Attrition of Credit Card Users

INFO 4907

Jonah Turner

Abstract

An important indicator in the success of credit card companies is the attrition rate of credit card users. Keeping the attrition rate of users sufficiently low decreases the number of customers switching to competitors and increases revenue from customer purchases. By analyzing the patterns of customers who stop utilizing their credit card, the company can identify customers in danger of leaving and implement measures to preserve their customer base.

Introduction

Three questions will be important in successfully analyzing the attrition of credit card customers.

* What features are most common among attrited customers?
* What are the characteristics of customers who use their credit card the most?
* Which features are the most different between the users who continue to use their credit cards versus those who cancel their credit cards?

In order to understand the most compelling features of attrited customers, it is necessary to understand the features of the most frequent users of credit cards. The second question is therefore to provide context to the analysis of the other questions.

Methods – Exploratory Data Analysis

The dataset was collected from Kaggle and consists of customer data from a major brand of credit card, including over 10,000 observations. The dataset was first cleaned of two erroneous features included in the download that would have had no bearing on the results of the analysis. This was conducted in R using the dplyr package in the Tidyverse.

The exploratory data analysis was conducted to show the validity of the three questions posed. It was first necessary to show that there were enough customers in both subsets from which to draw conclusions. The following graph shows the distribution of customers. Chart, bar chart

Description automatically generated

While there are far more existing customers than there are attrited customers, the almost 2,000 attrited customers provide a sample large enough to draw some conclusions.

Examining the possible differences in factors between the two subsets of customers will inform the visual analysis. The first feature examined was the total amount of transactions on the card. There is a clear difference in the distributions.

Chart, diagram, box and whisker chart

Description automatically generated

Average utilization ratio is another feature that may have different distributions among customer subsets. The distribution of this feature is shown in the following graph. Chart, box and whisker chart

Description automatically generated

The Exploratory Data Analysis confirms enough useful differences in the data among the subsets of customers to continue with a more thorough visual analysis.

Visual Analysis

Chart, treemap chart

Description automatically generated

The treemap above shows some of the most common characteristics among attrited customers. The treemap is colored by gender, as the features of men and women who have attrited are significantly different. The females, of whom there are more by count, are more likely to be in the lowest income bracket of this dataset, with many more highly educated women attritioning than in the other education brackets. This phenomenon is made more significant by the demographics of education, with far fewer people by percentage reaching a graduate school level of education. This suggests strongly that highly educated women are far more likely to abandon their credit cards than other demographics.

Men experience a similar trend in education but a different trend in income. The most educated men also abandon their credit cards in higher numbers. This is even more significant in the male populace, as fewer men than women by count reach the highest levels of education. By income, there are more men in higher income brackets who abandon their credit cards than in the lower income brackets.

Chart, bubble chart

Description automatically generated

The packed bubbles chart above shows differences between strong customers. The size of the bubbles corresponds to the average utilization rate of the customers – a good metric for customer strength. The color of the bubbles shows gender, with females being blue and men orange. Income bracket and education levels are labeled on the bubbles. This graph shows that average utilization rate is higher in general among women. The lower income brackets are far more present, with incomes under $40,000 and between $40,000 and $60,000 being represented most frequently. Education seems to be less of an indicator than income, with all education levels represented in the strongest customers.

Chart, treemap chart

Description automatically generated

This visualization highlights several important differences among the features of attrited and existing customers. The average utilization rate of existing customers is almost double that of attrited customers. The average total revolving balances of the two subsets follows a similar ratio, with existing customers carrying almost $600 more on average. The number of transactions is not as heavily skewed, but existing customers utilize their cards far more frequently than attrited customers did. The average number of months inactive in the past 12 months does not provide as large of a difference between the two groups, but attrited customers tended to go longer without using their cards than did existing customers.

Results

The dataset yielded significant results that can be used to answer the three questions presented. Female attrited customers are often in the lower income brackets and are highly educated, while men are more frequently in higher income brackets and similarly educated. The difference in income brackets is the most striking between men and women, with fewer men in the lowest income bracket leaving their credit cards and exactly the opposite phenomenon occurring in women.

Strong customers show less variance between men and women in average utilization rate. Strong customers are frequently in the lower income brackets with varying levels of education. Women in general have a slightly higher utilization rate than men, and, in general, lower levels of education similarly have slightly higher utilization rates. The most defining characteristic of utilization rate is income bracket, with the lower income brackets having a significantly higher utilization rate than the higher income brackets. The utilization rate drops significantly for the brackets above $60,000.

The differences between attrited and existing users are clear. Attrited customers have far lower utilization rates than existing customers. Attrited customers also have lower revolving balances and total transaction counts than existing customers. Attrited customers have slightly longer average months inactive in a 12 month period than existing customers, but the relationship is not as strong as the other features.

Discussion

There is a measurable difference in many features between attrited and existing customers. The visualizations show enough evidence for and would be strengthened further by a classification algorithm that can help to predict the features of a leaving customer. By being able to predict customer attrition, a company could increase customer retention by implementing measures specific to customer demographics that experience higher attrition rates.