Lending Club Case Study

Surender Yadaram

Lending Club – Case Study

Company Details

- Lending club is consumer finance company which specializes in lending various types of loans to urban customers. This company 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.

Problem Statement

- The credit loss is the amount of money lost by the lender when the borrower refuses to pay or runs away with the money owed. In other words, borrowers who default cause the largest amount of loss to the lenders. In this case, the customers labelled as 'charged-off' are the 'defaulters'.

- In other words, the company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilise this knowledge for its portfolio and risk assessment.

Analysis Approach
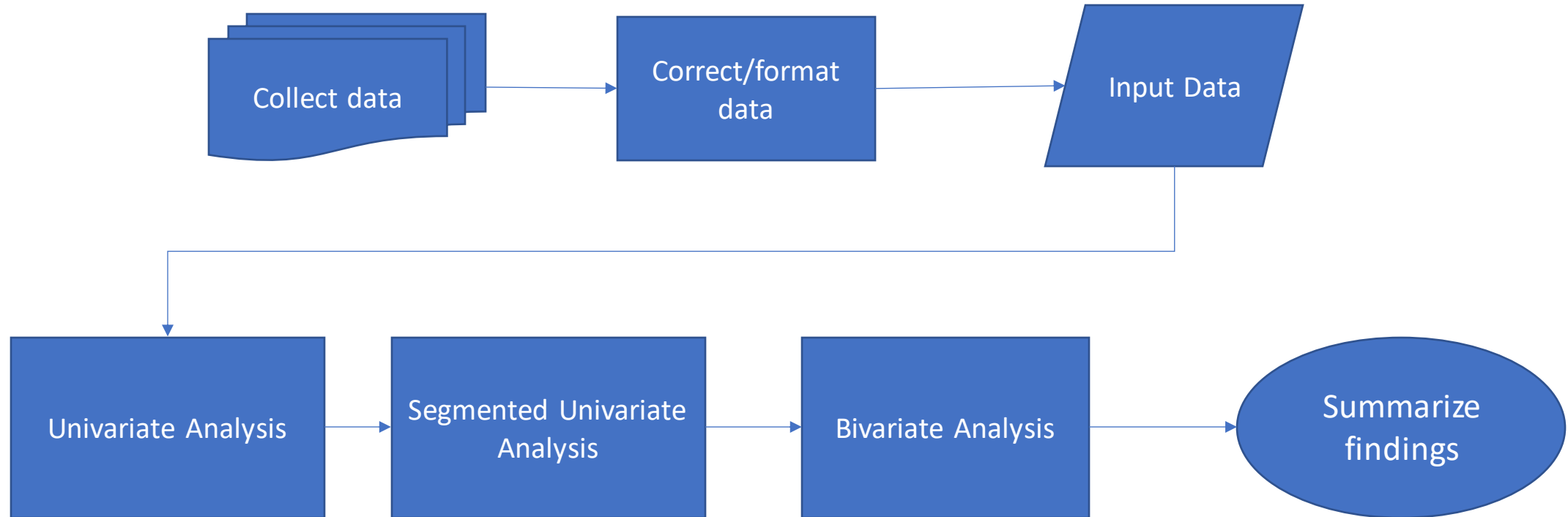
Data Understanding and Cleaning

- The Loan status "Current" is ignored as we don't know their future behavior of defaulting yet.

- The columns with missing data more than 60% are dropped from the analysis

- "int_rate" is converted to float and Loan status to numeric categoric values ("Fully Paid" to 0 and "Charged Off" to 1) for easy analysis

- Maximum loans were given for term of 36 months and 50% of the loans are amount less than 10,000.

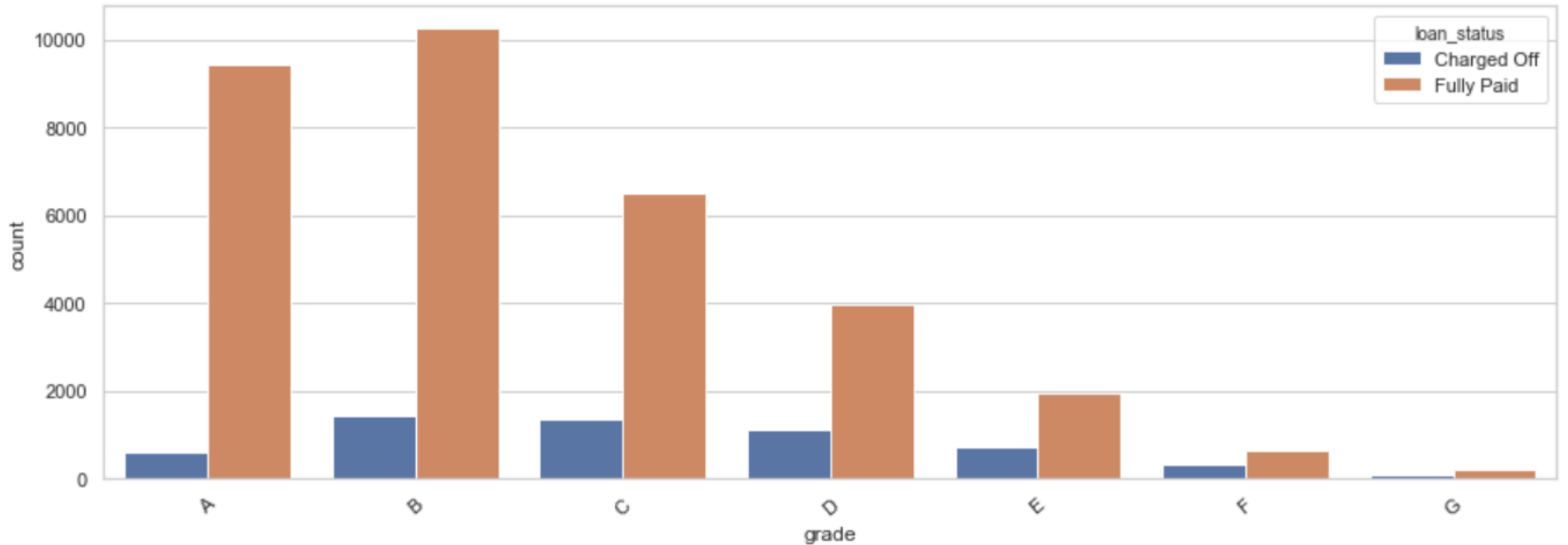- Amount wise, 84% of total amount is given for "Fully Paid" and 16% for the "Charged Off" customers

Boxplot grouped by loan_status
loan_amnt
[loan_status]
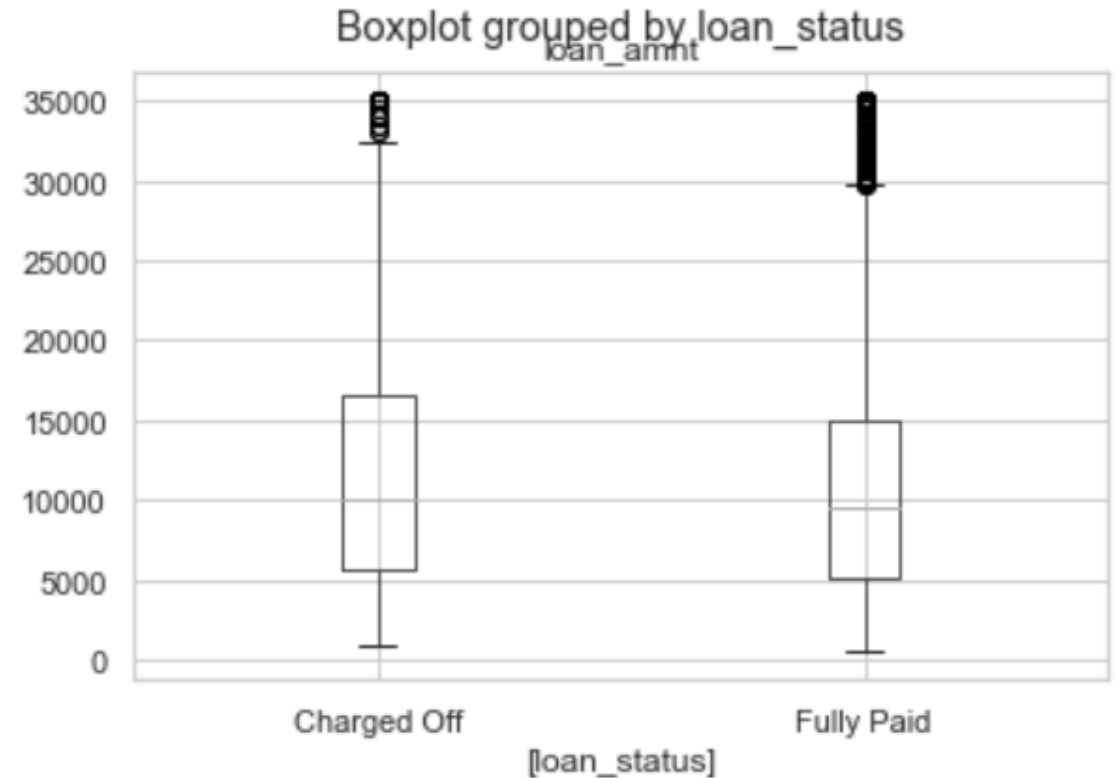
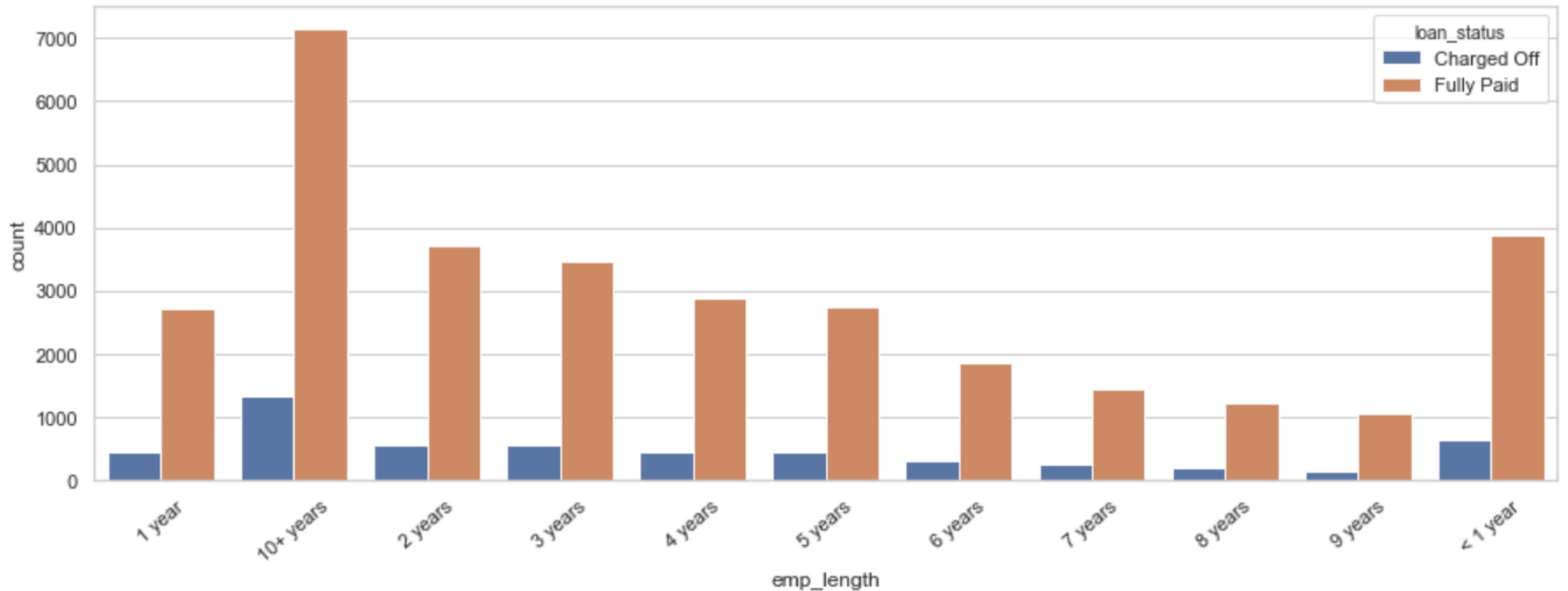# Loan Distribution by "Purpose"
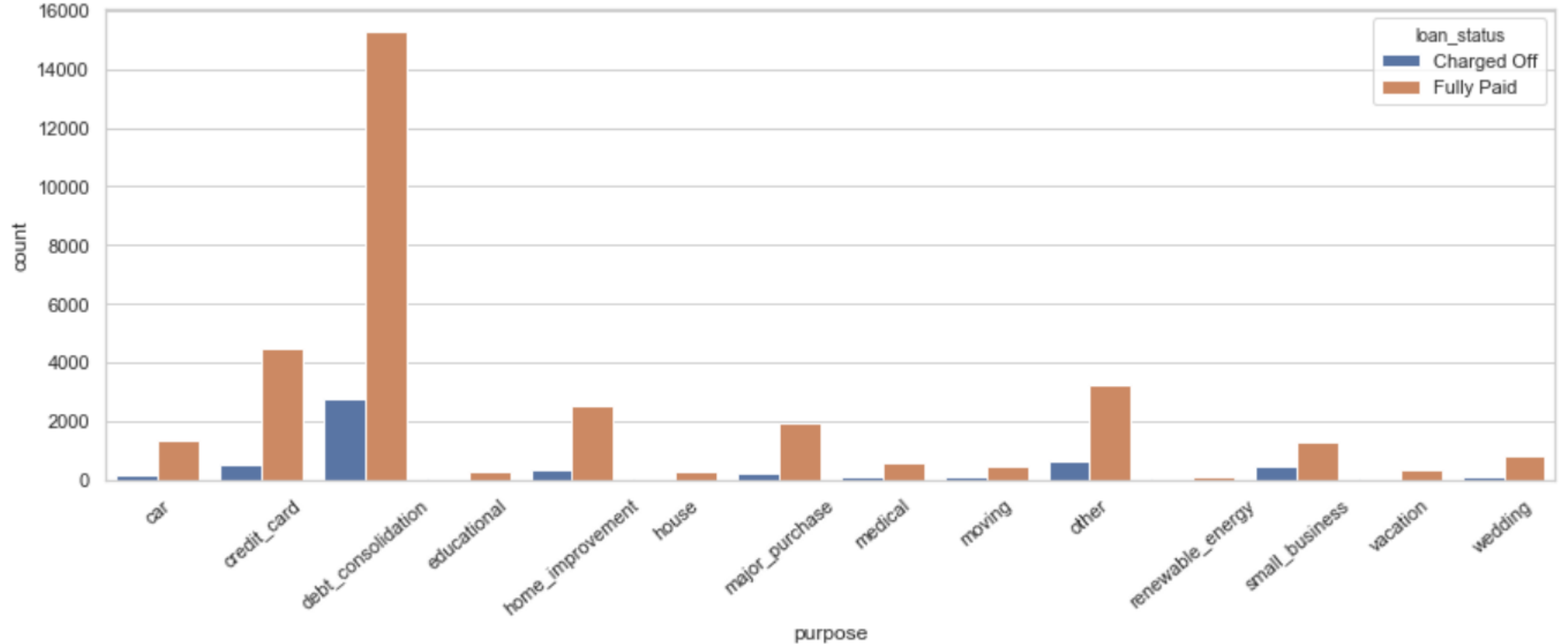
Most of the loans were given for "debt_consolidation" and "credit_card" purposes

Loan Distribution by "Grade"

As expected most of the loans were given for A and B customers and number of loans decreases as grade decreases

Loan Distribution by "Length of Employment"

Most of the loans were given to people having 10+ years emp_length. Interestingly for less than 10 years of emp_length, no of loans are decreasing with increase of emp_length

# Loan Distribution by "Home Ownership" and "Verification status"
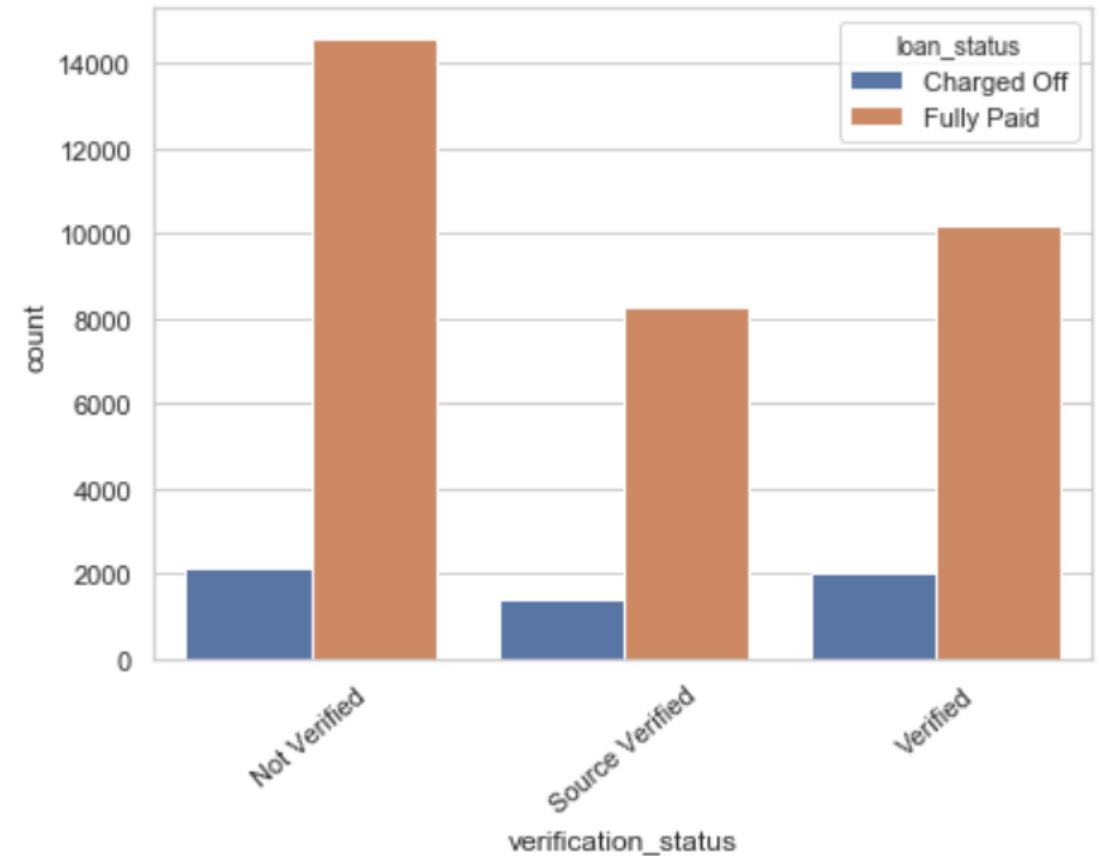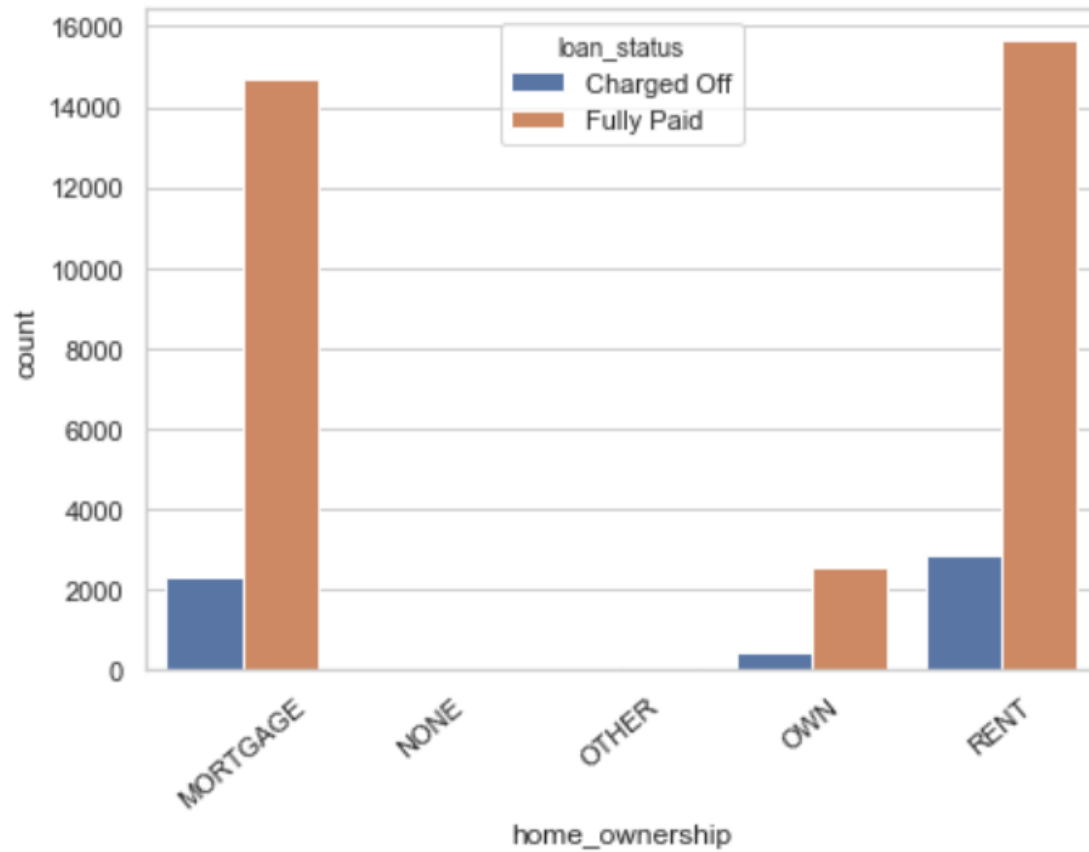
✓ Distribution of loans with home_ownership is in the order Rent, Mortagage, Own and etc
✓ Distribution of loans with verification_status is in the order Not Verified, Verified and Source Verified

Overall Loan Status
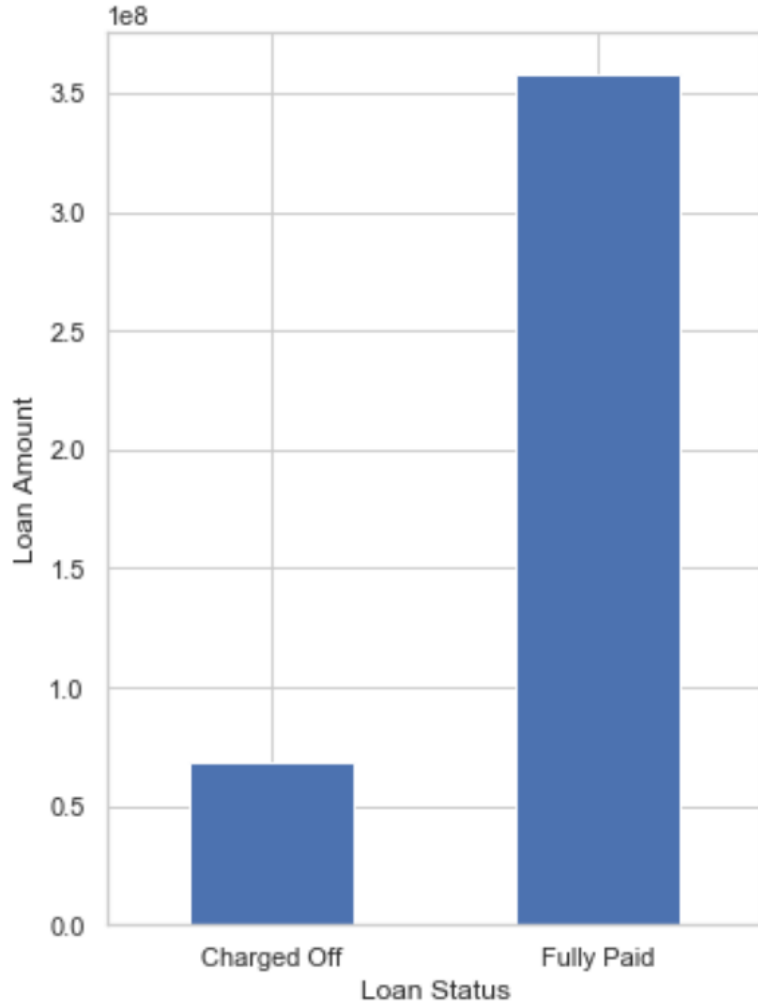
- LC earned 19.4% profit on "Fully paid" loans where as only 56.5% recoveries on the money lent to "Charged off" customers

- Around 14.6% of the total loans provided are defaulted.

- Further analysis is performed based on this data to figure out the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default.

- This can be utilize as knowledge for portfolio and risk assessment.

Loan Amount Distribution
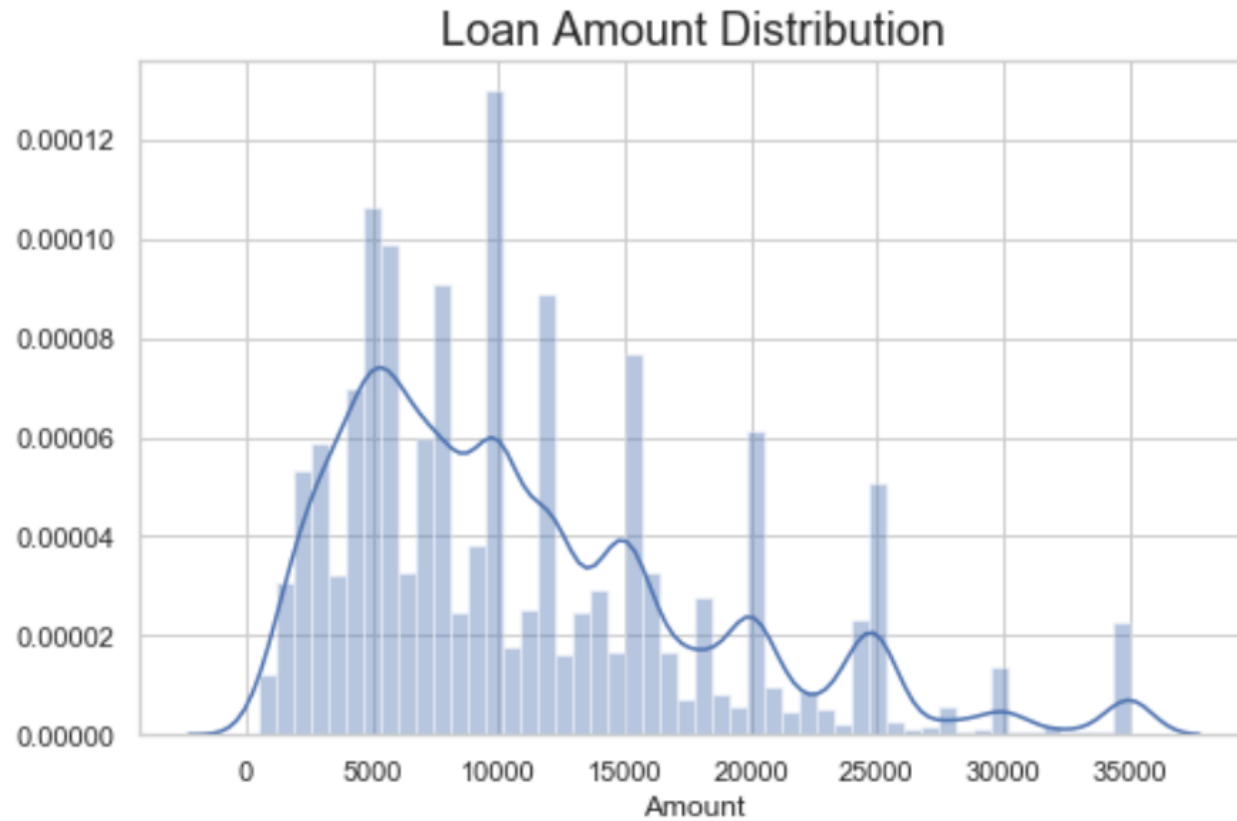
- It is observed that the amount observes a jump after every 5000 interval.

- Costumers are more inclined towards rounded amount.

Defaulters based on Home Ownership
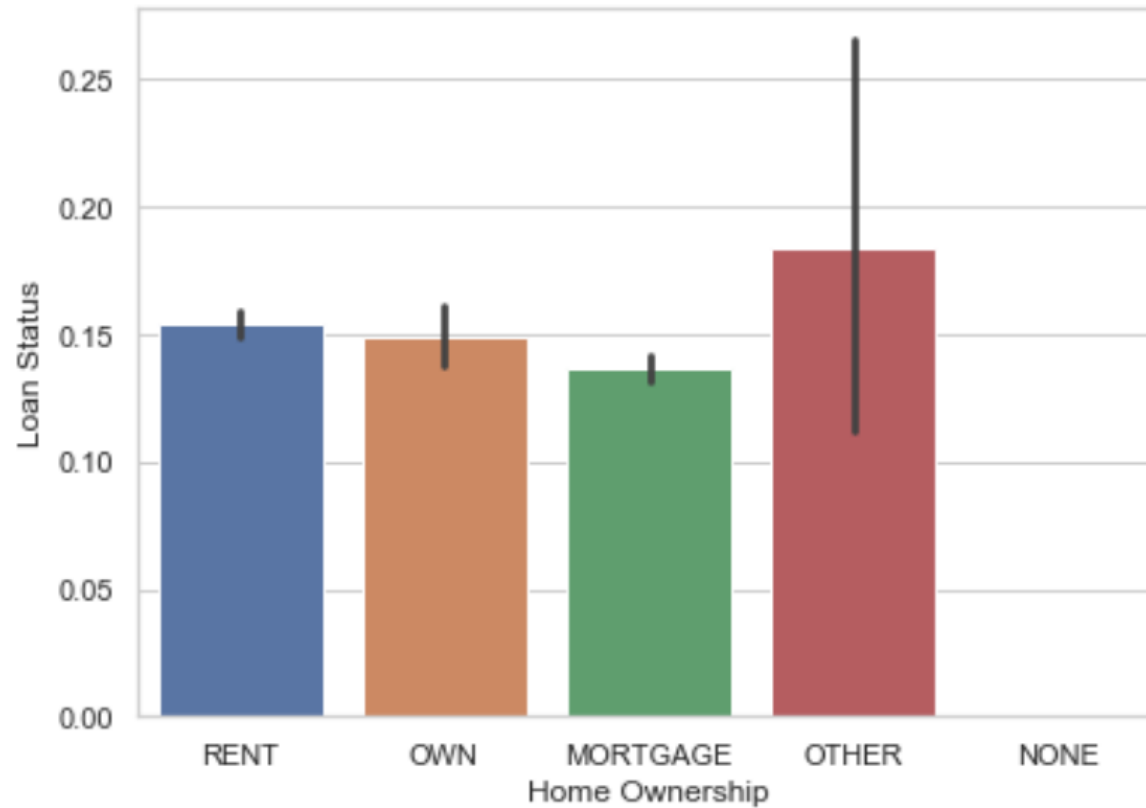
- People having either own or mortgage home are not part of defaulter as compared to others.

- Though observation wise, "Others" category is performing worst in case of home ownership we cannot conclude it due to lesser number of data points

Defaulters based on Loan Purpose
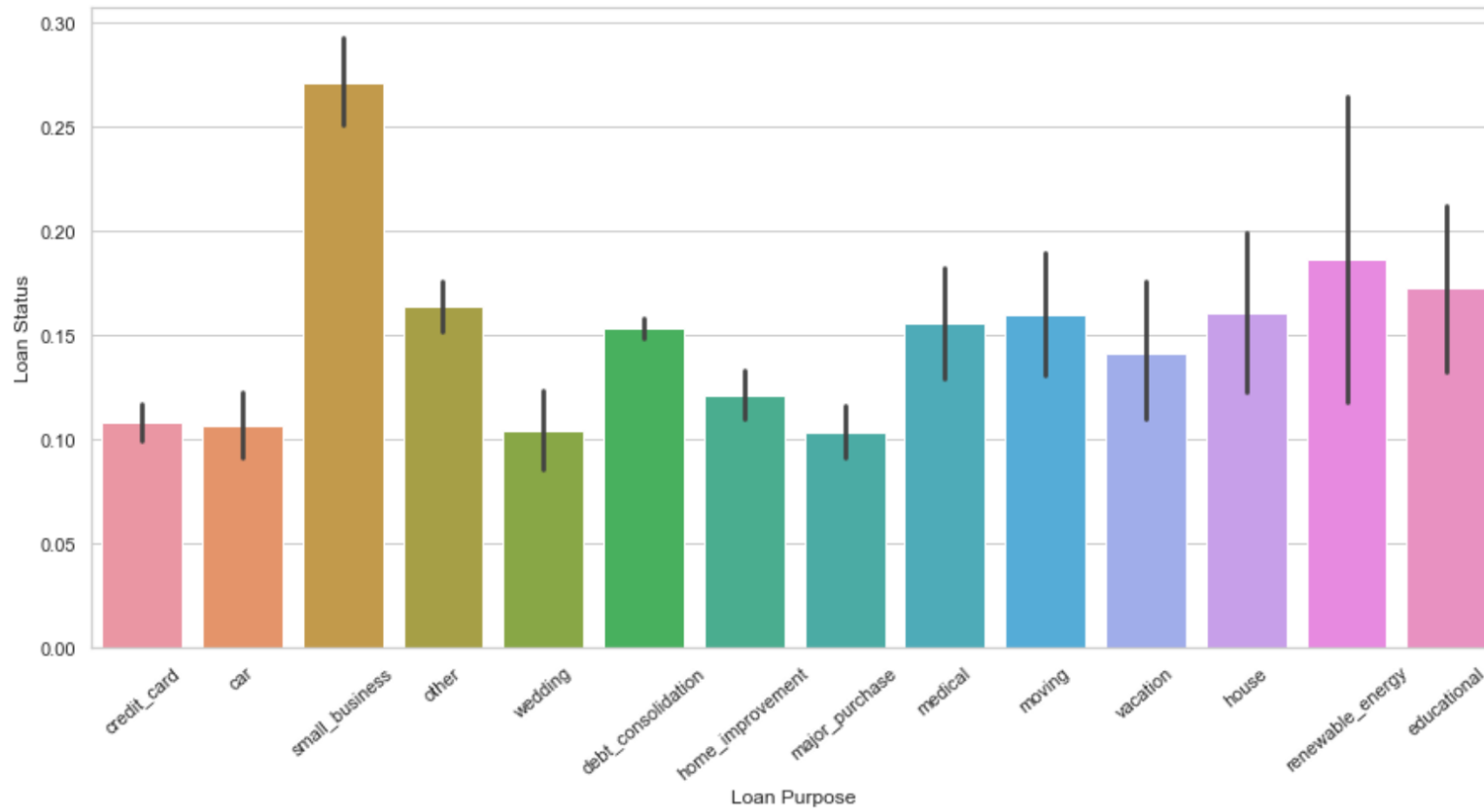
- Loan taken for the purpose of small business is performing the worst as compared to other purposes for which loan is taken.

- Organization has to be very careful while extending loan for small business purpose.

Defaulters based on Employment Length
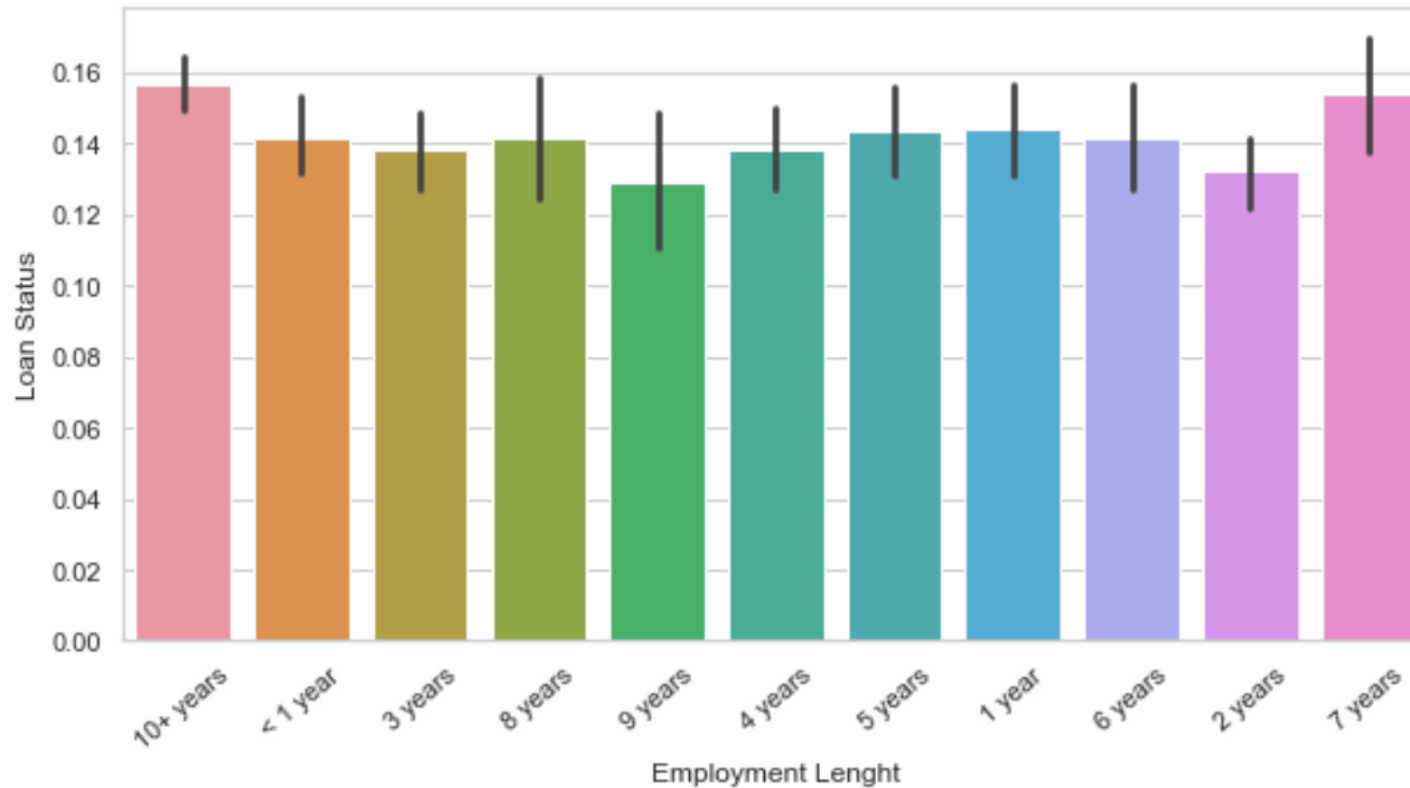
- It is observed that there is no significant change in defaulters based on their employment length.

- This factor does not play any vital role to identify a defaulter.
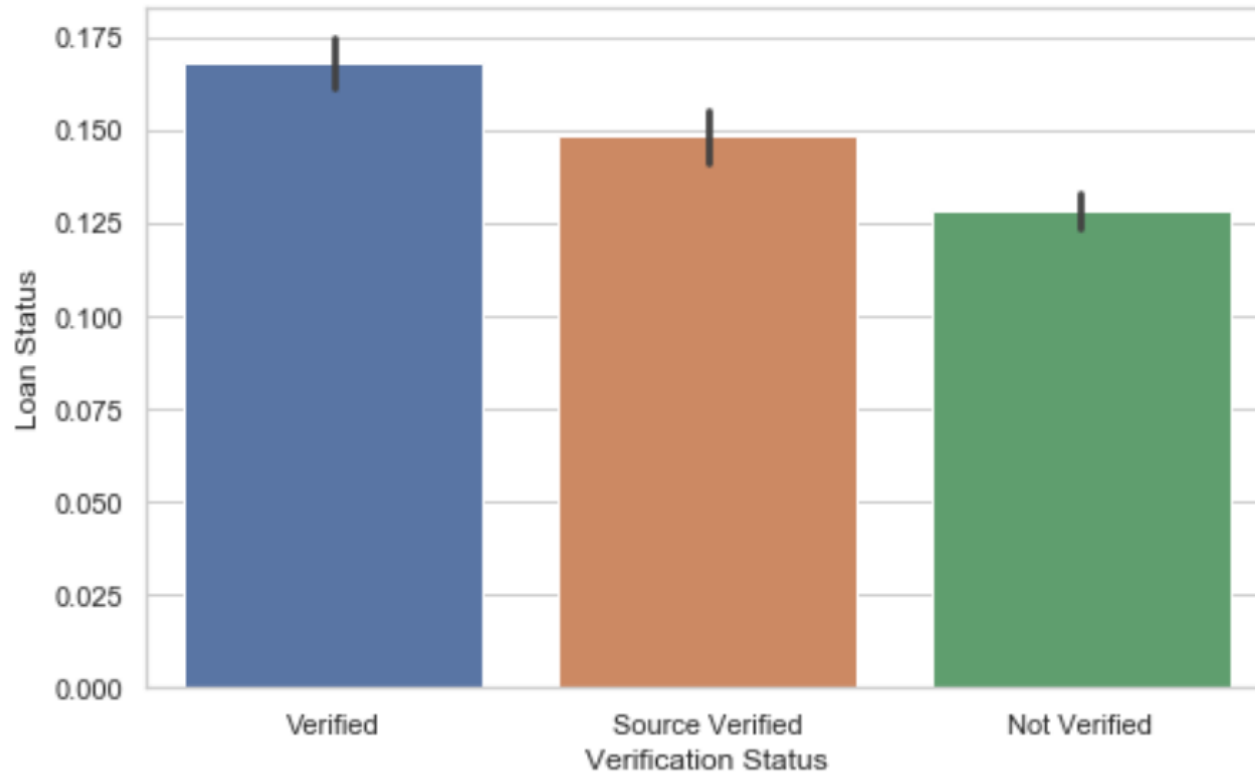
Defaulters based on Grade
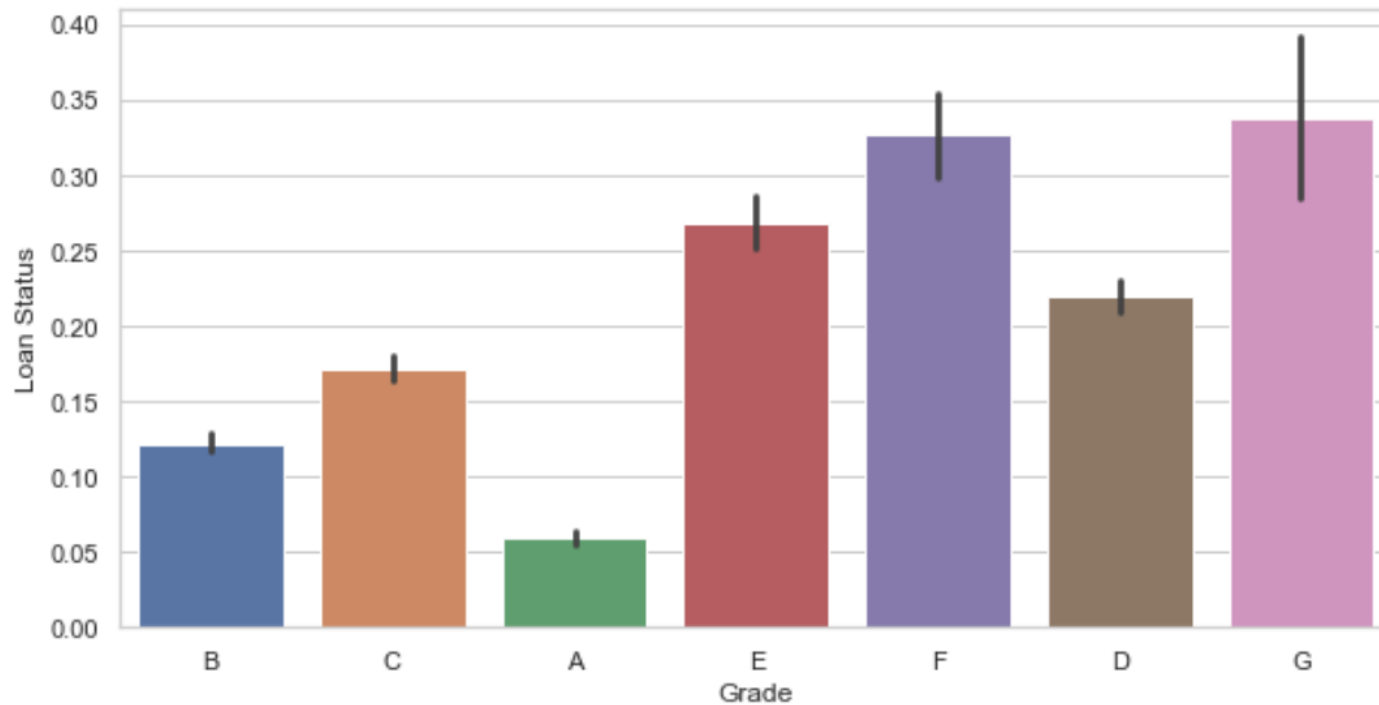
- Grading system is working as designed.

- We have less defaulters with grade A as compared to B and so on.

- This factor will not play any significant role further.
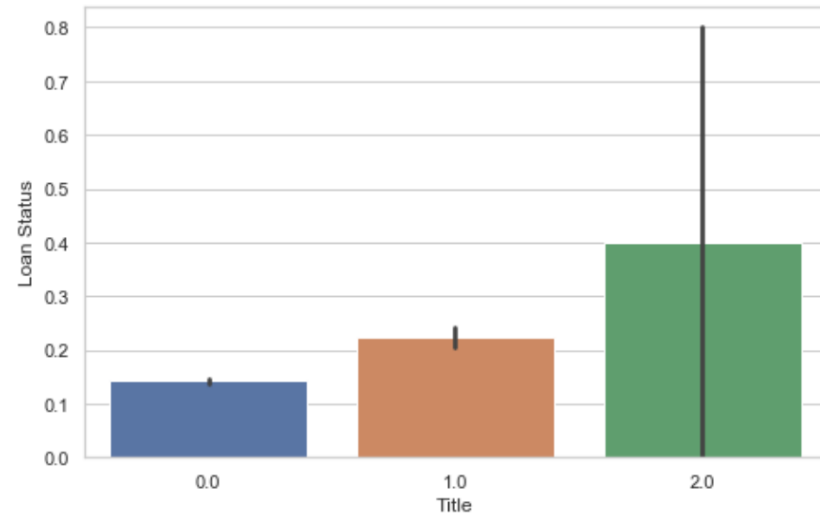
Defaulters based on Interest Rate
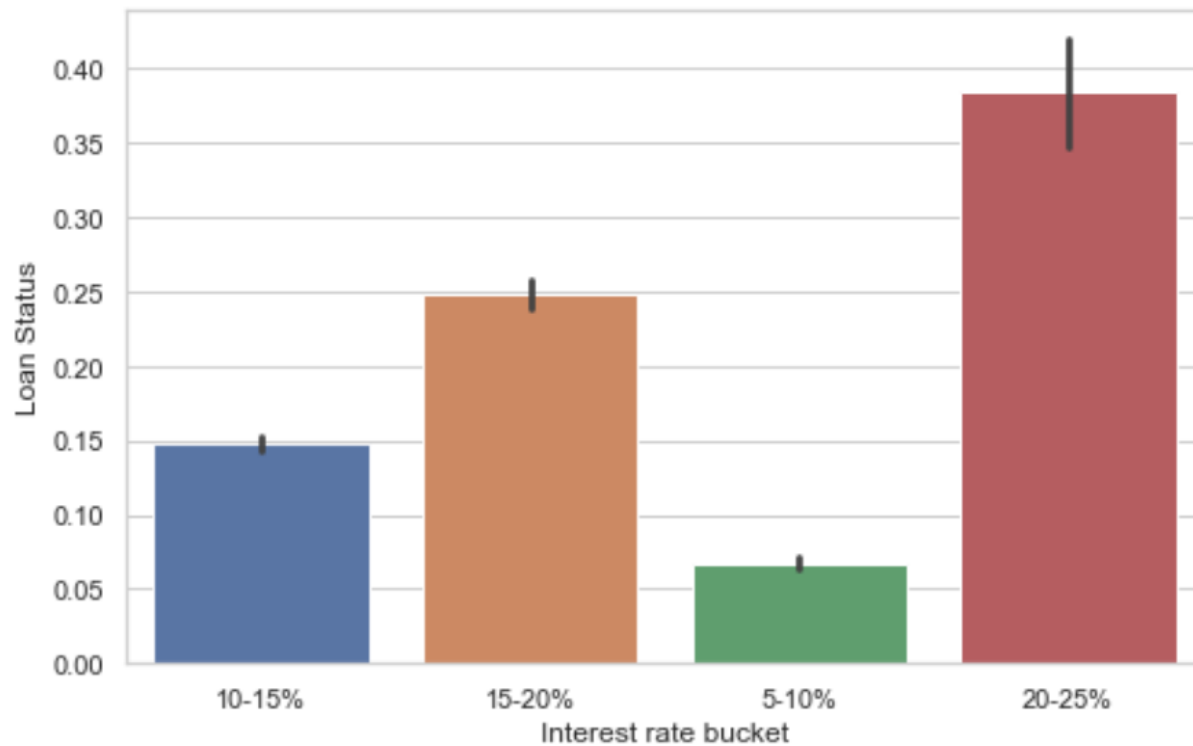
- As the interest rate increases, defaulters increases as well.

- Higher interest rate would be provided to people with low grade.

- Proper checks to be in place before providing loans with higher interest rate.

Defaulters based on Verification Status

- It was observed that the loans which were verified have highest number of defaulters.

- Clearly verification process needs to be analysed and corrective measure needs to be adopted immediately.
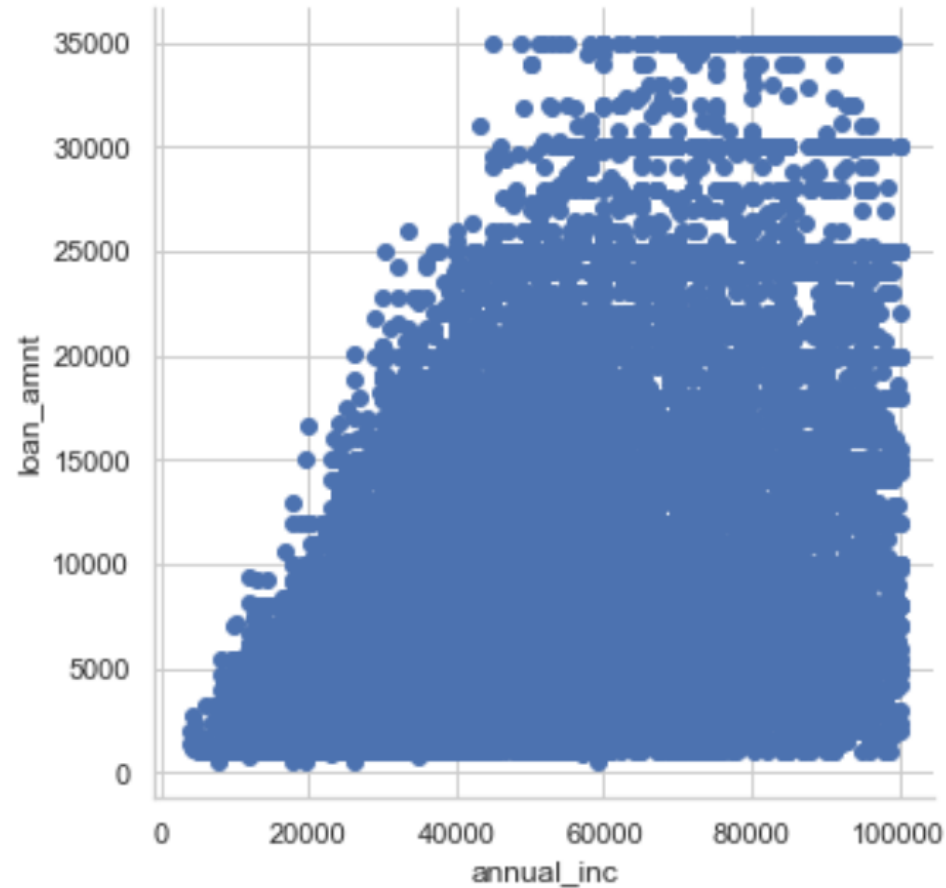
Defaulters based on Prior Bad Records

- There are instances where people having prior bad records have been provided loan with low interest rate as well as high amount.
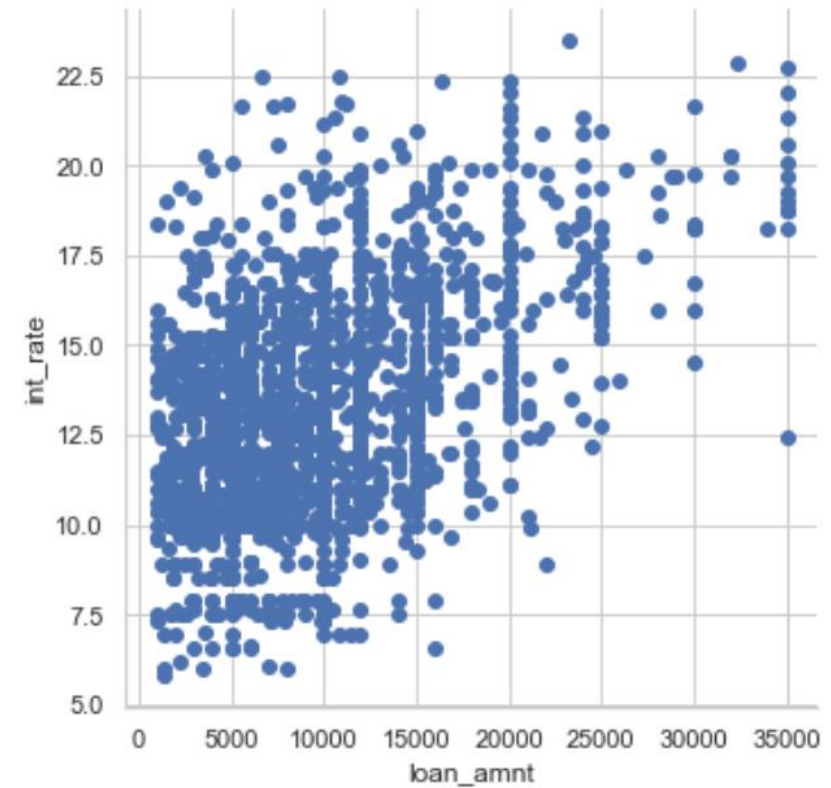
Graph plotted with people having greater than 0 bad records

Defaulters based on Annual Income

- It was observed that people having lower annual income were provided with higher amount of loan.

- These kind of loan has higher possibility of getting defaulted in near future.

- Not more than 50% of annual income should be provided in form of loan

Conclusions

1. Verification process needs to be rechecked. Presently verified loans are defaulted more.

2. Loan should be provided to customers who either has own or mortgage home

3. Higher amount or lower interest rate should not be provided when loan purpose is small business.

4. Stop providing higher loan and lower interest rate to customers who has prior bad history.

5. Higher loan amount should not be provided to low annual income customers. Loan amount should not exceed 50% of total annual income.