Visualization Project

Credit Card Customers

Deepmalya Dutta

Roll No: MDS202218

DG1107: Visualization

Email: deepmalya@cmi.ac.in

Chennai Mathematical Institute

30-10-2022

Abstract

A credit card is a payment card issued to users to enable the cardholder to pay a merchant for goods and services based on the cardholder's accrued debt. In today's digital age, it has become an intergate part of day-to-day transactions and money exchange. In this project, we explore the trend and usage of credit card users among different demographics. We use visualization techniques in R to find patterns in the credit card usage dataset and following this, we attempt to find relations between Credit Card Usage, Utilization, Revolving Balance, and other variables. Finally, we develop a dashboard to address the above-posed questions.

Keywords : Credit Card, R, Ggplot2, Visualization.

Table of Contents

1. *Introduction:*

A credit card is a thin rectangular piece of plastic or metal issued by a bank or financial services company that allows cardholders to borrow funds with which to pay for goods and services with merchants that accept cards for payment. Credit cards impose the condition that cardholders pay back the borrowed money, plus any applicable interest, as well as any additional agreed-upon charges, either in full by the billing date or over time. It works like a loan, but instead of receiving money beforehand, you get a set credit limit on the card. the issuing authority sets a credit limit on the amount that can be spent before you have to pay back a specified amount to the lender at the end of the month.
Credit cards come with some benefits:

- You can use a credit card to make a large purchase which can be repaid in smaller installments over some time
- A credit card is relatively safer to carry as against cash and is accepted mostly everywhere
- If you have a good credit score, you can avail of additional cash back, lower interest rates, and many more exciting offers

Having such benefits, the usage of credit cards is increasing in the modern age. In this project, we want to analyze the credit card user base and attempt to find different trends.

2. *DataSet Description:*

This data set consists of 10,000 customers mentioning their age, salary, marital status, credit card limit, credit card category, etc. There are nearly 18 features. The data is sourced from [1].

- **CLIENTNUM** : Client number. An unique identifier for the customer holding the account
- **Attrition_Flag** : If the account is closed then 1 else 0 - Customer activity variable
- **Customer_Age** : Customer's Age in Years - Demographic variable
- **Gender** : M = Male, F = Female - Demographic variable
- **Dependent_count** : Number of dependents - Demographic variable
- **Education_Level** : Educational Qualification of the account holder (example: high school, college graduate, etc) - Demographic variable
- **Marital_Status** : Married, Single, Divorced, Unknown - Demographic variable
- **Income_Category** : Annual Income Category of the account holder (< \$40K, \$40K - 60K, \$60K - \$80K, \$80K-\$120K, > \$120K) - Demographic variable
- **Card_Category** : Type of Card (Blue, Silver, Gold, Platinum) - Product Variable
- **Months_on_book** : Period of relationship with 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 in the last 12 months
- **Credit_Limit** : Credit Limit on the Credit Card
- **Total_Revolving_Bal** : Total Revolving Balance on the Credit Card
- **Avg_Open_To_Buy** : Open to Buy Credit Line (Average of last 12 months)
- **Total_Amt_Chng_Q4_Q1** : Change in Transaction Amount (Quarter 4 over Quarter 1)
- **Total_Trans_Amt** : Total Transaction Amount (Last 12 months)
- **Total_Trans_Ct** : Total Transaction Count (Last 12 months)
- **Total_Ct_Chng_Q4_Q1** : Change in Transaction Count (Quarter 4 over Quarter 1)
- **Avg_Utilization_Ratio** : Average Card Utilization Ratio.

3. *Objectives:*

Using the above dataset, we are trying to determine the following:

- How transaction amount, limit, revolving balance, and utilization are dependent on different demographics
- How different income groups spend their money on credit card transactions
- How changing card limit affects things
- How does having more revolving balance effects transactions
- Which groups should the card companies target to increase their business

The above objectives can be achieved using histograms, piechart and boxplots for univariate analysis, scatter plots and line plots for bivariate analysis, and gradient scatter plots for multivariate analysis.

4. *Graphical presentation of Key variables:*

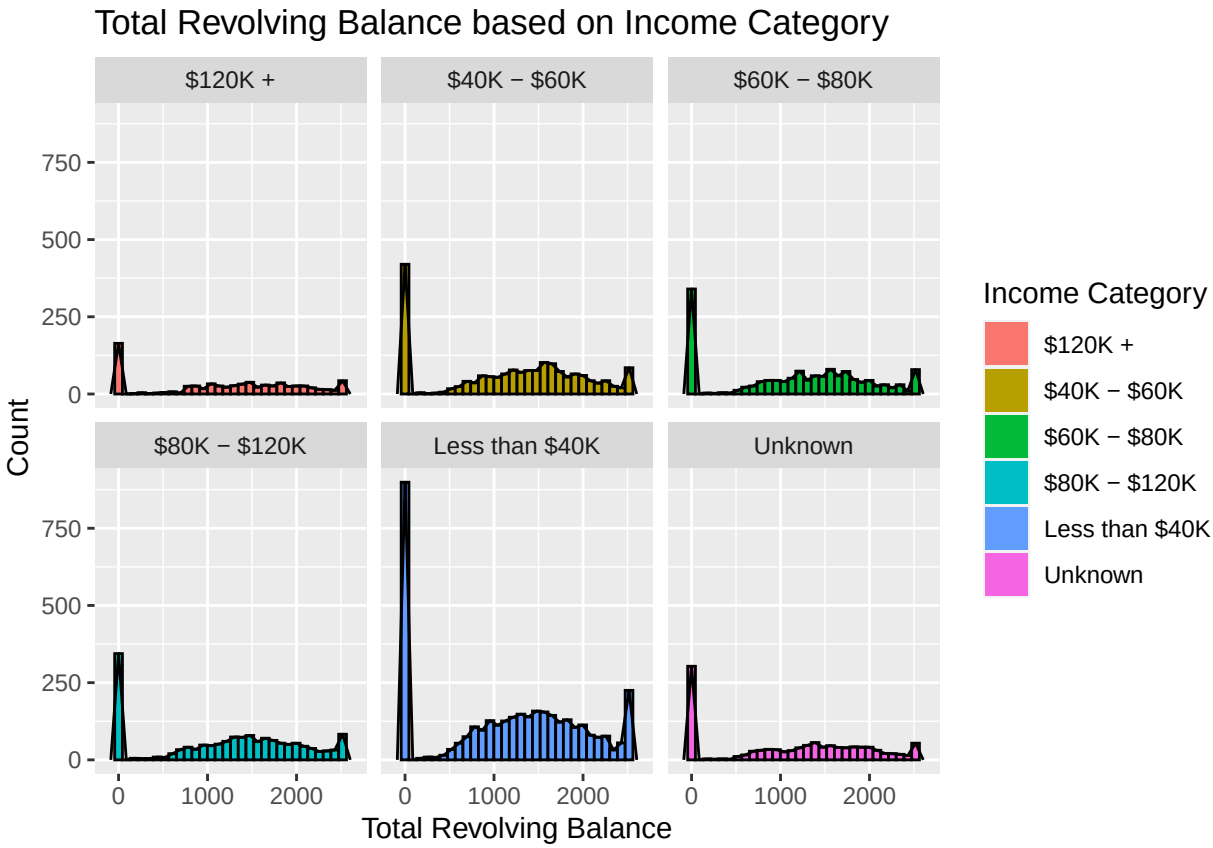

Figure 1: **Histogram Representation.** We see that for all the income categories, total revolving balance is centered around the median and at 0 it's at peak, that is most accounts have no total revolving balance.

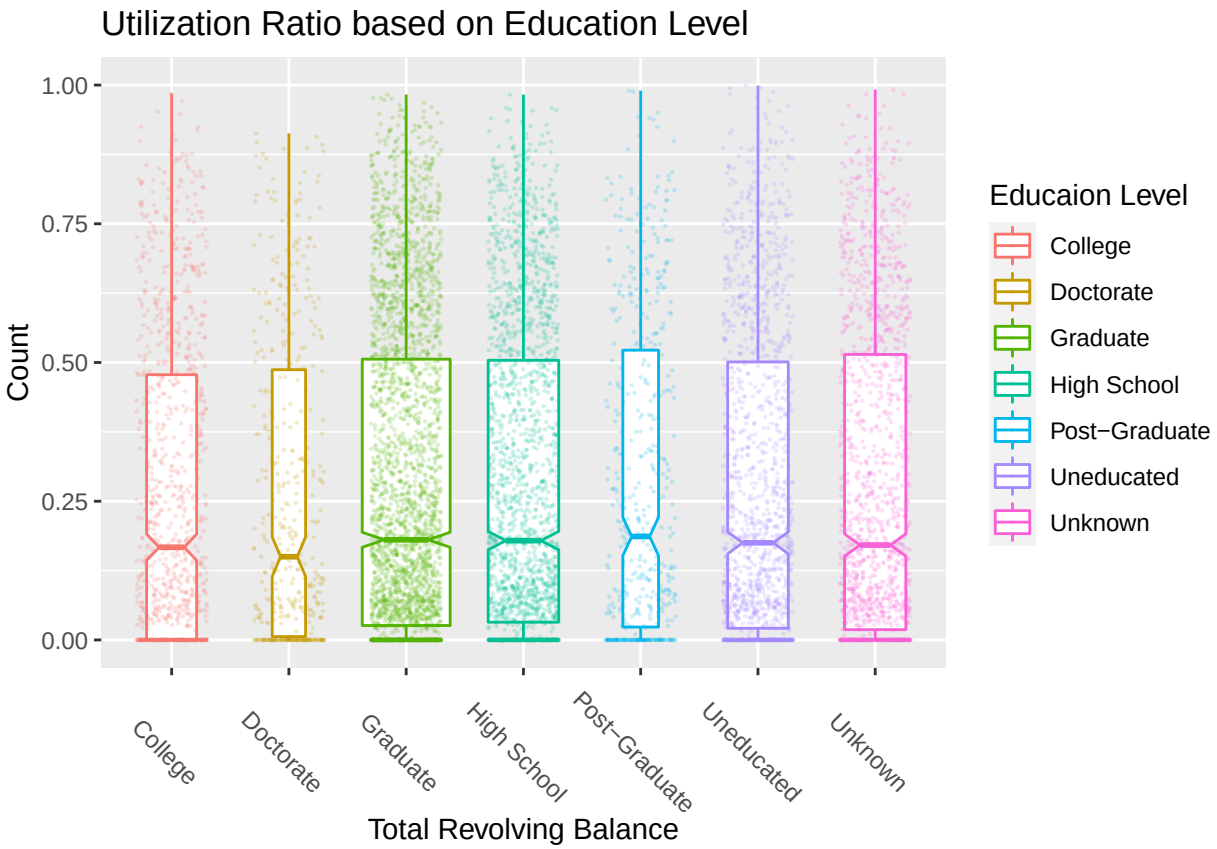

Figure 2: **Boxplot Representation.** We see that despite of having different education qualifications, the total revolving balance amount count is more or less same across different demographics. Also a large number of points are outliers.

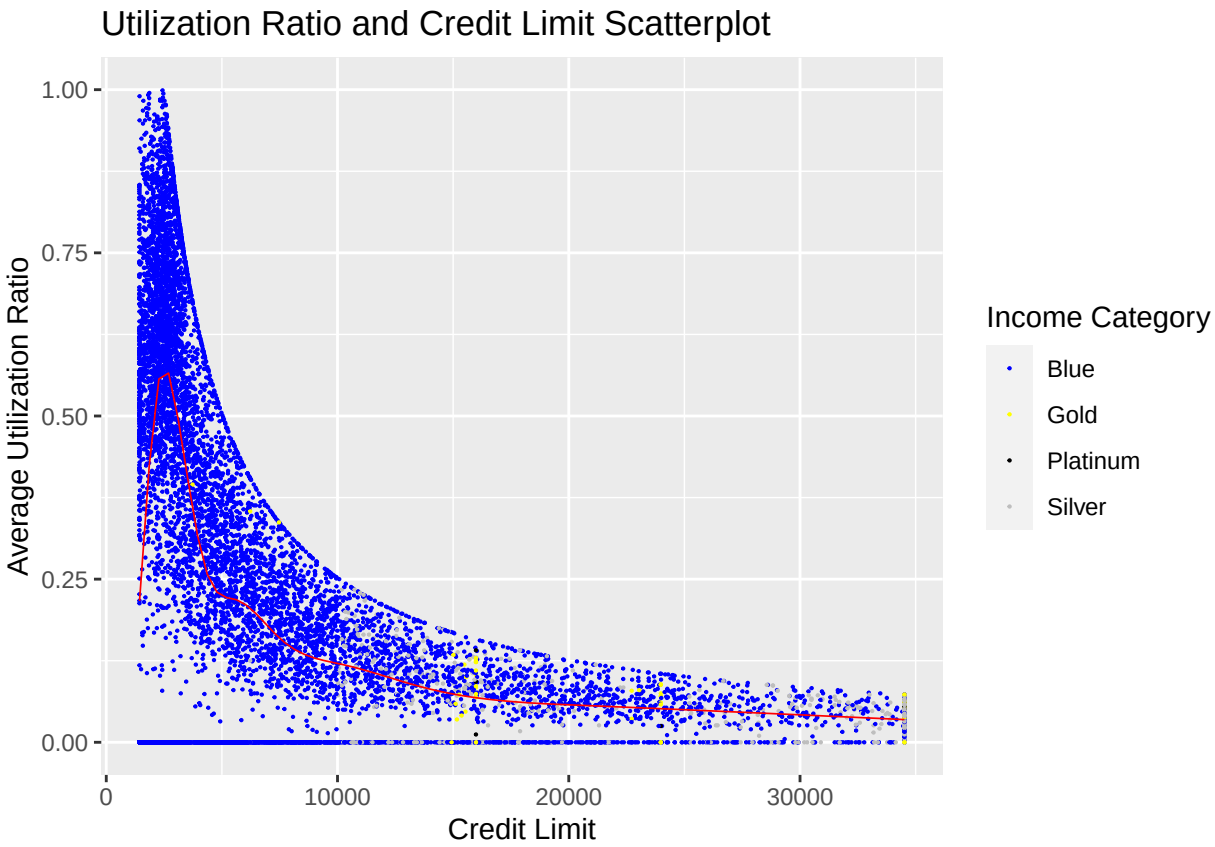

Figure 3: **Scatterplot Representation.** Here we can see as the credit limit is increasing, the average utilization ratio is decreasing, in a logarithmic relation. That is the higher the credit limit, the less customers are using it. Also we see the line peaks at around 2500, which is the credit limit, where utilization peaks.

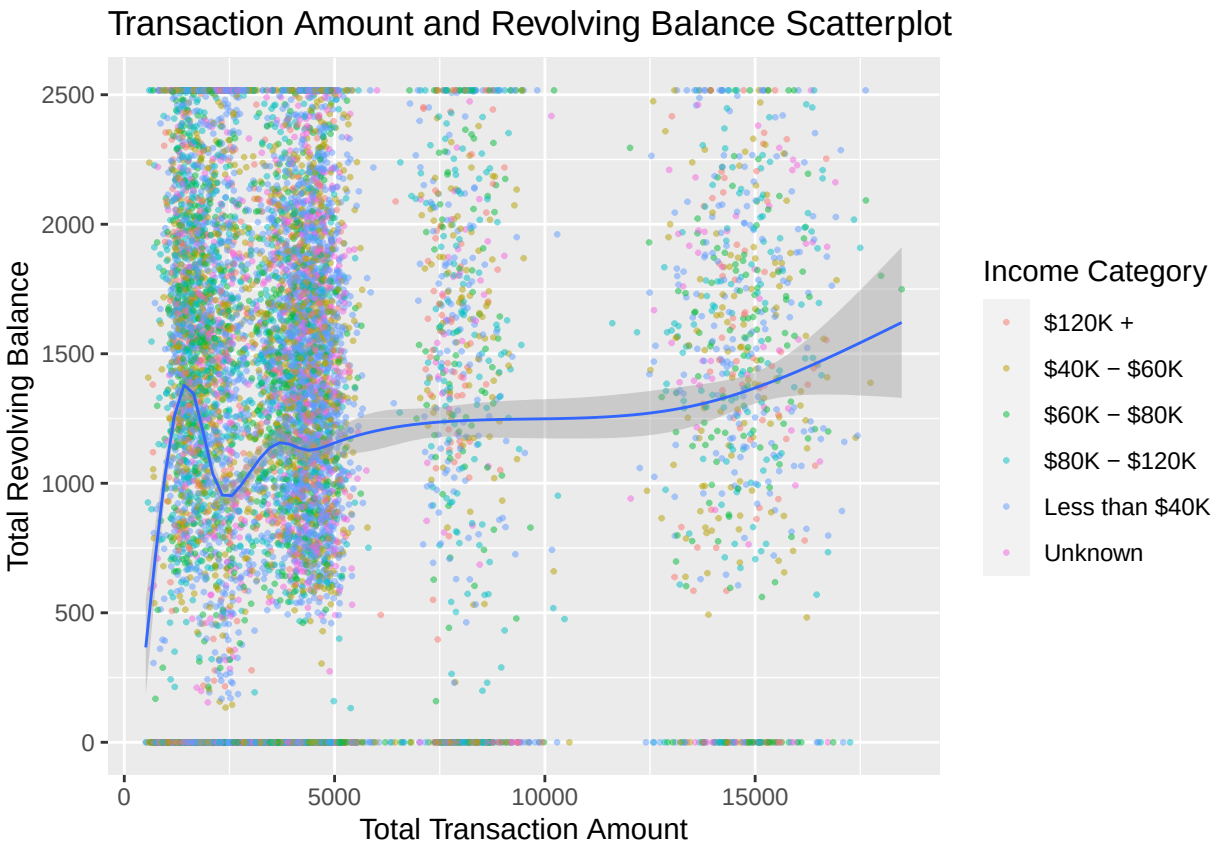

Figure 4: **Scatterplot Representation.** We see the total transaction amount is clustered around 3 points and most these are in between 0 to 5000. Also no matter what the transaction amount is, the total revolving balance stays around the median, with large points at 0 and max.

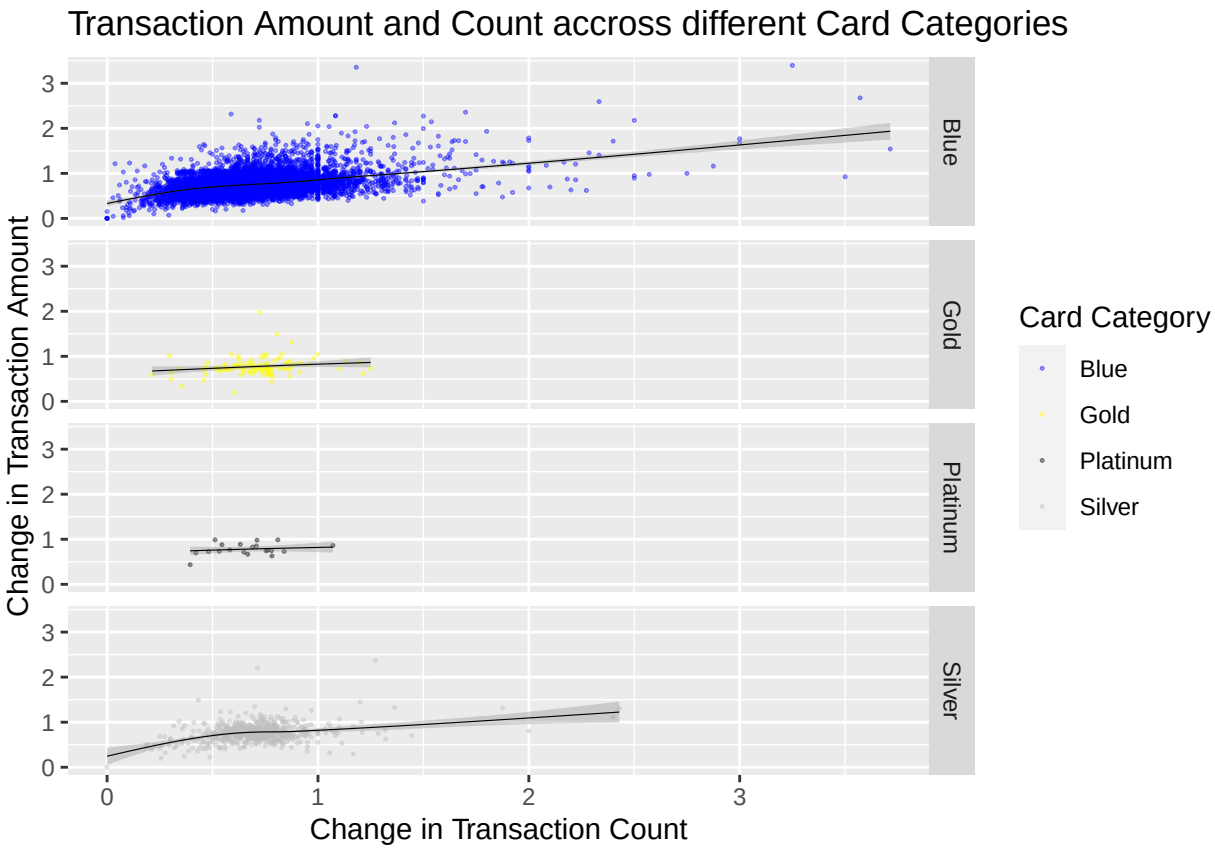

Figure 5: We see that across different cards, the transaction amount is hardly effected by the number of transactions made by the users.

5. *Summary of Analysis:*

- The total revolving balance is centered around the median and it tends to a normal distribution. Also, most people like to keep a zero-revolving balance.
- Despite having different education qualifications, the total revolving balance amount is more or less the same across different demographics. So we can say, having different educational backgrounds doesn't affect one's financial nature.
- The credit card limit and utilization ratio are in a logarithmic relation. That is the higher the credit limit, the fewer customers are using it. Also from the graph, we notice the utilization ratio peaks at around 2500 credit limit, which says, having such a credit limit would benefit the card issuing company.
- No matter what the transaction amount is, the total revolving balance stays around the median, with large points at 0 and maximum. So we can say the total revolving balance is independent of the transaction amount.
- We also notice that the transaction amount is pretty independent of the number of transactions. So users who spend more on credit cards, don't necessarily use them more than others.

6. *Conclusion*

In this project, we try to establish correlations between different variables in the dataset and how they are dependent on different demographics. We see how changing credit limit decreases the total utilization ratio, and how the transaction amount is independent of the transaction count. Though our analysis still is not completed we can conclude, having mid salary level, and lower credit card limit causes users to spend more on credit card transactions. Also, the spending is independent of most of the demographics discussed here. Now since the data has been collected from a single card company and a single city, there is obviously some bias to it. We will try to explore these biases and analyze these questions later in our dashboard.

Dashboard Link:

https://deepmalya.shinyapps.io/Credit-Card-Customers/

Tutorial Link:

https://www.youtube.com/watch?v=sqInA9XVcQo

References:

[1] Credit Card customers - SAKSHI GOYAL.
URL: https://www.kaggle.com/datasets/sakshigoyal7/credit-card-customers