Lending Club Case Study

Ayush Mehta

Abhishek Balajisingh
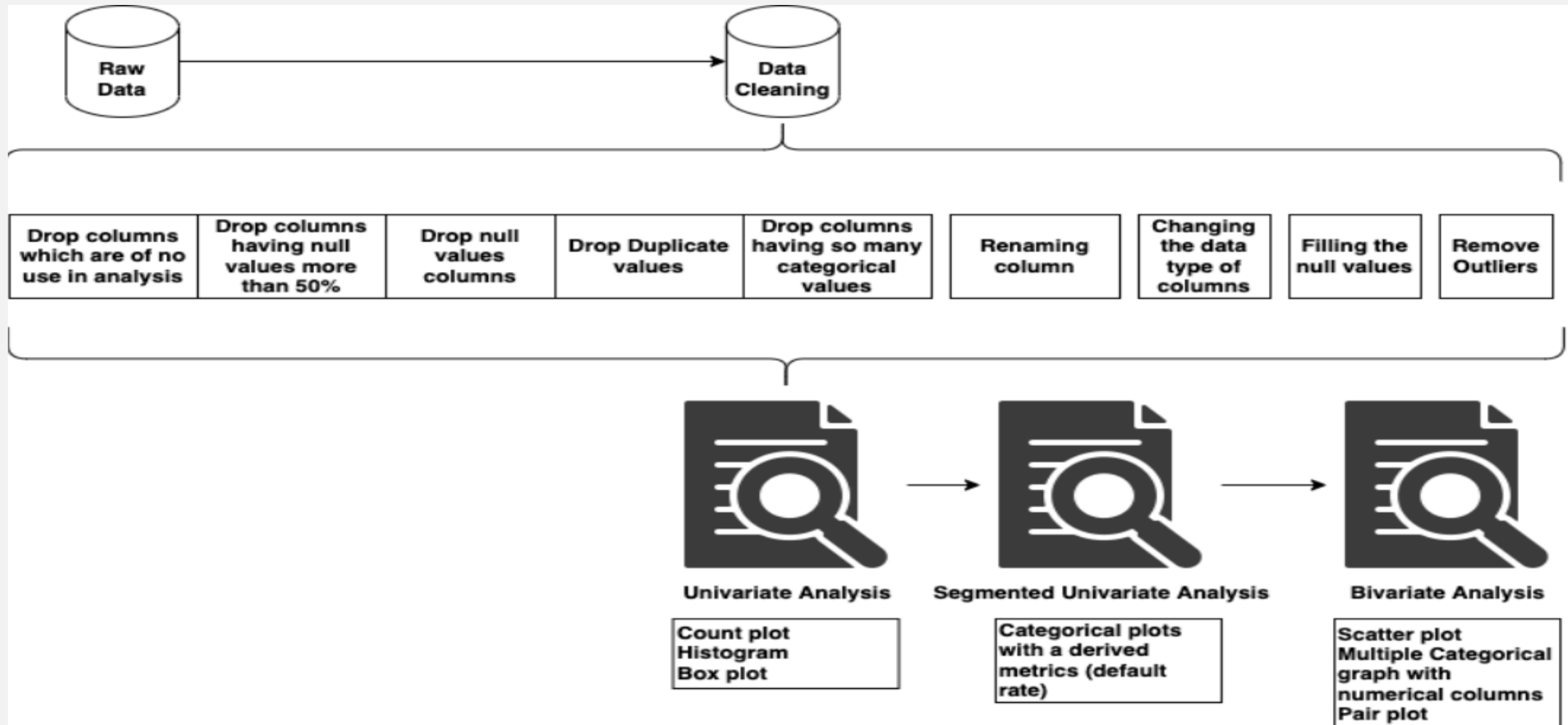
Lending Club Case Study

.

- A consumer finance company that specializes in providing urban customers with various forms of loans. When a company receives a loan application, it must decide whether or not to approve the loan based on the applicant's profile. The bank's decision is related with two sorts of risks:

- If the applicant is likely to repay the loan, the company will lose business if the loan is not approved.

- Approving the loan may result in a financial loss for the company if the applicant is unlikely to repay the loan, i.e. if he or she is likely to default.

- The data pertains to previous loan applicants and whether or not they 'defaulted.' The goal is to find patterns that can be used to make decisions.

Business Problem

.

- This firm is the world's largest online loan marketplace, allowing personal loans, commercial loans, and medical procedure funding. Through a quick internet interface, borrowers can readily acquire cheaper interest rate loans.

- Lending to 'risky' applicants, like most other lending organisations, is the most common source of financial loss (called credit loss). The amount of money lost by the lender when a borrower refuses to pay or flees with the money owed is referred to as credit loss. In other words, defaulting borrowers do the most financial harm to lenders. The 'defaulters' are the consumers who have been labelled as 'charged-off.'

- If these problematic loan applicants can be identified, the size of the loan can be reduced, reducing the amount of credit loss.

- The goal of this case study is to identify such applications using EDA.

- In other words, the organization needs to know the reasons (or driver variables) that cause loan default, i.e. the variables that are significant predictors of default. This knowledge can be used to the company's portfolio and risk assessment.

- To gain a better understanding of the field, you should perform some independent study on risk analytics.

Process followed

.

Drop columns which are of no use in analysis	Drop columns having null values more than 50%	Drop null values columns	Drop Duplicate values	Drop columns having so many categorical values	Renaming column	Changing the data type of columns	Filling the null values	Remove Outliers

Univariate Analysis

Count plot
Histogram
Box plot

Segmented Univariate Analysis

Categorical plots with a derived metrics (default rate)

Bivariate Analysis

Scatter plot
Multiple Categorical graph with numerical columns
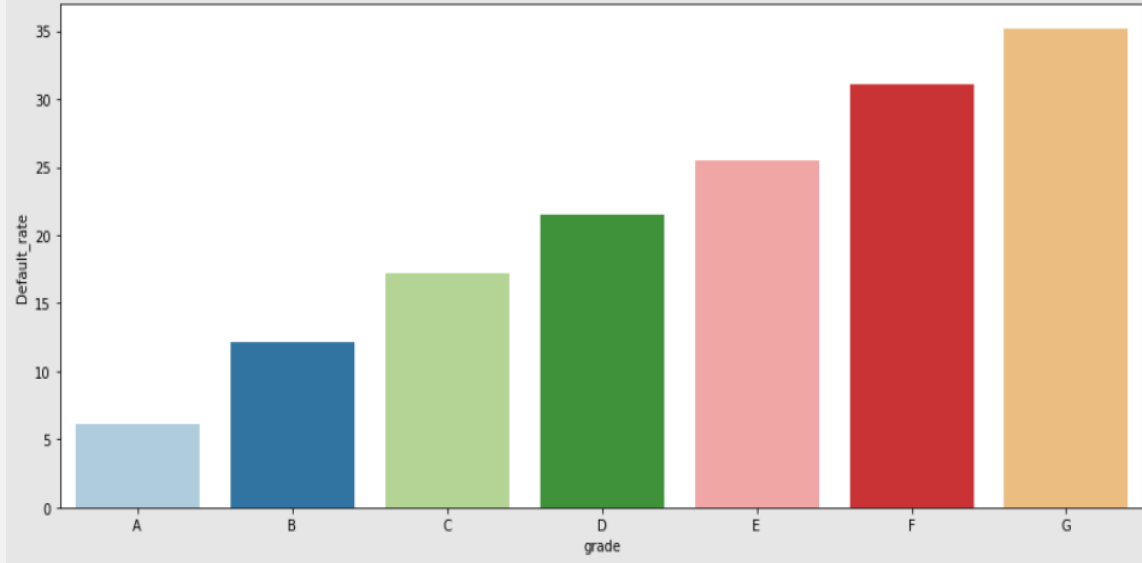Pair plot

Univariate & Segmented Analysis

Defaulters

.

Max Count:The most number of defaulters count are from 36 months term

- Max Defaulters Rate: But the defaulters are most likely to be the one from 60 months term as the average rate of defaulters is the highest for this 22%.¶
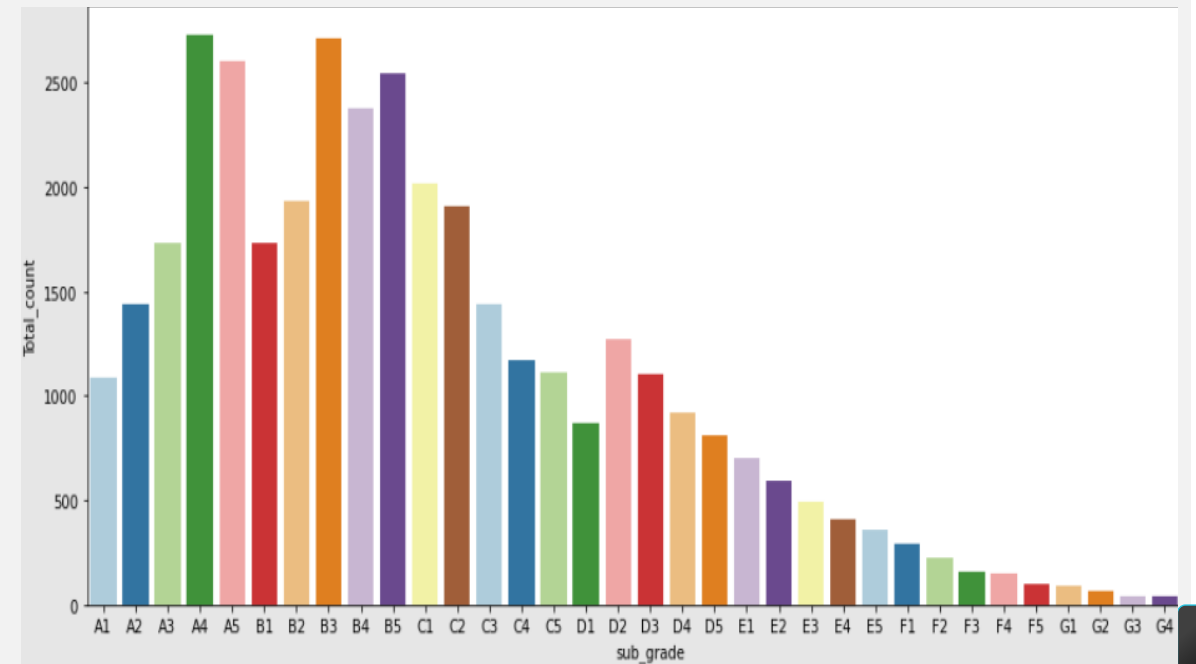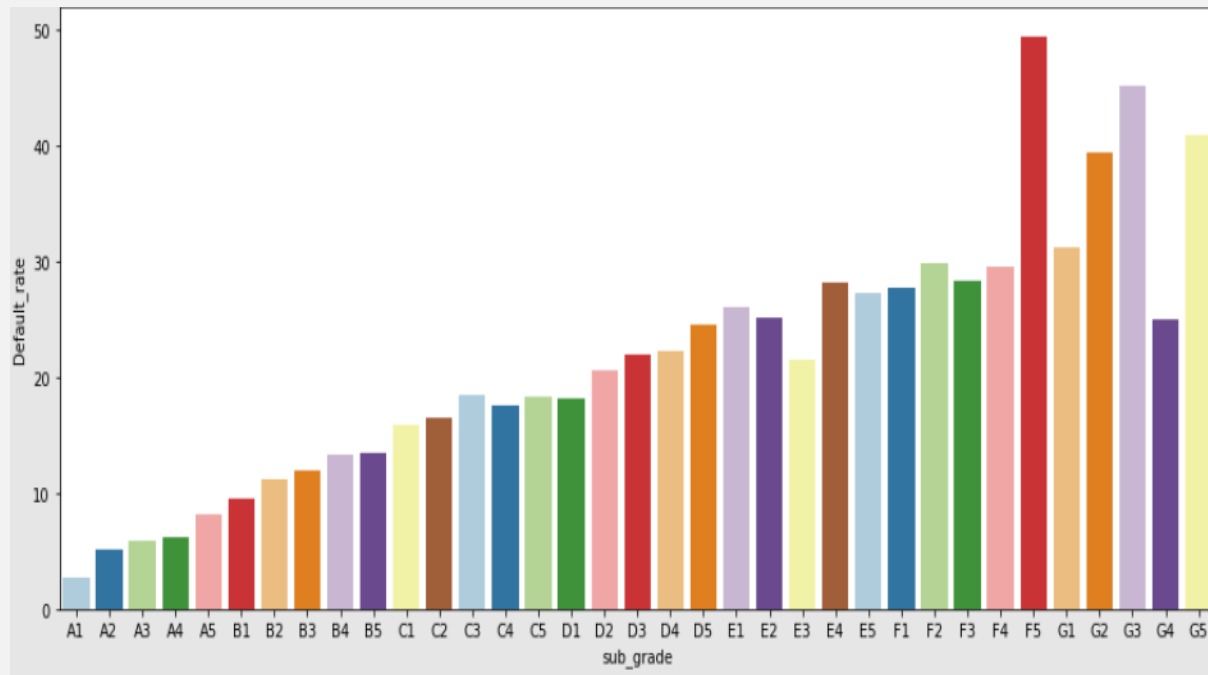
Grades

.
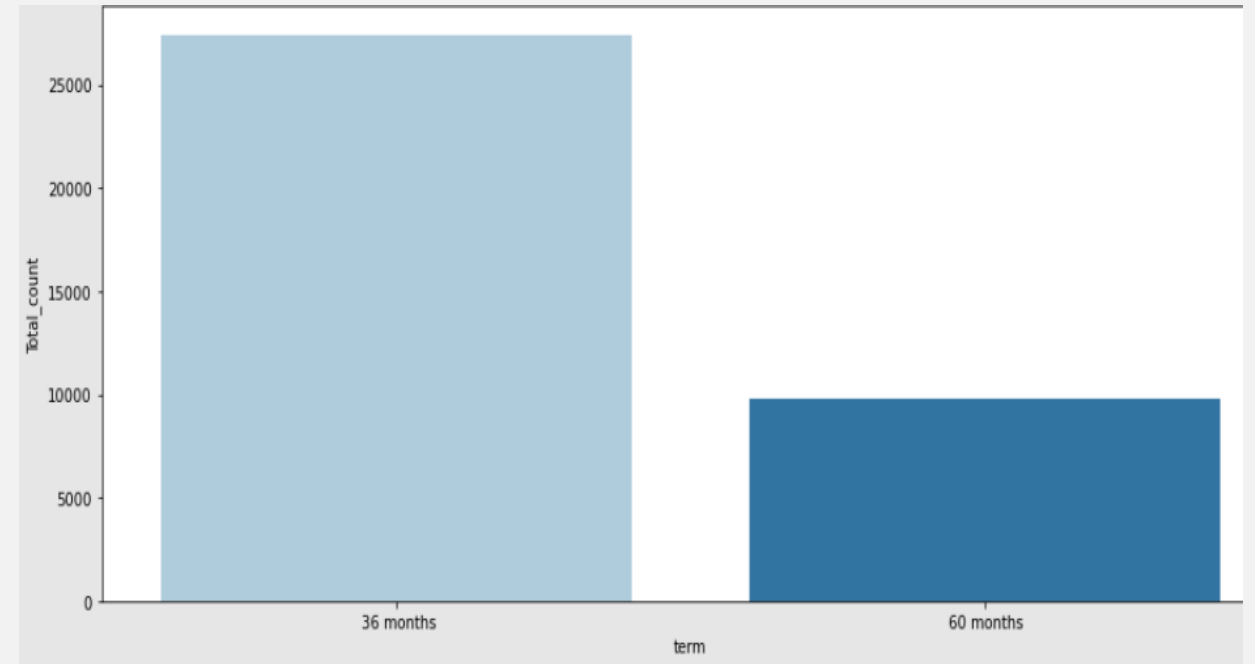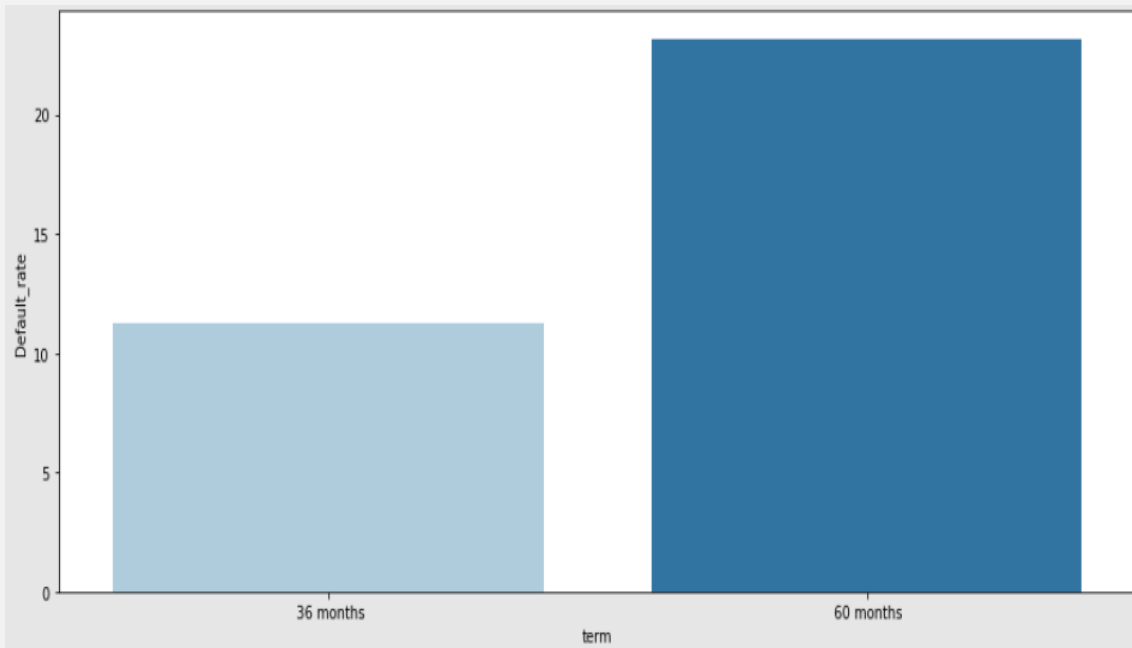
- Max Count: The most number of defaulters count are from grade B.

- Max Defaulters Rate: But the defaulters are most likely to be the one from grade G and F as the average rate of defaulters is the highest for this 31%. The rate for F is less if we exclude sub grade F5 because of which the F grade is contributing 30%.
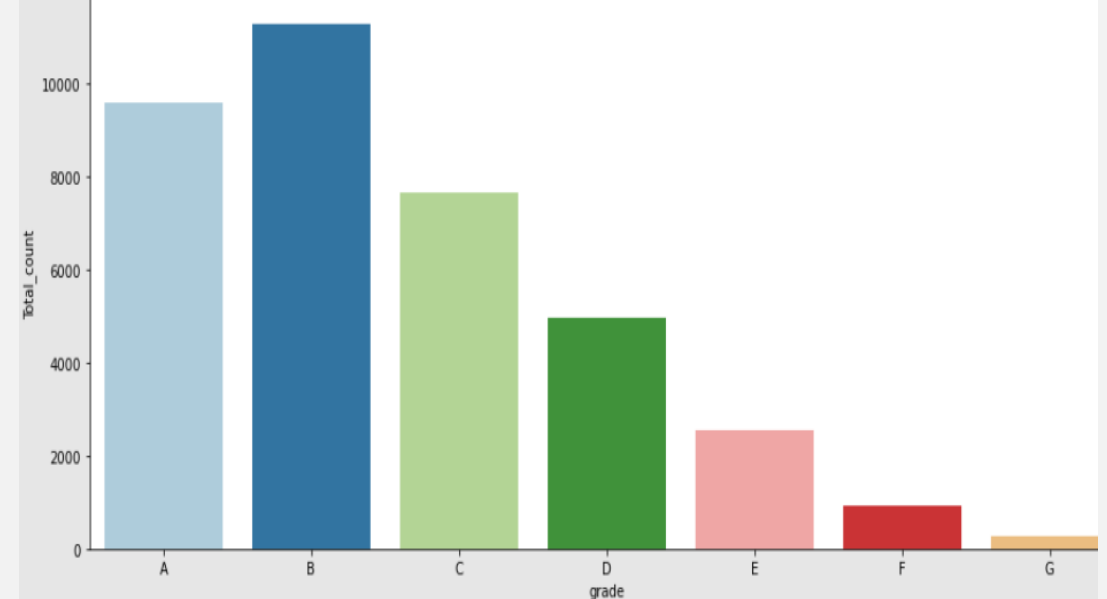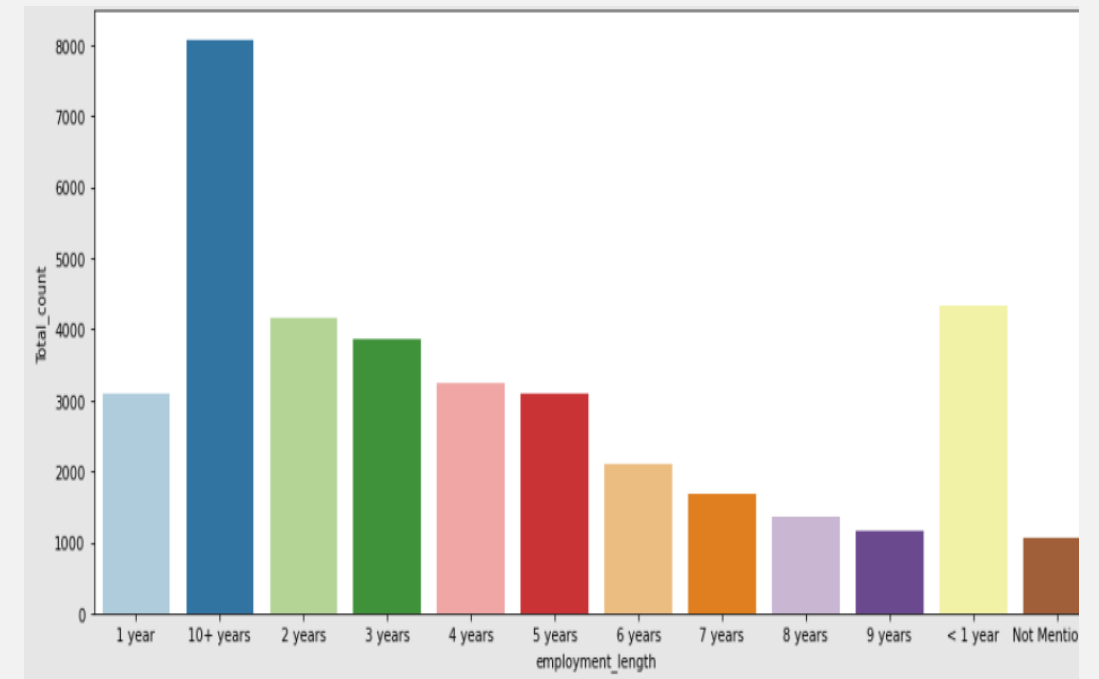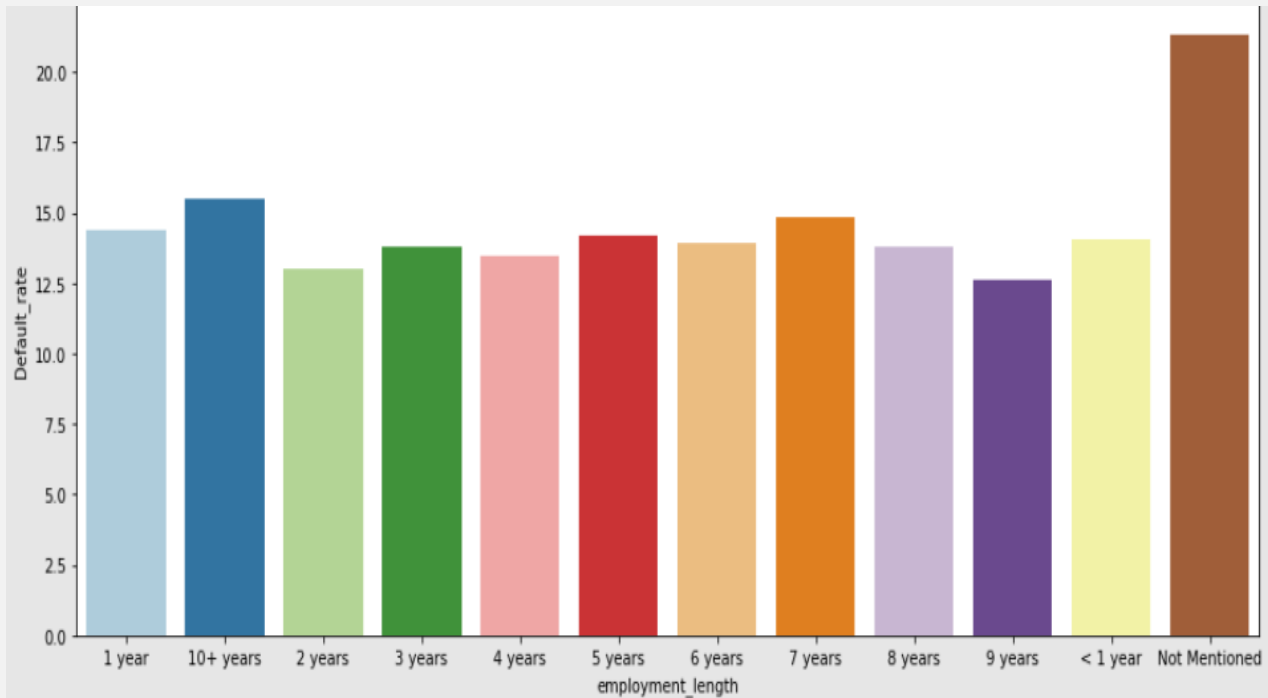
Sub-Grades

.

- Max Count: The most number of defaulters count are from Sub grade A4.

- Max Defaulters Rate: But the defulters are most likely to be the one from Sub grade F5 as the average rate of defaulters is the highest for this 45%.¶
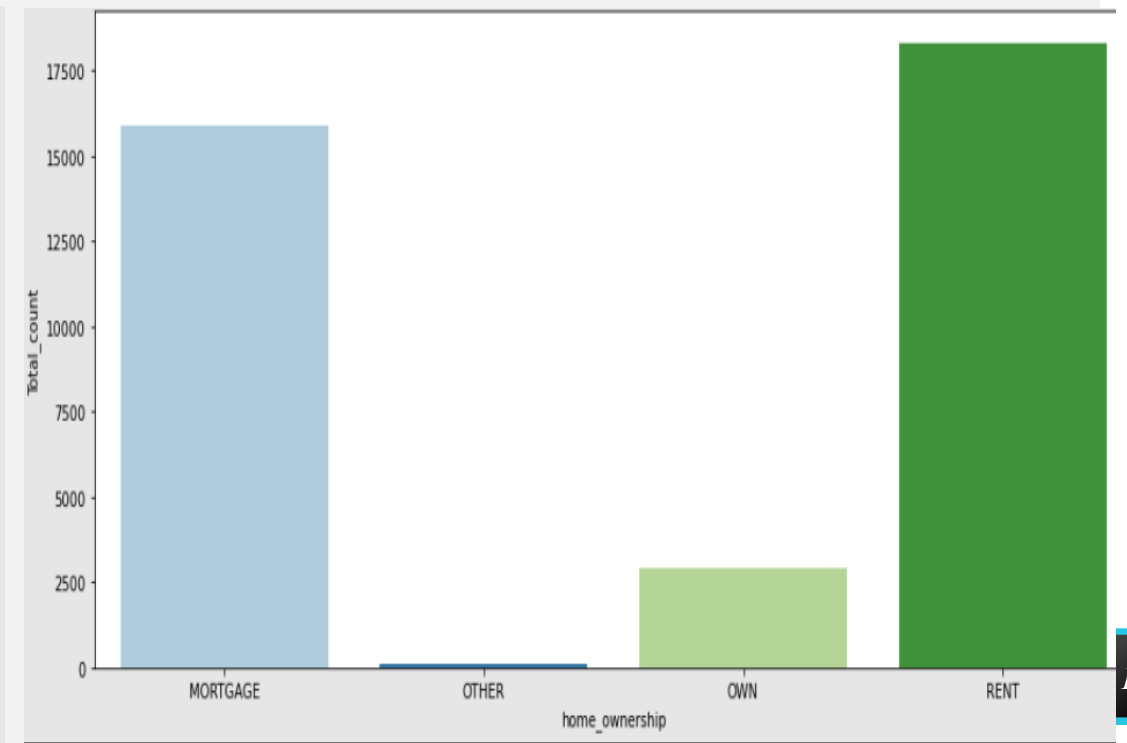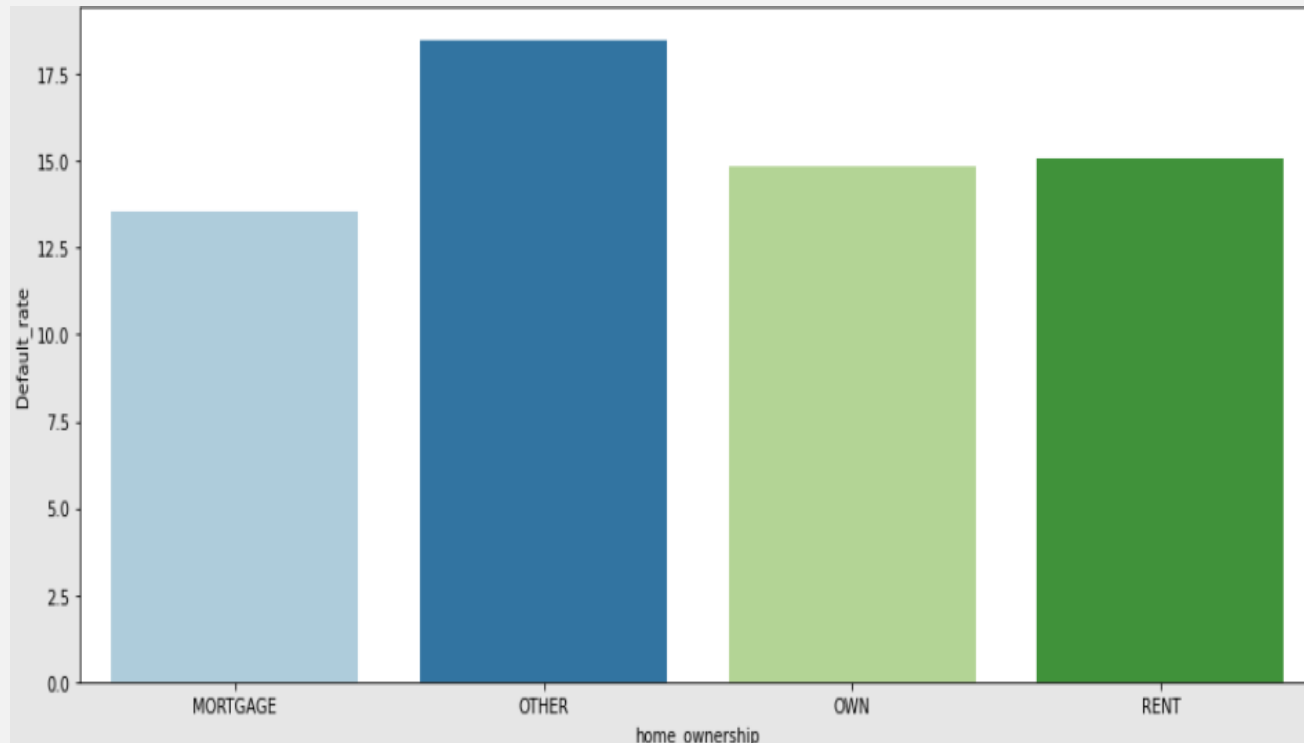
Employment Length

.

- Max Count: The most number of defaulters count are one having the length 10+ years.

- Max Defaulters Rate: But the defulters are most likely to be the one who haven't mentioned there employment length as the average rate of defaulters is the highest for this 21%.¶
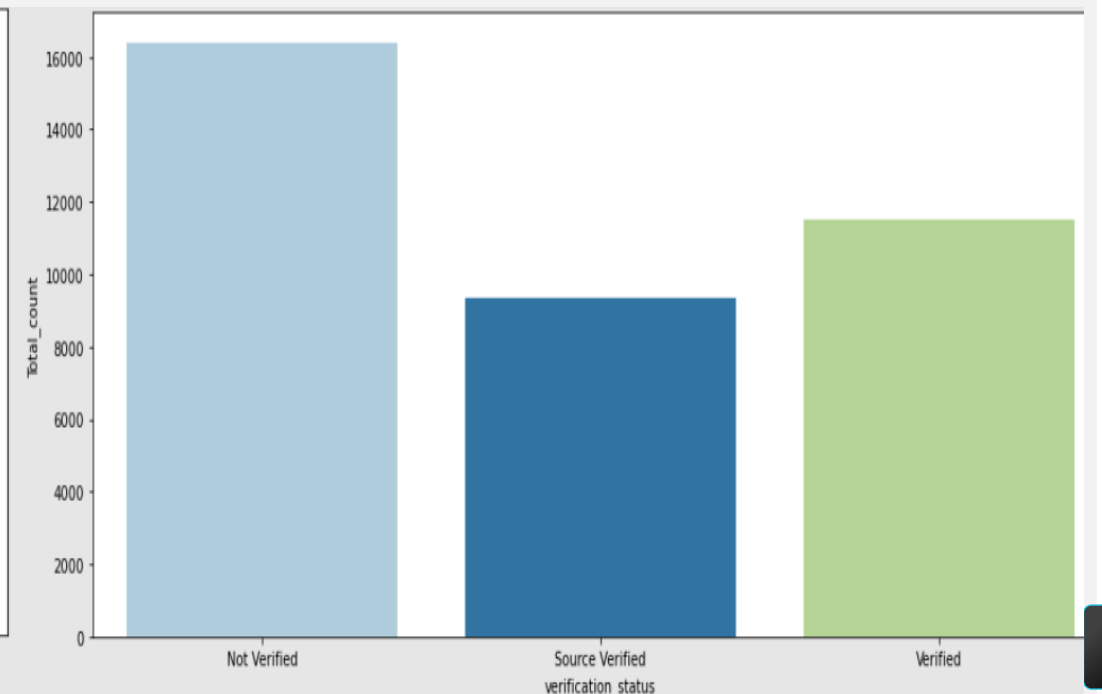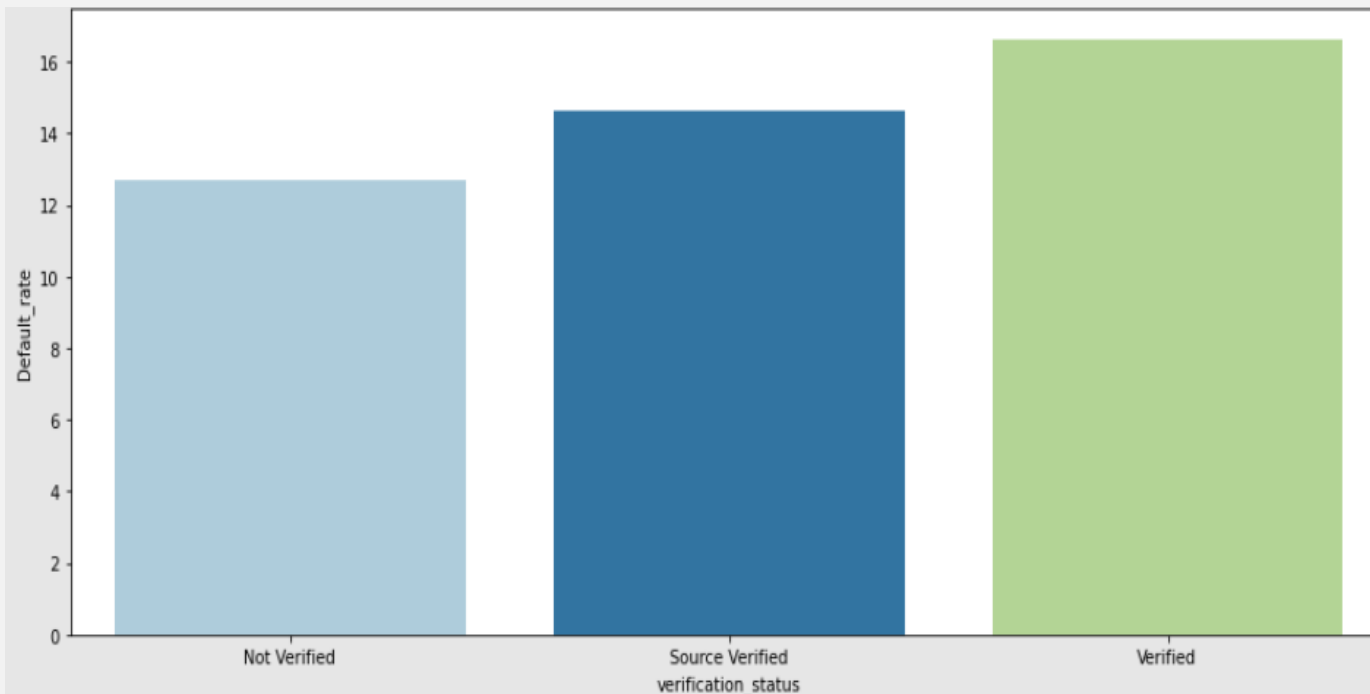
Home Ownership

.

Max Count: The most number of defaulters count are from the one who lives on rent.

Max Defaulters Rate: But the defaulters are most likely to be the one who mentioned other in there home ownership column as the average rate of defaulters is the highest for this 18%.¶
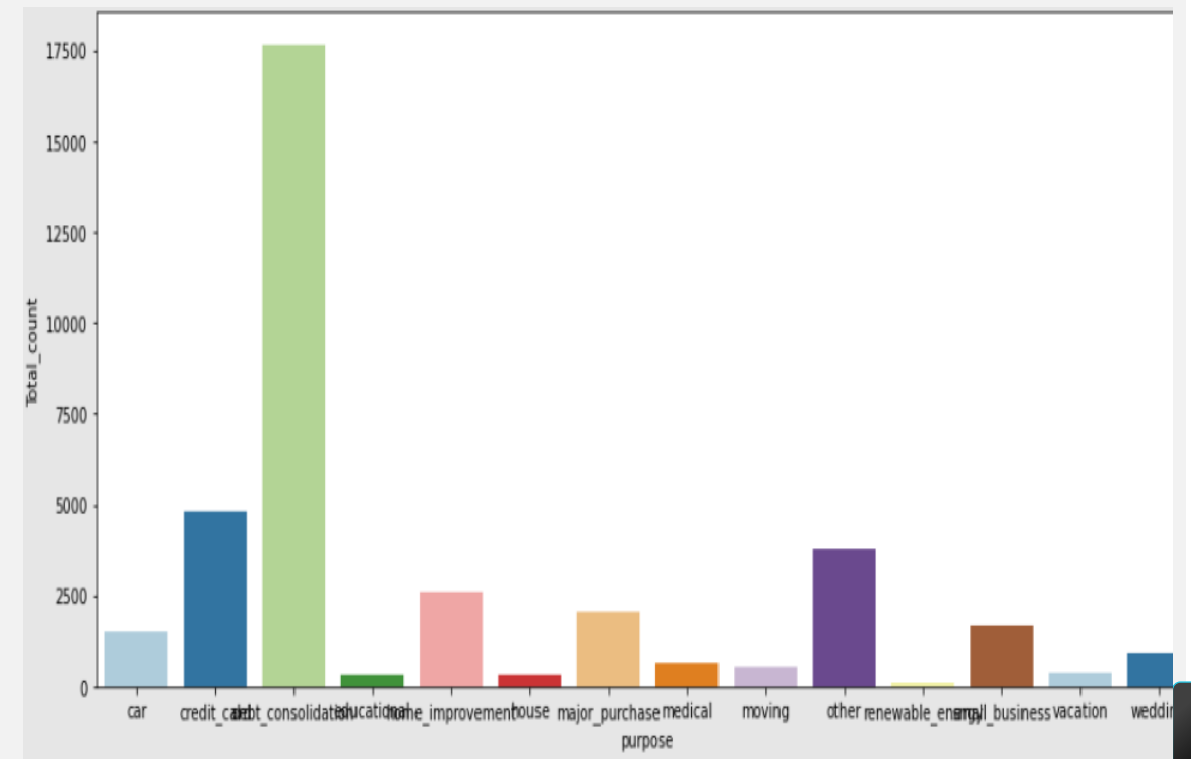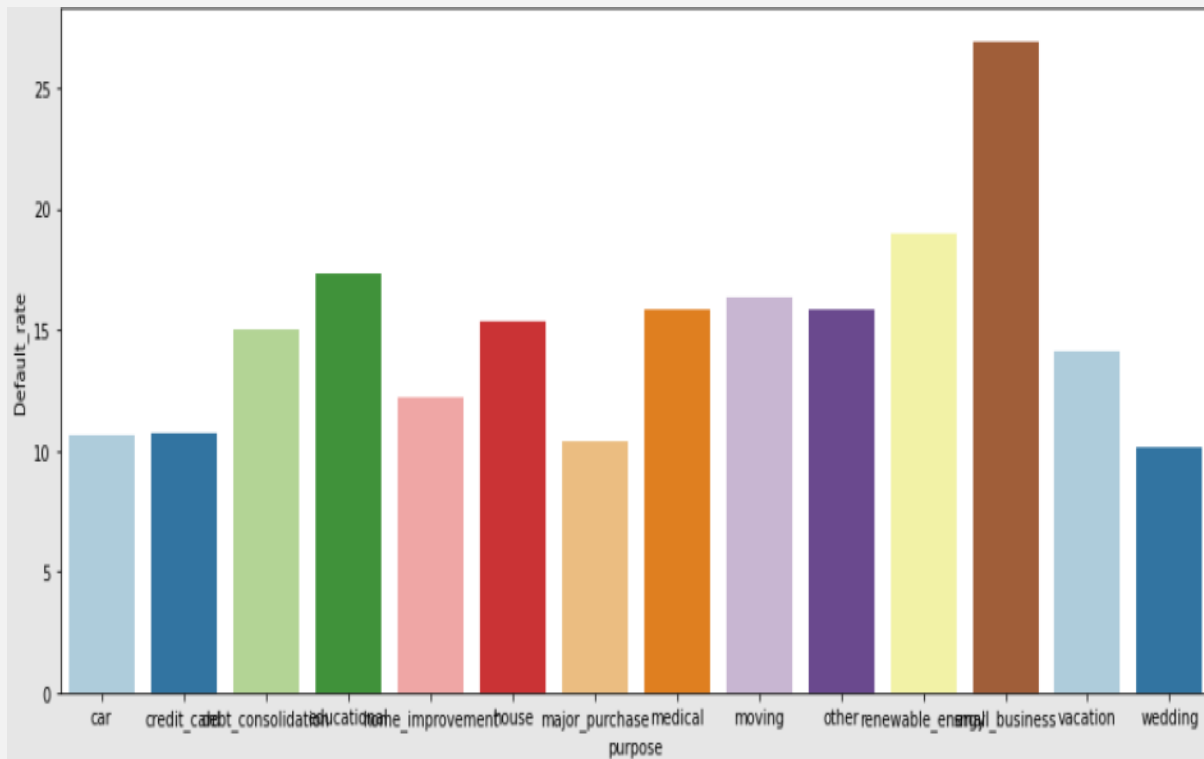
Verification Status

.

- Max Count: The most number of defaulters count are from the one who are not verified.

- Max Defaulters Rate: But the defaulters are most likely to be the one from the one who are verified as the average rate of defaulters is the highest for this 16%. Cannot comment much on it.¶
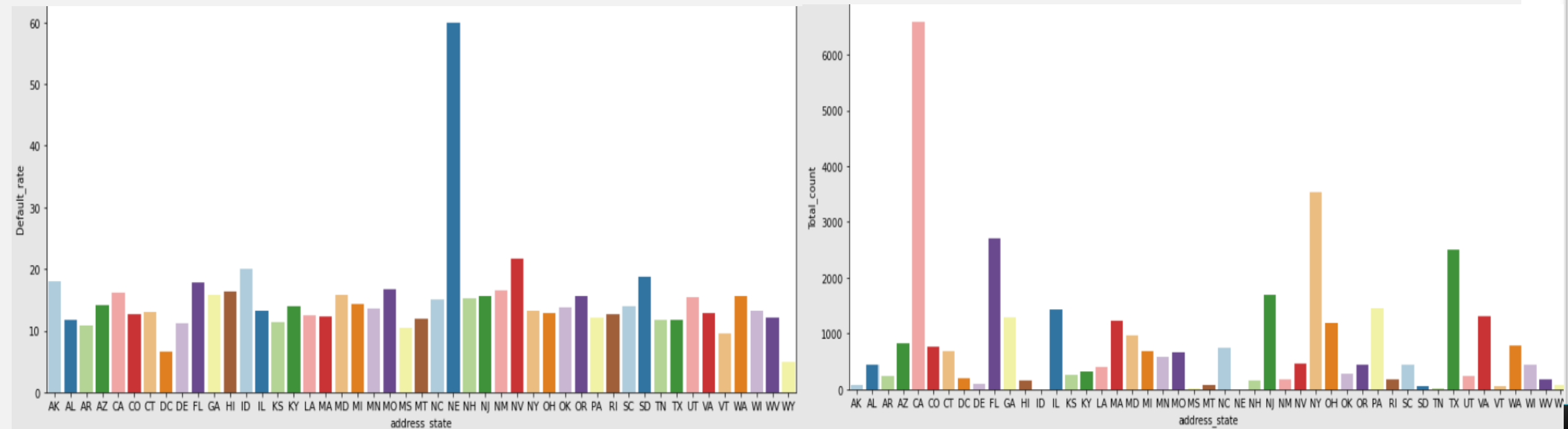
Purpose

.

- Max Count: The most number of defaulters count are the one who had taken loan for debt consolidation.

- Max Defaulters Rate: But the defaulters are most likely to be the one who had taken loan for small businessess as the average rate of defaulters is the highest for this range 25%.

Address State

.

- Max Count: The most number of defaulters count are from state CA.

- Max Defaulters Rate: But the defaulters are most likely to be the one from state NE as the average rate of defaulters is the highest for this range 60%. This is a big number we have to take this into a serious consideration.¶
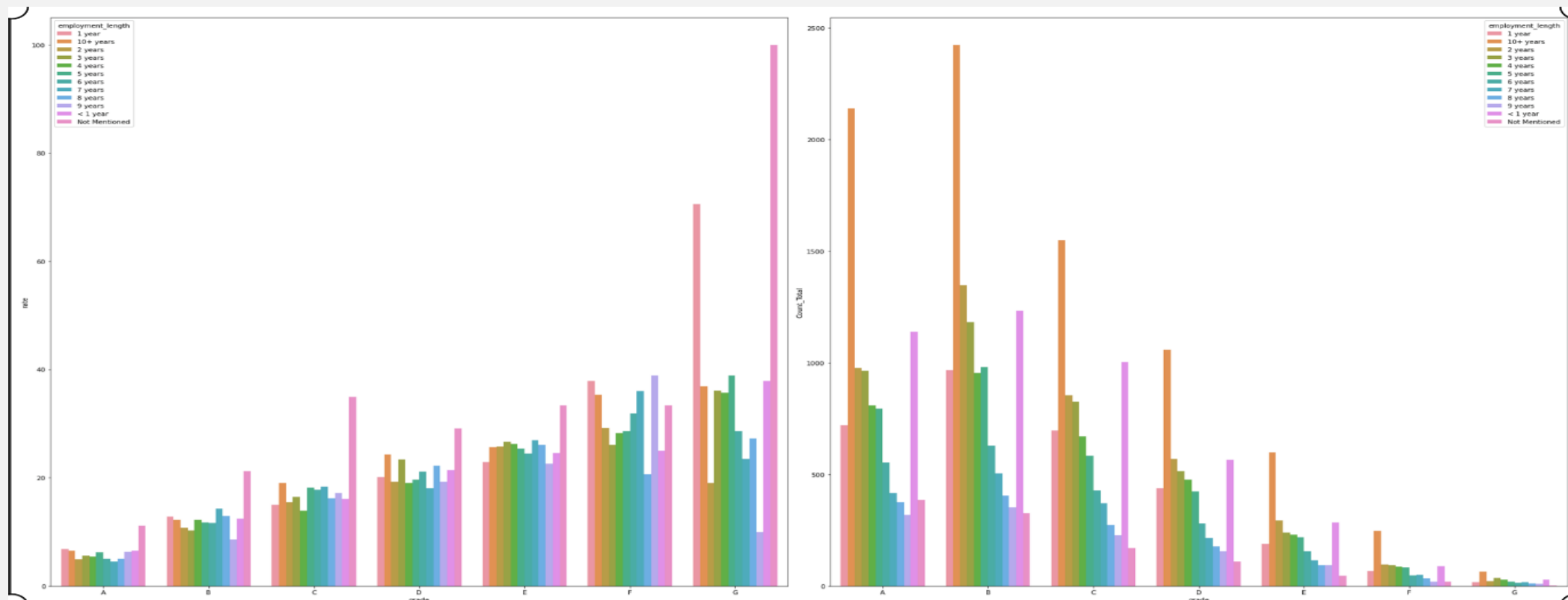
Bivariate Analysis

Employment length

.

- A grade people are the highest who haven't mentioned there employment length.
- There are 100% defaulters who haven't mentioned there employment length

Grade and purpose

.

Highest default rate is from the one who had taken loan for opening small businesses and mostly high grade people take loan for this purposes. Mediocre grade people take loan for purposes like education and house.

16
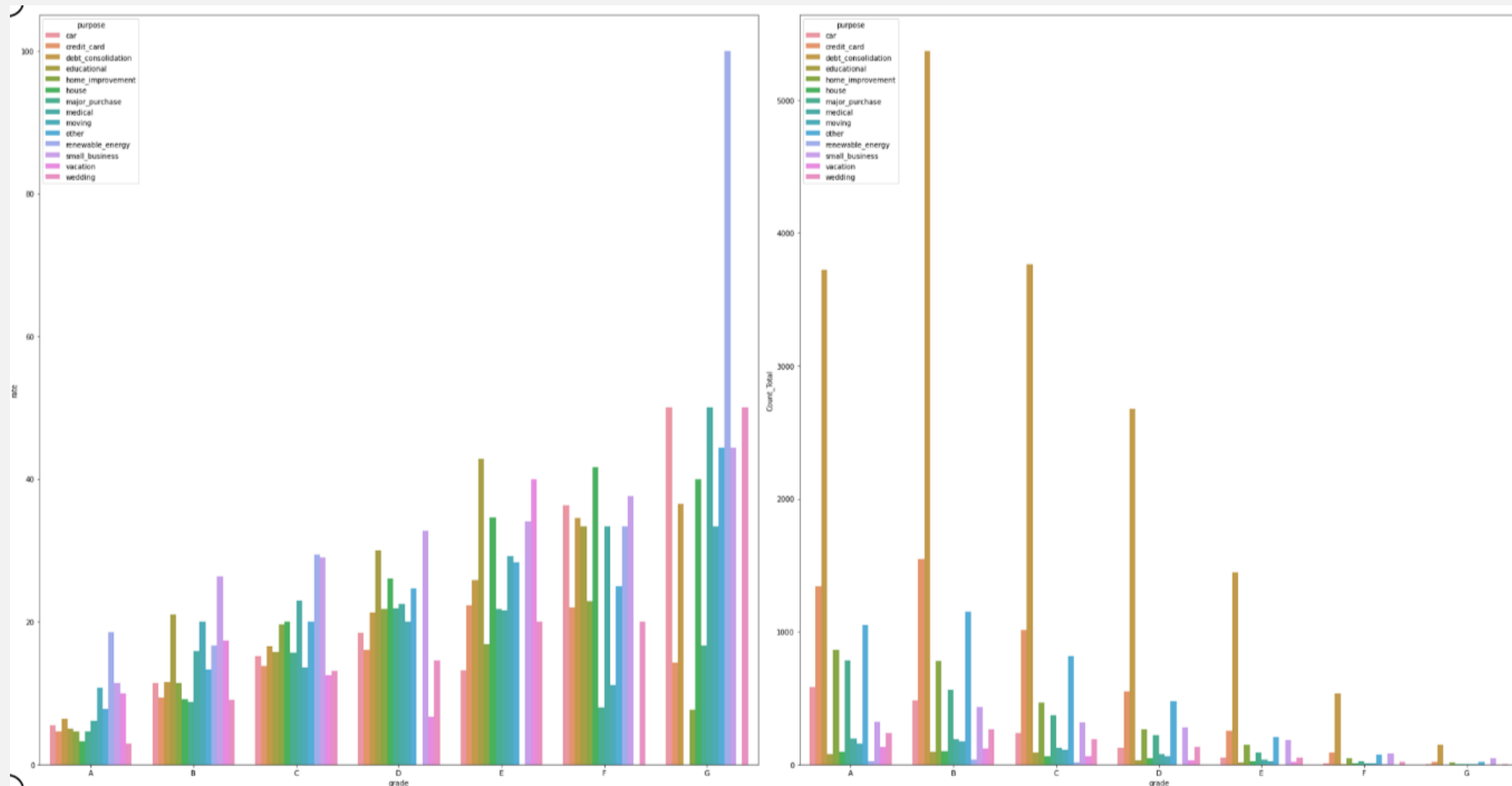
Address state & Home ownership

PA, MO, AR are major states where people had filled home ownership = other and from our analysis we know that the one who had filled home ownership as other are most likely to be defaulters.
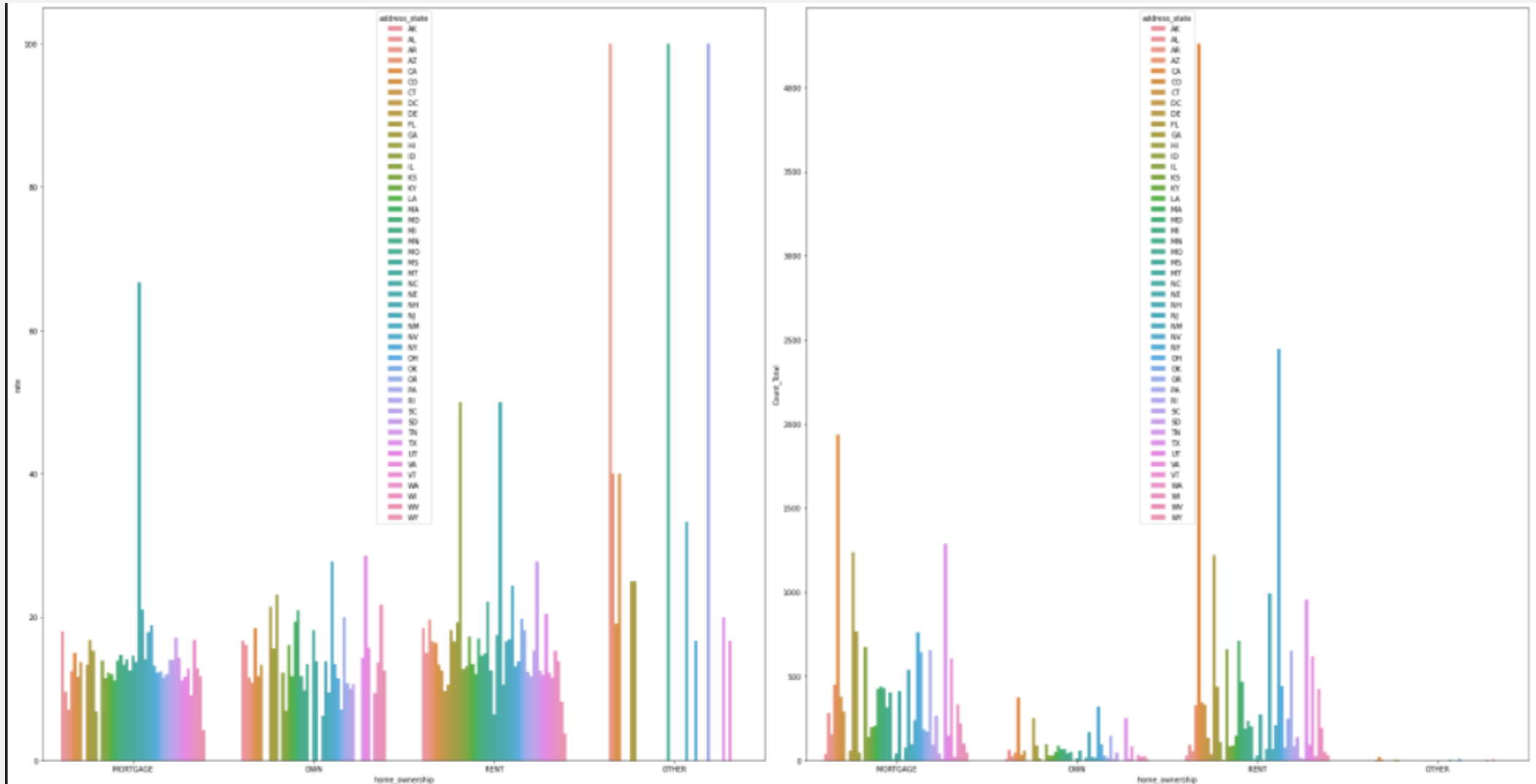
Interest rate & grade

.

It has been seen that loan are provided at high interests to low grade people as compared to the high grade people.

Issue date & Grade

The rate of defaulters for grade A has been decreased in a span from 2007 to 2011 and the number of A grade people taking loan has been increased.

Recommendations

Recommendations

.

- Interest rate increases as per the grade. May be this is one of the reason for the increase in the defaulters in the grade G as the interest rate is highest for this grade 20-25%. We need to reconsider the rate of interest according to some different factors not by the grades.

- 2. There should be through analysis of the people who are applying loans for opening small business. The background check is usually be done but we also need to do the detail analysis of the business plans which the one is asking loan for because from our analysis mostly defaulters belongs to this category and from our bivariate analysis we can clearly see that mostly high grade people apply for this type of loan so majority there background is great but not there plans for the business which ultimately leads to the loss.

- 3. It should be strict to mention there employment length as most of the defaulters are coming from the one who haven't mentioned there employment length and also a detail background check of there employability as there can be cases where the show fake employment.

- 4. Either stop or be very careful before taking any application from the state NE, it has the highest defaulters rate and all apply for only 36 months terms.

Interesting Facts

Interesting Facts

1. For sub grade A1 who has taken loan for 60 months terms are least likely to be defaulters with 0% defaulters.

2. For sub grade G3 who has taken loan for 36 months terms are most likely to be defaulters with 100% defaulters.

3. For grade G who has not mentioned there employment length are most likely to be defaulters with 100% defaulters.¶

ThankYou

👤 Ayush Mehta and Abhishek Balajisingh