ISyE 412 Fundamentals of Industrial Data Analytics

Investigating the Factors That Affect Credit Card Application

Group 7: Jai Khanna, Manavv Shah, Sanjana Parikh

Date: 12/10/2021

Table of Contents

1. Introduction

Credit cards have gone from a mere gimmick a few years back to a major part of the American's financial lives today. Currently, 70.2% of American households have at least a general purpose credit card. Before applying for a credit card, every individual considers their likelihood of being approved for one. Credit card approval decision is dependent on several factors including applicants' personal information such as gender, age, income and data such as education, marital status and number of children. The goal of the report is to analyze the dataset to draw relationships between variables to define which factors most influence the approval process.

The questions we would like to answer in this report are:

a. What is the relationship between income and a credit card application?
b. Does gender bias exist in the credit card approval process?
c. What is the impact of marital status on a credit card application?
d. What is the impact of the number of children on a credit card application?
e. How does past payment history affect a credit card application?

2. Analysis

2.1 Dataset

To investigate the factors that affect a credit card approval, the *credit card approval* dataset (*Credit Card Approval - With Target*) was analyzed, which provides various factors that are taken into account when approving credit card applications. This dataset contains 537,668 observations for 18 different variables. The variables provide insight into the various factors used during the approval process. The data set also consists of a target variable that aids in highlighting a successful application. Table 1 below states the list of the variables from the dataset along with their definition and format of response.

Table 1: Variables in the credit card approval dataset

Variable	***Definition***	***Response Format***
Gender	Gender of the applicant	M/F
Own Car	Own a car	Yes/No
Own Realty	Own a property	Yes/No
Children Count	Number of Children	Count (0,1,2…)
Total Income	Annual Income in Dollars	Numerical Value in $
Education Type	Education Level	High School/Graduate..

Family Status	Marital Status	Single/Married/Separated
Housing Type	Way of Living	With parents/House/Apt
Date of Birth	Age	MM/DD/YYYY
Days Employed	Duration of work days	Number of days
Mobile Number	Is there a Mobile Number	Yes/No
Work Phone	Is there aWork Phone	Yes/No
Phone	Is there a Phone	Yes/No
Email	Is there an email	Yes/No
Job	Job Type/Occupation	Manager/Staff/Driver
Begin Months	Record Months - Month of extracted data	
Status	Payment Status	

Table 1 highlights all the variables present in the dataset, however, out of the 18 variables a couple variables such as presence of phone and email do not significantly affect the approval process. Hence, some of the variables were omitted from the dataset during the cleaning process. The final variables used in the analysis are as follows:

1. Income
2. Gender
3. Education Type
4. Marital Status
5. Number of Children
6. Status

2.2 Softwares Used

The softwares used for this report are R, Tableau and Excel. Tableau is used for mathematical calculations, plotting graphs and other data visualizations. Microsoft Excel is used to analyze and understand the data since the dataset file is a *.csv* file i.e., credit_card approval.csv and R was used to draw up a correlation matrix to disregard any fields that have no effect on the outcome.

2.3 Approaches

The areas that are emphasized in this project are:

1. Relationship between salary and chance of approval
2. Gender discrimination between approvals

3. Minimum threshold to achieve a credit card approval
4. Relationship between marital status, education type and chance of approval
5. Effect of payment history on credit card approval

These areas will aid in analyzing the data and developing the Tableau dashboard to reach final conclusions.

3. Tableau

We used Tableau to create different graphs as an aid for visualization. For our approaches, we made use of bar graphs, box plots, combined charts and packed bubble charts. Additionally, our original dataset of 500,000 rows proved to be overwhelming for Tableau so we narrowed it down to a dataset of 100 applicants. For some approaches, we used the original dataset as a method to reinforce our results from the narrowed down database.

3.1 Relationship between Salary and an Approved Application

Upon research, annual income is considered to be one of the most important factors in the credit card approval process. We wanted to explore further into the details of what factors or threshold of income can be considered adequate for a good chance of getting approved for a credit card. We started with plotting a bar graph of each of the 100 applicant's income. This is shown in Fig 3.1A.

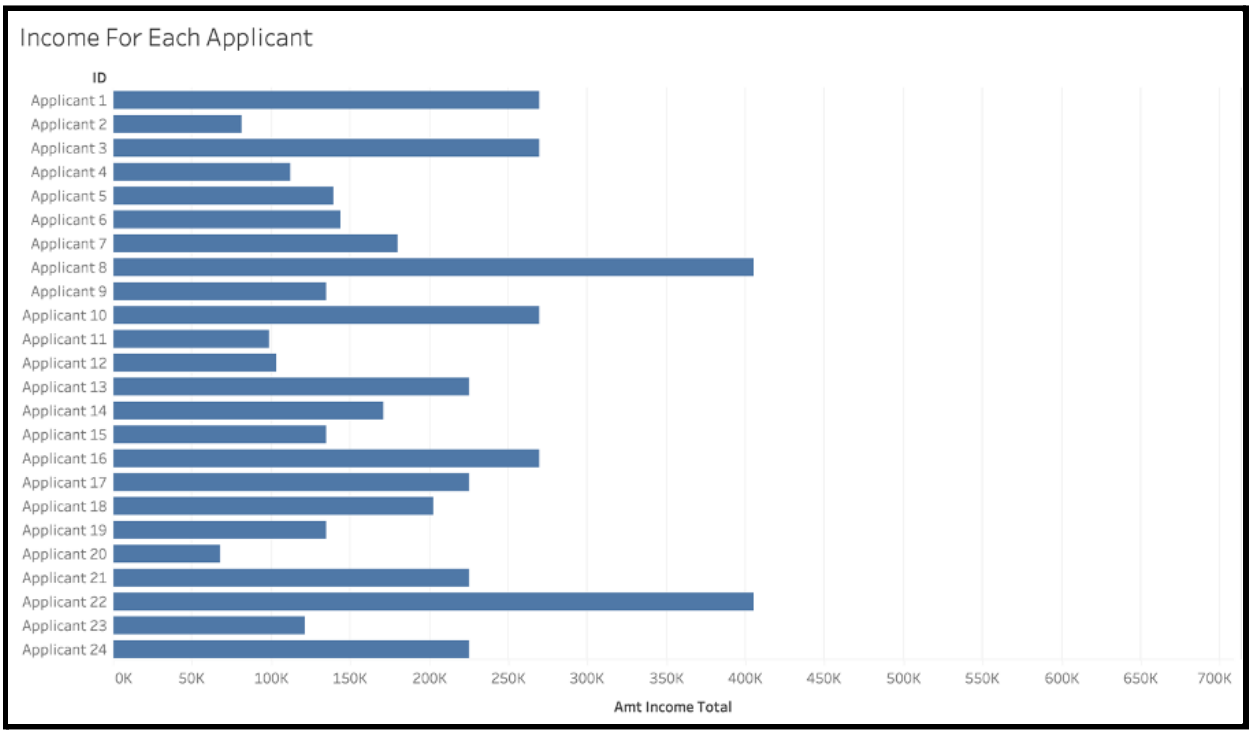

Figure 3.1A Bar Graph of the 24 Applicant's Income

Figure 3.1A gives us a good starting point on analyzing the relationship between income and an approved application. Since the bar graph of 100 applicants doesn't provide more information, our next step is to get the statistics on the thresholds on a 'good' income for an applicant. For the next analysis, we were able to use our original dataset of 500,000 rows to get a better understanding of the thresholds and range for the income of our applicants.

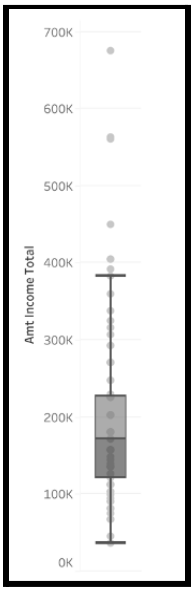

Figure 3.1B Boxplot for Income

Boxplot Statistics	
Upper Whisker	$382,500
Median	$171,000
Average	$98,045
Lower Whisker	$36,500

Figure 3.1C Boxplot Statistics

We calculated a boxplot and found the average income to be $98,045 with the lower whisker on $36,500. From our analysis we concluded that one needs at least an annual income of $36,500 to have a good chance at receiving a credit card.

3.2 Relationship between gender and approval

Gender discrimination is one of biggest challenges today. It exists in many different sectors and we wanted to check if either gender is discriminated against in the process of receiving a credit card. Based on our dataset, we calculated the number of males approved for a credit card vs the number of females approved for a credit card.

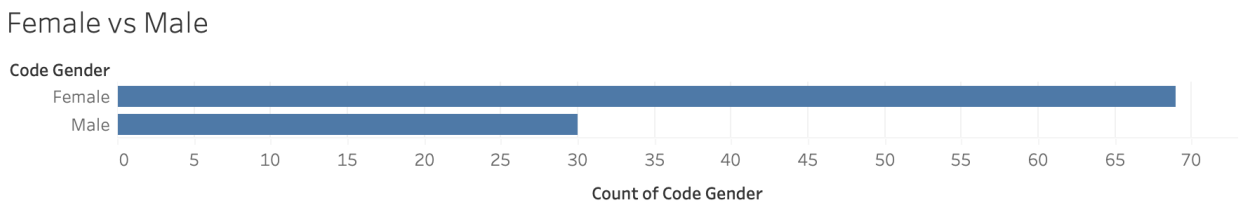

Figure 3.2A Bar Graph of Male vs Female Approved for Credit Card

As we can see from the bar graph above, 69 females were approved for a credit card compared to 30 males who were approved. However, this analysis is not sufficient. We are unable to accurately visualise if gender discrimination exists by looking at a narrowed down dataset of 100 applicants. To further strengthen our analysis, we used the full dataset of 500,000 and compared the education levels of male and female, we got the following graph:

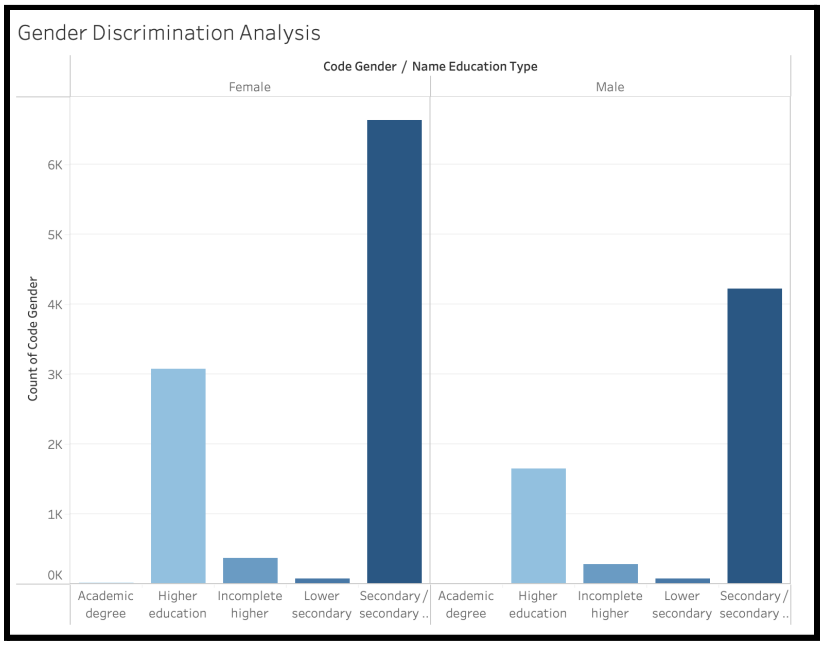

Figure 3.2B Gender Bias Graph with Education and Gender Features

The bar chart depicted above, shows us that gender discrimination does not exist in the process of a credit card application. For various education levels, females are approved equally for credit cards as males, if not more.

3.3 Relationship between marital status, number of children and approval

Another challenge we wanted to test was whether the marital status and number of children affect a credit application. We conducted a packed bubble chart and can conclude that the primary factor is the number of children. The more children you have, the less likely you are to get a credit card application. This could be because for those candidates, their income may not be adequate to support the increase in costs that comes with more children. The marital status plays a secondary role as applicants who were married demonstrated a higher chance of getting a credit card approved.

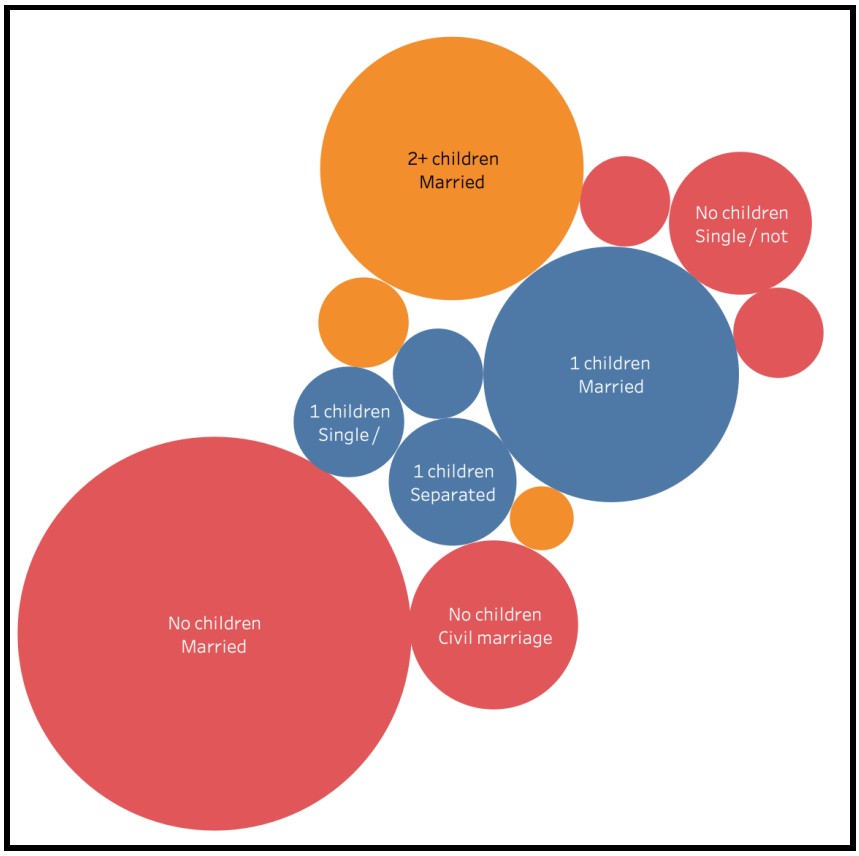

Figure 3.3A Packed Bubble Graph of Marital Status and No. of Children

We have an orange bubble at the top of Figure 3.3A which shows that a sizable number of applicants that have more than 2 children were successful in getting a credit card. As this contradicts our initial analysis, we calculated the income of those specific applicants and found them to be well above the average income. Therefore, we can conclude that our initial assumption was correct and the orange bubble could be considered an anomalous result for this approach as a higher income takes precedence in a credit card application.

3.4 Effect of payments due on credit card application

When applying for a credit card, an applicant's past history of payments is a vital source of information that the credit card company can use. Information such as loan repayments, mortgage, current assets, rental payments and many more can indicate how suitable an applicant is to pay off their credit card debt. Our dataset consists of a 'Status' feature which outlines different categories of late payments and how they affect the data. There are 8 categories according to Figure 3.4A that the credit card issuer can place you under. Category 0 means that payments are 1-29 days past due, 1 means 30-59 days past due, 2 means 60-89 days past due, 3 means 90-119 days past due, 4 means 120-149 days past due, 5 means bad debts, C means payments were paid off that month, X means no loans were taken that month.

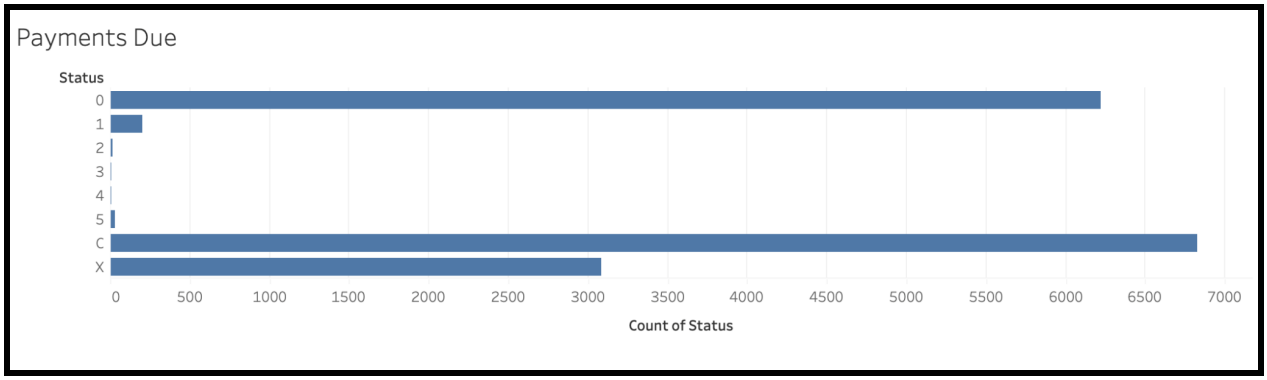

Figure 3.4A Bar Graph of Late Payments

We can interpret from this graph that the majority of the applicants who were approved were under the 0, C or X category and the rest were close to 0. The number of applicants in each category were:

- Category 0: 6,222
- Category 1: 199
- Category 2: 17
- Category 3: 4
- Category 4: 4
- Category: 29
- Category C: 6,828
- Category X: 3,080

As we can see, out of 16,354 applicants whose past payments were accessible, 98.6% of the applicants fall under the categories 0, C or X. This analysis provides conclusive evidence about the importance of past loan repayments. In order for an applicant to be approved for a credit card, they should have no outstanding payments for more than 30 days. An outstanding payment of more than 30 days may indicate that the applicant has no means of paying the loan back which will deem them unsuitable for a credit card.

4. Results and Conclusion

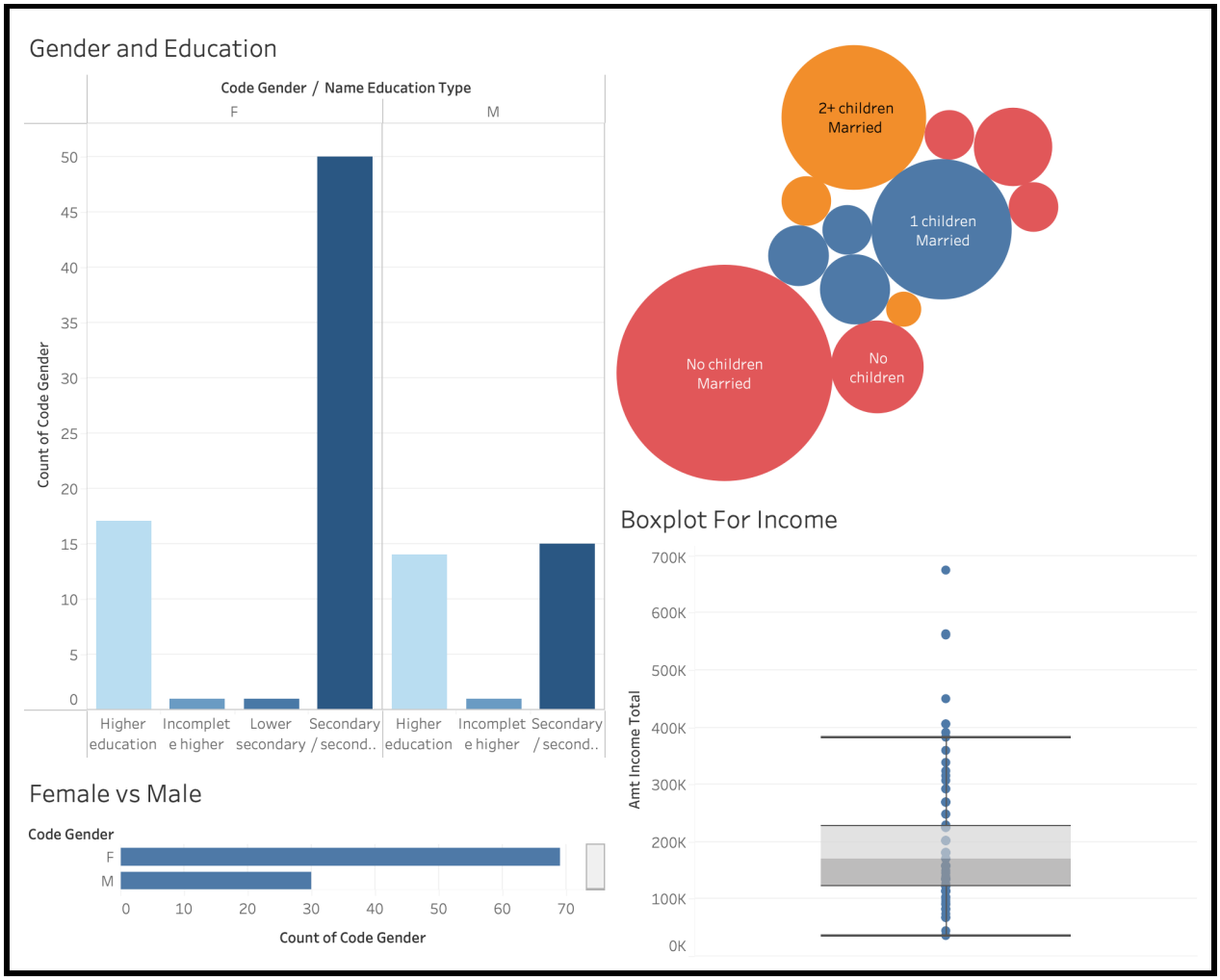

Figure 4 Snapshot of Tableau Dashboard

Figure 4 shows our final dashboard which consists of all the graphs described in the report. Since our dataset was very large, we first started with using a correlation matrix or a heat map, by using R, to figure out which variables do not affect the credit card application. This led us to remove mobile number, phone and email fields from consideration. Next, we utilized a dataset of 100 applicants as it proved to be more efficient and translated well into Tableau. This allowed us to utilize our initial dataset of 537,668 observations when we needed to bolster our initial assumptions with a more robust dataset. From our approaches, we can conclude that there are various factors that play an essential role in the credit card approval process. In our Approach 3.1, we investigated the relationship between income and credit card approval, we saw a strong correlation between the 2 variables. We reinforced our initial hypothesis by using a larger dataset to show us a complete overview of the income statistics. We concluded that at least an annual income of $36,500 is required to stand a good chance at receiving a credit card application. Next,

in Approach 3.2, we tackled one of the main issues faced in today's world - gender discrimination in the credit card approval process. First, we calculated the number of men and the number of women who receive a credit card from our dataset of 100 applicants. This gave us a foundation to dive deeper into our analysis. Next, we used our larger dataset to figure if the education level of men and women play a role in getting a credit card. We plotted a gender discrimiantion graph (Fig 3.2B) which proved that there is no discrimination as women and men are equally likely to get a credit card despite of their education status. Finally, in Approach 3.3, we checked if there is a relationship between the marital status and the number of children, against the chance of getting a credit card. As depicted in Fig 3.3A, we saw the number of children and the marital status play a crucial role in the credit card approval process. The packed bubble graph showed us that as the number of children increases there is a relatively lower chance of getting a credit card. This could be interpreted as a liability because the applicants in our dataset may not be financially stable to support the increasing costs. Additionally, married applicants were more likely to get a credit card compared to other types of marital status when the number of children were kept constant.

5. References

- https://budgeting.thenest.com/percentage-americans-credit-cards-30856.html