

# **Lending Club Case Study**

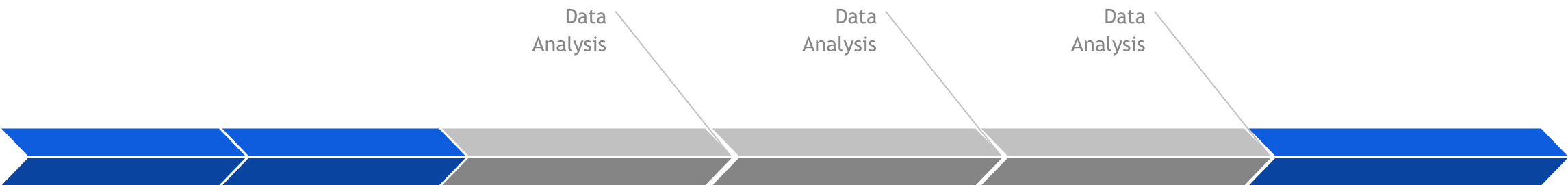
Group Members:

Raghuram Hindupur

Pallabi Paul

## **ABSTRACT**

- Lending club is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures.
- Borrowers can easily access lower interest rate loans through a fast online interface.
- The objective of analysis is to use the information about past loan applicants and find whether they 'defaulted' or not.



### Data Cleaning

Removing the null valued columns, unnecessary variables and checking the null value percentage and removing the respective rows.

### Data Understanding

Working with the Data Dictionary and getting knowledge of all the columns and their domain specific uses

### Univariate Analysis

Analysing each column, plotting the distributions of each column.

### SegmentedUnivariate Analysis

Analysing the continuous data columns with respect to the categorical column

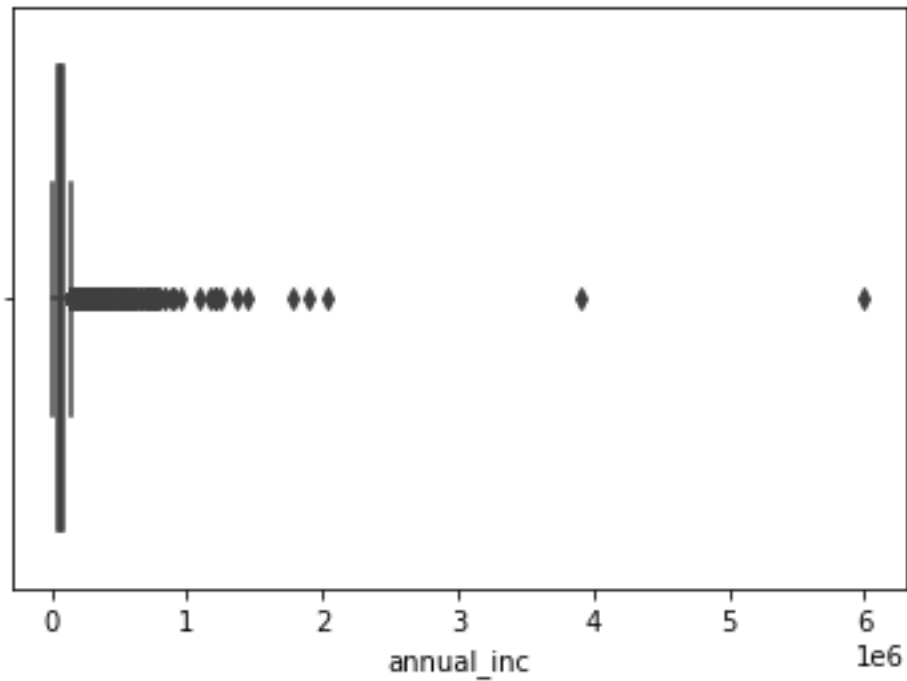
### Bivariate Analysis

Analysing the two variable behaviour like term and loan status with respect to loan amount.

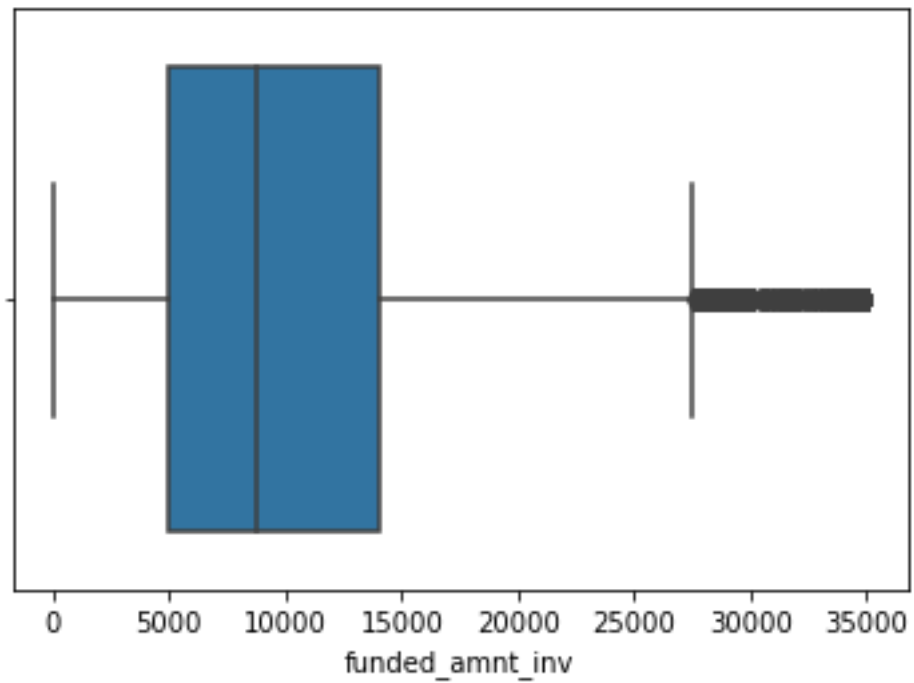
### Recommendations

Analysing all plots and recommendations for reducing the loss of business by detecting columns best which contribute to loan defaulters.

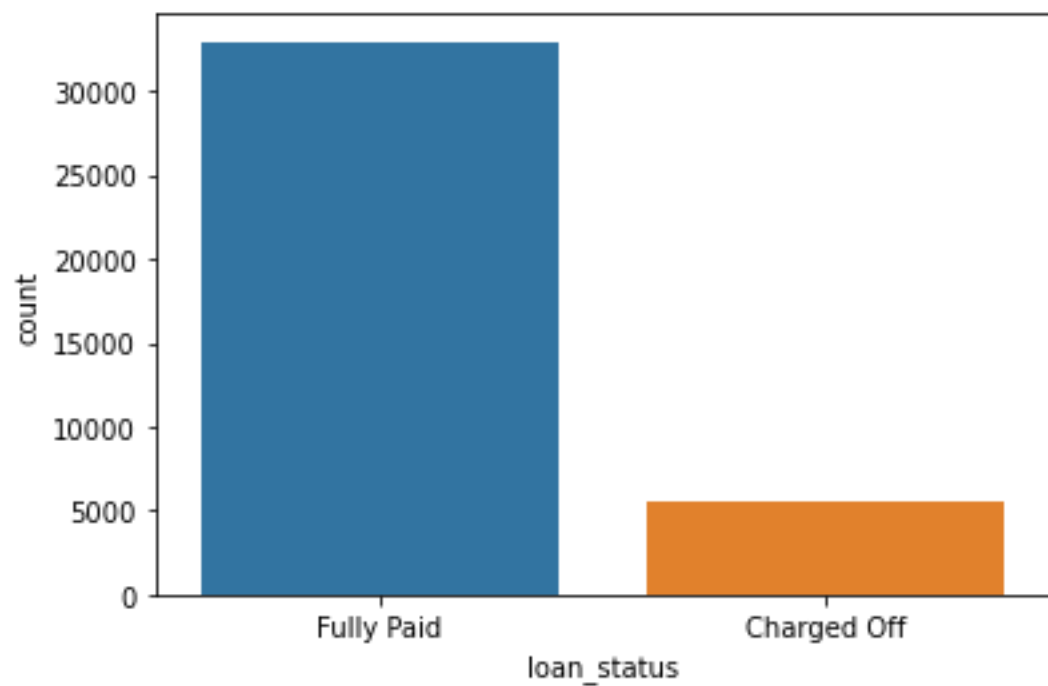
# Univariate Analysis



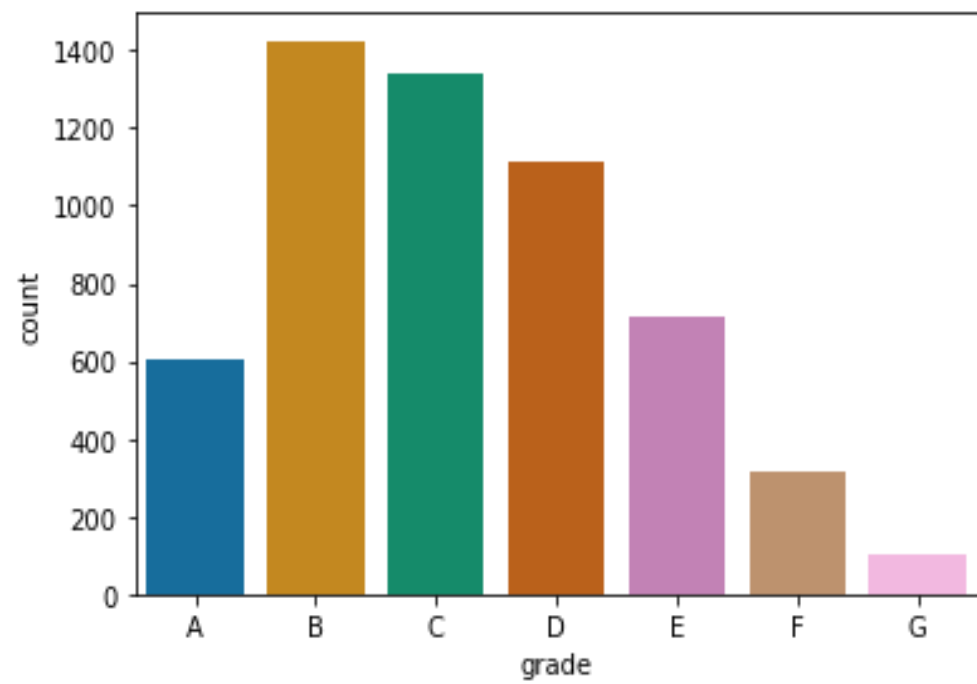
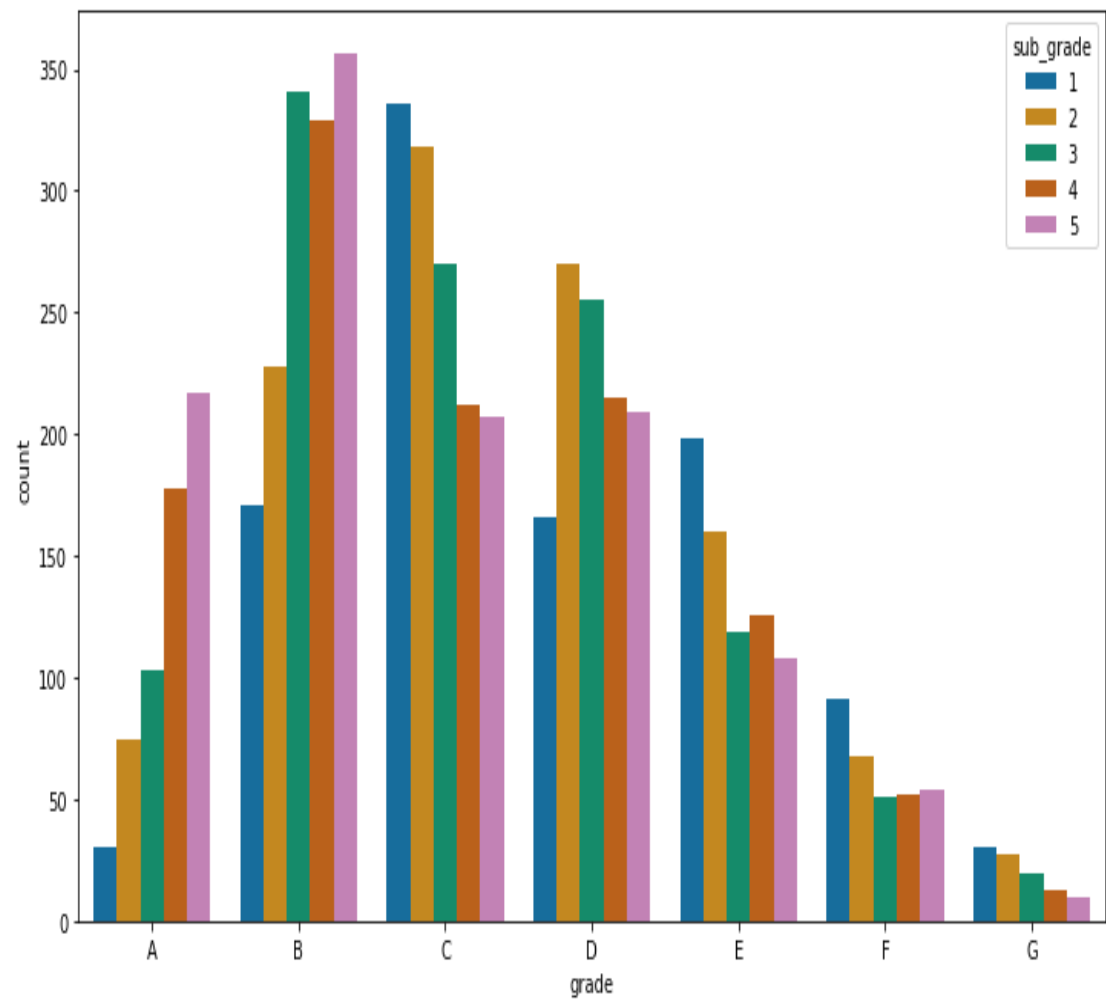
- Removing outliers from annual Income up to 95% to make it easy for plot visualization.



- Checking for outliers for funded amount from the investor for any loan rejection or approval



We are Visualizing the Categorical data for the loan status for customers who are fully paid and for the customer who is charged off.

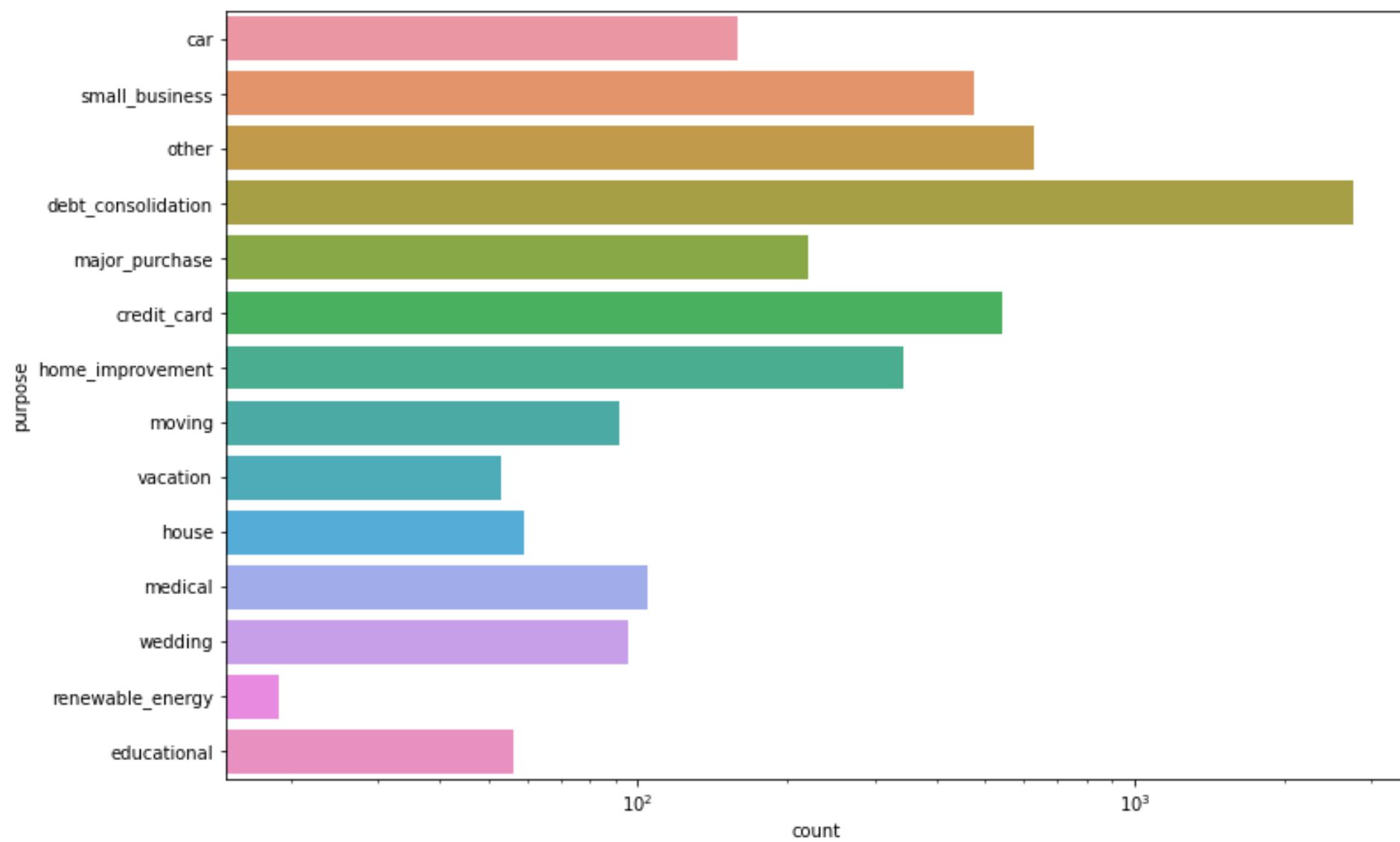


- We are visualizing the defaulter data that are charged off by the investors. From the above plot, we can see Grade B has the maximum number of charged-off customers.

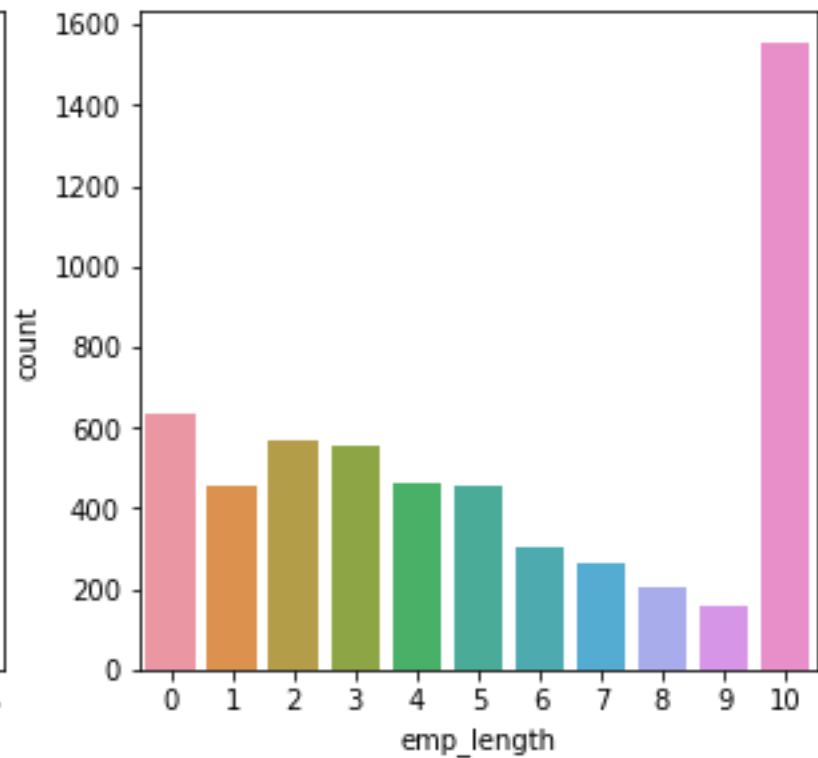
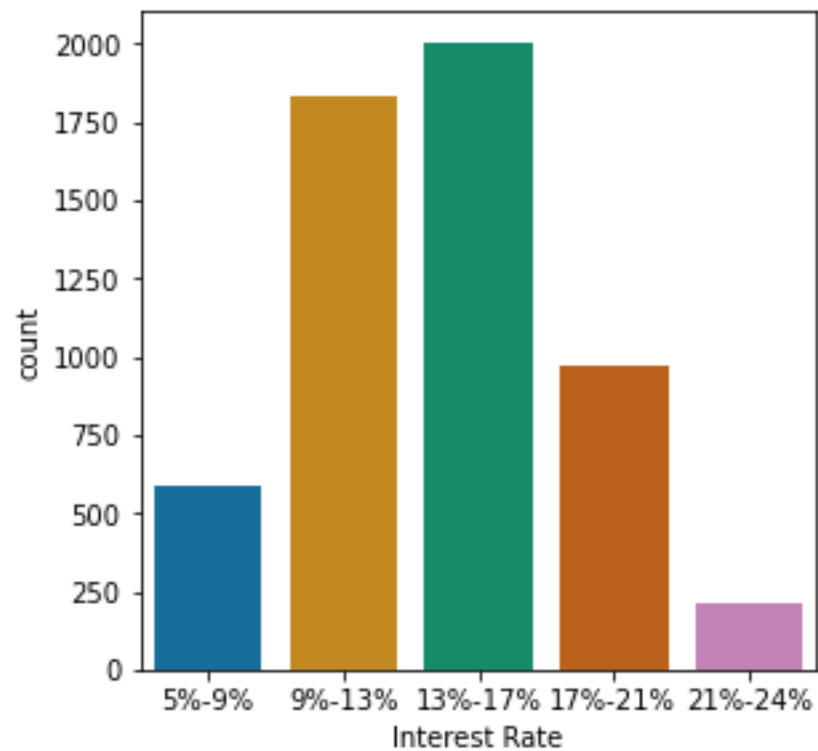




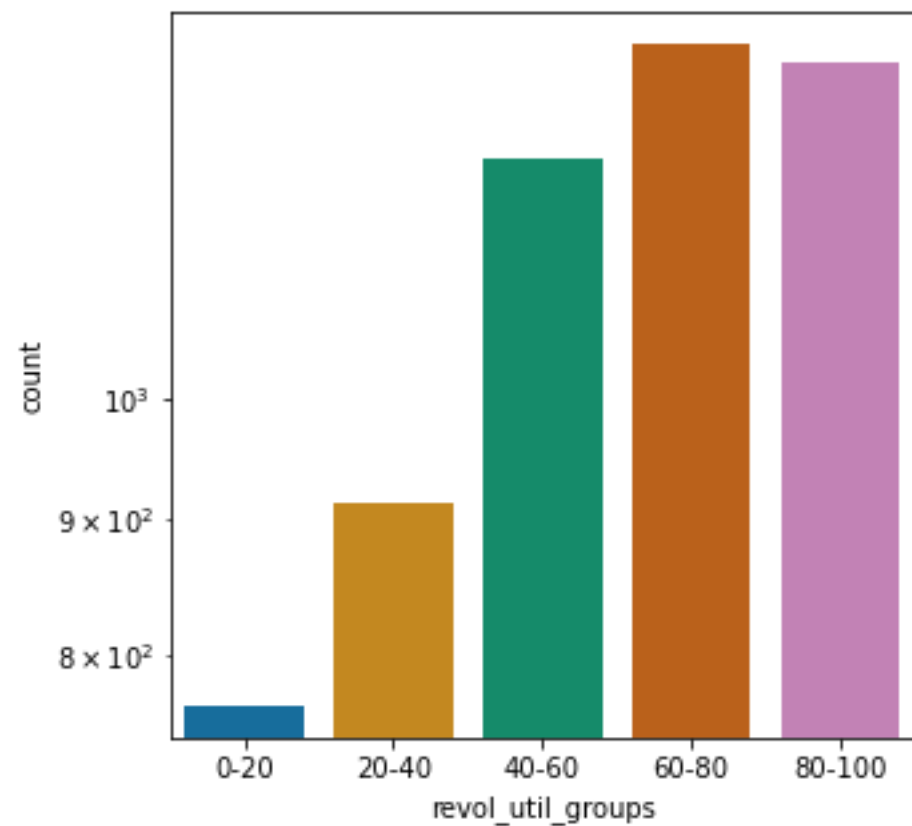
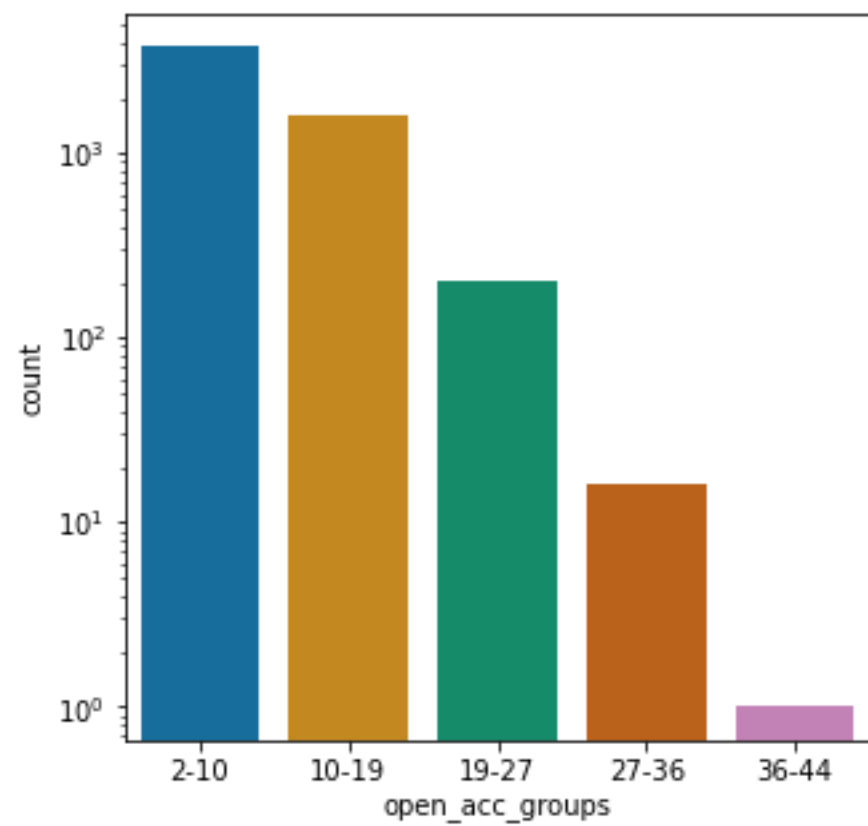
From this plot, we can observe that customers who are staying in rental homes have high chances of being charged off.

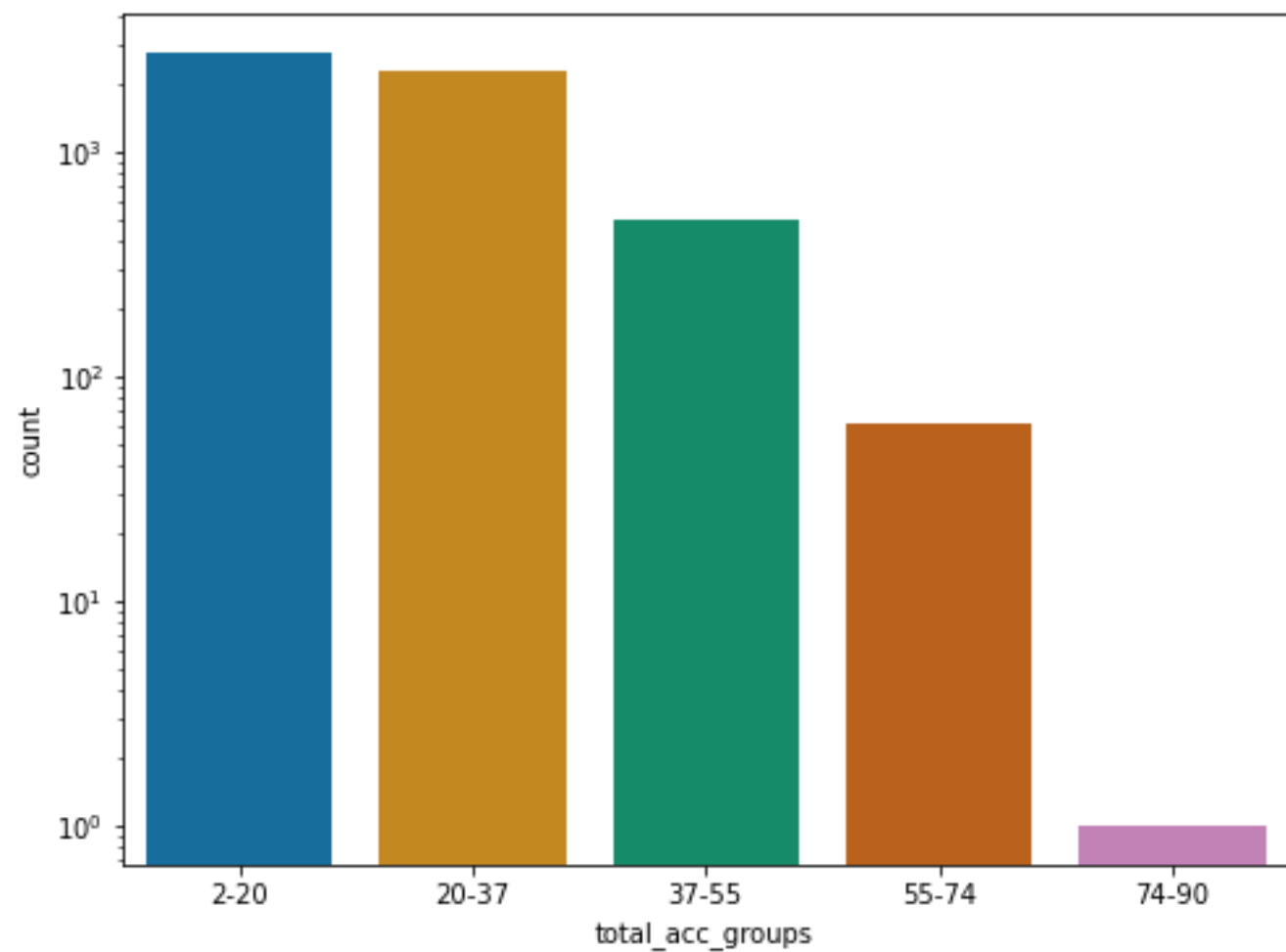


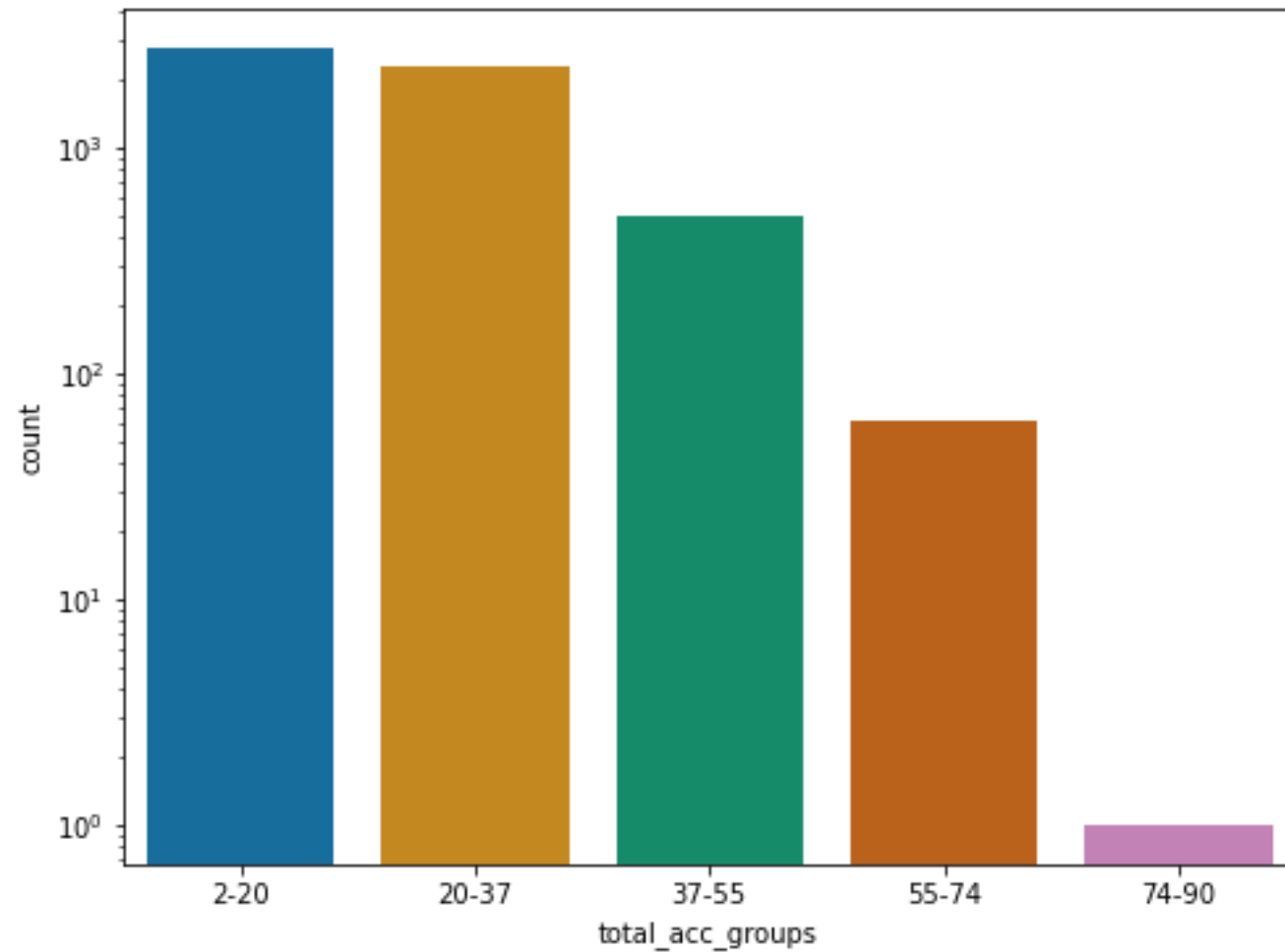
- From the above plot, we can infer the purpose of the loan from the customers is due debt\_consolidation which means to repay the existing the loans.



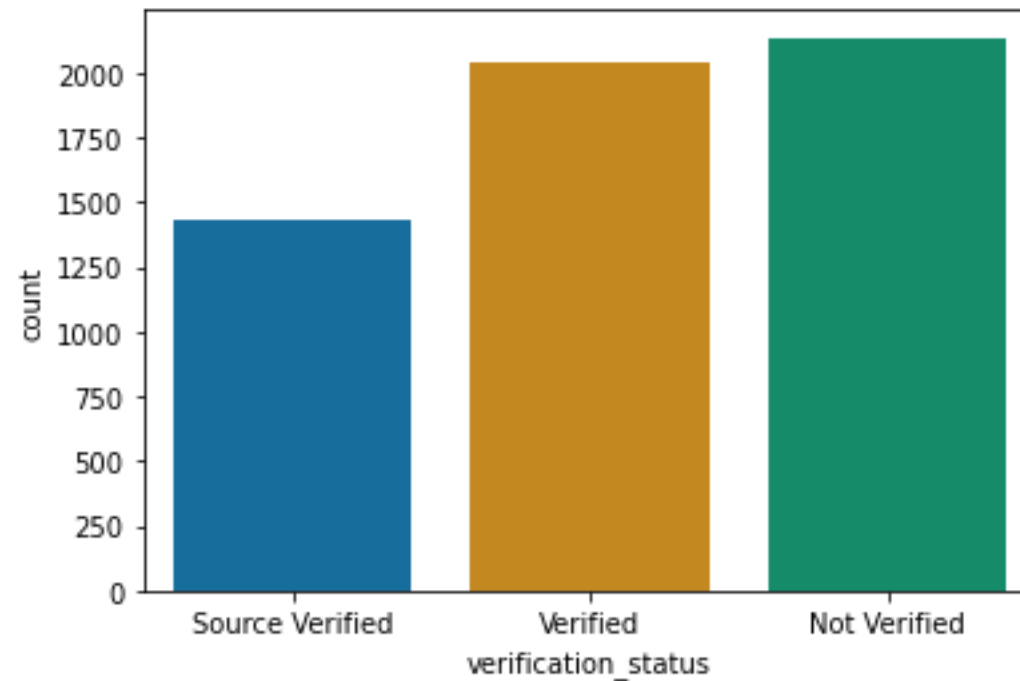
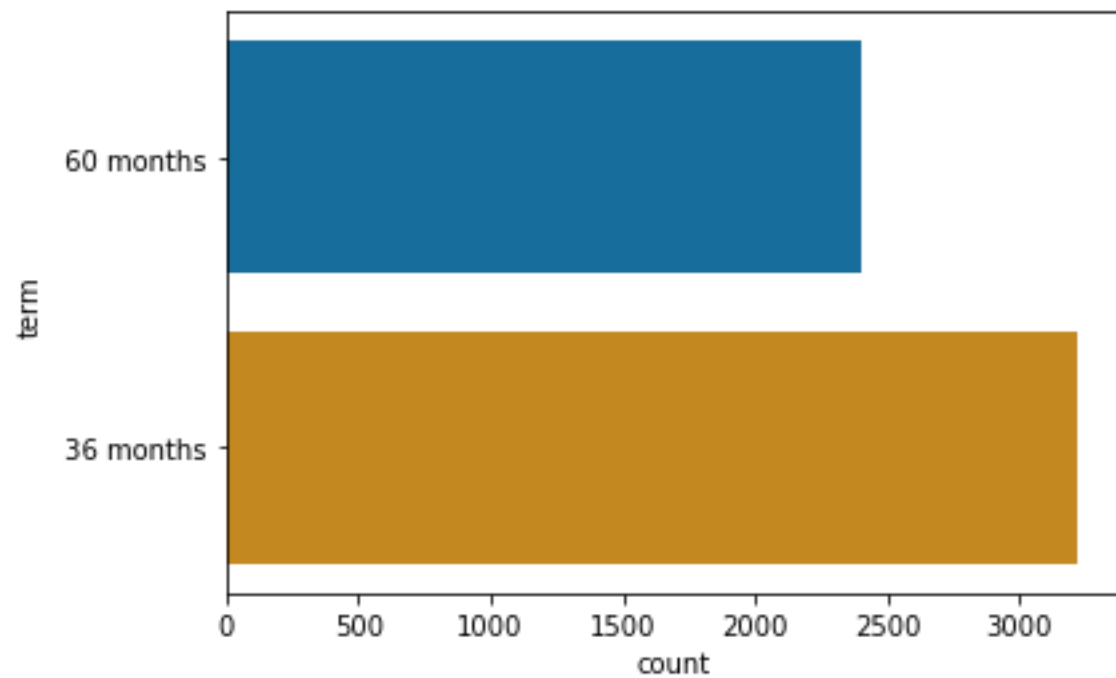
Maximum number of customers who are charged off is having around 10 years of Employment length and when interest rate falls under 13 – 17 %





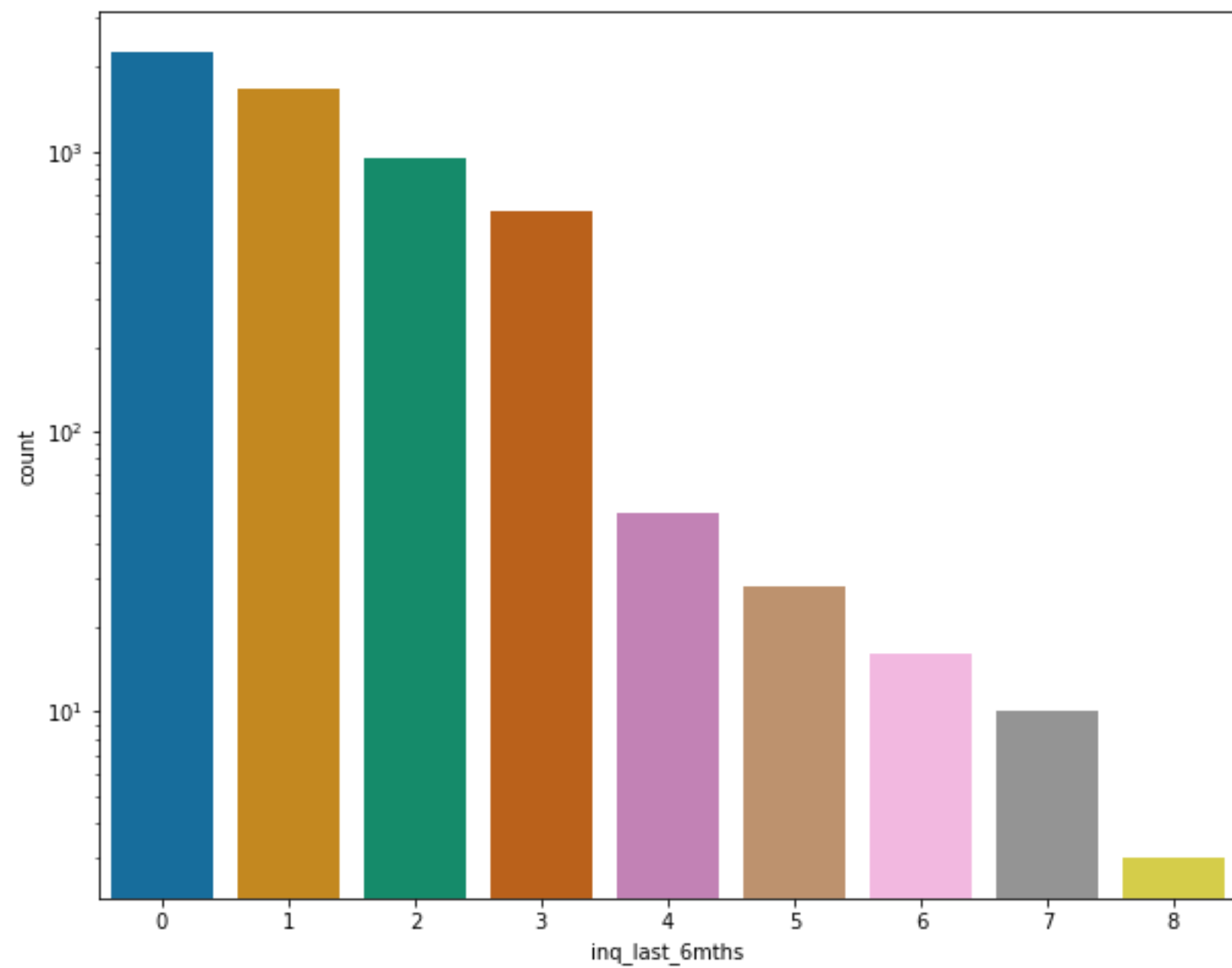


- From the plot1 and plot 3 borrowers who is having open\_acc and total\_acc 2 -10 has been charged off
- From the plot 2 borrowers who is having rev\_util 60 – 80 has been charged off
- From the plot 4 borrowers who is having annual\_inc between 3 – 31k has been charged off

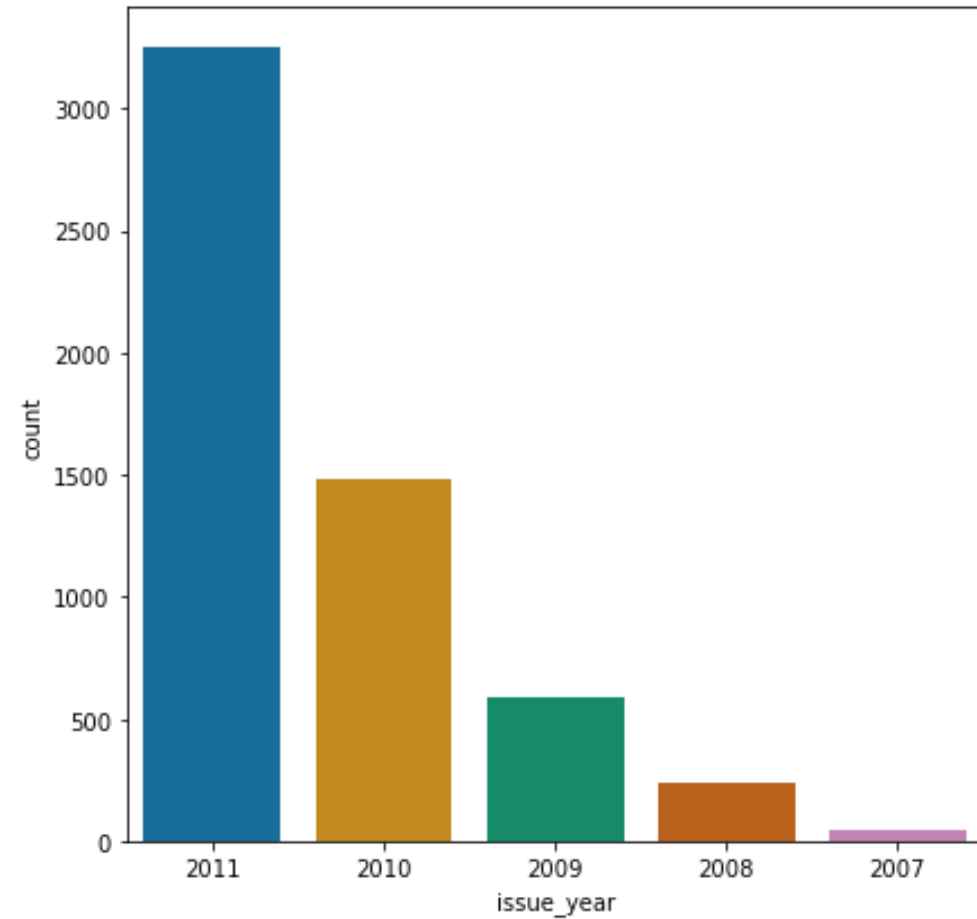
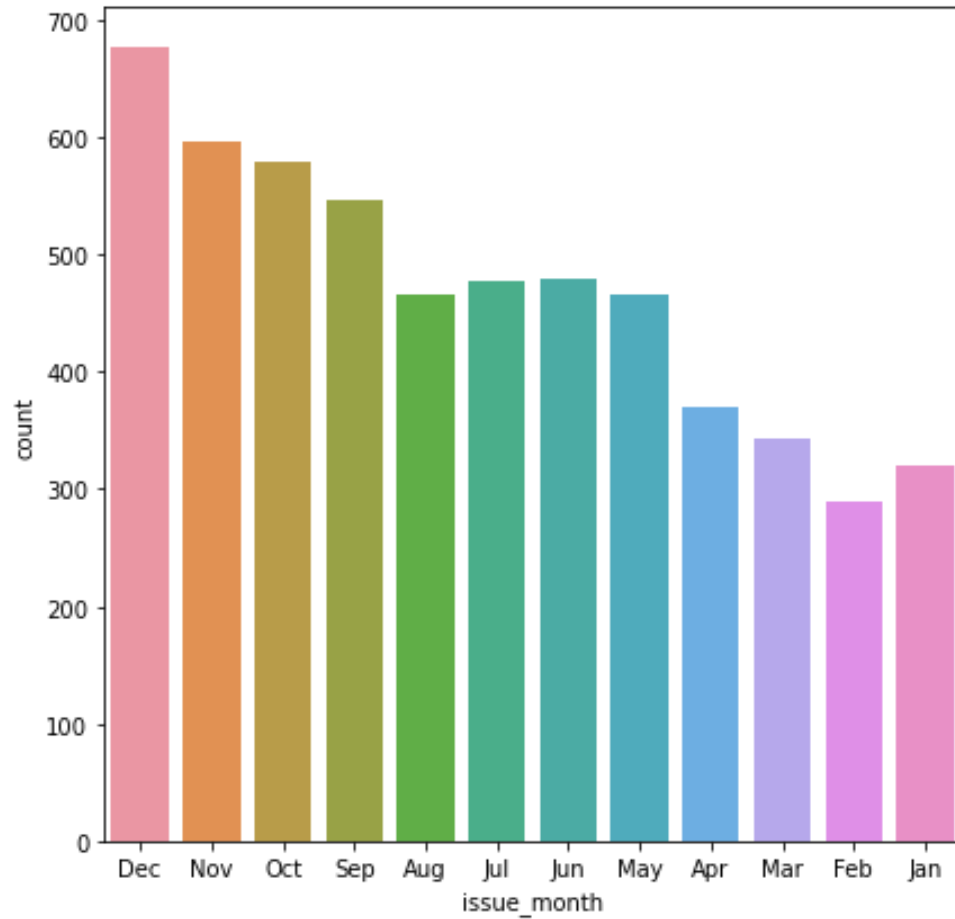


- From the plot1 borrowers who is having loan term for 36 months is having high chances to be charged off
- From the plot 2 borrowers whose source is not verified is having high chances to be charged off





- The number of inquiries in past 6 months (excluding auto and mortgage inquiries). The maximum cases of a customer to be defaulter when the number of inquiries in the past 6 months is 0.



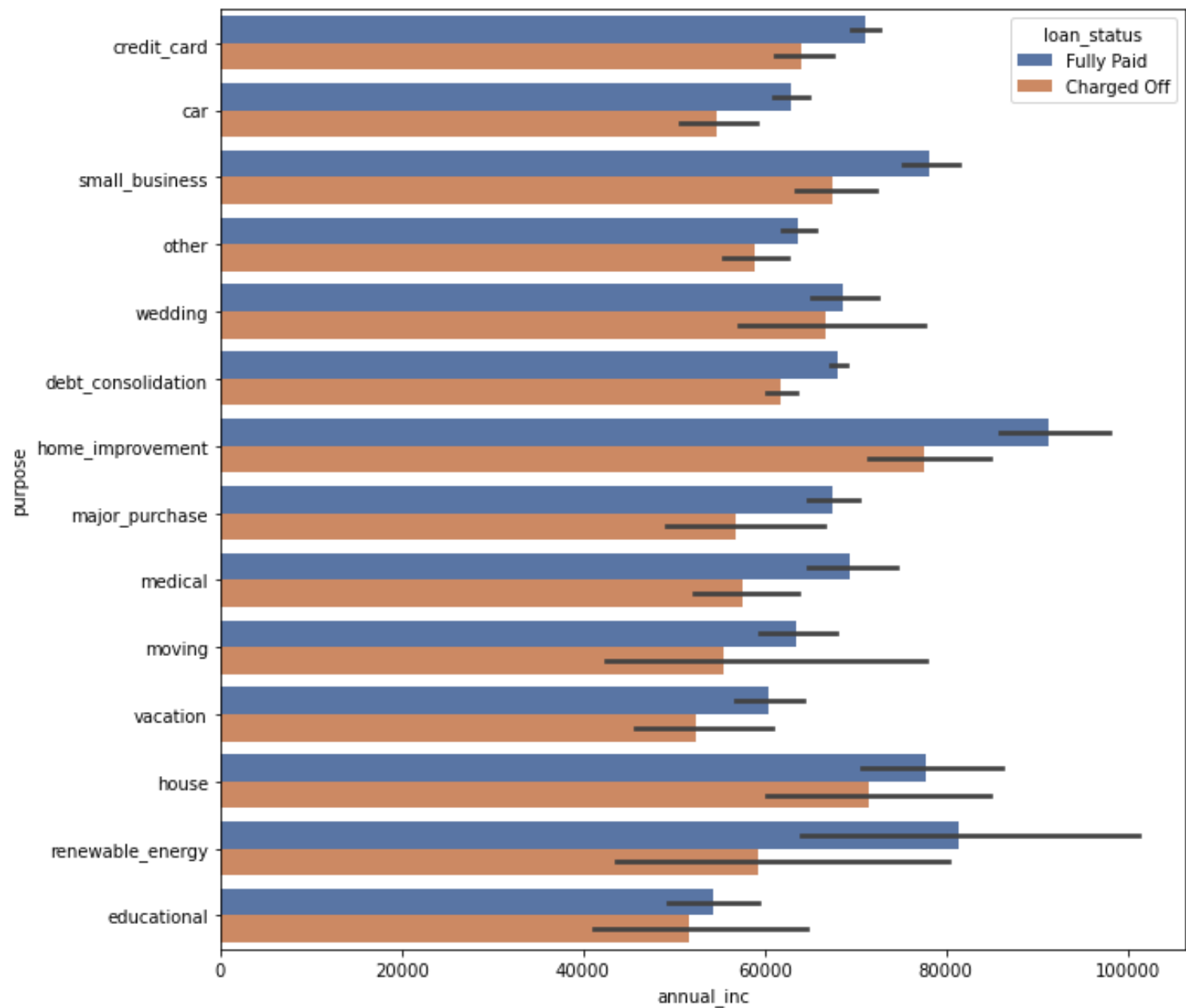
- The maximum number of defaults occurred when the loan was issued in Dec. Loan issued in the year 2011 were also charged off high as compared to other years

### **Consolidate Univariate Analysis Observation:**

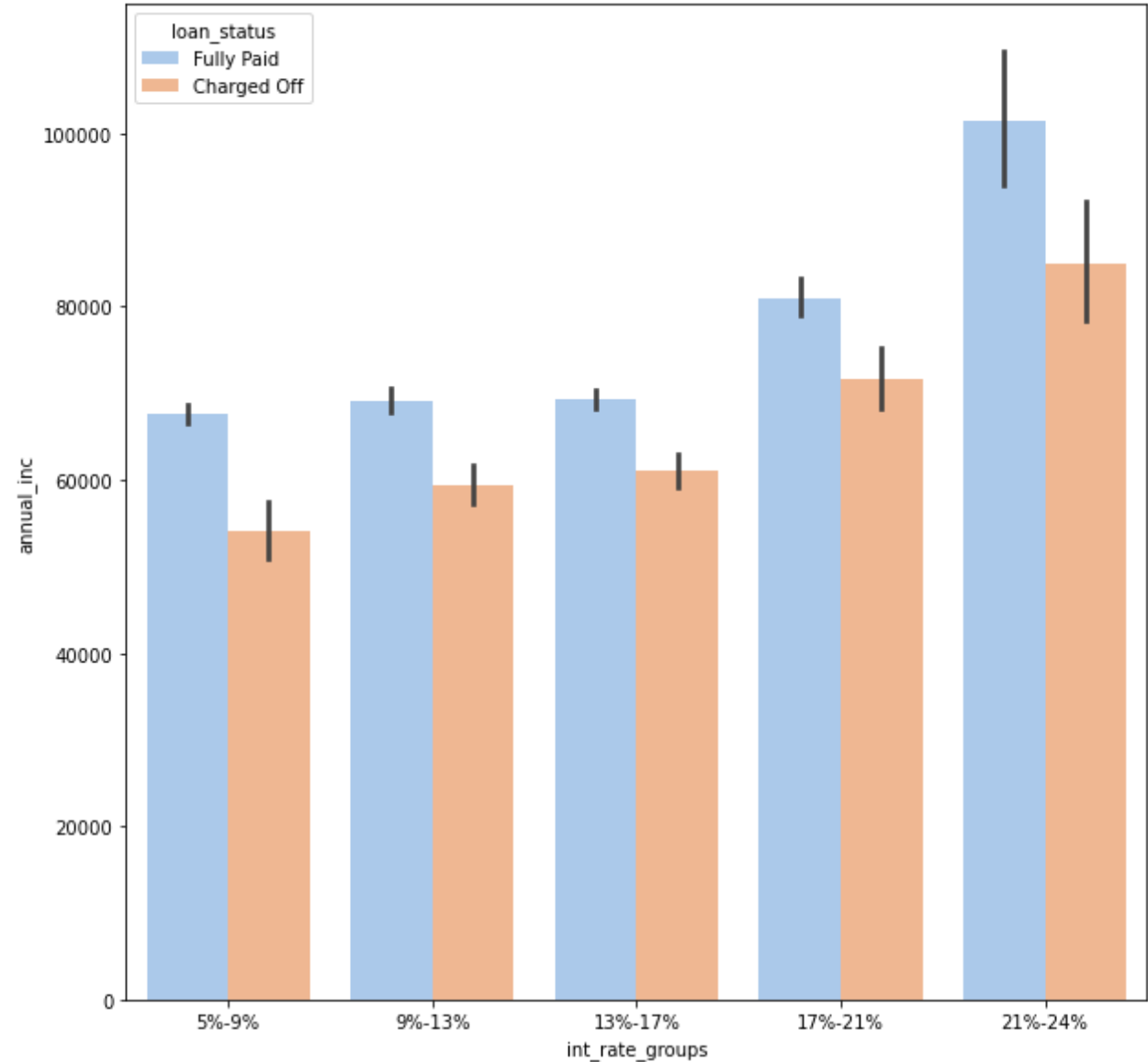
1. The applicant's homeownership is 'RENT'
2. The purpose of the loan is Debt clearance
3. The applicants who receive interest at the rate of 13-17%
4. Applicants who have an income of range 31201 - 58402
5. Applicants who have 2- 10 open\_acc
6. Applicants with employment length of 10
7. When funded amount by investor is between 5000-10000
8. Loan amount is between 5429 - 10357
9. Data is between 12-18
10. When monthly installments are between 145-274
11. Term of 36 months
12. When the loan status is Not verified
13. When the no of inquiries in the last 6 months is 0
14. When the number of derogatory public records is 0
15. When the purpose is 'debt\_consolidation'
16. Grade is 'B'
17. And a total grade of 'B5' level.
18. Also there is a very interesting observation from the date issued. The late months of a year indicated the high possibility of defaulting.

The high number of loan defaults in 2011 could be due to the financial crisis in the USA (Assuming the data is of US origin)

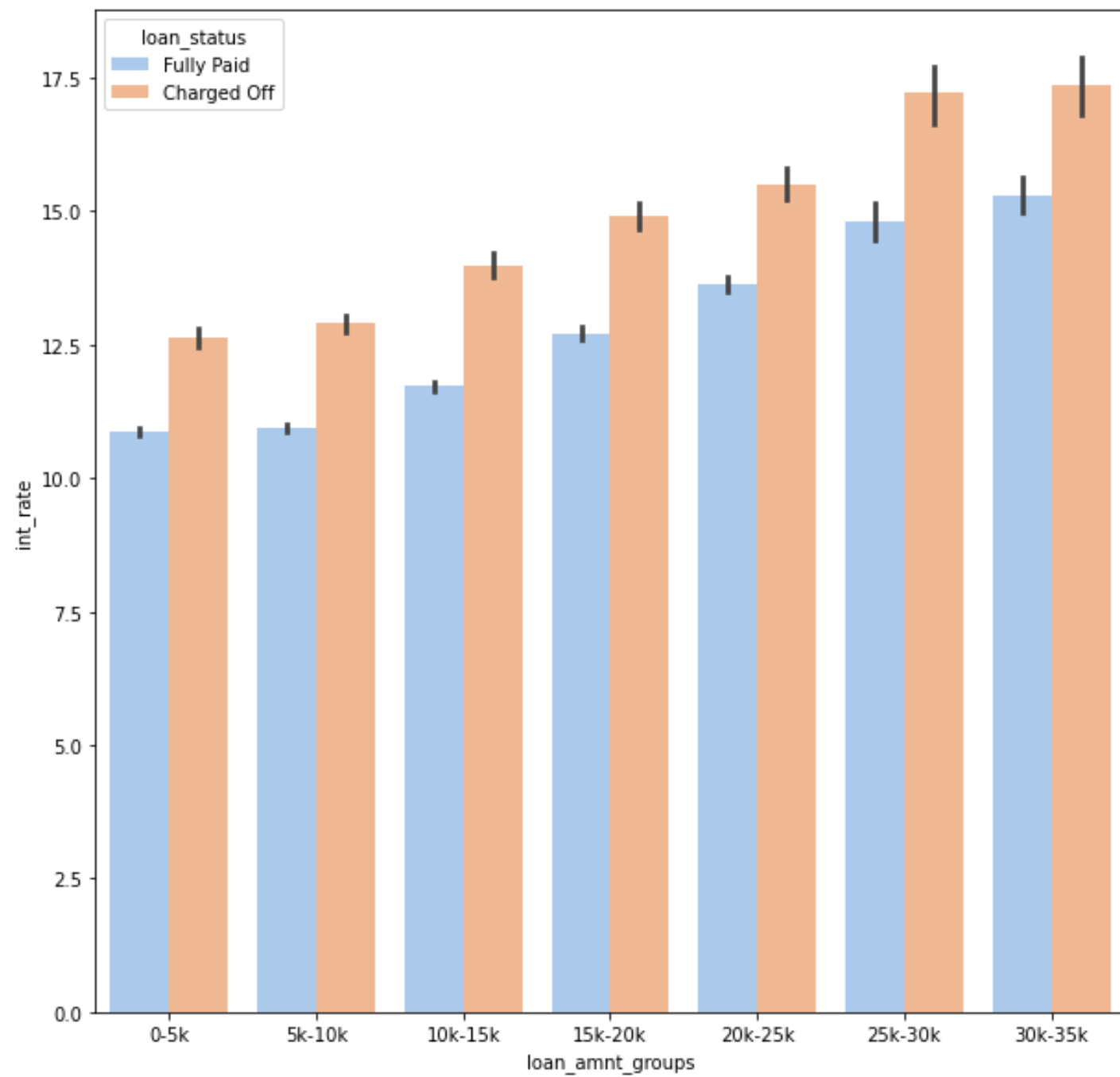
## BI-VARIATE ANALYSIS

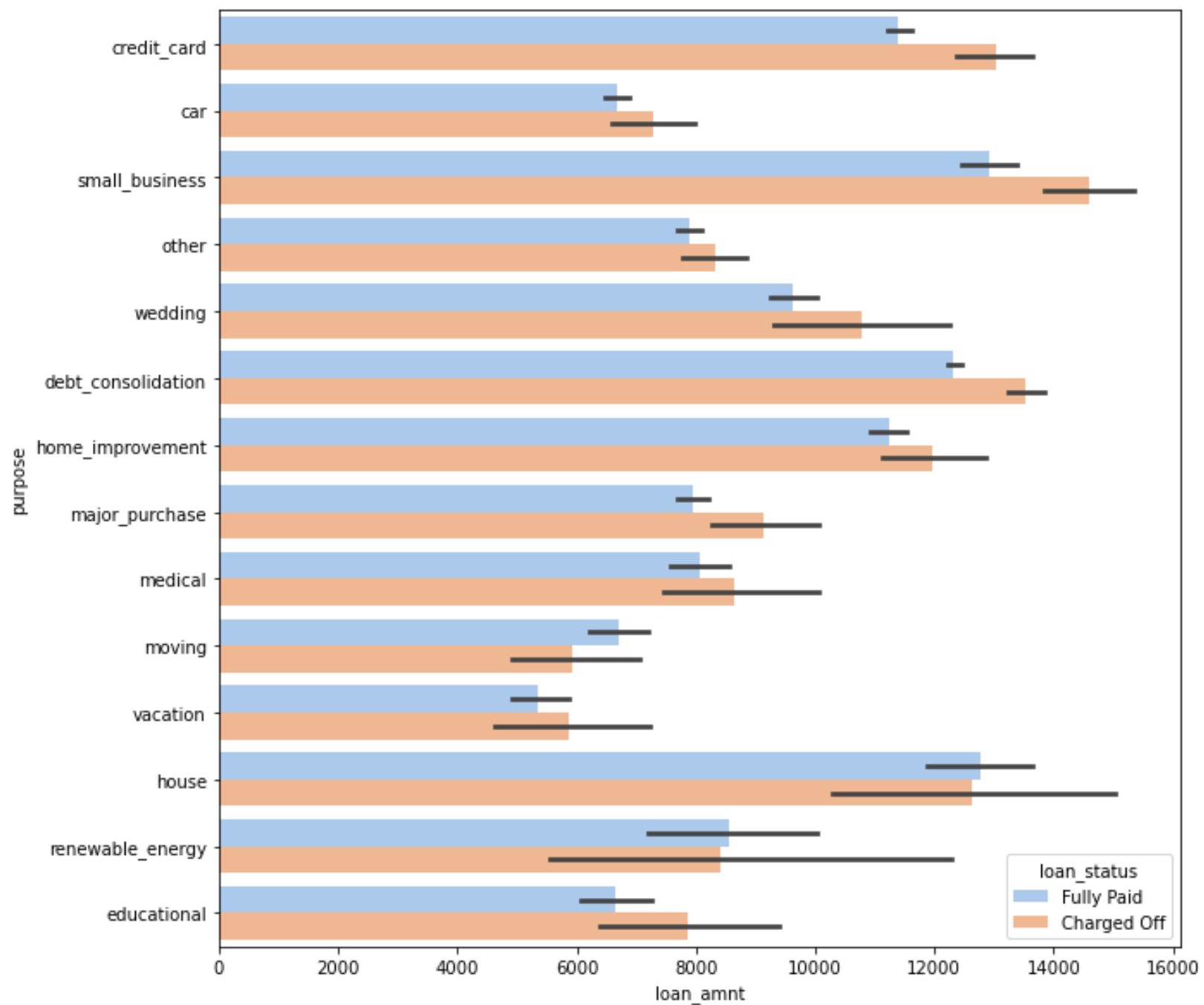


- Applicants with higher salaries applied for home improvement, small business, house, and renewable energy.

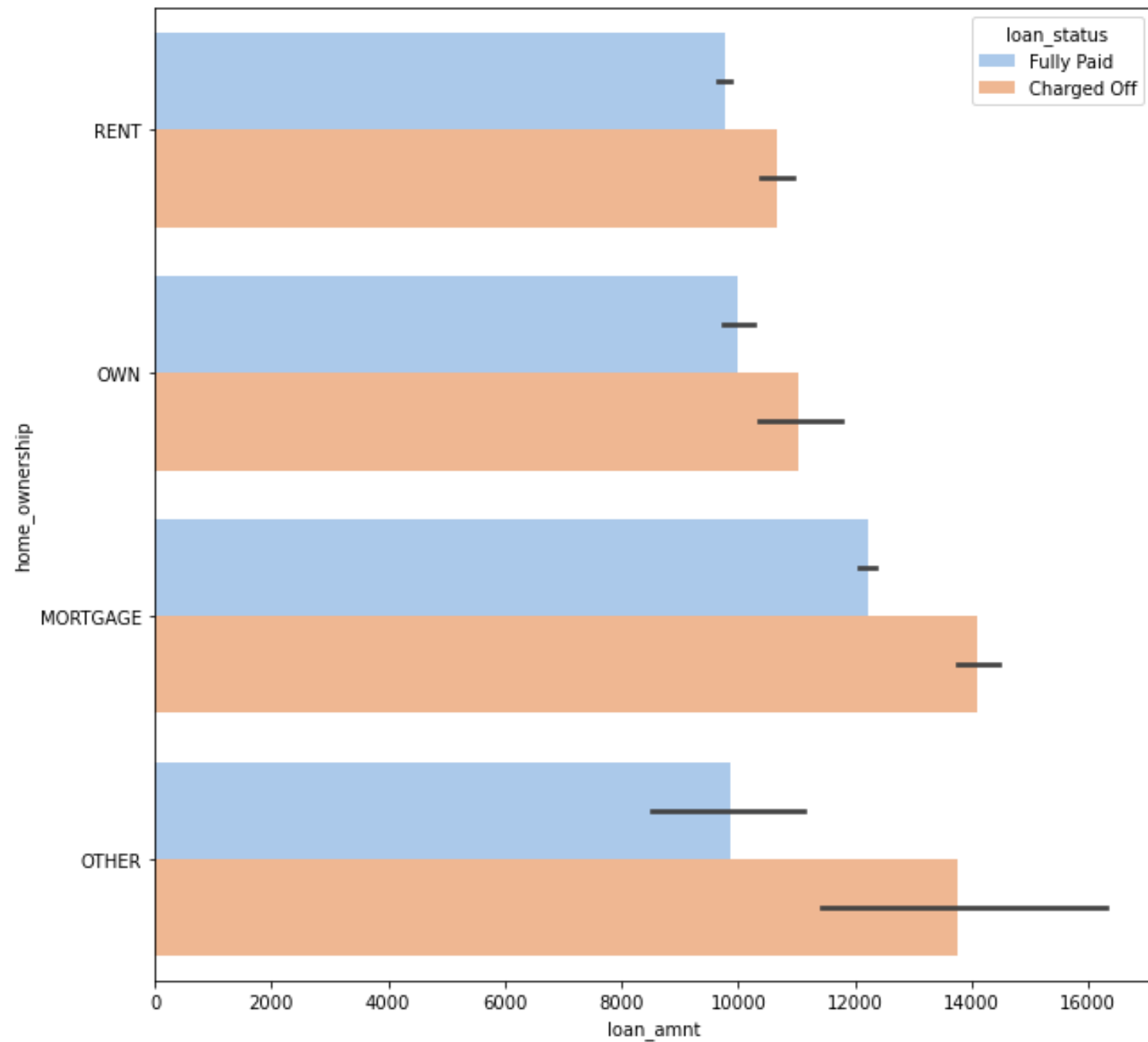


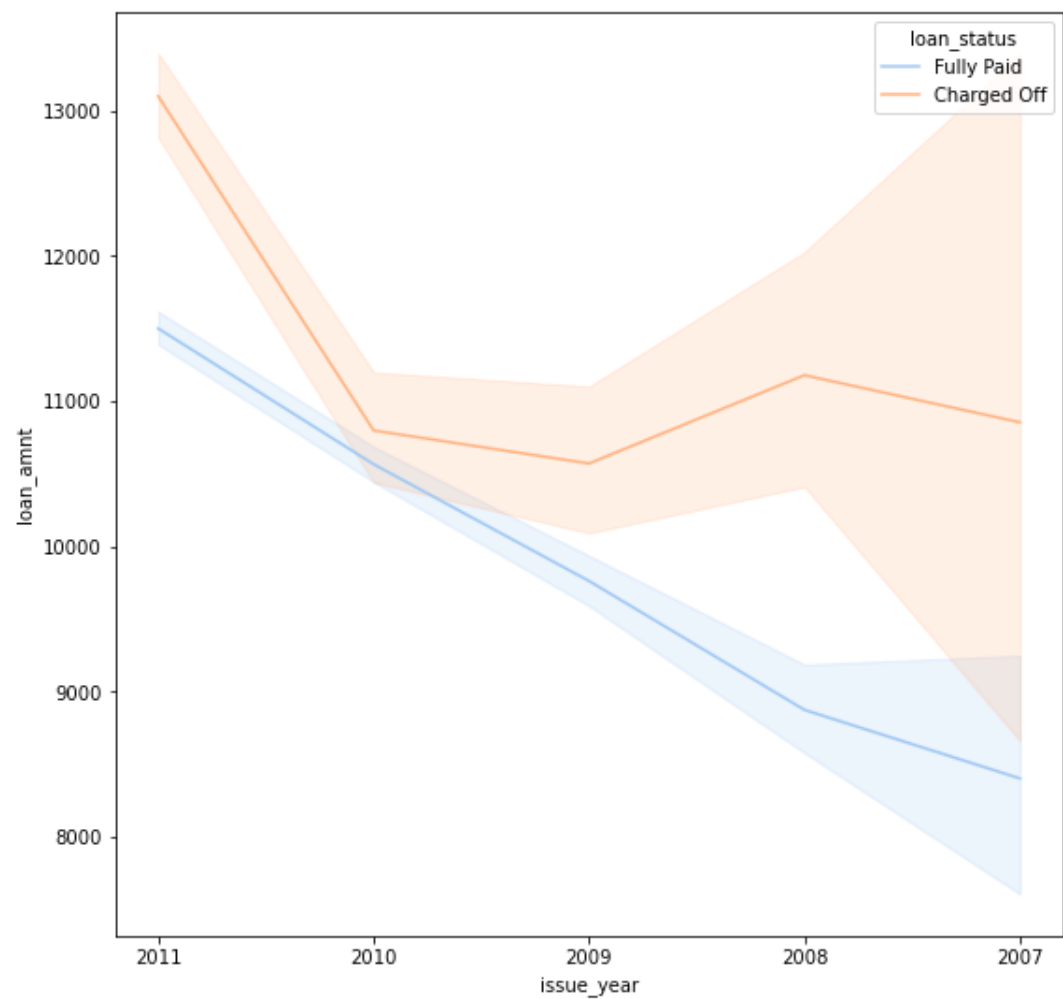
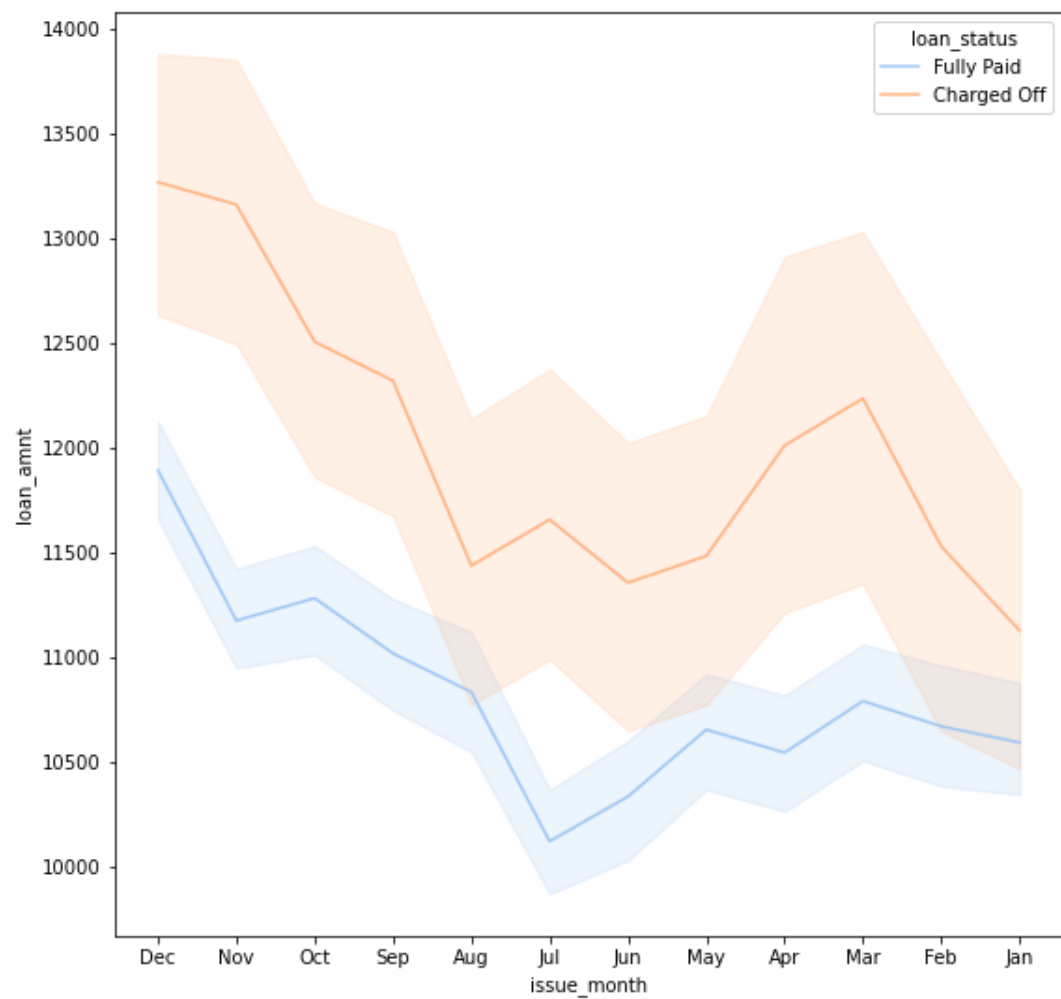
The probability of defaulter is when Applicants who receive interest at the rate of 21-24% and have an income of 70k-80k

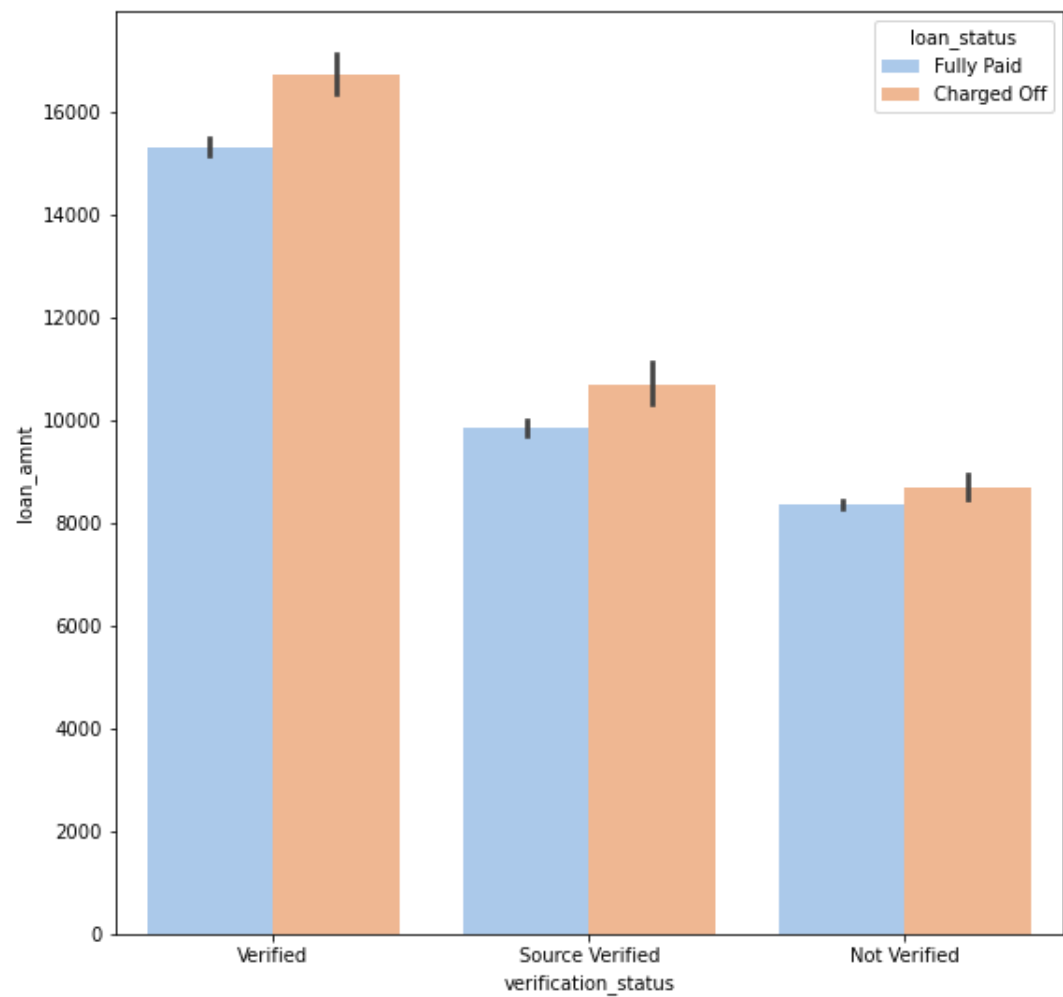
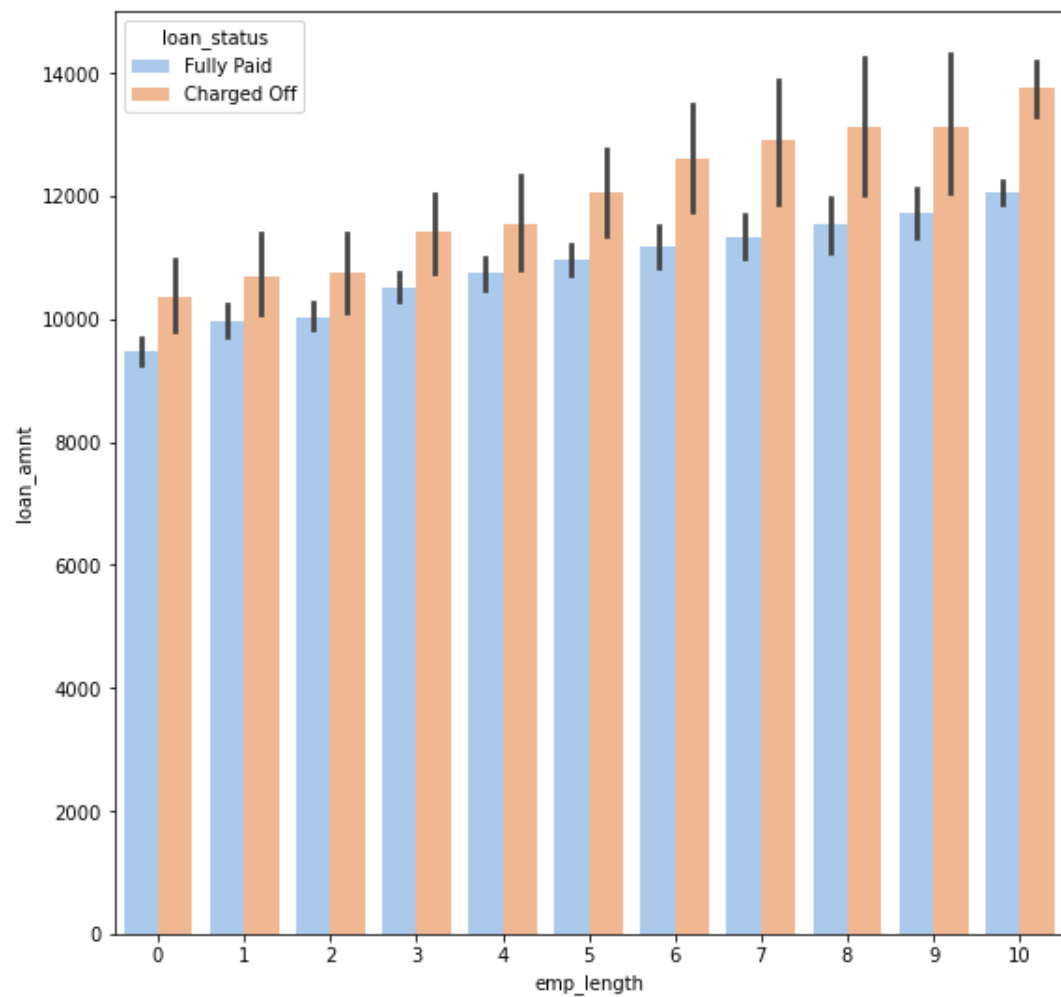


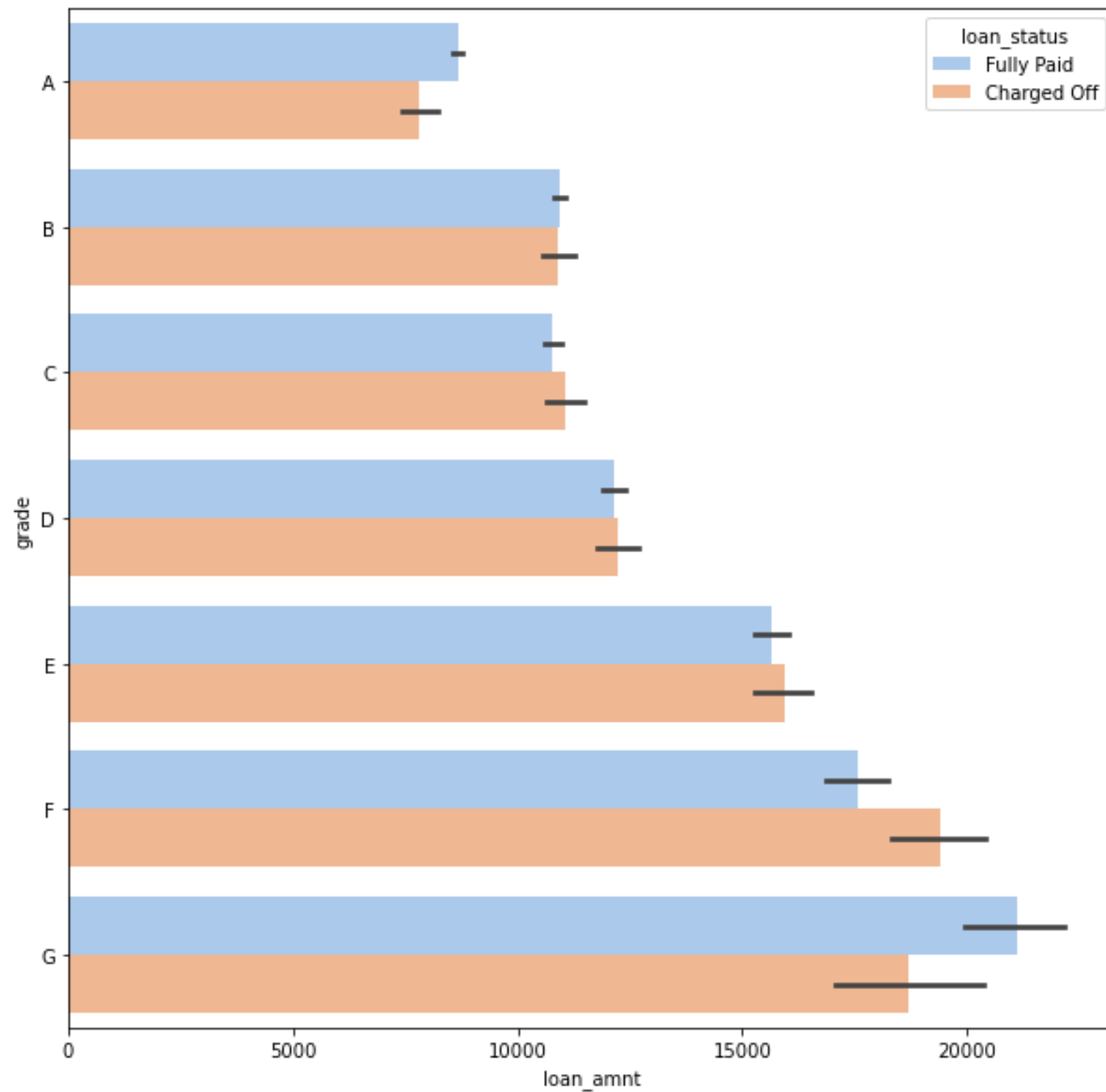


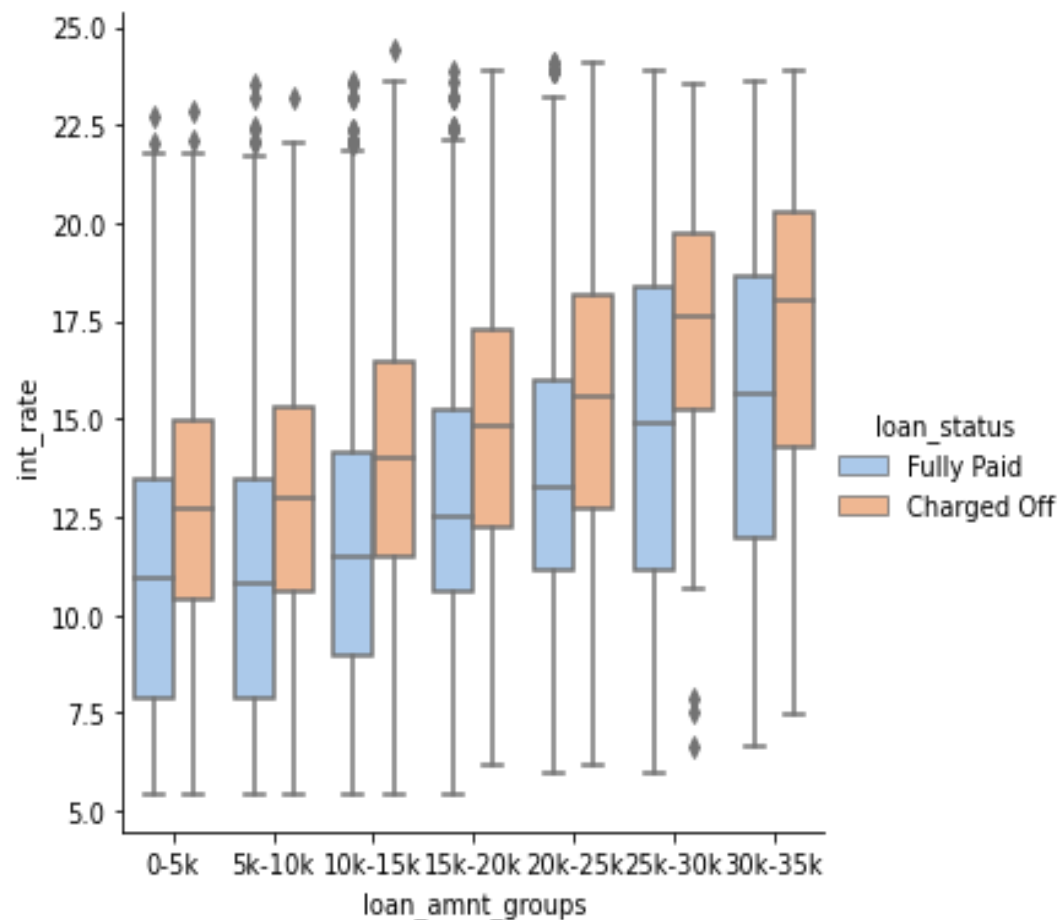
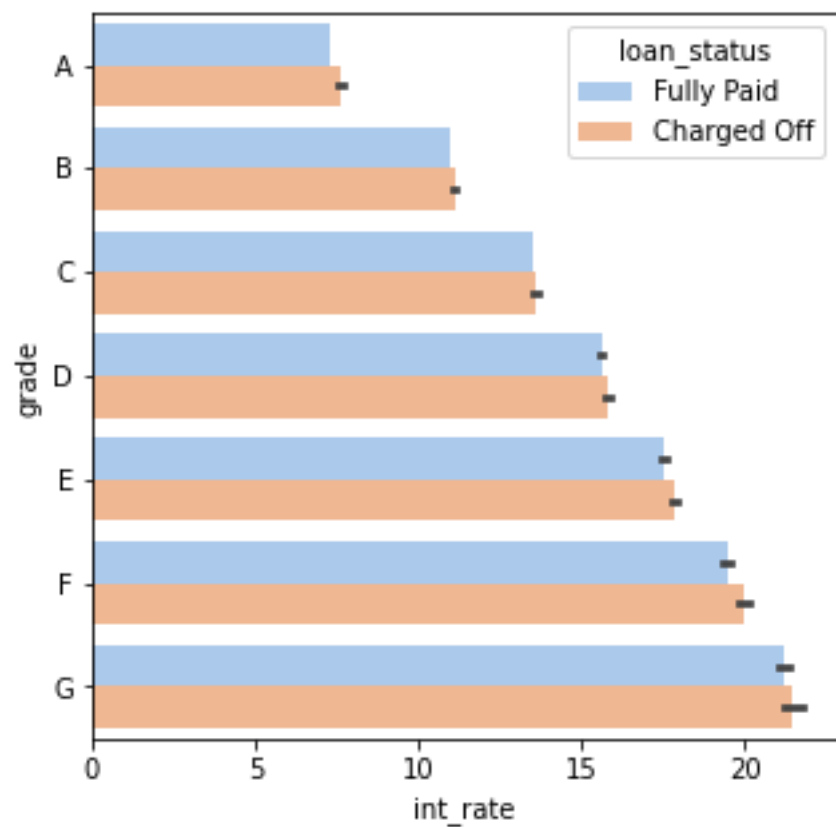


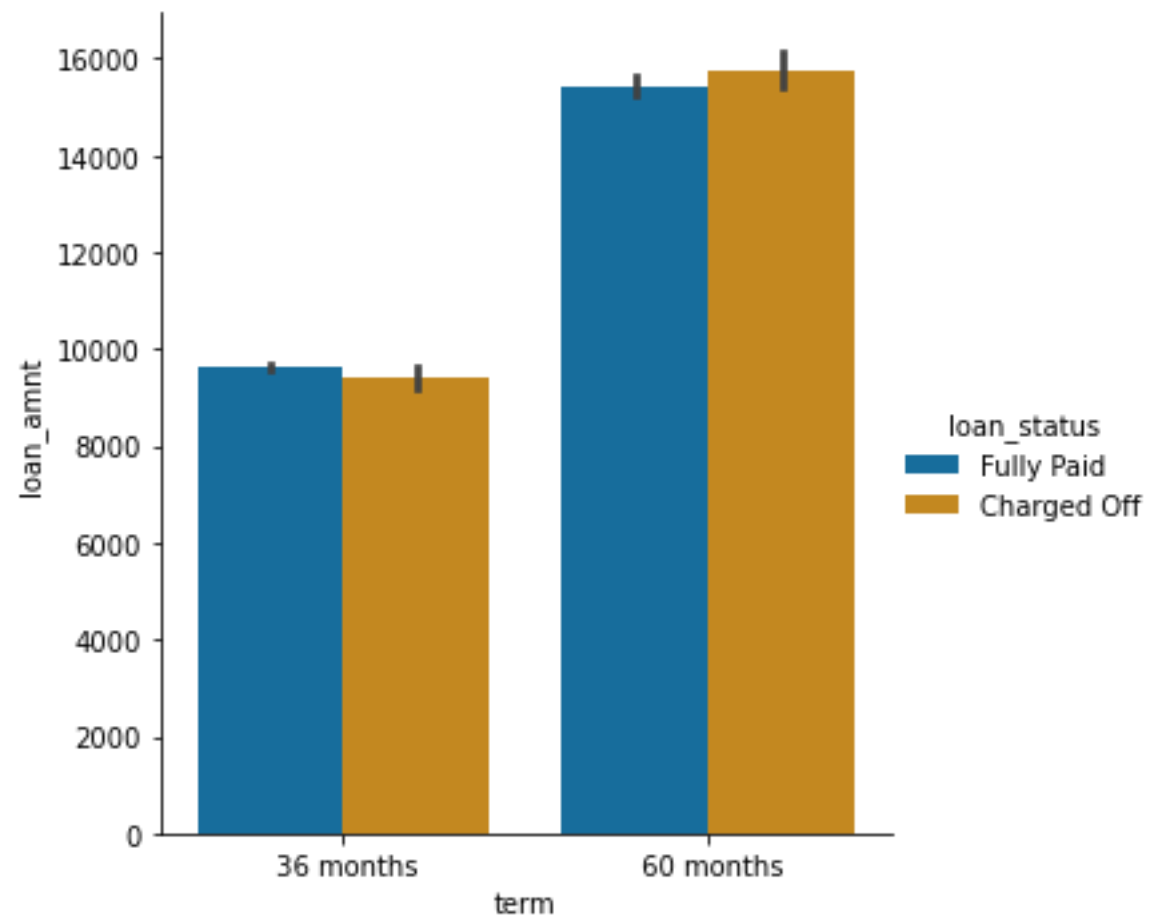












## Consolidated Bi- Variate observations:

The above analysis with respect to the charged-off loans. There is the probability of defaulting when :

1. Applicants taking a loan for 'home improvement' and have an income of 60k -70k
2. Applicants whose homeownership is 'MORTGAGE and have an income of 60-70k
3. Applicants who receive interest at the rate of 21-24% and have an income of 70k-80k
4. Applicants who have taken a loan in the range 30k - 35k and are charged an interest rate of 15-17.5 %
5. Applicants who have taken a loan for a small business and the loan amount is greater than 14k
6. Applicants whose homeownership is 'MORTGAGE and have a loan of 14-16k
7. When a grade is F and loan amount is between 15k-20k
8. When employment length is 10yrs and loan amount is 12k-14k
9. When the loan is verified and the loan amount is above 16k
10. For grade G and interest rate above 20%

## Conclusions

- Lending club should reduce the high-interest loans for 36 months tenure, they are prone to loan default.
- Grades are a good metric for detecting defaulters. The lending club should examine more information from borrowers before issuing loans to Low grades (Grade B and Sub Grade 5).
- Small business loans default more. Lending clubs should stop/reduce issuing the loans to them.
- Borrowers with mortgage homeownership are taking higher loans and defaulting to the approved loans. The lending club should stop giving loans to this category when the loan amount requested is more than 14K – 16K.
- Borrowers who receive interest at the rate of 21-24% and have an income of 70k-80k are prone to loan default