



Lending Club Case study

Problem Statement

- Lending loans to 'risky' applicants is the largest source of financial loss(called credit loss). 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.
- Lending Club wants to find such applicants who will risky applicants before giving them any loans

Analysis Approach

- To understand the consumer attributes and loan attributes influencing towards the tendency of the default of the applicant, and by identifying those attributes it will easier to reject risky applicants.

- Post clean up of data we found the following qualitative and quantitative variables which is influencing the default of the applicants :

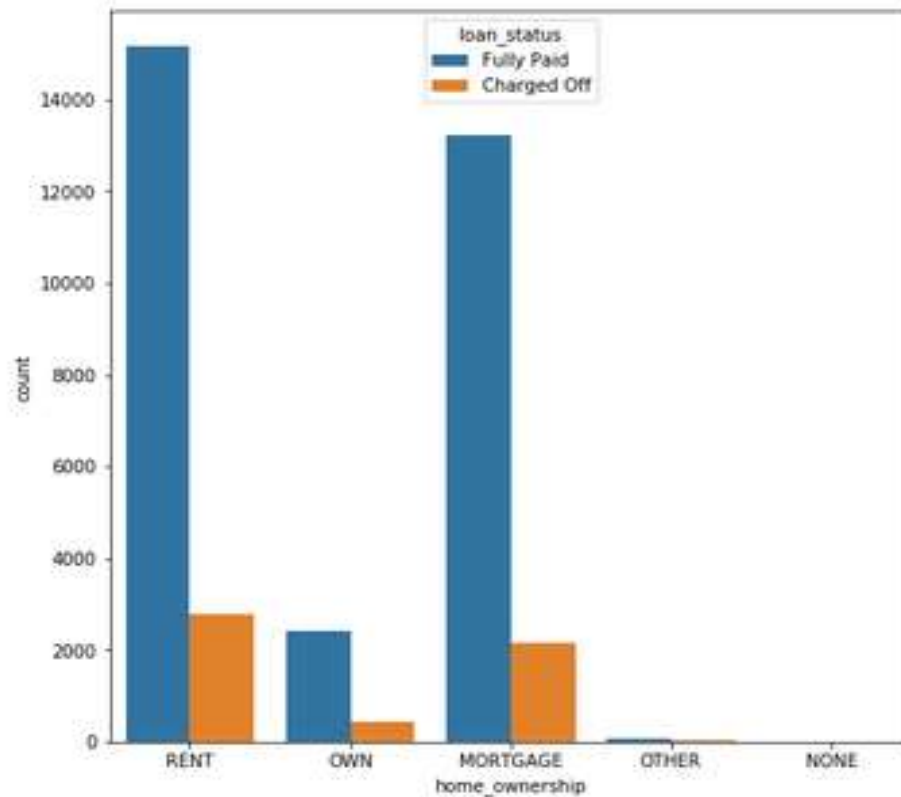
▫ 1.

S.No	Qualitative Variables
1.	<u>verification status</u>
2.	<u>term</u>
3.	<u>grade</u>
4.	<u>pub_rec bankruptcies</u>
5.	<u>pub_rec</u>
6.	<u>sub_grade</u>
7.	<u>home_ownership</u>
8.	<u>loan_status</u>
9.	<u>purpose</u>

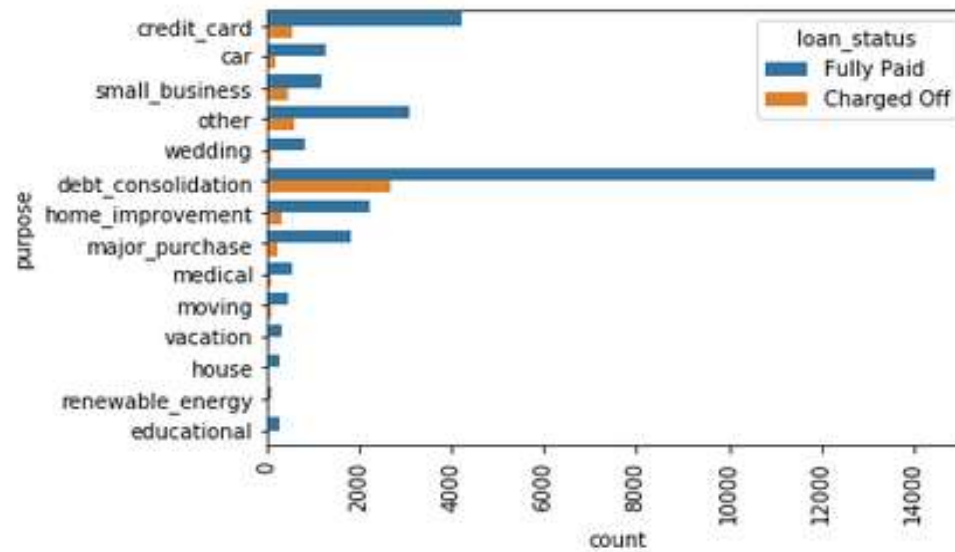
S.No	Quantitative Variables
1.	<u>emp length</u>
2.	<u>dti</u>
3.	<u>annual inc</u>
4.	<u>int rate</u>
5.	<u>funded amnt inv</u>
6.	<u>funded amnt</u>
7.	<u>loan amnt</u>

Univariate Analysis

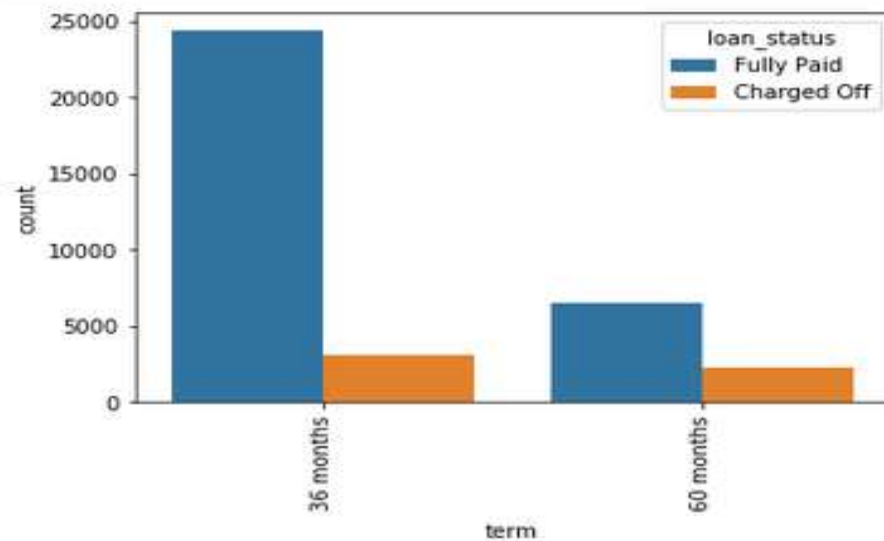
- In this Analysis we will try to see mostly likely reasons for the loans to default, it is only possible with Fully Paid and Charged Off Data.
- 1.) Home Ownership



- Purpose

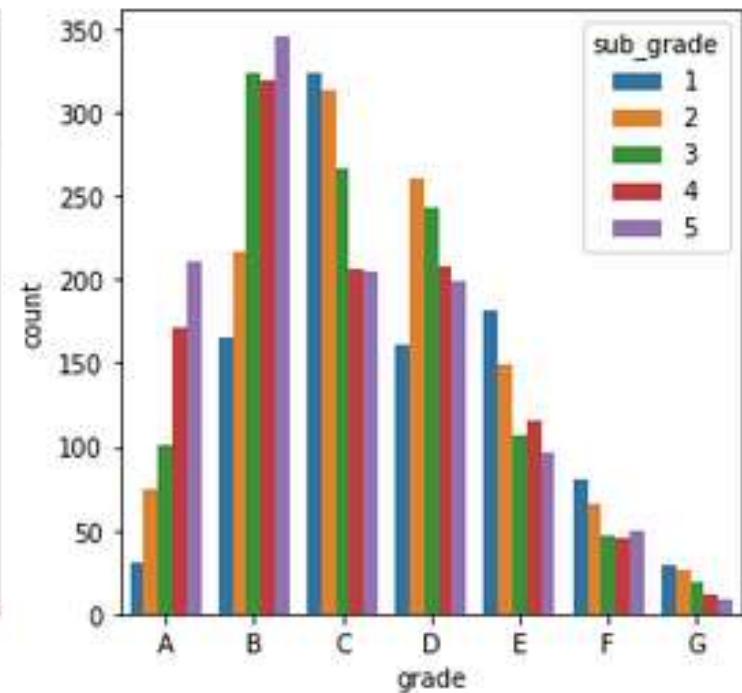
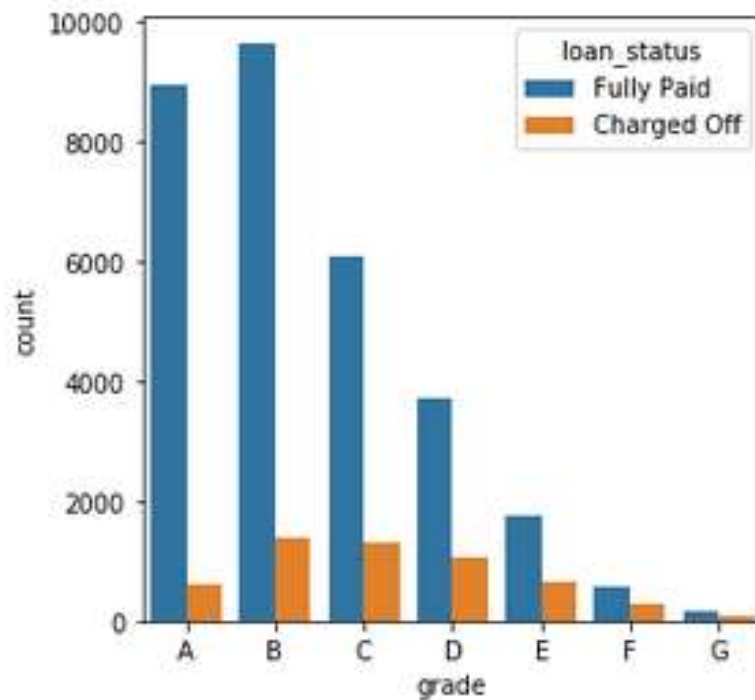


- term



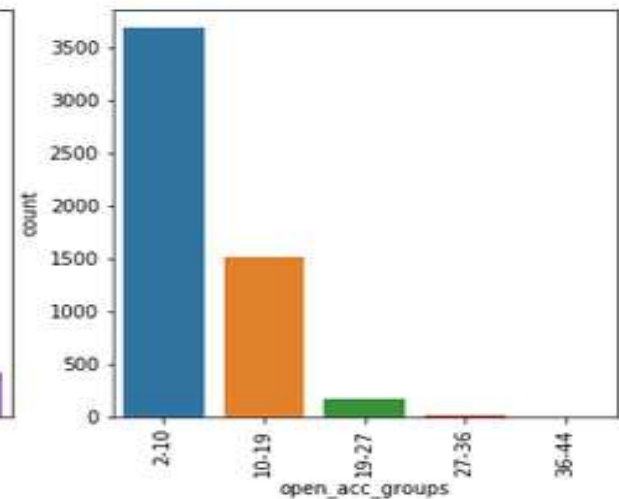
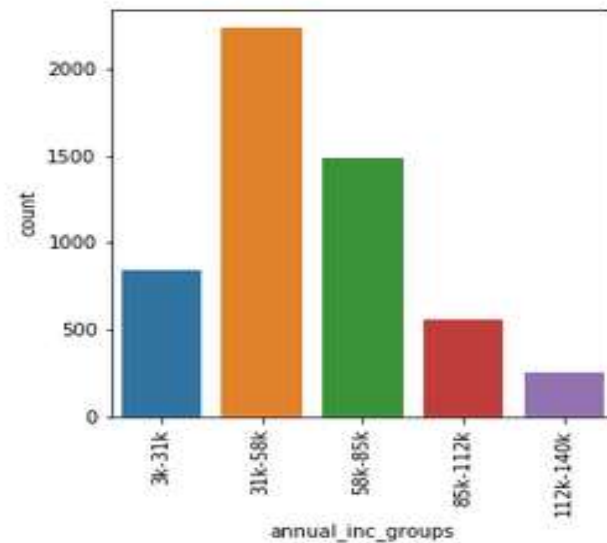
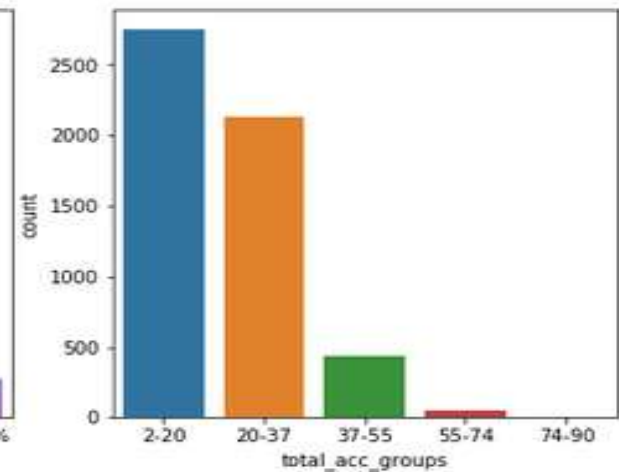
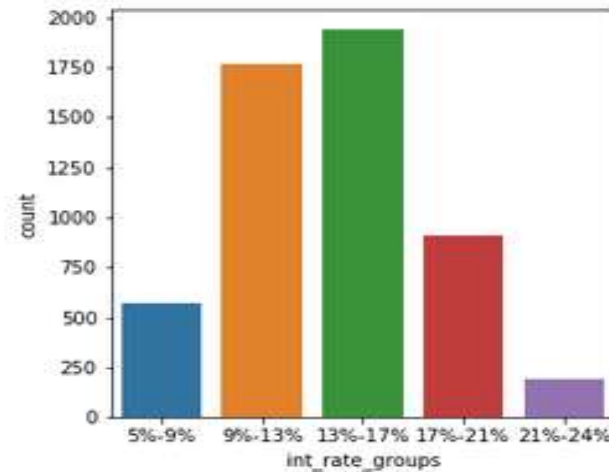
Segmented Univariate Analysis

- Grade

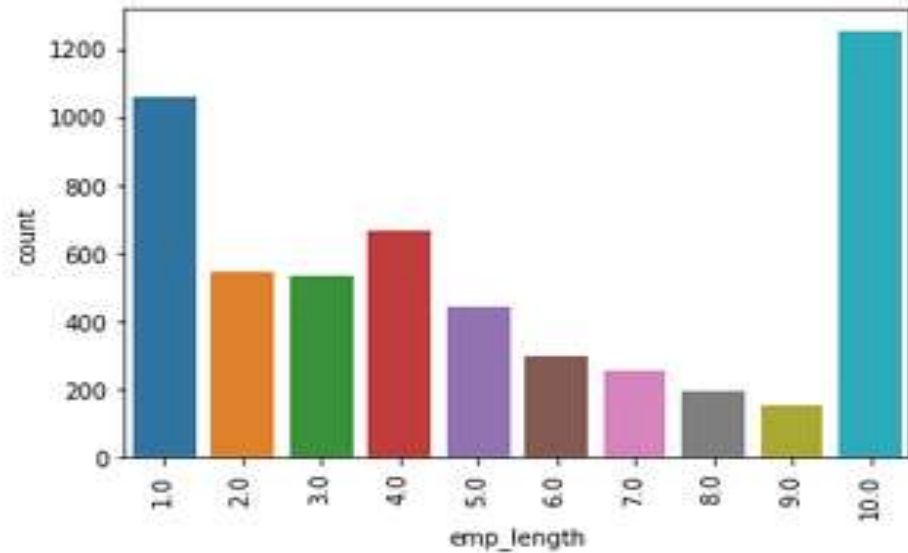


Numerical Variables converted to categorical variables

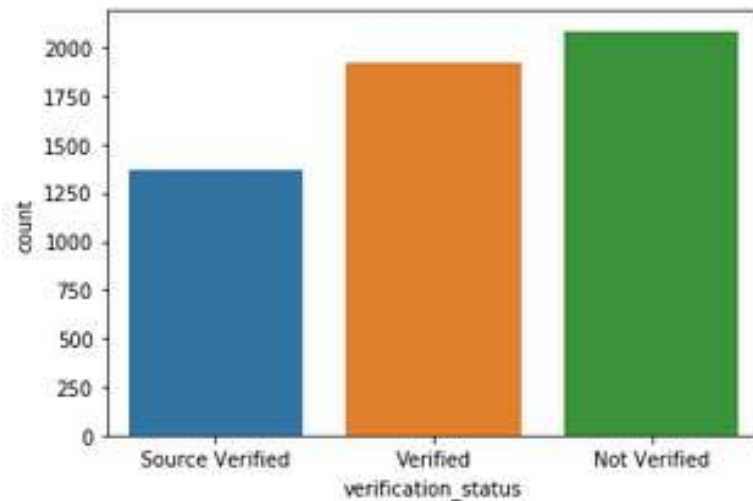
- Interest rate groups
- Total Account groups
- Annual Income groups
- Open Account groups



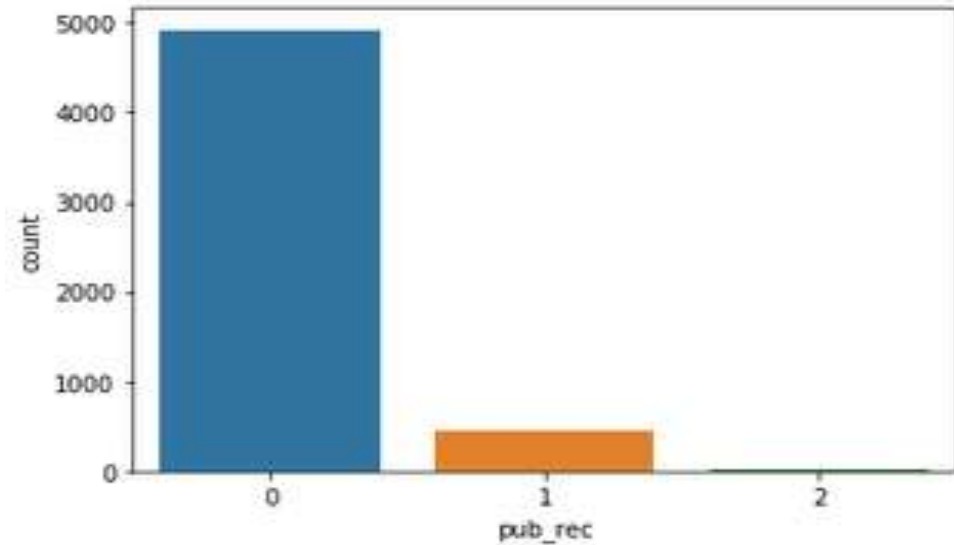
- Emp length



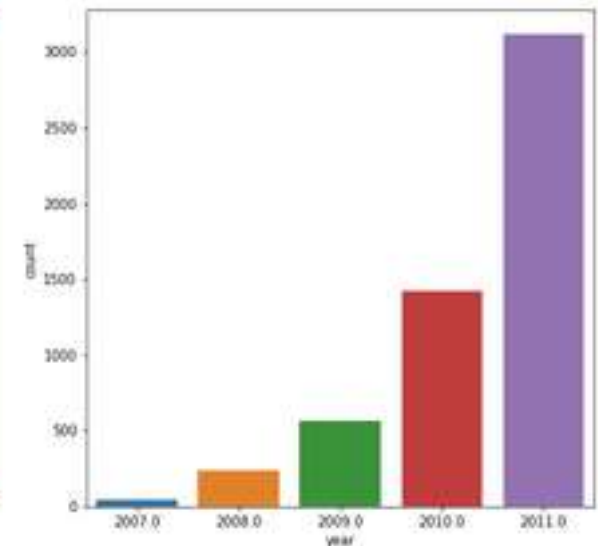
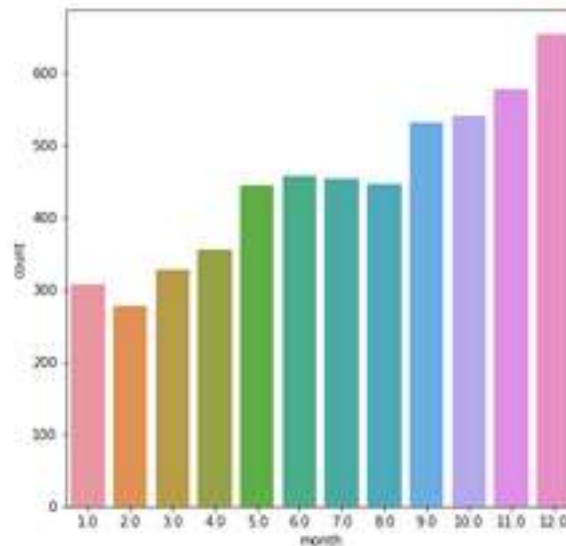
- verification_status



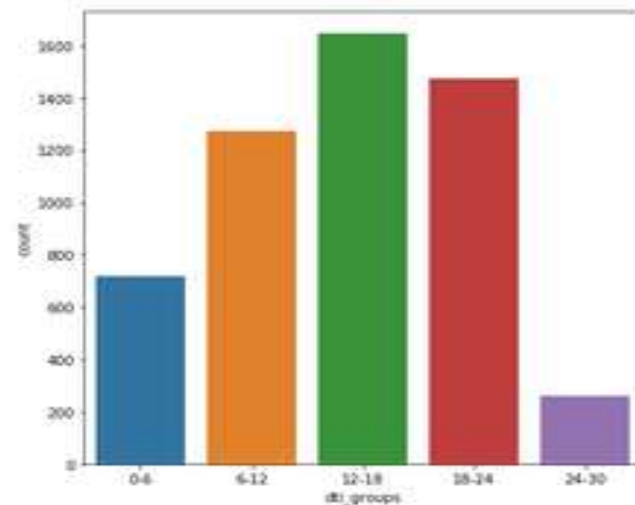
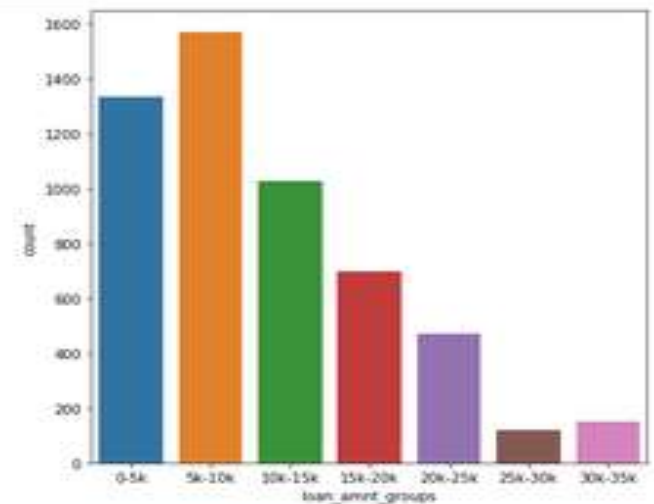
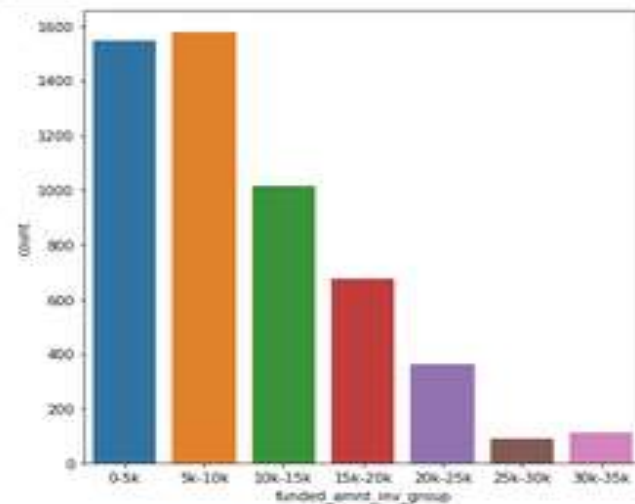
- Public records



- Month And Year



- Funded_amnt_inv_group
- Loan_amnt_groups
- dti_groups



Observation

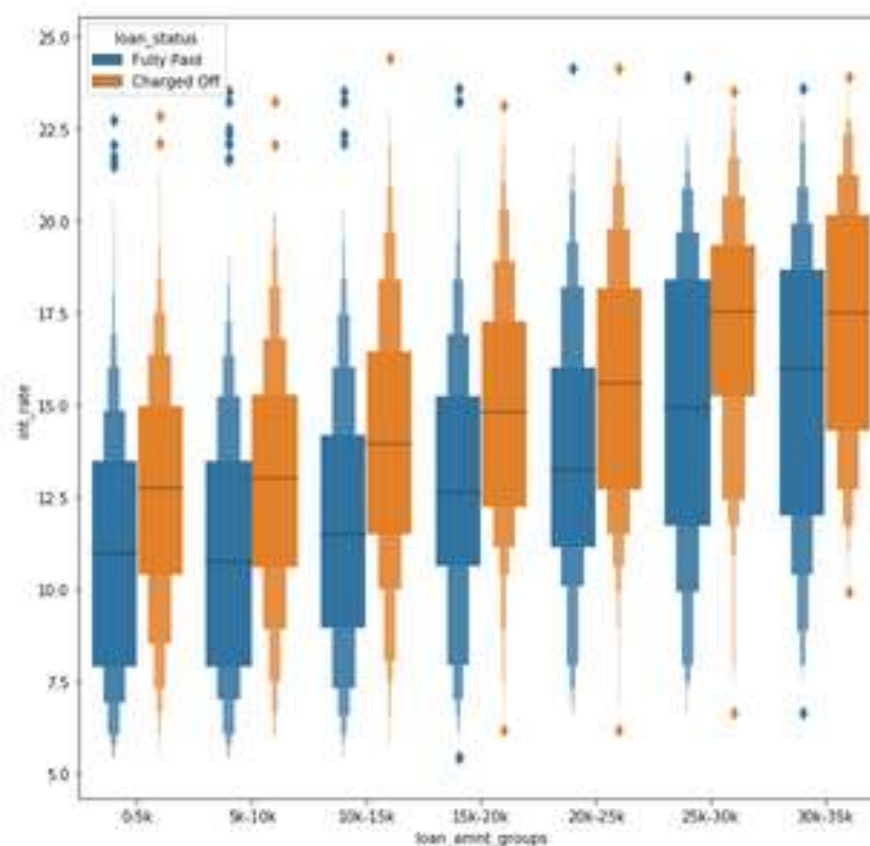
The above analysis with respect to the charged off loans for each variable suggests the following. There is a more probability of defaulting when :

- Candidate having house_ownership as 'RENT'
- Candidate who use the loan to clear other debts
- Candidate who receive interest at the rate of 13-17%
- Candidate who have an income of range 31201 – 58402
- Candidate who have 20-37 open_acc
- Candidate with employment length of 10
- When funded amount by investor is between 5000-10000
- Loan amount is between 5429 – 10357
- dti is between 12-18
- Term of 36 months
- When the loan status is Not verified
- When the number of derogatory public records is 0
- When the purpose is 'debt consolidation'
- Grade is 'B'
- And a total grade of 'B5' level.

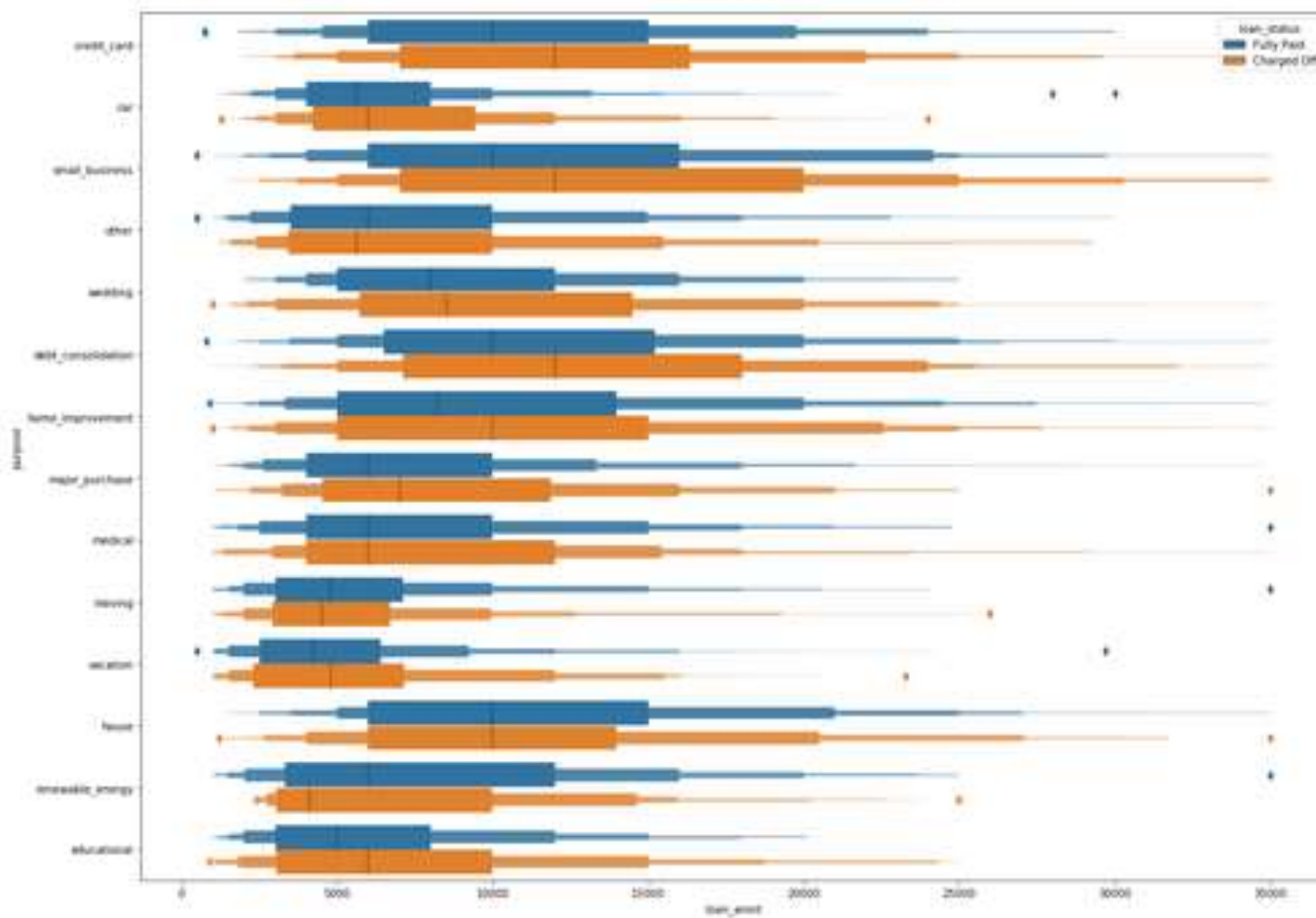
Bivariate Analysis

Analysing loan Amount

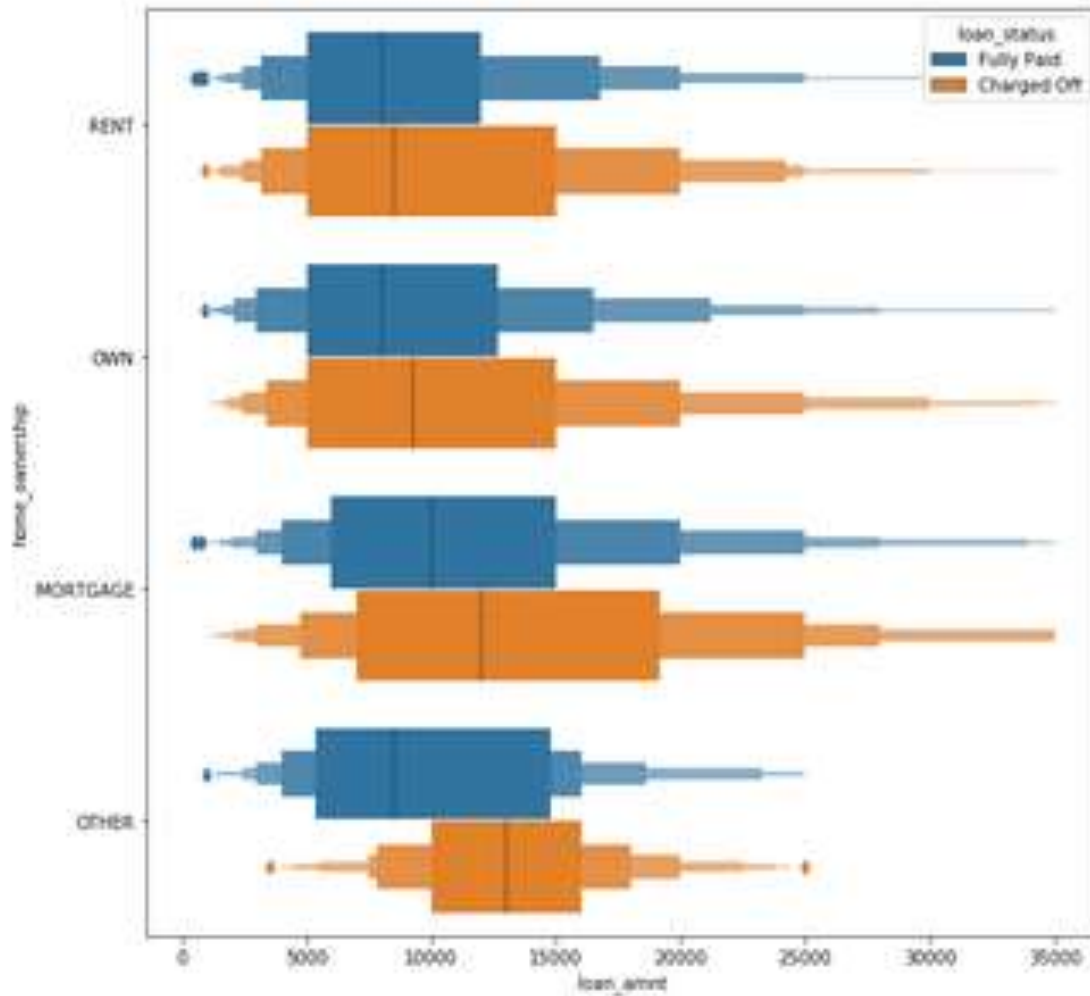
- `loan_amount_groups` vs interest rate ::



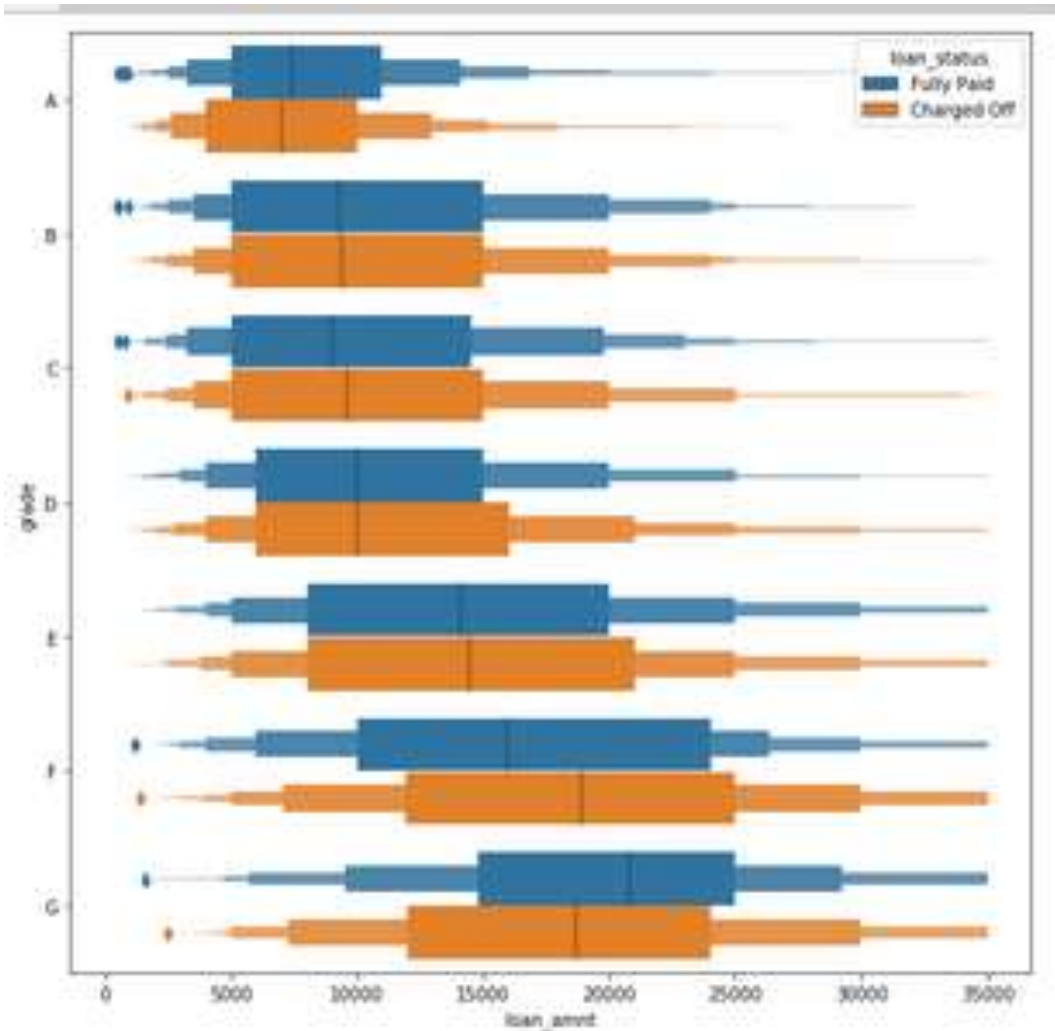
loan_amnt Vs Purpose



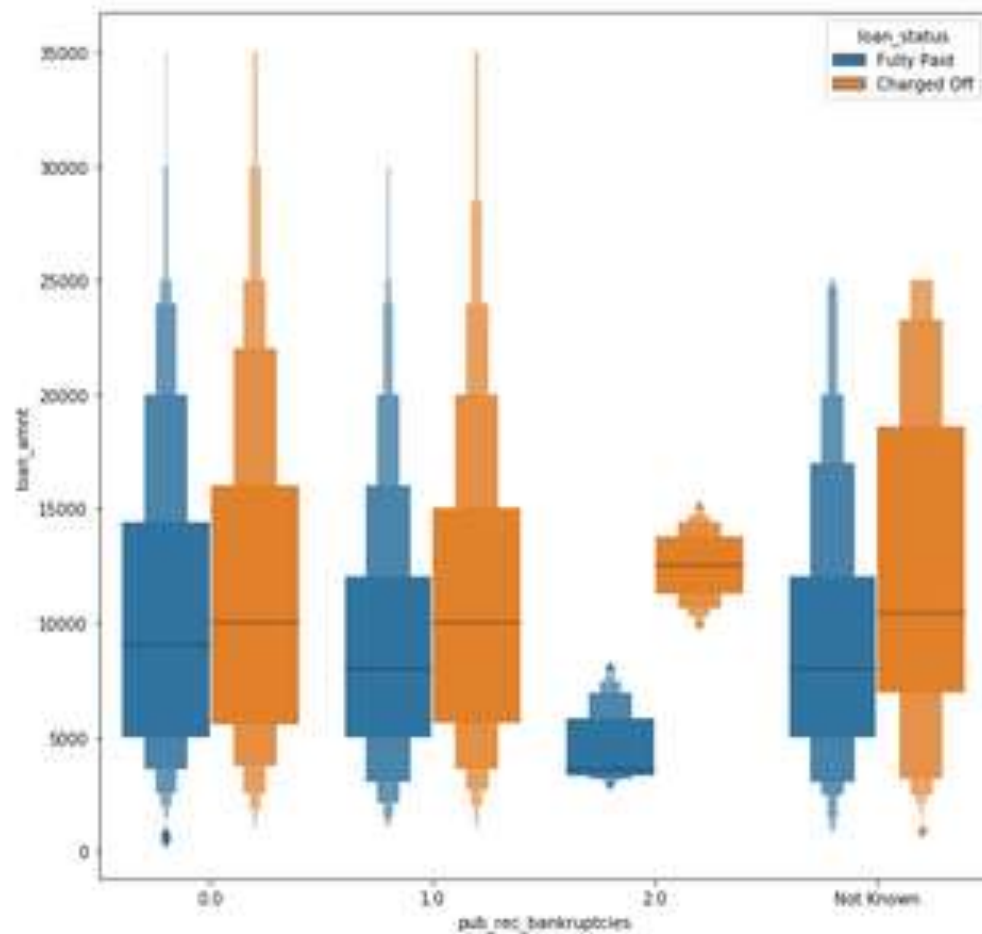
loan_amnt Vs Home Ownership



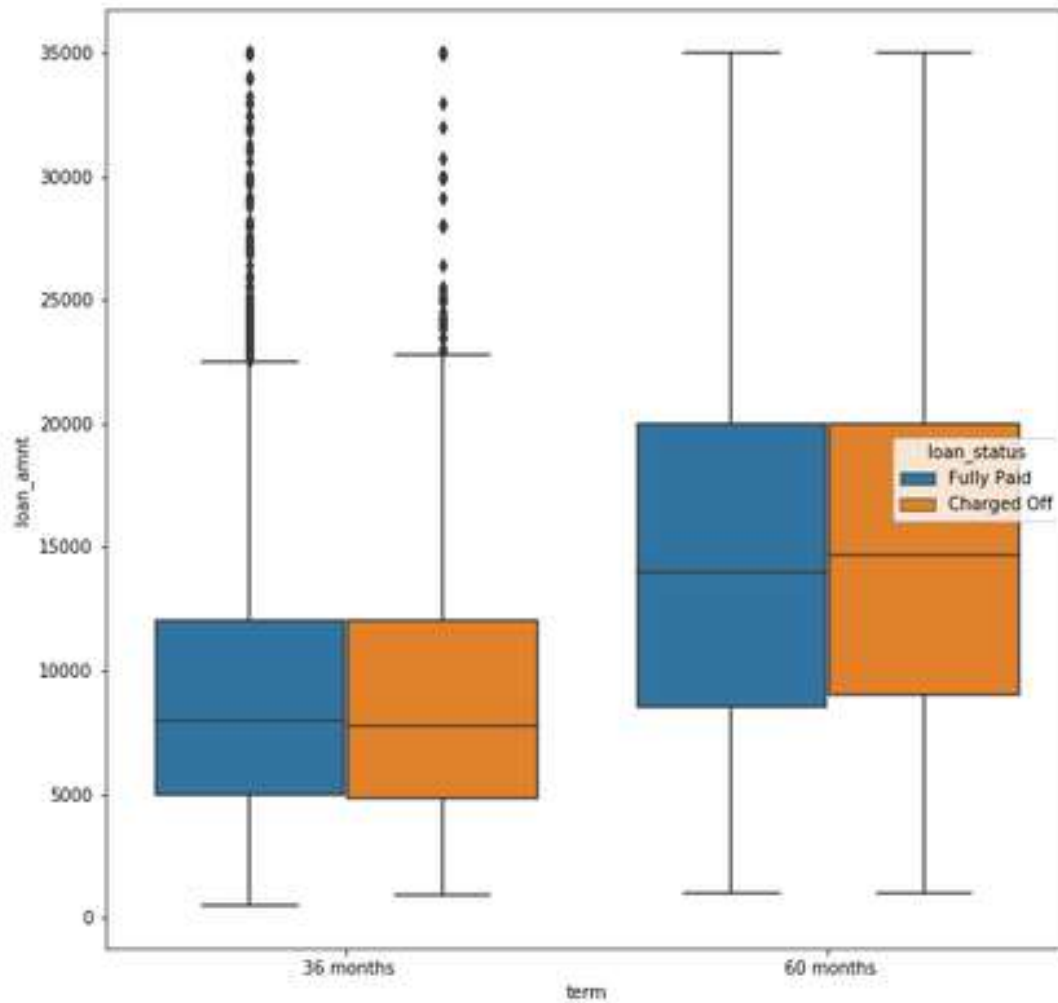
loan_amnt Vs grade



loan_amnt Vs Pub_Rec_bankruptcies

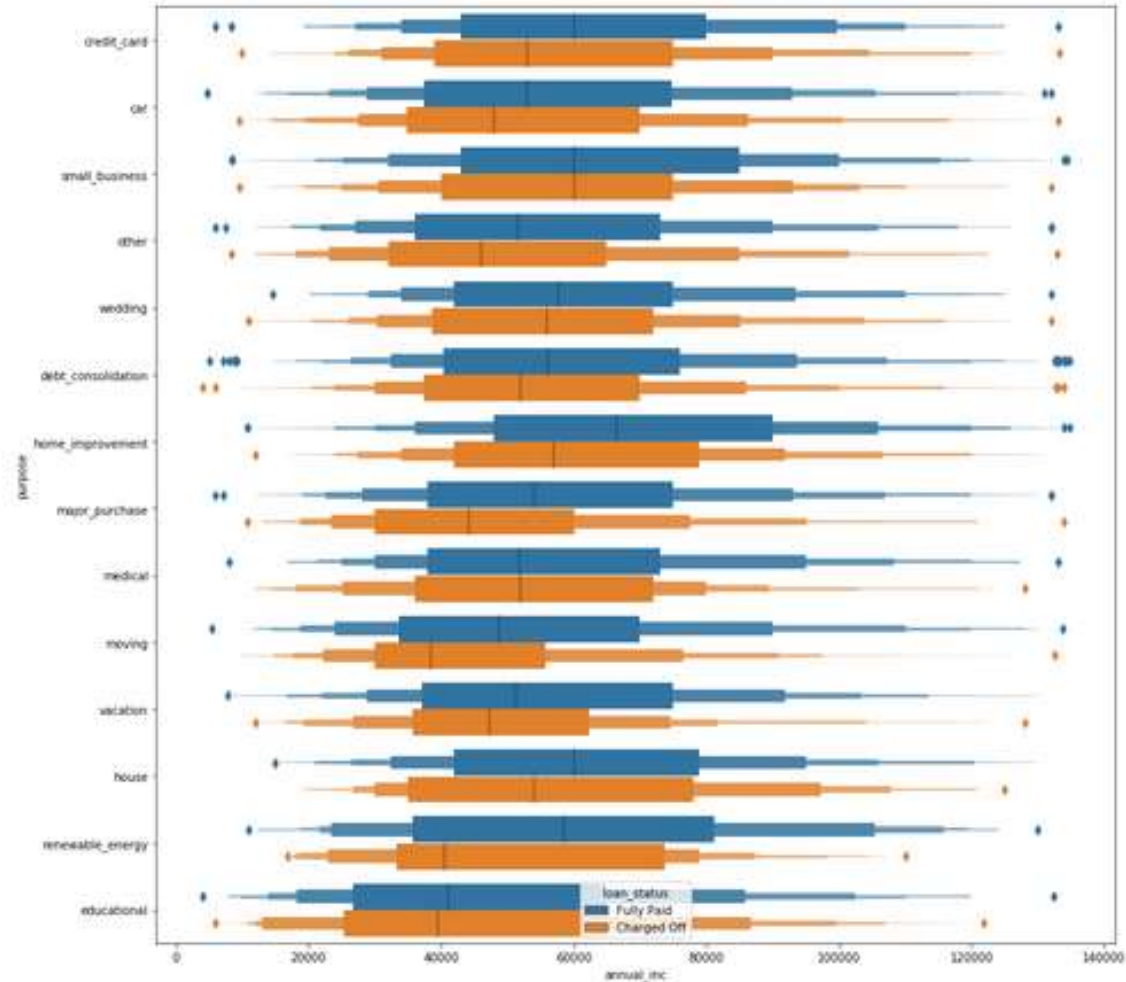


loan_amnt vs term

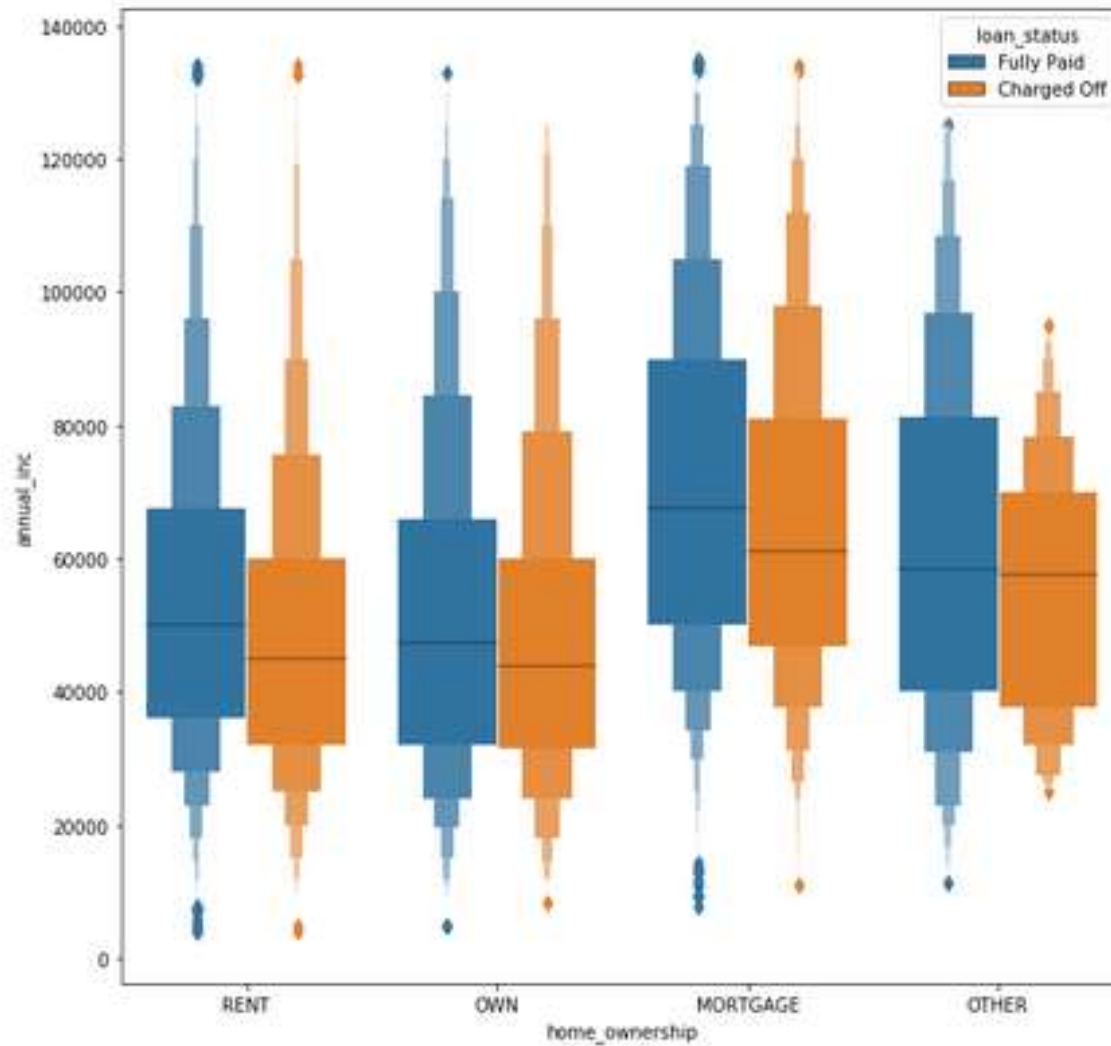


Analysing Annual Income

Annual Income Vs Purpose



Annual Income Vs home Ownership



Bivariate Observation

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