

**School of Informatics & IT**

**SGUS: Upskill in DaTA SCIENCE  
Specialist Diploma in Data Science for Business**

**AY 2021/2022 INTAKE #3**

**Data-driven Storytelling (CDS1S03)**

**Assignment 2**

Submitted by

Student Name: Wong Poh Yeng, Joey

Admin No: 2073306I

Group: 1

Declaration of Work of Originality

**SGUS: Upskill in DaTA SCIENCE**

**Specialist Diploma in Data Science for Business**

**Data-driven Storytelling (CDS1S03)**

**AY2021/2022 Intake #3 – Assignment 2**

**Group: 1**

**Submitted by: Wong Poh Yeng, Joey**

**Date: 2 July 2021**

**“By submitting this work, I am declaring that I am the originator(s) of this work and that all other original sources used in this work has been appropriately acknowledged. I understand that plagiarism is the act of taking and using the whole or any part of another person’s work and presenting it as my/ our own without proper acknowledgement. I also understand that plagiarism is an academic offence and that disciplinary action will be taken for plagiarism.”**

**Name and Signature of student: Wong Poh Yeng, Joey**

**2a. Introduction**

**Key issue behind your change**

Moneycard is a well-known credit card company headquartered in the United States. The company has recently been affected by serious issues. The number of inactive customer accounts are increasing, and many customers are cancelling their credit card services with them.

Losing customers not only makes it harder for the company to grow, but it also leads to significant loss of profit. It may even cause investor to switch to another company.

Therefore, the company’s top priority right now is to identify potential churners and contact them before they cancel their accounts. To address the issues, the key management has delegated the task of analyzing the situation to the business analysis team.

In two weeks, a meeting will be held at Moneycard's headquarters, and the following key management will be present:

* Chief Executive Office
* Chief Finance Office
* Branch Managers

The business analysis team uses tableau to understand the data and determine how to solve the problem.

To better understand the situation and determine how to resolve it, the BA team needs to analyze the followings:

**How to solve the problem**

First of all, our team need to understand the behavior pattern of inactive customer accounts and when they will cancel their account if it is not being used.

Secondly, we must understand our customers and ascertain the following:

1. Which customers are churning?
2. When are they cancelling their account?
3. Why are they cancelling their account?

We use Tableau to analyze the data on gender, marital status, ages, education level, card types, and income level to determine which customers are churning.

To identify when and why they are likely to cancel their account. We must understand the followings:

1. Their customers' life expectancy with the company
2. Which revolving balance brackets do they belong to?
3. Customers spending pattern.
4. The credit limit granted to customers.

**Central insights**

From the storyboard we gathered that a large proportion of our customers who leave are:

1. Female, married women between the ages of 41 and 50 with a graduate degree.
2. People with a blue card and a graduate degree who make less than $40,000 per year.
3. Customers who are given a credit of less than $2,000 are more likely to cancel their accounts.
4. The customers' life expectancy with us is only 36 months, and typically they will leave upon 36 months with us.
5. Customers with less than 250 revolving balances and infrequent card use are more likely to have their account cancelled.
6. Customers who make fewer purchases and spend less money are more likely to leave us.

**2b. Data Attributes**

This dataset was downloaded from Polymall. It contains data of more than 10,128 credit card accounts with around 21 variables of different types as of a time point.

Data description is as below:

|  |  |  |
| --- | --- | --- |
| Variable | Type | Description |
| Clientnum | Num | Client number. Unique identifier for the customer holding the account |
| Attrition\_Flag | char | Internal event (customer activity) variable - if the account is closed then 1 else 0 |
| Customer\_Age | Num | Demographic variable - Customer's Age in Years |
| Gender | Char | Demographic variable - M=Male, F=Female |
| Dependent\_count | Num | Demographic variable - Number of dependents |
| Education\_Level | Char | Demographic variable - Educational Qualification of the account holder (example: high school, college graduate, etc.) |
| Marital\_Status | Char | Demographic variable - Married, Single, Unknown |
| Variable | Type | Description |
| Income\_Category | Char | Demographic variable - Annual Income Category of the account holder (< $40K, $40K - 60K, $60K - $80K, $80K-$120K, > $120K, Unknown) |
| Card\_Category | Char | Product Variable - Type of Card (Blue, Silver, Gold, Platinum) |
| Months\_on\_book | Num | Months on book (Time of Relationship) |
| Total\_Relationship\_Count | Num | Total no. of products held by the customer |
| Months\_Inactive\_12\_mon | Num | No. of months inactive in the last 12 months |
| Contacts\_Count\_12\_mon | Num | No. of Contacts in the last 12 months |
| Credit\_Limit | Num | Credit Limit on the Credit Card |
| Total\_Revolving\_Bal | Num | Total Revolving Balance on the Credit Card |
| Avg\_Open\_To\_Buy | Num | Open to Buy Credit Line (Average of last 12 months) |
| Total\_Amt\_Chng\_Q4\_Q1 | Num | Change in Transaction Amount (Q4 over Q1) |
| Total\_Trans\_Amt | Num | Total Transaction Amount (Last 12 months) |
| Total\_Trans\_Ct | Num | Total Transaction Count (Last 12 months) |
| Total\_Ct\_Chng\_Q4\_Q1 | Num | Change in Transaction Count (Q4 over Q1) |
| Avg\_Utilization\_Ratio | Num | Average Card Utilization Ratio |

We used Excel to fulfill the following tasks:

1. check that there was not duplication of credit card account numbers.
2. check that there were no missing values on each column.
3. clean up some inconsistencies in the data, such as ensuring that the credit limit column numbers are rounded up and that the Total Amt Chng Q4 Q1, Total Trans Ct, and Avg Utilization Ratio columns only have two decimal places.
4. The Attrition Flag column initially shows the categorical value (existing and attired customers); We have created additional “numerical” columns for this variable (0 & 1).
5. We used Excel to group the customer age column for better visualization of the age range.
6. additional binning was performed in tableau on the following columns: average utilization percentage, credit limit, number of months with bank, total revolving balance, total transaction amt, Total\_Trans\_Ct.

**2c. Effective Visualisation and Insights**

A line chart, bar chart, scatter plot, and dotted chart were used to plot this project.

|  |  |
| --- | --- |
|  |  |
|  | We used Bar graphs to compare the data in gender, marital status, revolving balance, credit limit.  It was used because it is easy to compare sets of data between different groups.  For effective visualization, we used the following techniques.   * Consistent colors were used throughout. * Darker color was used to highlight important data point. Grey color was pushes unimportant datapoint into background. * The annotation was used and the same color text was used to tie with the Graphs |

|  |  |  |
| --- | --- | --- |
|  | | We also used Bar graph when comparing the Attrition and existing customers.  The bar was colored with red color and also the same color of the highlighted text in the annotation because we want to draw the audience's attention to the 6% difference. |
|  | | Scatterplot was used to show the amount of customer spending patterns, as well as the normal tendency and outliers.  We used this graph because we wanted to compare a large number of data points and we can clearly see the patterns. |
|  | | We used this chart when comparing two variables:   1. Ages vs education 2. Card type vs Income category   From these graphs, we can clearly see the relative sizes and the highest groups. |
|  | Some techniques we used for better visualization are as follows:   * Color contrast with darker color to direct the audience's attention to key points. * the same color text was used in the annotation to tie with the Graphs | |
|  | Line chart was used because we wanted to see the number of attrited and existing customers changes over the months of inactivity.  Some techniques we used for better visualization are as follow:   * the same color text was used in the annotation to tie with the Graphs * Direct labels were used so that audience no need to reference to the legend | |

**2d. Storyboard and Narrative Model**

The storyboard's layout is as follows:

|  |  |  |
| --- | --- | --- |
| Problem |  | |
| Analysis | |  |
|  | |  |
| Solution | |  |

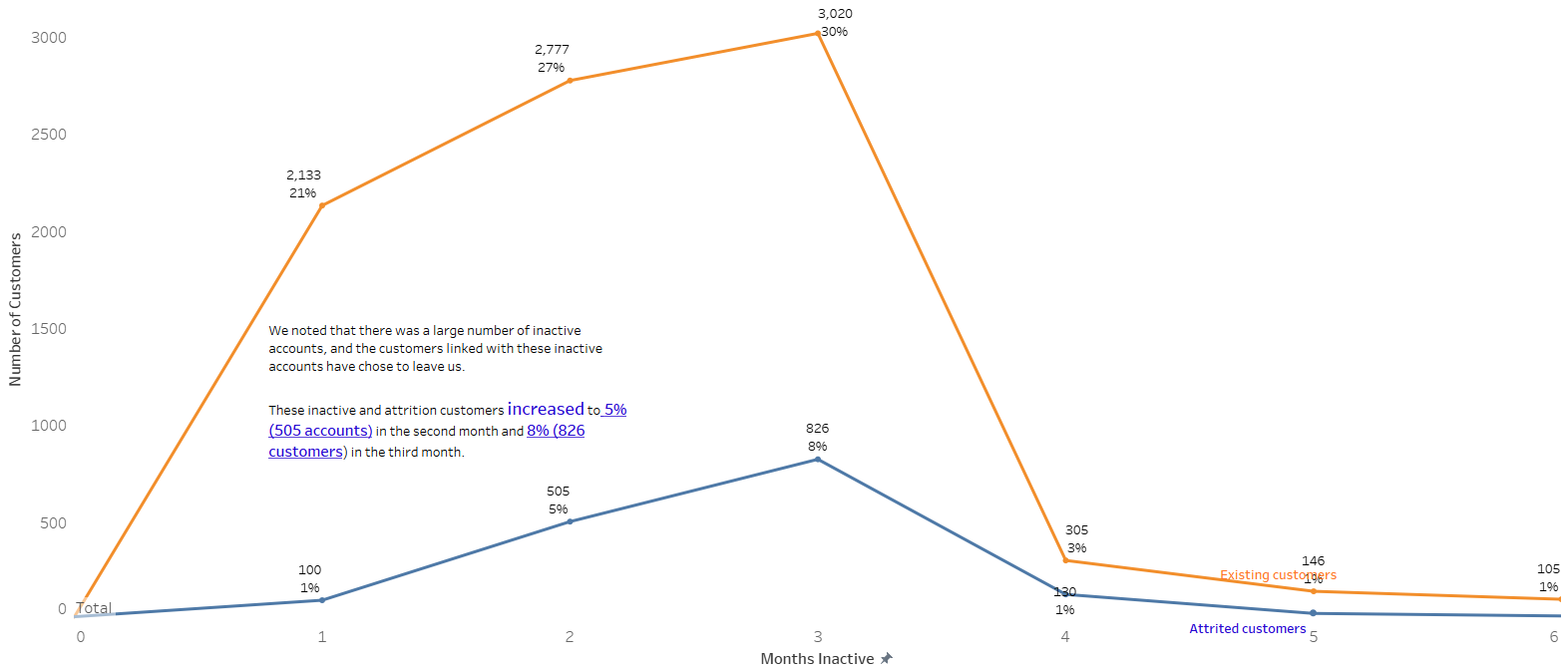
**Storytelling Narrative Model**

For our data-driven story, we used Freytag's pyramid because it is more useful in terms of dramatic structure.

It was divided into five sections: introduction (exposition), rising action, climax, falling action, and resolution. Start by setting the scene and then lead the audience to the main insight (climax), followed by a recommendation or resolution.

After introductions and setting the background of the company, we started the story as follows:

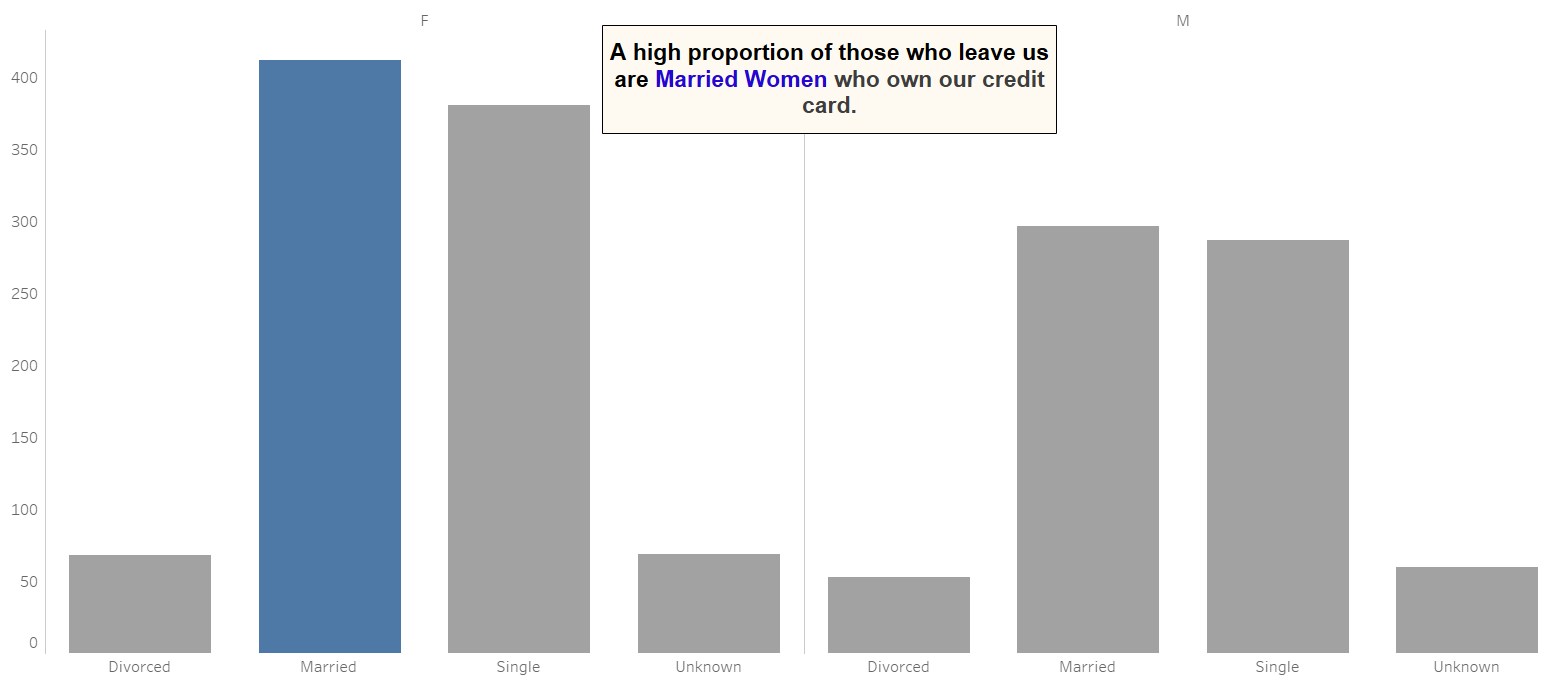
1. **Exposition**

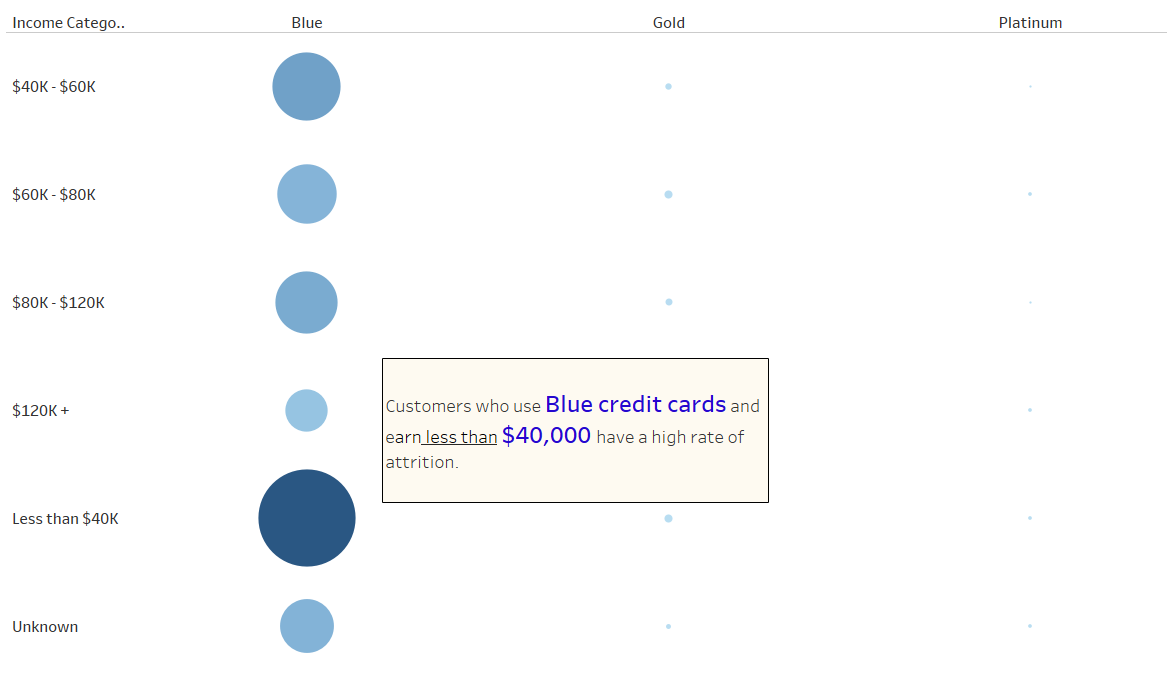
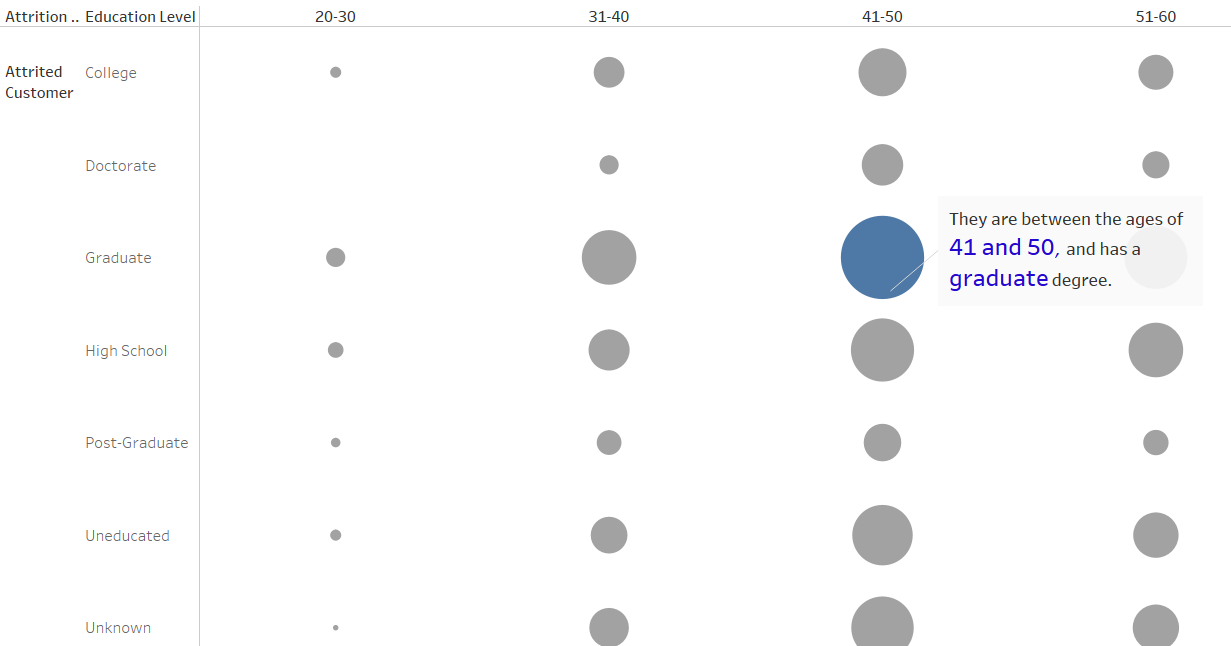


We began to introduce tension to the audience by stating that there were a large number of inactive accounts. These customers have a tendency to leave us, especially after 2 or 3 months of inactivity.

These inactive and attrited customers increased significantly to 5% (505 accounts) in the second month and 8% (826 customers) in the third month.

1. **Rising Action**

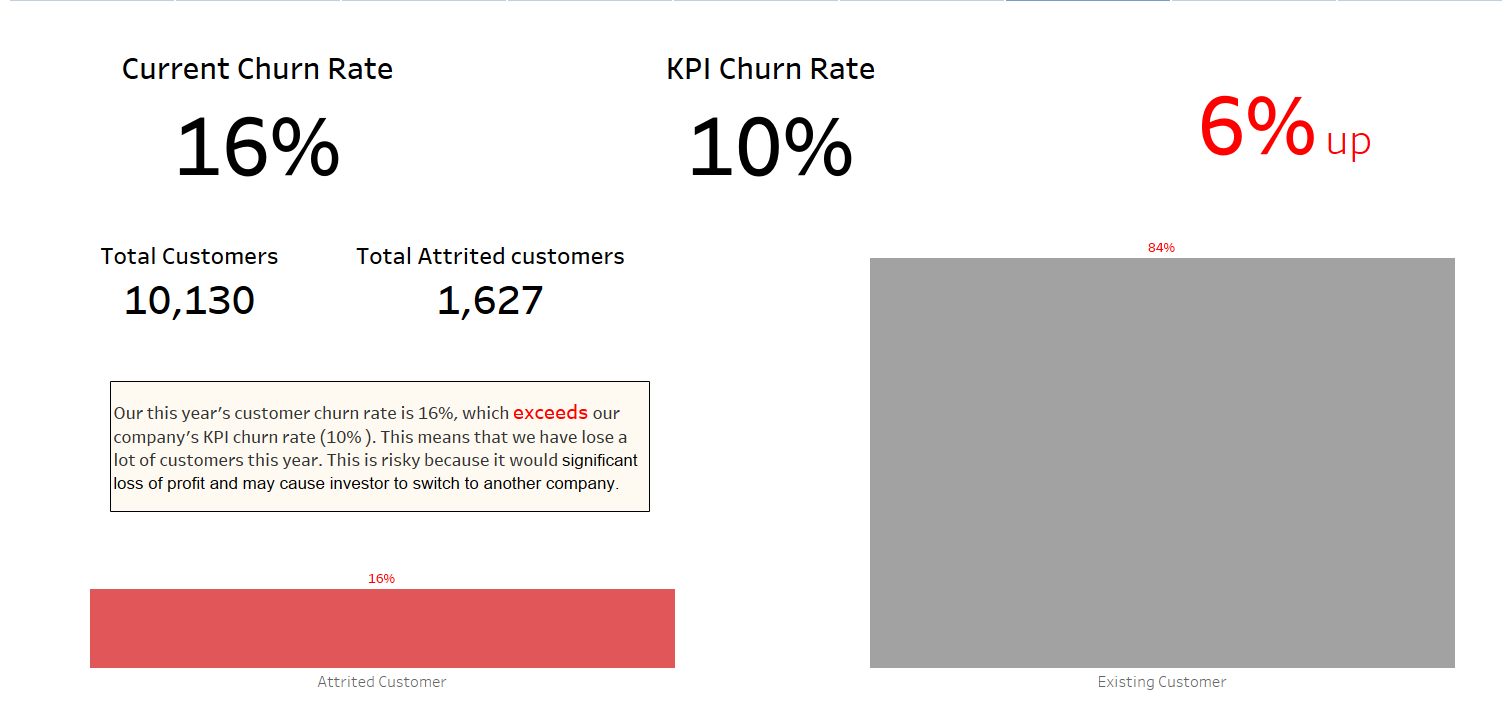
****

** **

After looking at the data, we discovered that a large proportion of these inactive accounts belong to Married Women who own our credit card, are between the ages of 41 and 50, and hold a graduate degree.

These customers also hold blue credit cards and earn less than $40,000 per year.

1. **Climax**

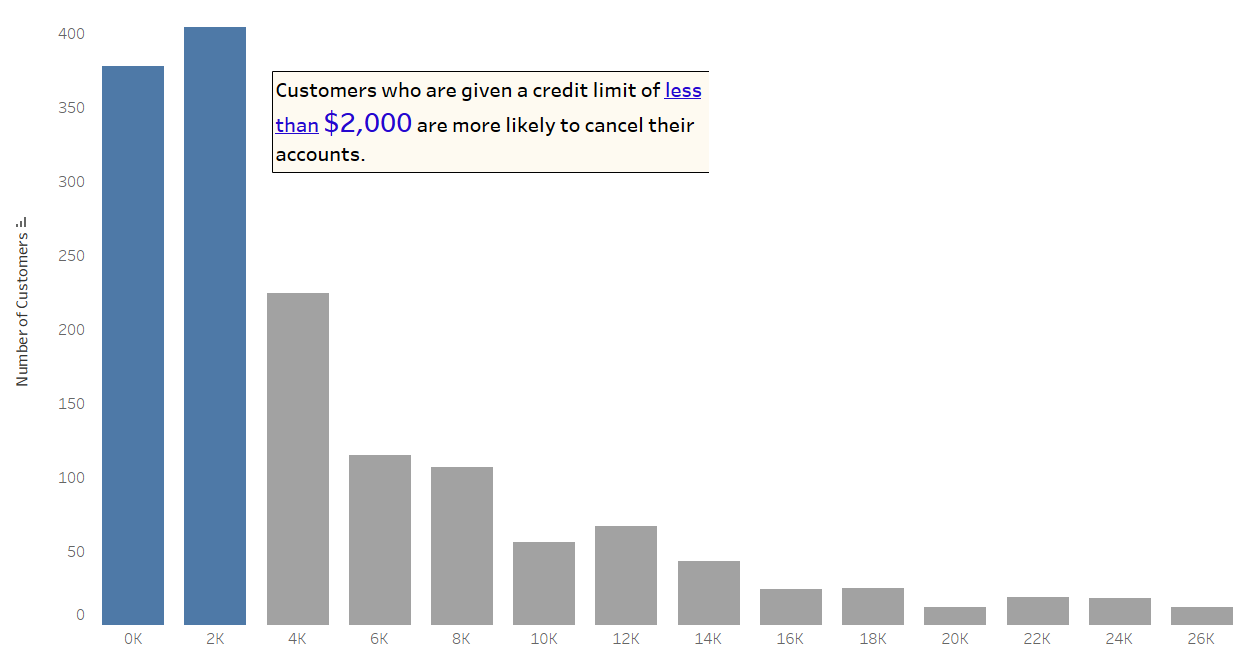
****

These inactive customers will eventually leave us. This year, our churn rate reached 16%, exceeding our KPI churn rate of 10%.As a result, we have lost a significant number of customers this year.

This is risky because it would result in a significant loss of profit and could lead to investors switching to another company.

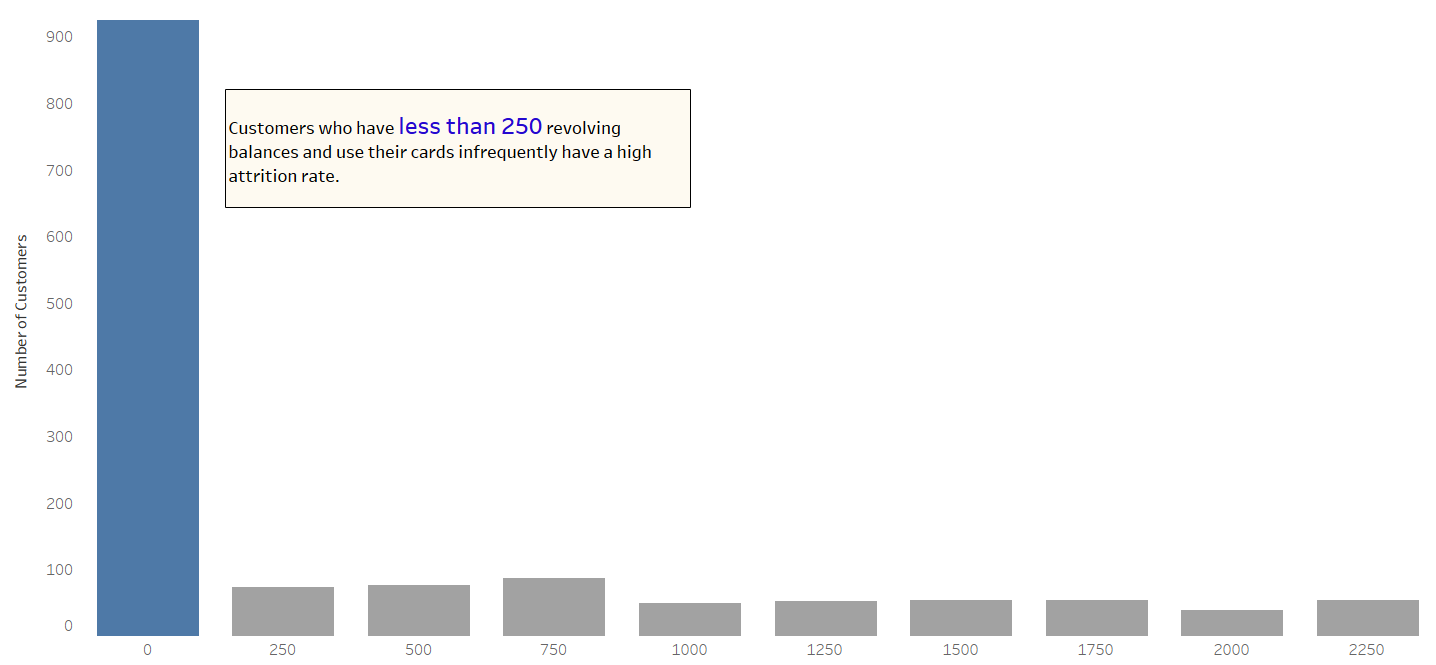
Therefore, we must act quickly or risk losing a significant number of customers the following year.

1. **Falling Action**

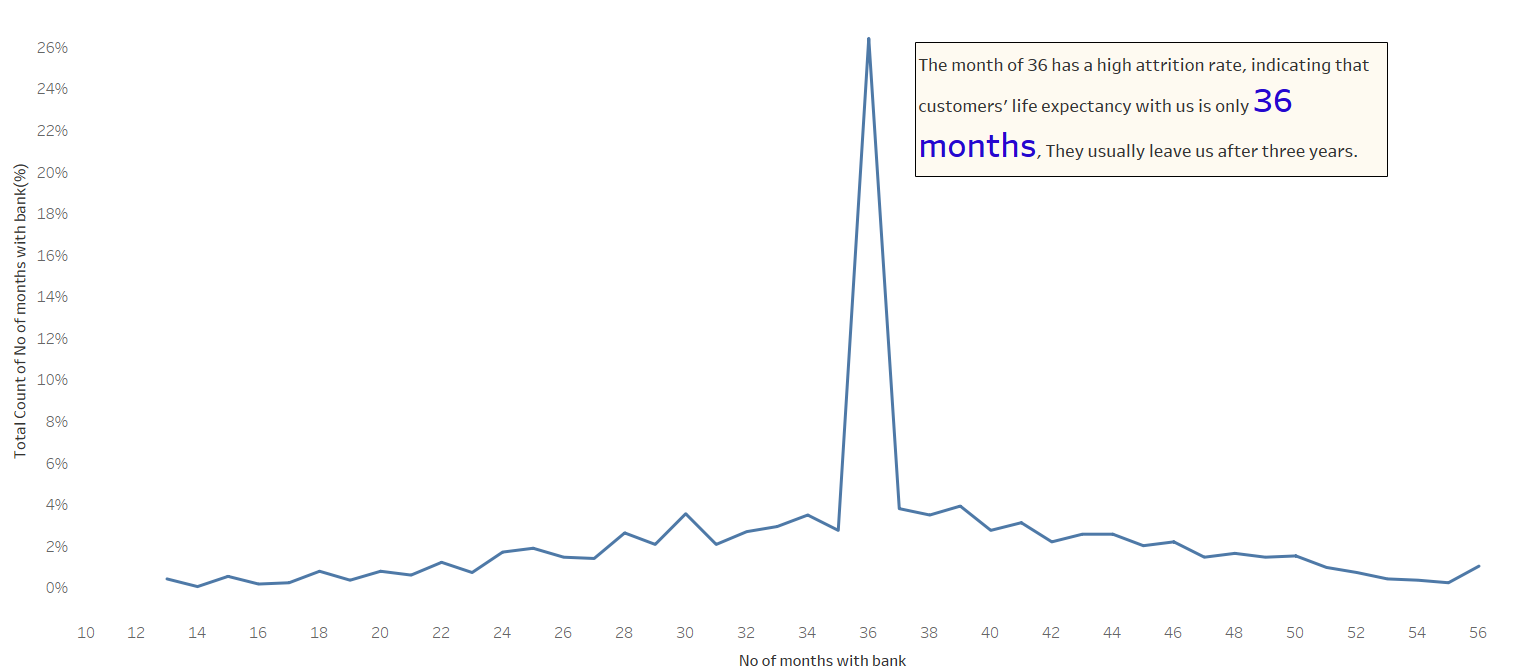


After digging deeper, we discovered that customers who are given a credit limit of less than $2,000 are more likely to cancel their accounts.

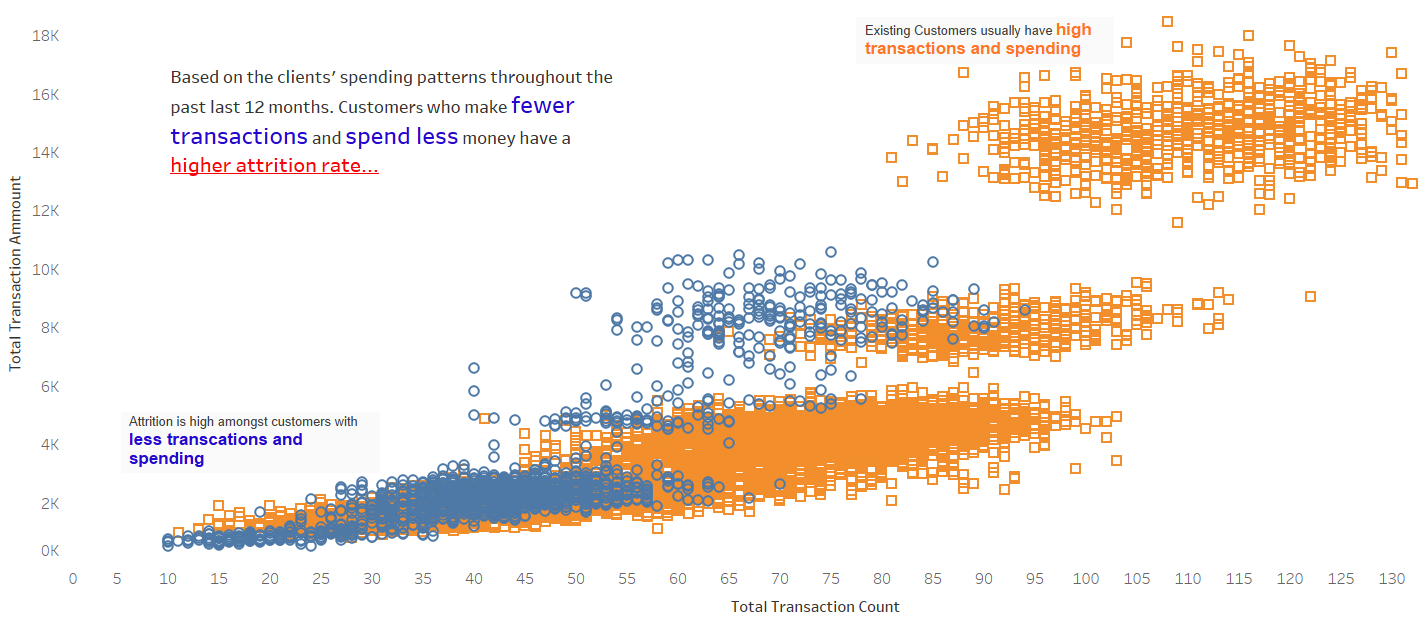
**4.. Falling Action**



Those customers who have less than 250 revolving balances and use their cards infrequently also have a high attrition rate.



In month 36, there was a high attrition rate, indicating that the customers' life expectancy with us is only 36 months. They usually leave upon 3 years with us.



We also explored clients' spending habits over the past year and discovered that customers who make fewer transactions and spend less money are more likely to leave us.

1. **Resolution**

To address this issue we must form a special task force to execute the followings:

* To keep track of inactive customer accounts.
* Analyze the creditworthiness of the customers and increase credit limits for customers with lower credit limits (under $2000).
* Providing more rewards, bonuses, and discounts to encourage them to use their credit balance and do more business with us.
* Design a retirement plan specifically for married women between the aged of 41 and 50.
* Consider introducing a variety of payment plans and interest rates for this group of people.
* Cold call and then follow up with customers who have been with us for nearly 36 months.

**2e. Conclusion with a call-to-action**

|  |
| --- |
| Loss of customers not only makes it more difficult for the company to grow, but it also results in significant profit loss, and our investor may switch to another company.  To avoid losing more customers, we must act quickly and form a special task force to address this issue. |