Lending Club Case Study

By -

Reetesh V

Sanjay P

Table of Contents

- Problem Statement

- Graphical Analysis

- Conclusion

Problem Statement

1. Loan accepted: If the company approves the loan, there are 3 possible scenarios described below:

1. **Fully paid**: Applicant has fully paid the loan (the principal and the interest rate)
2. **Current**: Applicant is in the process of paying the instalments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.
3. **Charged-off**: Applicant has not paid the instalments in due time for a long period of time, i.e. he/she has **defaulted** on the loan

2. Loan rejected: The company had rejected the loan (because the candidate does not meet their requirements etc.). Since the loan was rejected, there is no transactional history of those applicants with the company and so this data is not available with the company (and thus in this dataset)
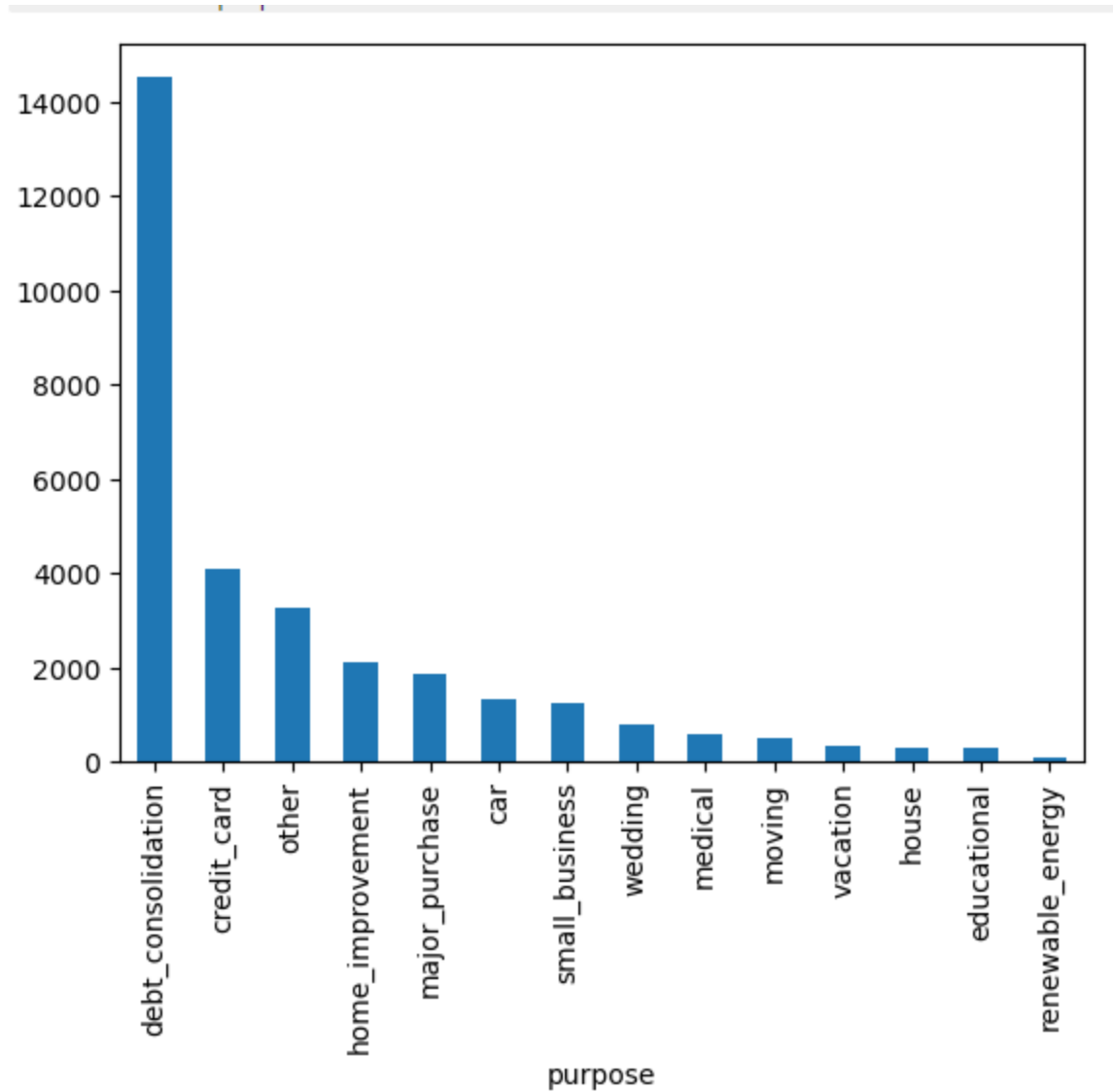
From the below bar plot we can see the people whose salary was ranging more than 80k was opting for huge loan amount.
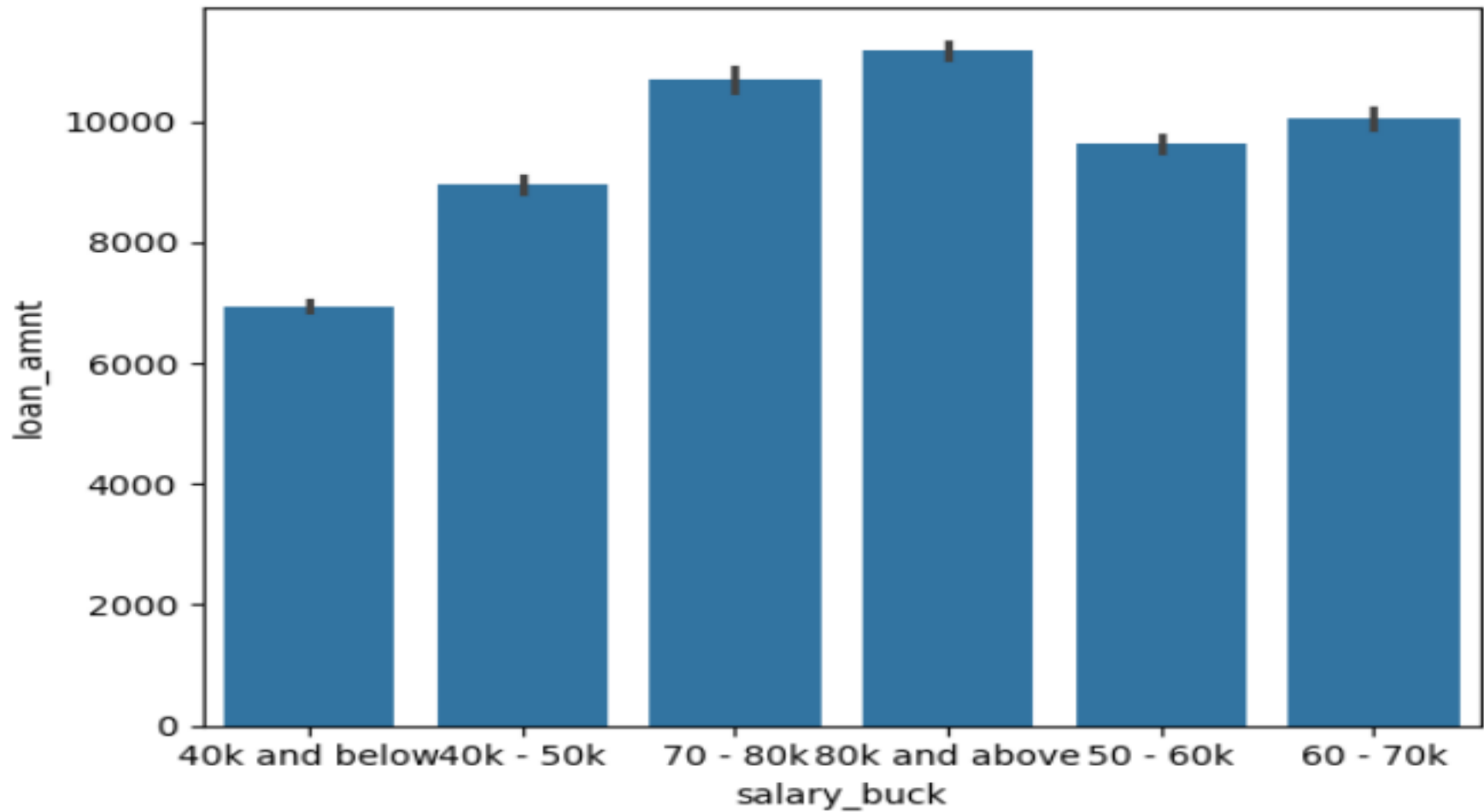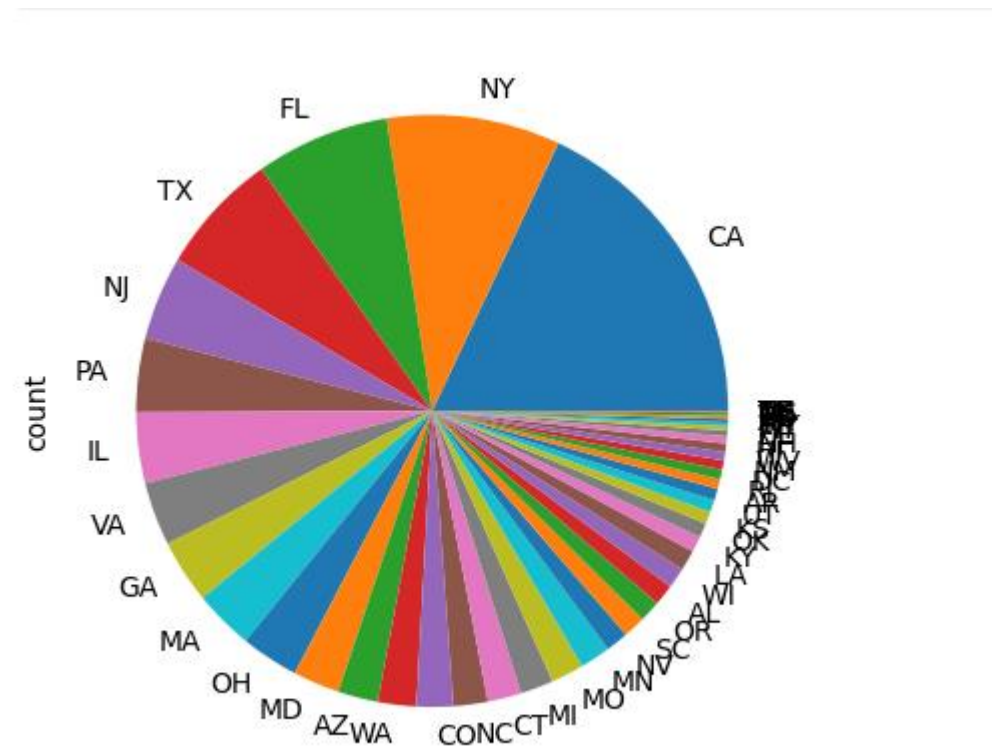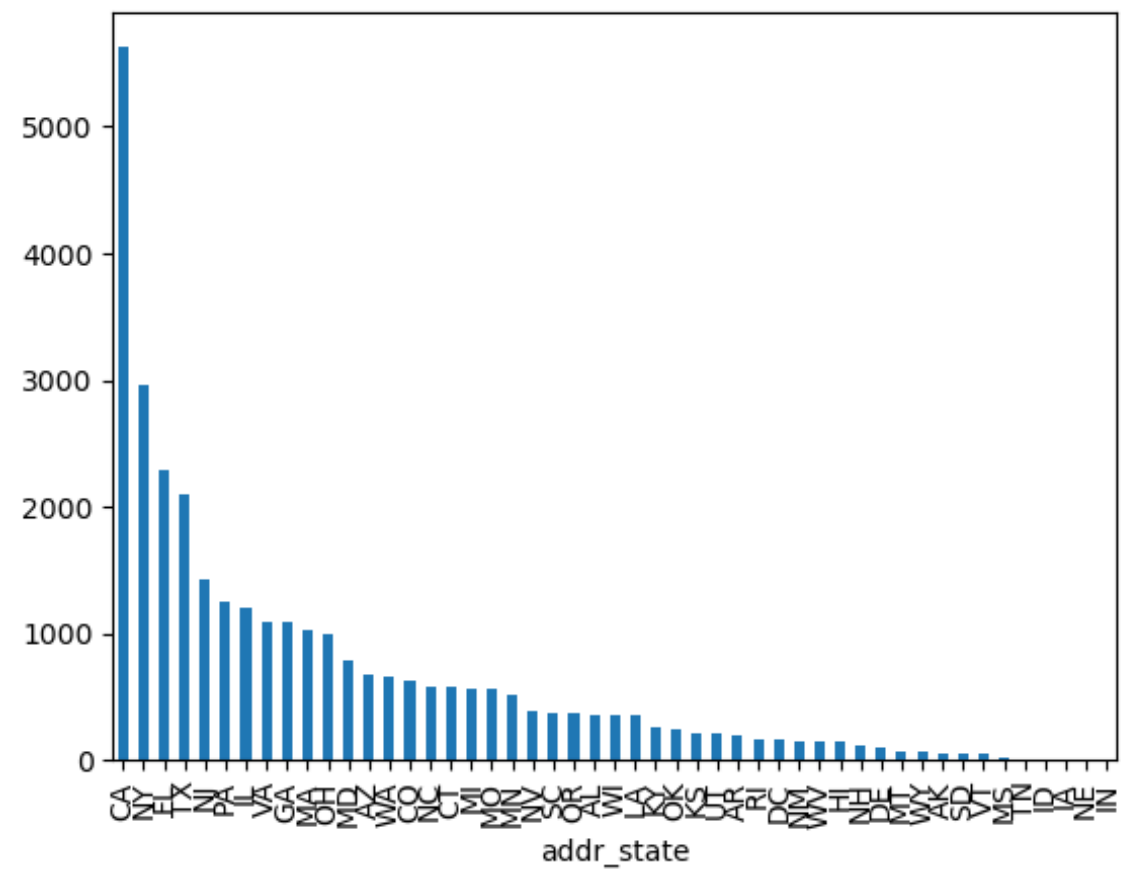
From the below plot we can say that the people who are staying in the rental house were more likely to make more number of loans
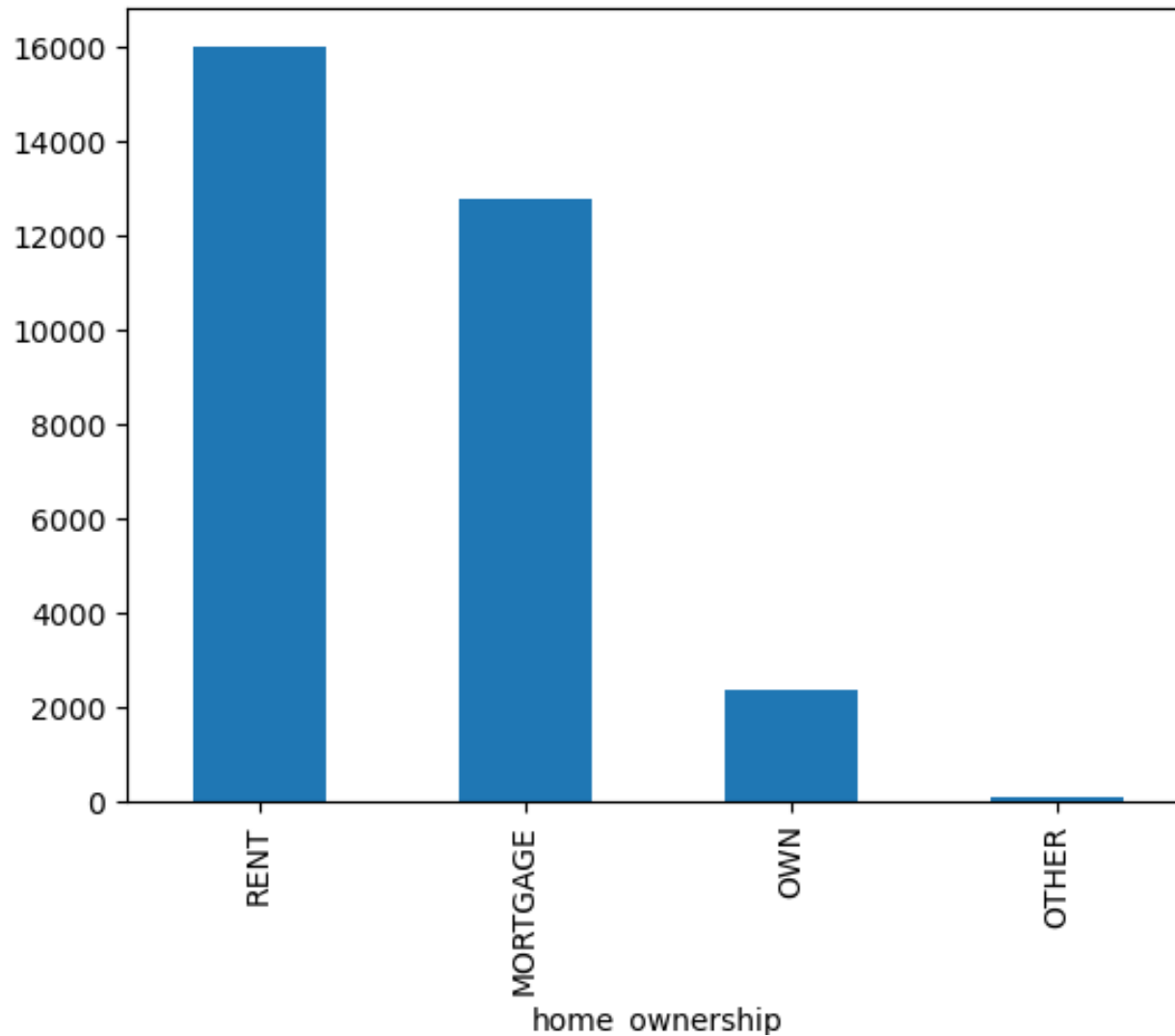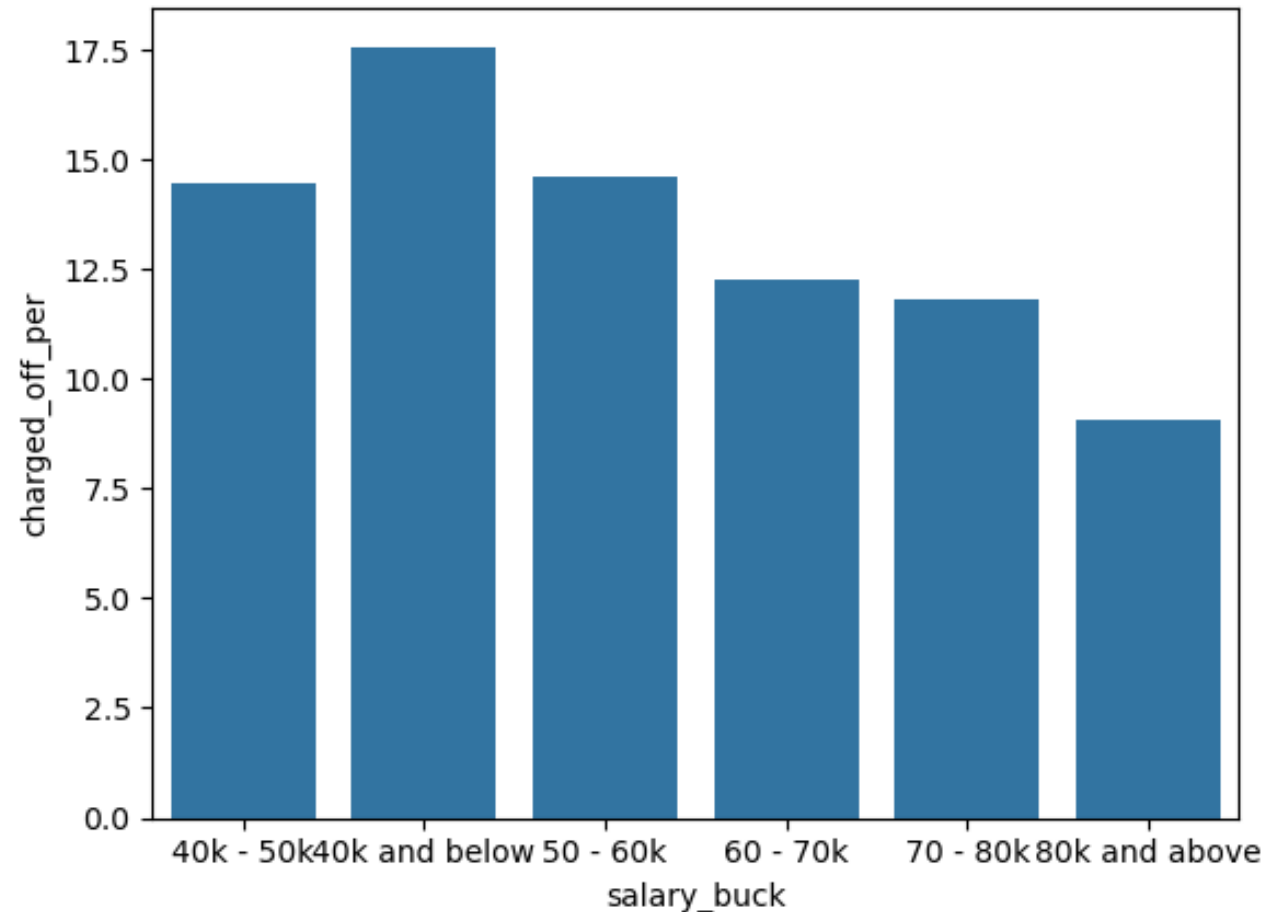
From the below graph we can see that majority of the people purchased the loan because of debit consolidation

From the below graph we can observe that majority of the people who applied for the loan is from CA state
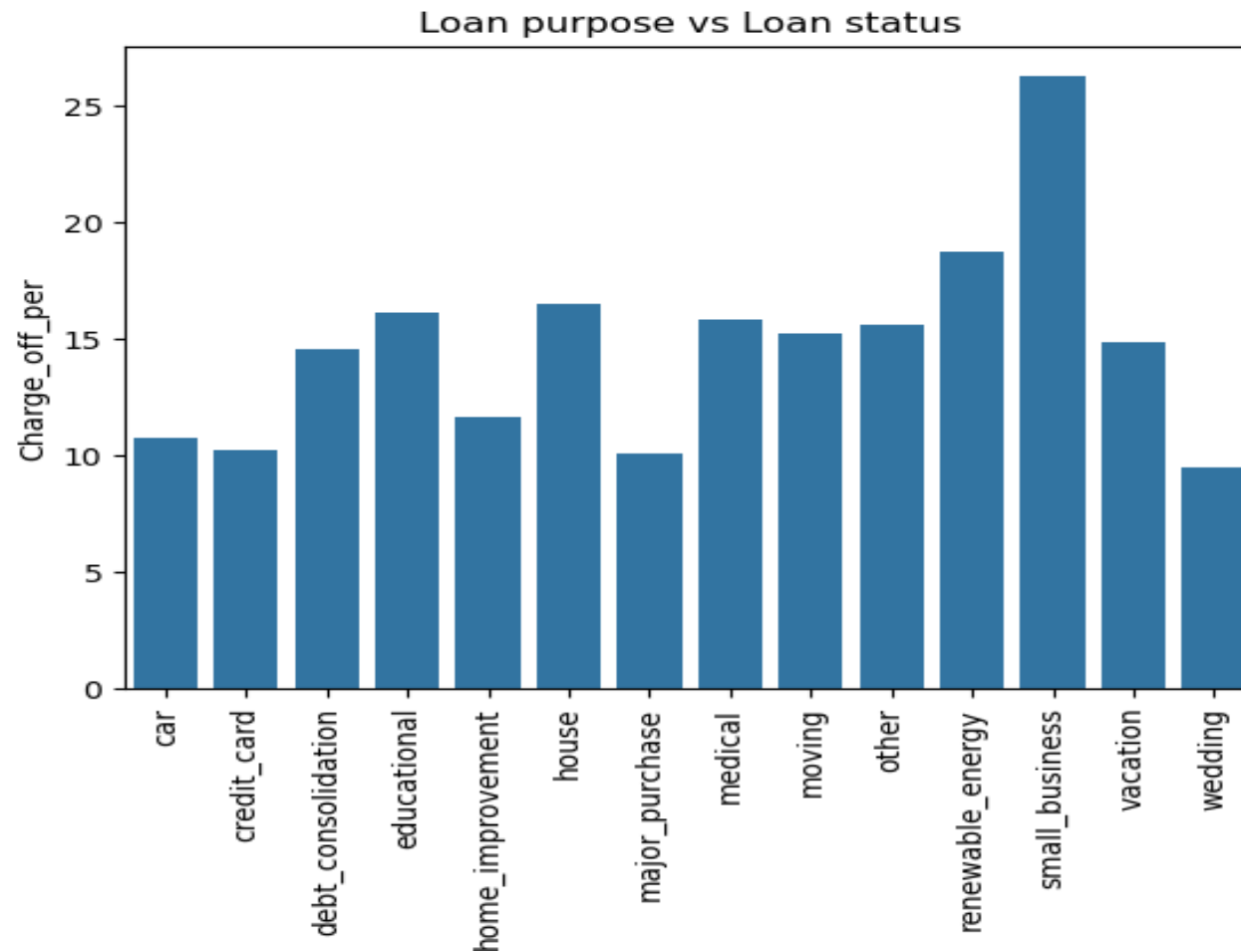
From the below plot we can see that the people whose salary is less than 40k has more chances of getting defaulted and the people whose salary is more than 80k and above has less chances of getting defaulted or getting charged off
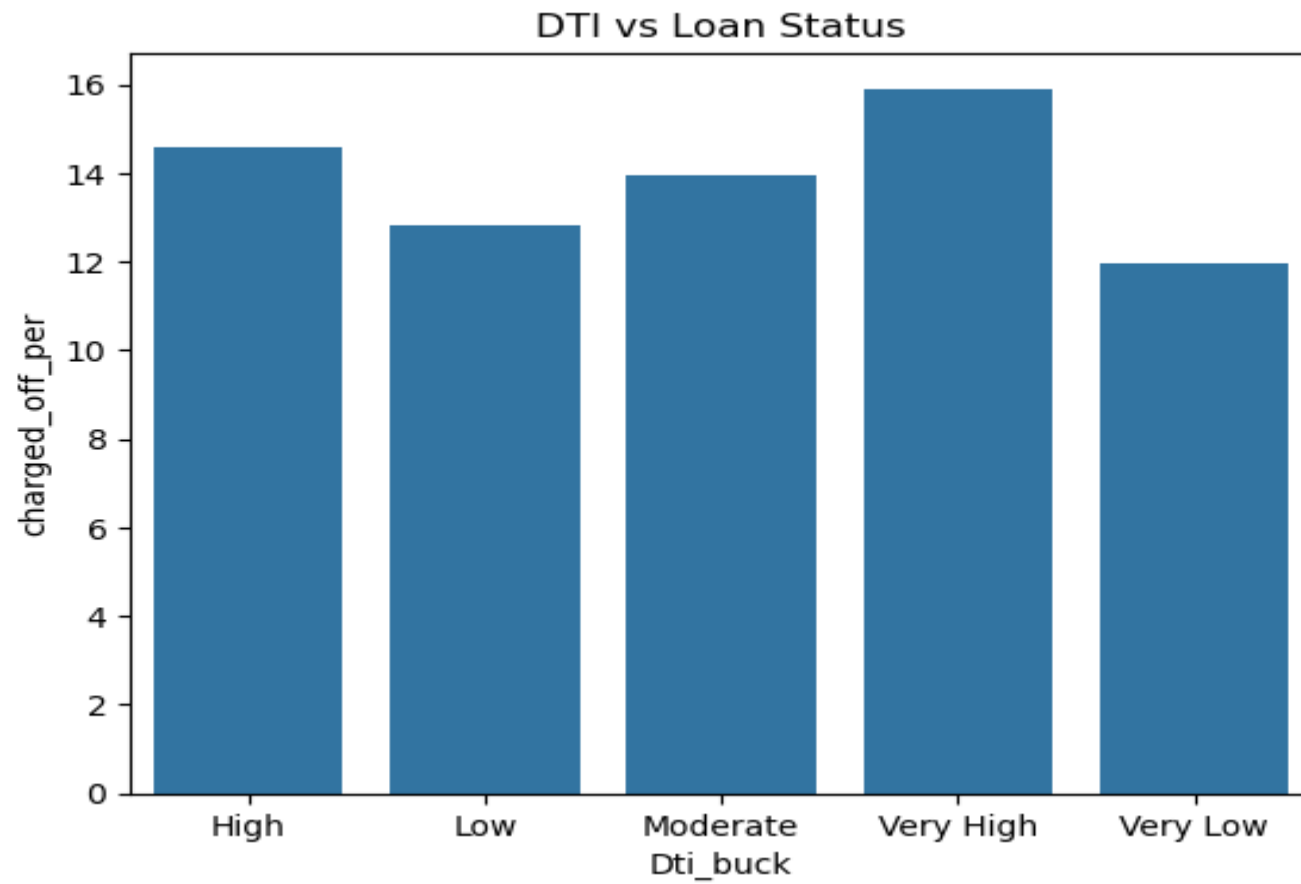
From the below analysis we can see that the people who has more interest rate has more chances of getting charged off compared to others
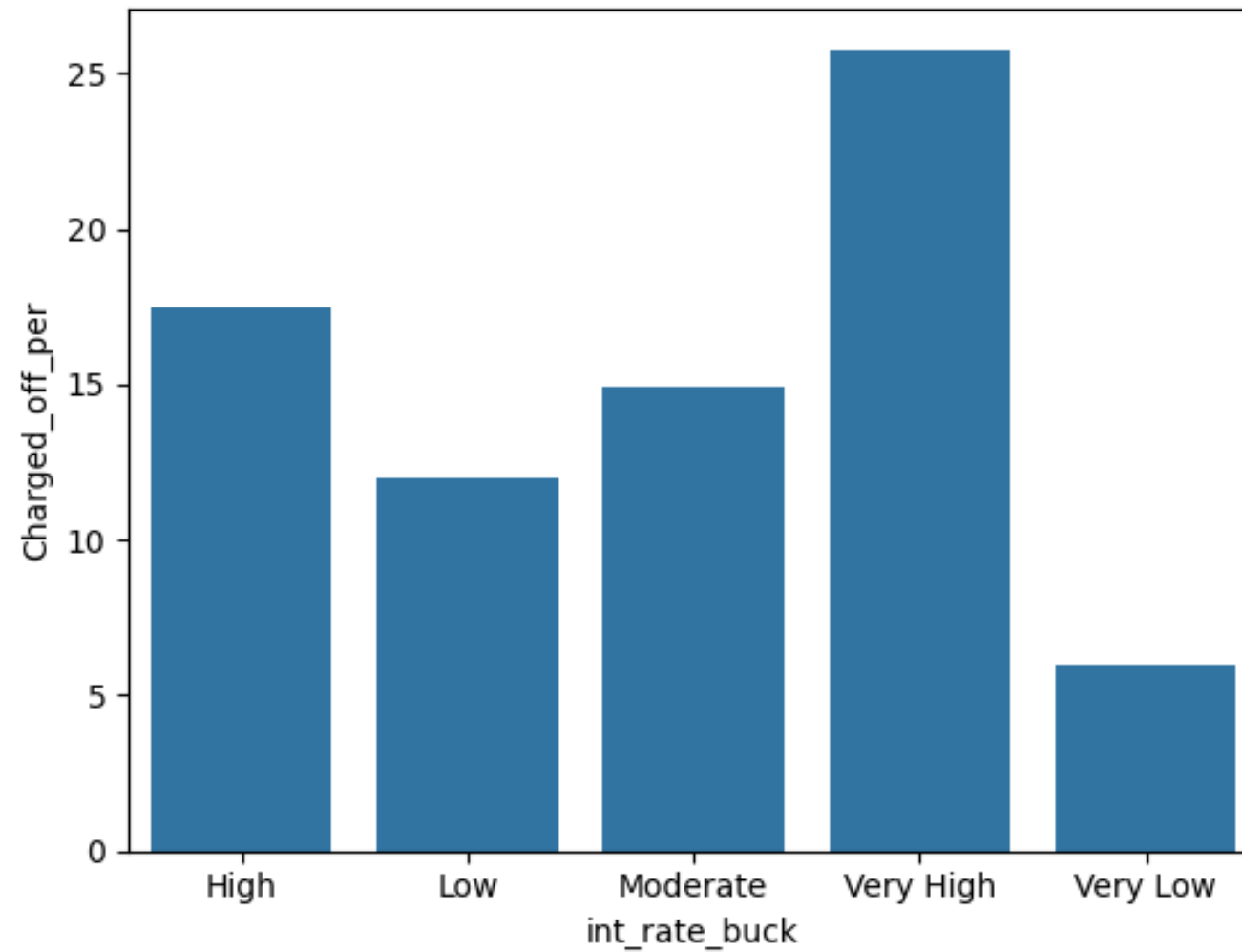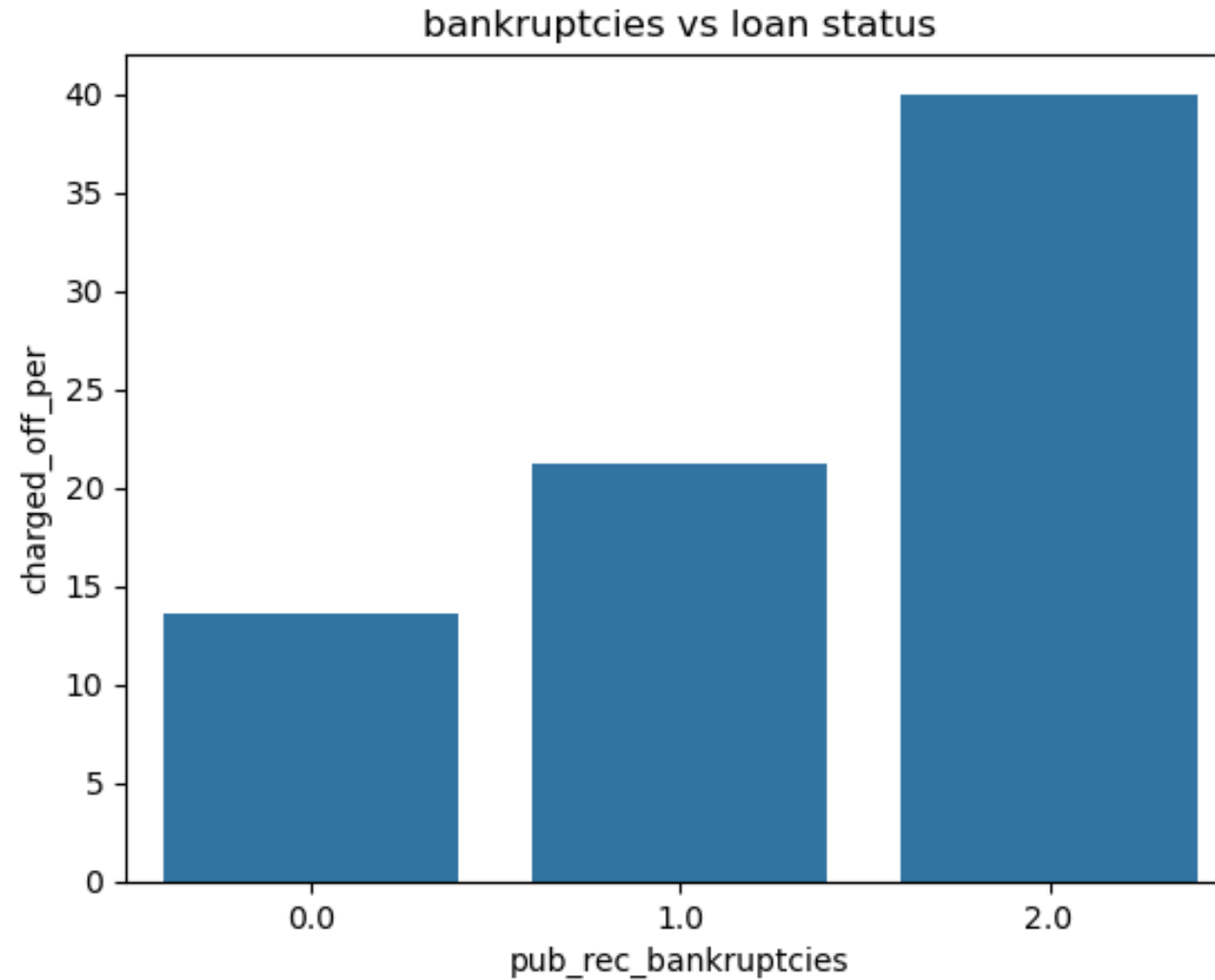
From the below graph we can see that the people who had took the loan for small business is more likely to get charged off
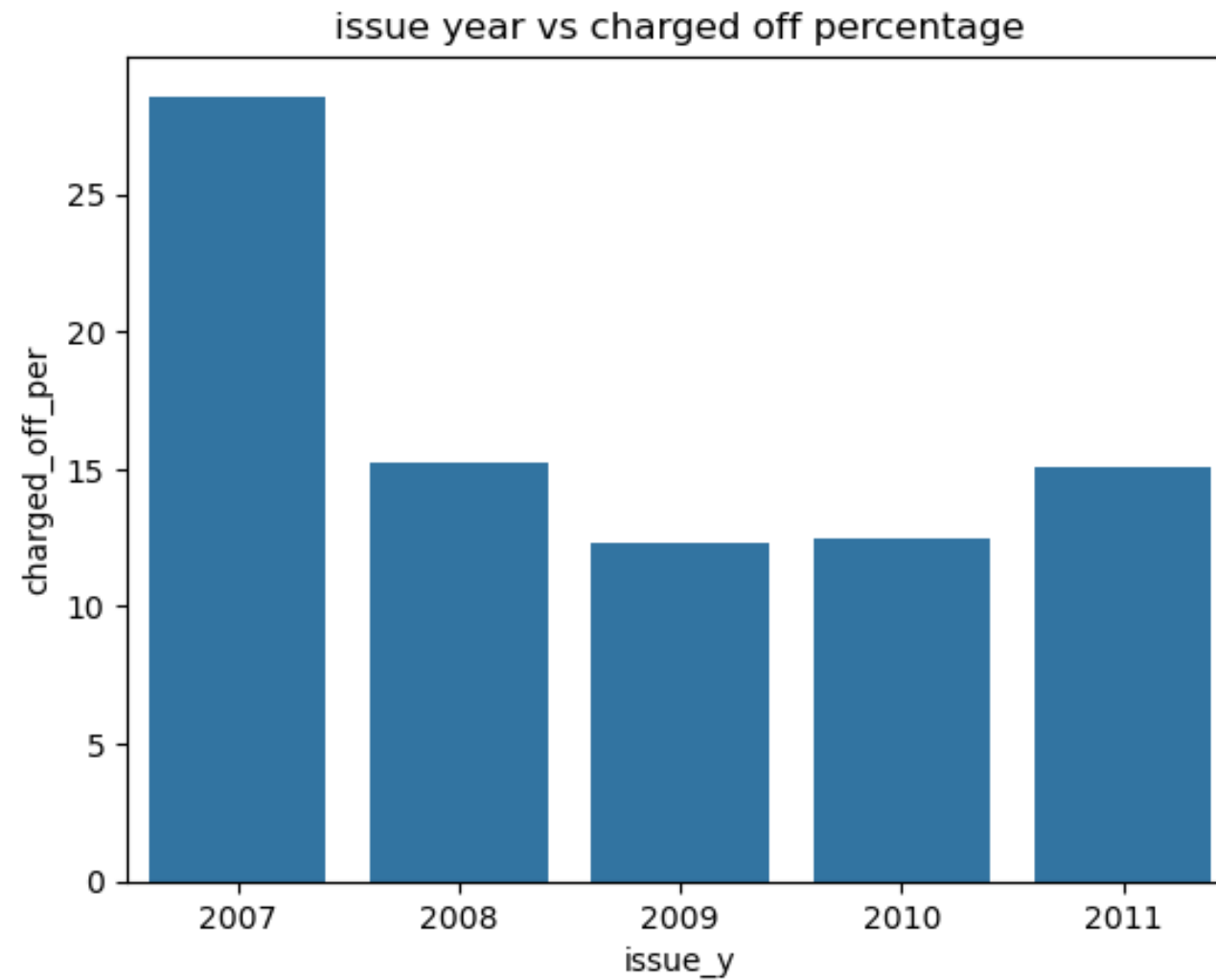

Loan purpose vs Loan status

People who has a very high DTI are more likely to get charged off

From the below graph we can see that the people having 2 bankruptcies are more likely to get charged off
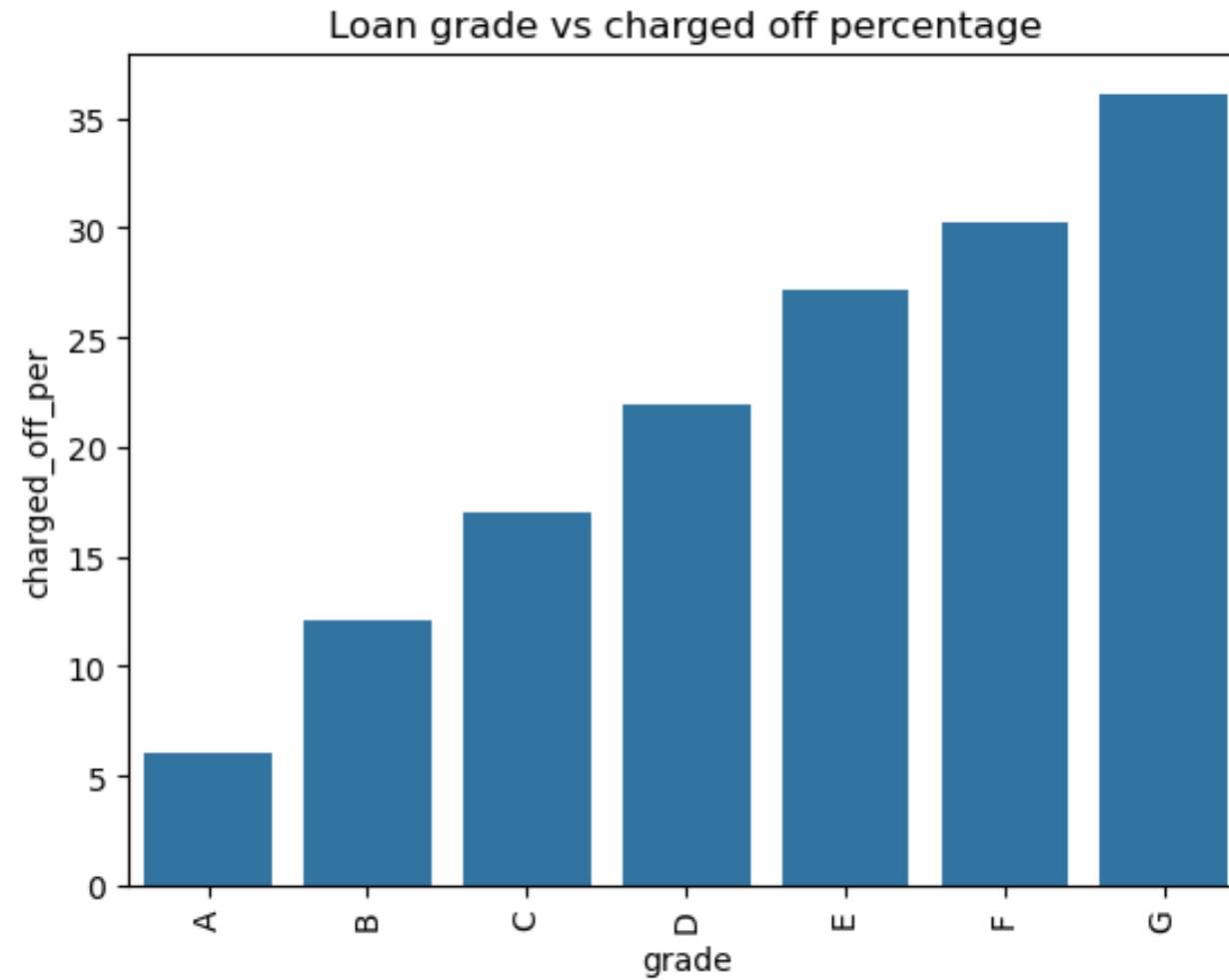
From the below graph we can see the people who took the loan in 2007 has more charged off percentage
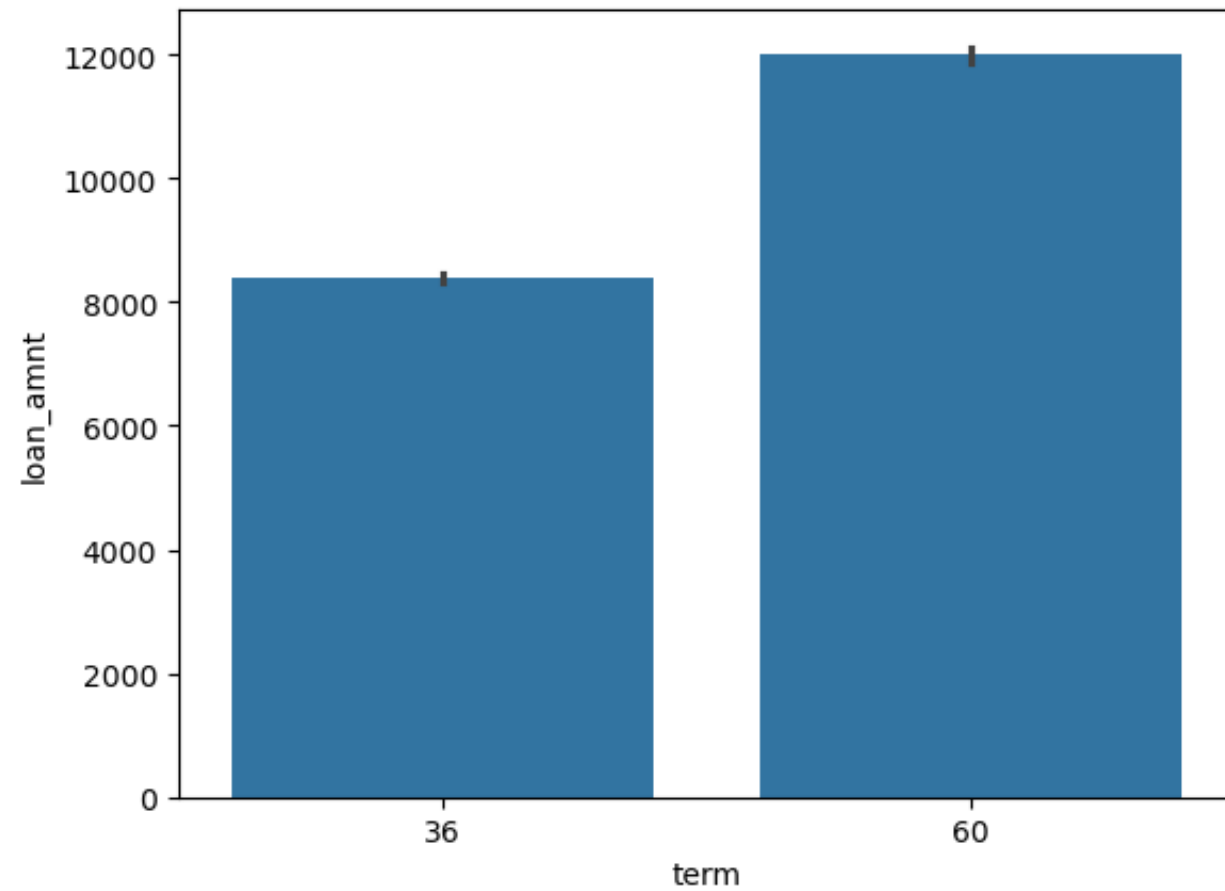

issue year vs charged off percentage

State address TN has more number of charged off loans

state vs charged off percentage
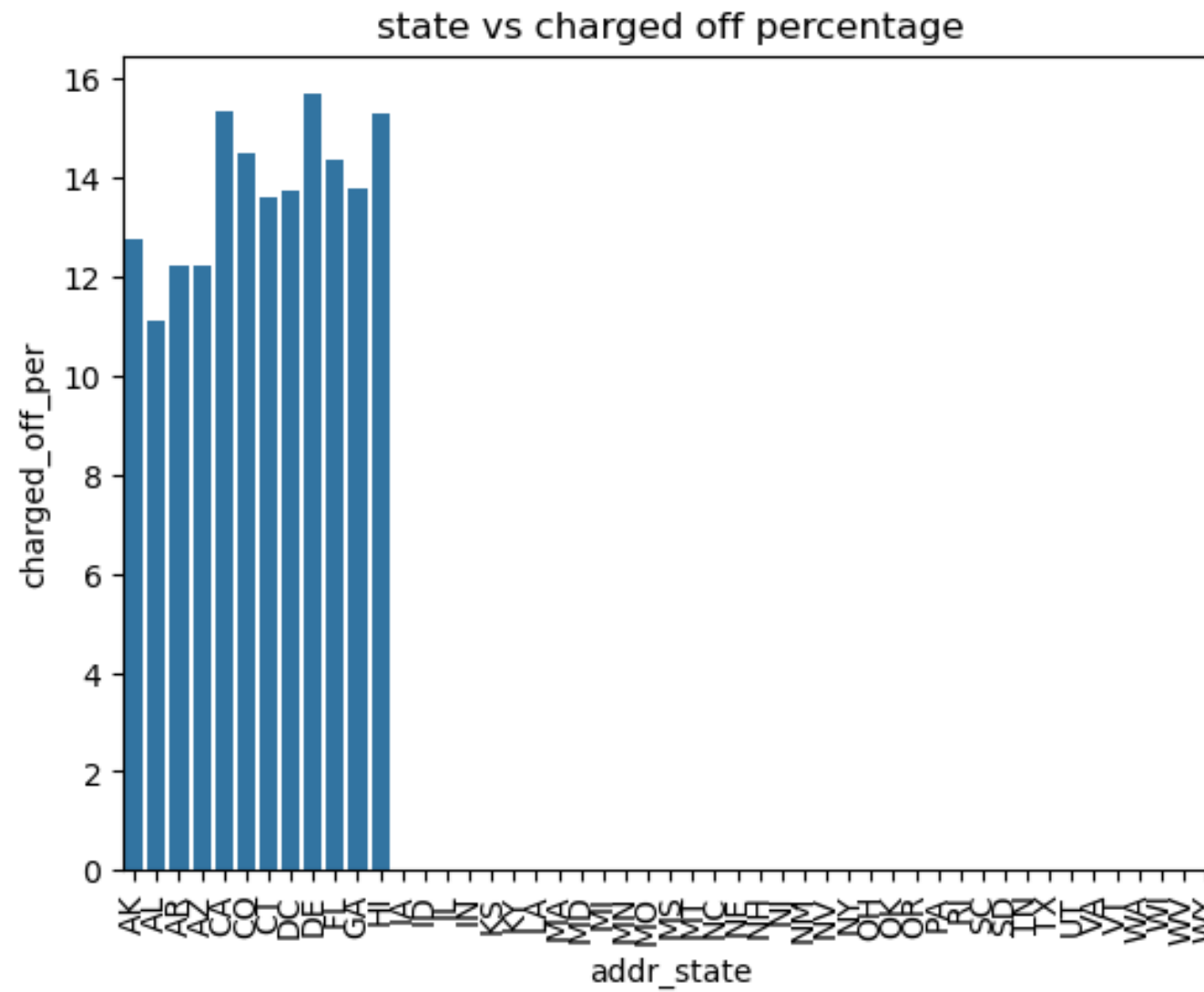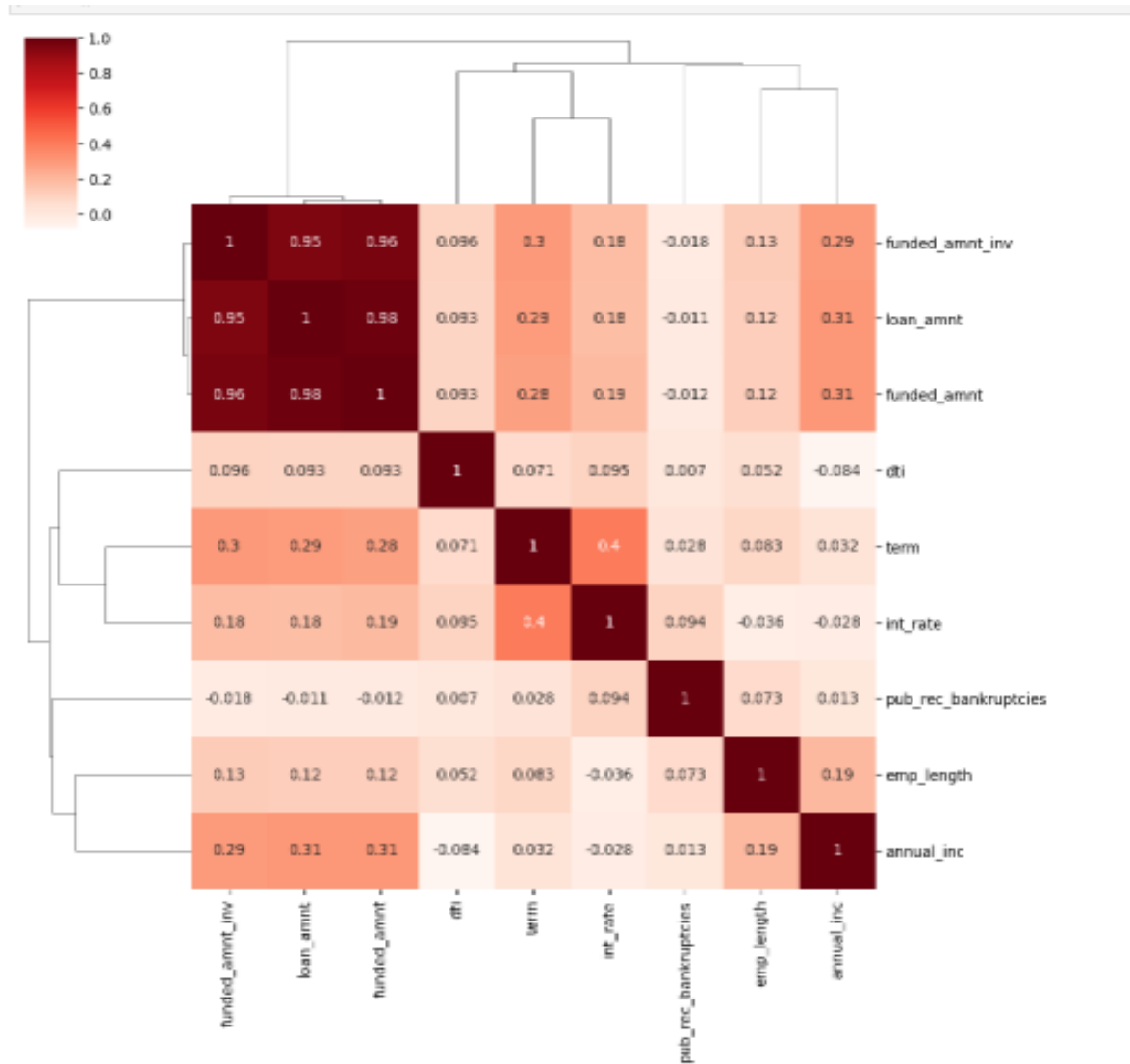
G grade loan has more number of charged off loans

Loan grade vs charged off percentage

From the below graph we can see that the term increases as the loan amount increases

Cluster Map

Conclusion

From the above cluster map we can conclude

Positive correlation :-
- Term has positive correlation with interest rate
- Funded amount inv has strong correlation with loan amount
- Loan amount has strong correlation with funded amount

Negative Correlation
- Bank ruptcies has weak correlation with funded amount inv, loan amount and funded amount
- Interest rate has a weak correlation with annual income
- DTI has weak correlation with annual income
- Employee length has weak correlation with interest rate