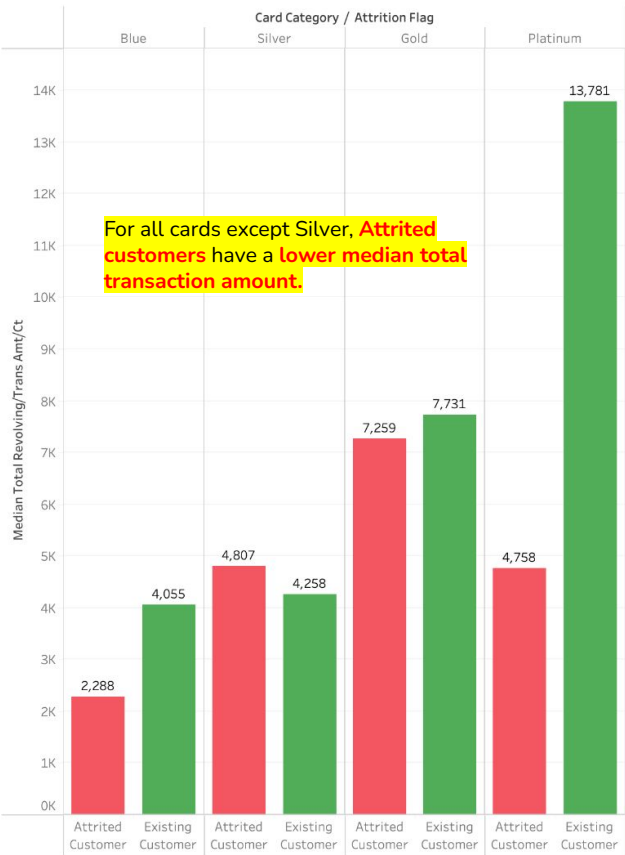
All Cards: Attrited vs Existing boxplot
Total Transaction Count



Spending habits of Attrited vs Existing customers

[View on Tableau Public](#)

All Cards: Attrited vs Existing by median
Total Transaction Amount



For all cards except Silver, Attrited customers have a lower median total transaction amount.

- Card Category
- (All)
 - Blue
 - Gold
 - Platinum
 - Silver

Total Revolving/Trans Amt/Trans Ct

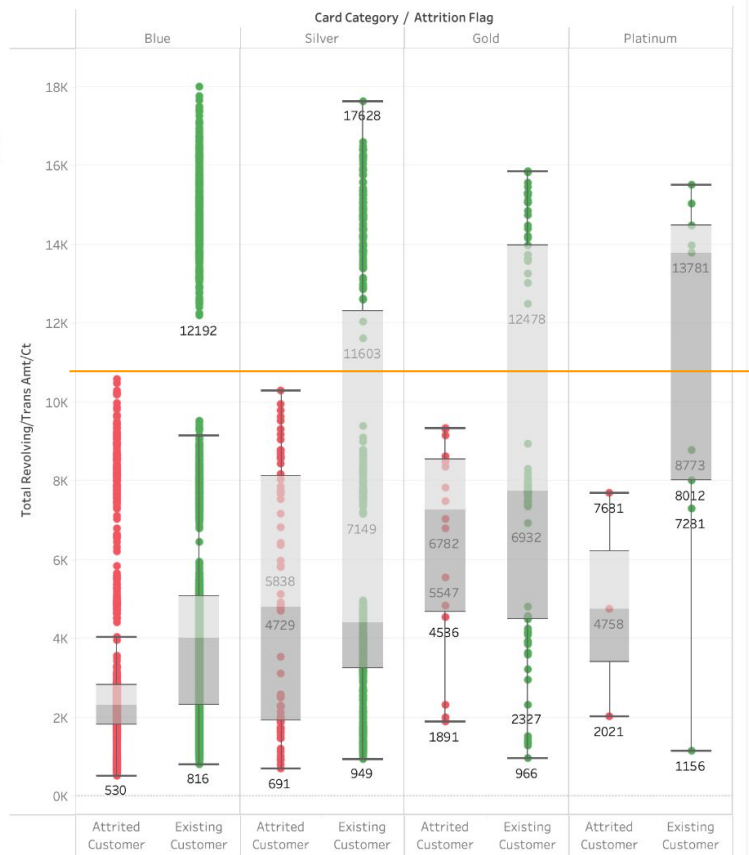
Total Trans Amt

- Attrition Flag
- Attrited Customer
 - Existing Customer

→ Across all cards, attrited customers are below \$11K in total transaction amount.

We should encourage customers to use their cards to achieve higher total transaction amounts.

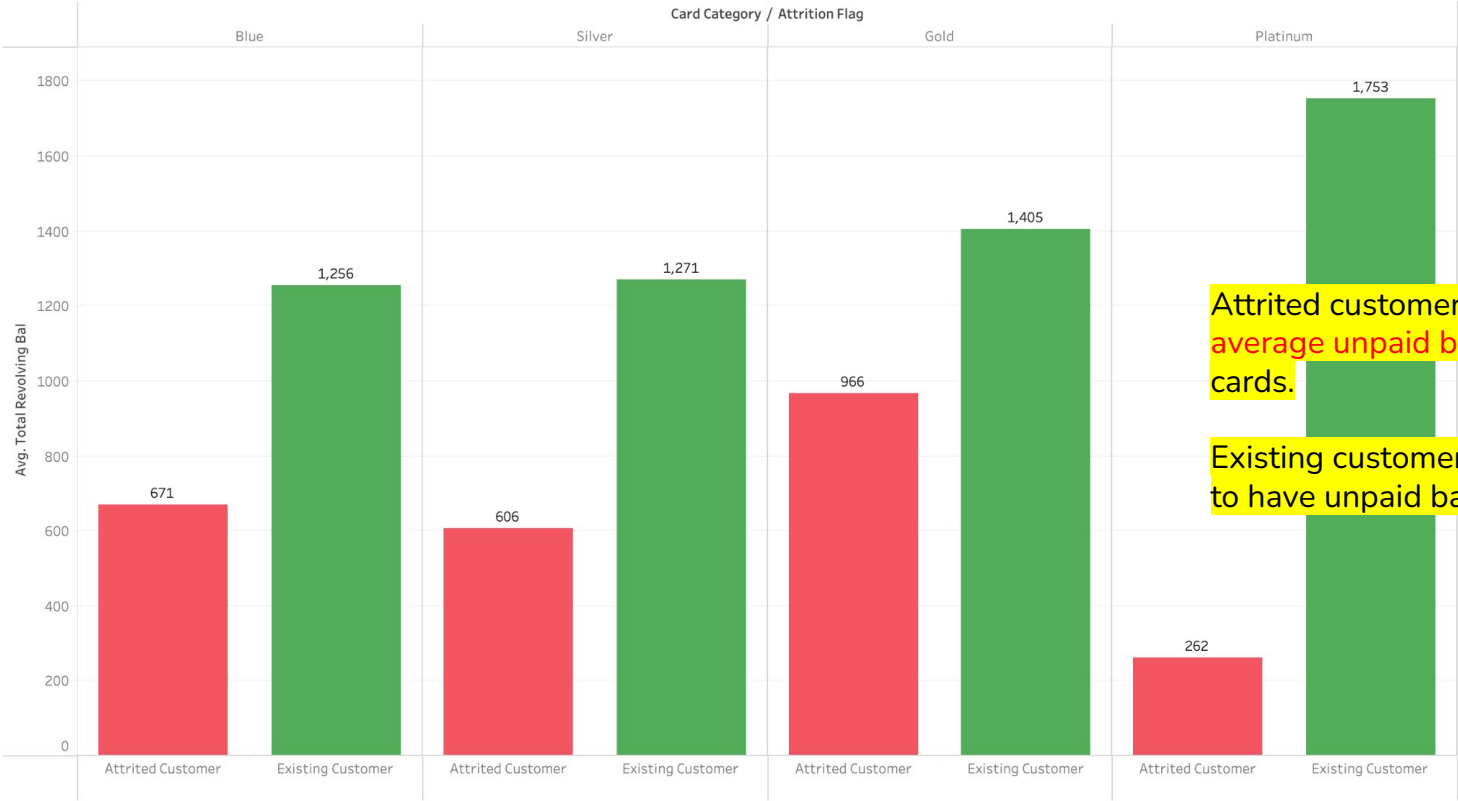
All Cards: Attrited vs Existing boxplot
Total Transaction Amount



Attrited vs Existing customers: Total Revolving Balance

[View on Tableau Public](#)

All Cards: Attrited vs Existing by total revolving balance (**Unpaid balance carried over to next month**)

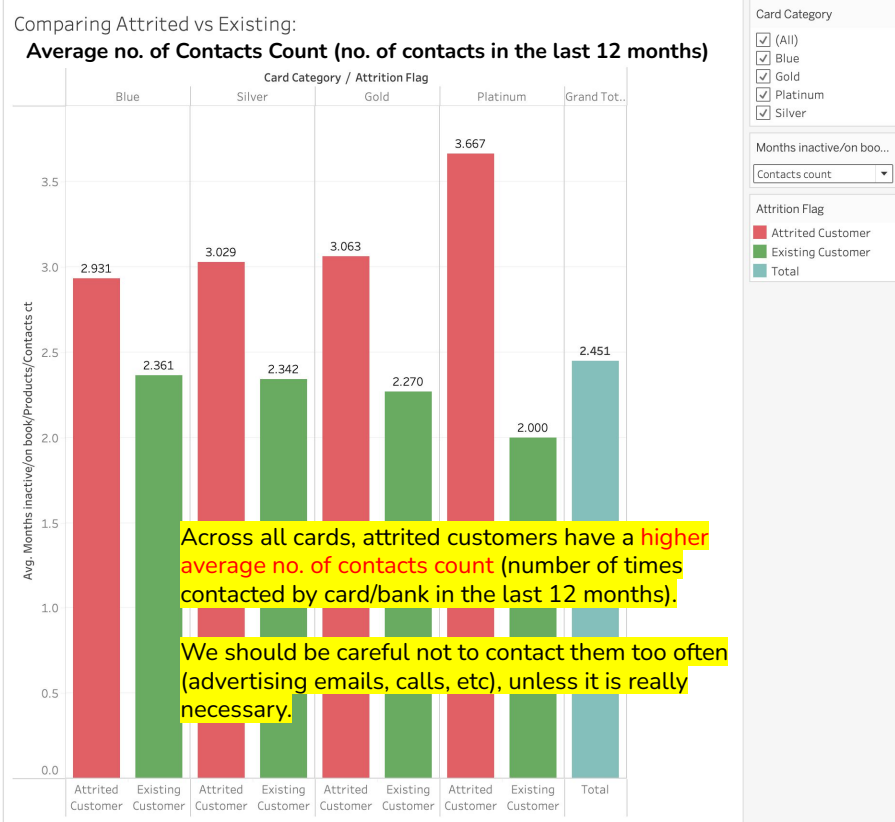
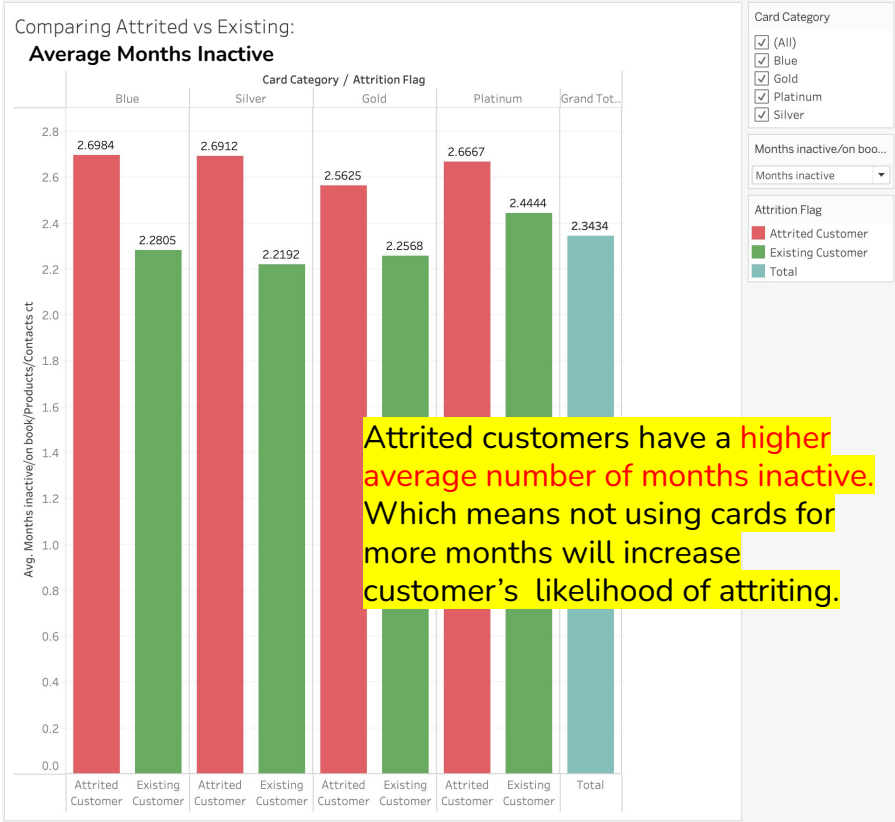


Attrited customers have a lower average unpaid balance on their cards.

Existing customers are more likely to have unpaid balances.

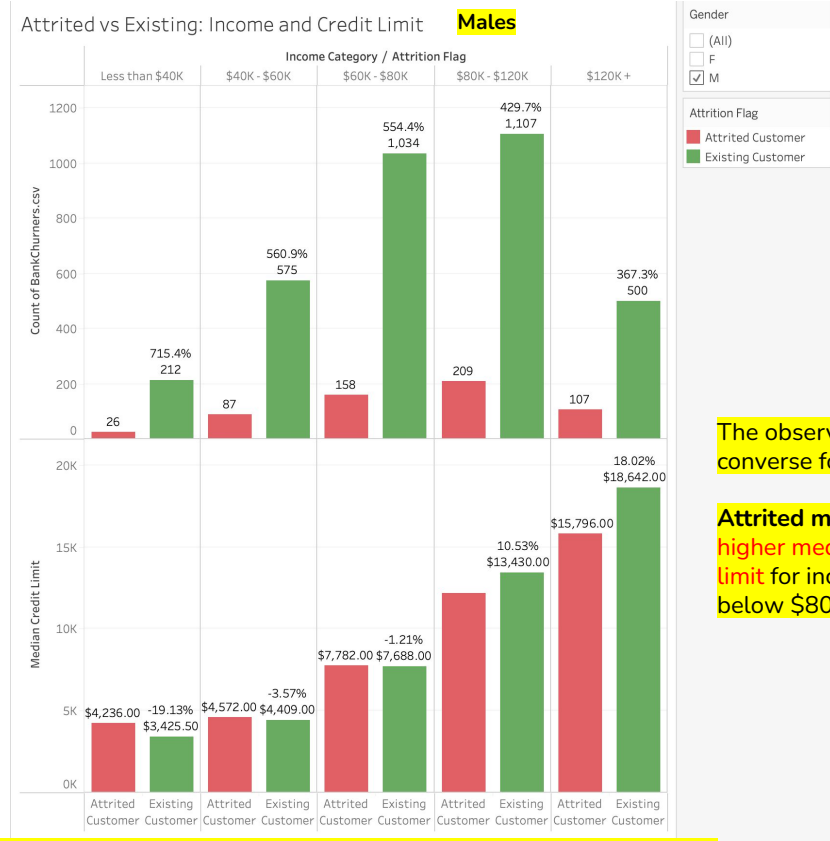
Attrited vs Existing customers:

[View on Tableau Public](#)



Attrited vs Existing customers:

[View on Tableau Public](#)



In this case, the **count of females attrited** in the lower than \$40K income group is high (490) compared to other groups. And female customers stand for a **higher portion (53.2% of attrited customers)**.

There is room for **increasing the credit limit** for females in the less than \$40K income group via A/B testing to check if it helps the attrition rate.

Recap problem statement:

Identify the high attrition risk customer groups so that the bank can provide the customer with better services and reverse churn.



Summary of findings:

Blue and **Silver** card customers stand for a collective **96.87% (\$3.65m)** of the **attrited total transaction amount**.

Hence, might have a **bigger business impact** if our remedial measures are focused on them.

Blue card attrited group:

- Slightly **more female customers** (53.2%)
- Ranking by income category, **less than \$40K** (43.68%), followed by **\$40K-\$60K** (18.68%)
- Slightly **more married customers** (44.75%) than single (39.68%).

Silver card attrited group:

- **More male customers** (73.53%).
- By income, **\$80K-130K** is the largest attrited group (30.88%) followed closely by **less than \$40K** (27.94%).
- **More single customers** (52.94%) than married (29.41%)

Recap problem statement:

Identify the high attrition risk customer groups so that the bank can provide the customer with better services and reverse churn.

Summary of findings:

Across all cards, **attrited customers:**

- ↓ Lower median **total transaction count**
- ↓ Lower median **total transaction amount** (all cards except Silver)
- ↓ Lower average **revolving balance**
- ↑ Higher average **number of months inactive**.

In the **less than \$40K** and **\$40K-\$60K** groups, **attrited females** have a **lower median credit limit** than existing females, and even attrited males in the same income groups.

Recommendations:

Promotions or rewards to **increase customer's transaction count and transaction amount**.

After approx. **2.5 months of inactive months**, customers are more likely to **churn**. Focus promotions to engage inactive customers near that period.

There might be room to **increase the credit limit for females** in the less than \$40K and \$40K-\$60K income groups.

Thank you!