Lending Club Case Study

VAIBHAV J

LAVANYA M

Overview

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

Aim	Objective	Approach
To identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss.	To understand the **driving factors** behind loan default, i.e. the variables which are strong indicators of default.	Identification of such applicants using EDA is the aim of this case study.

Data for analysis

➤ Dropped null columns as they will not help in analysis.

➤ Dropped loan status **Current** as it will not be considered for analysis.

➤ Outliers are handled where it can affect the analysis,

➤ Some more data is derived from existing columns to understand data more clearly.
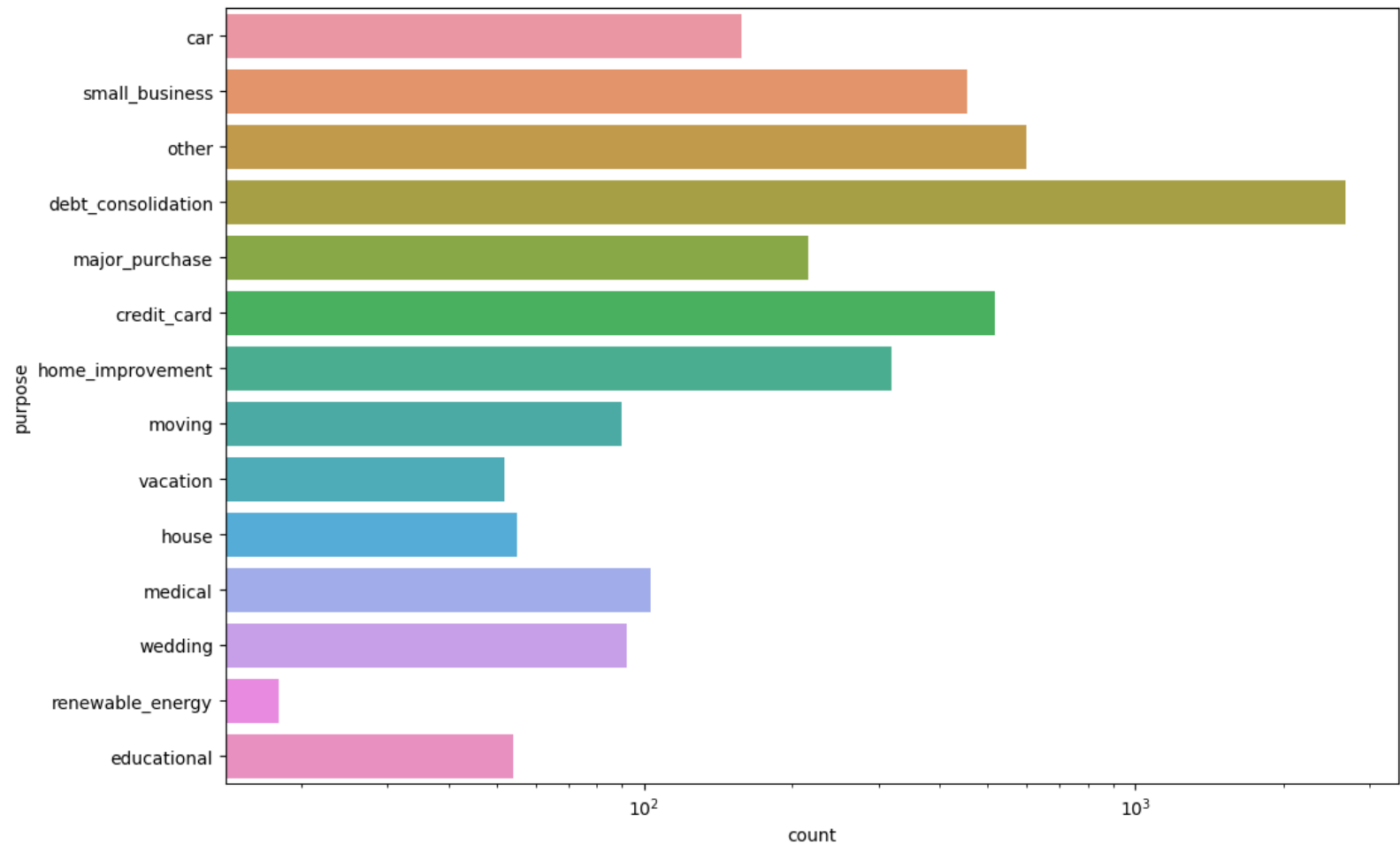
Univariate Analysis Results

Presenting observations and finding after preforming univariate analysis on cleaned data.

Analysis Result

Analysis :- By purpose

Findings/Observation:

People who's purpose of the loan is debt consolidation have high probability of defaulting loan.
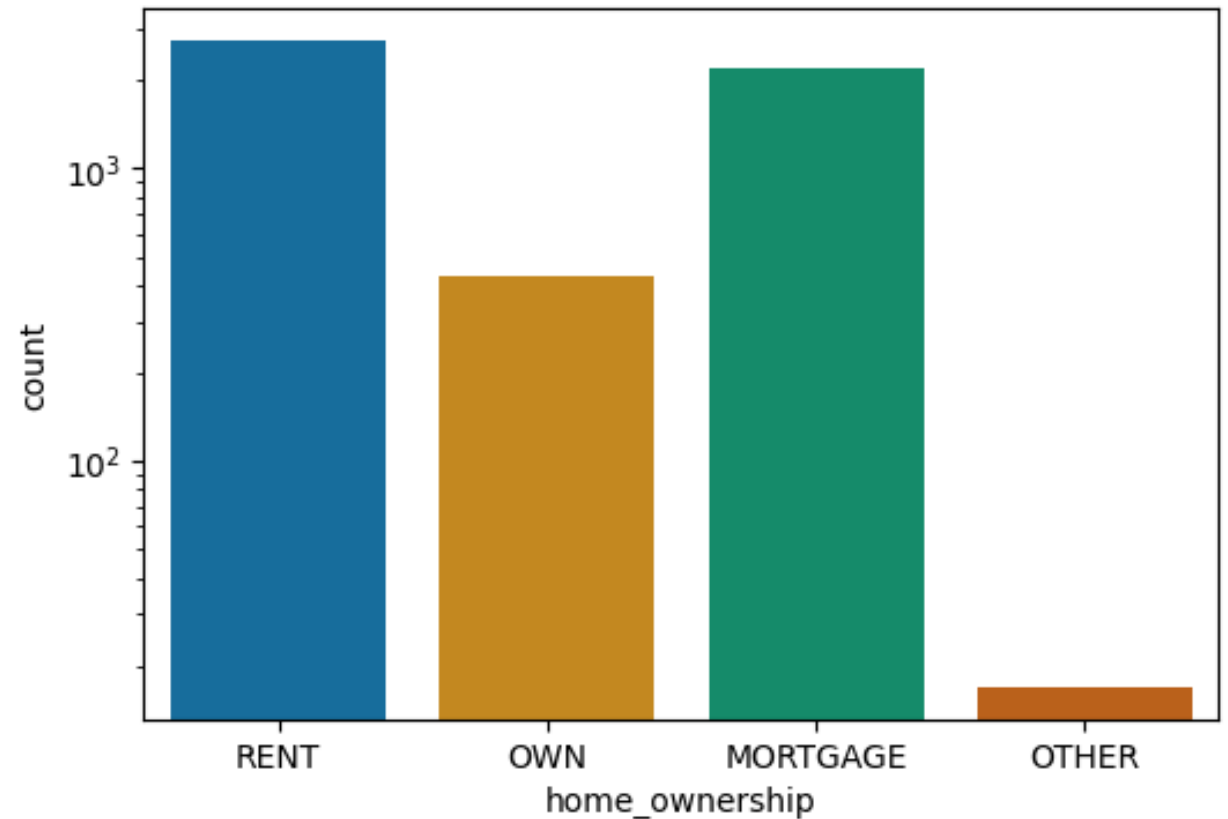
Analysis Result

Analysis :- By home ownership

Findings/Observation:

Applicants having house ownership as 'RENT' have high probability of defaulting loan.
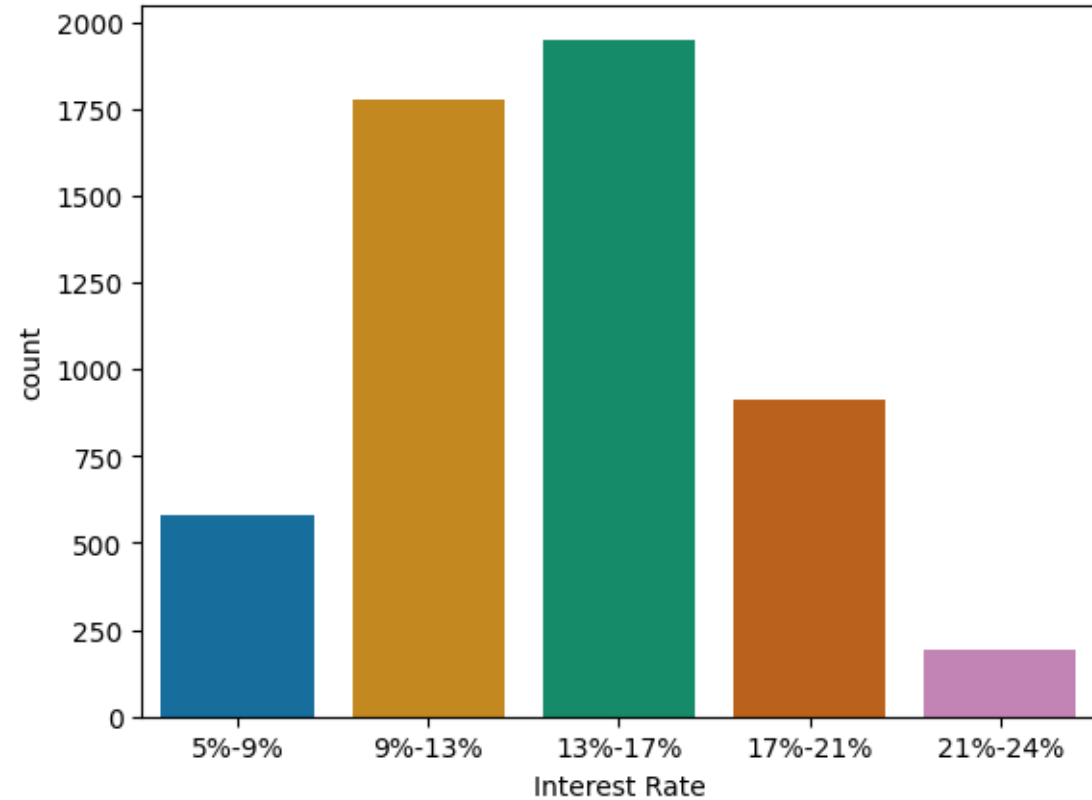
Analysis Result

Analysis :- By Interest Rate

Findings/Observation:

Applicants having interest rate of 13-17% is most likely to default loan.
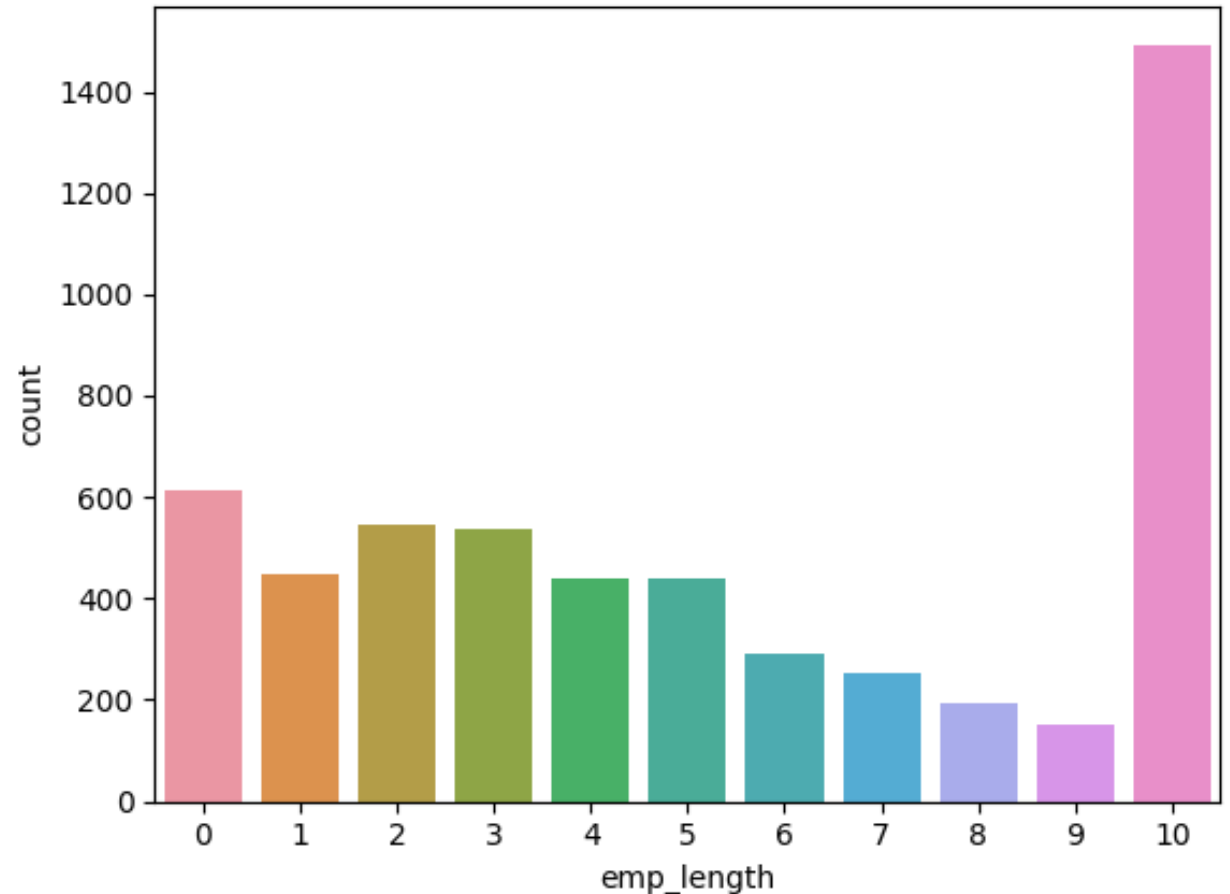
Analysis Result

Analysis :- By having employment length

Findings/Observation:

Applicants having employment length of 10 are most likely to default loan.
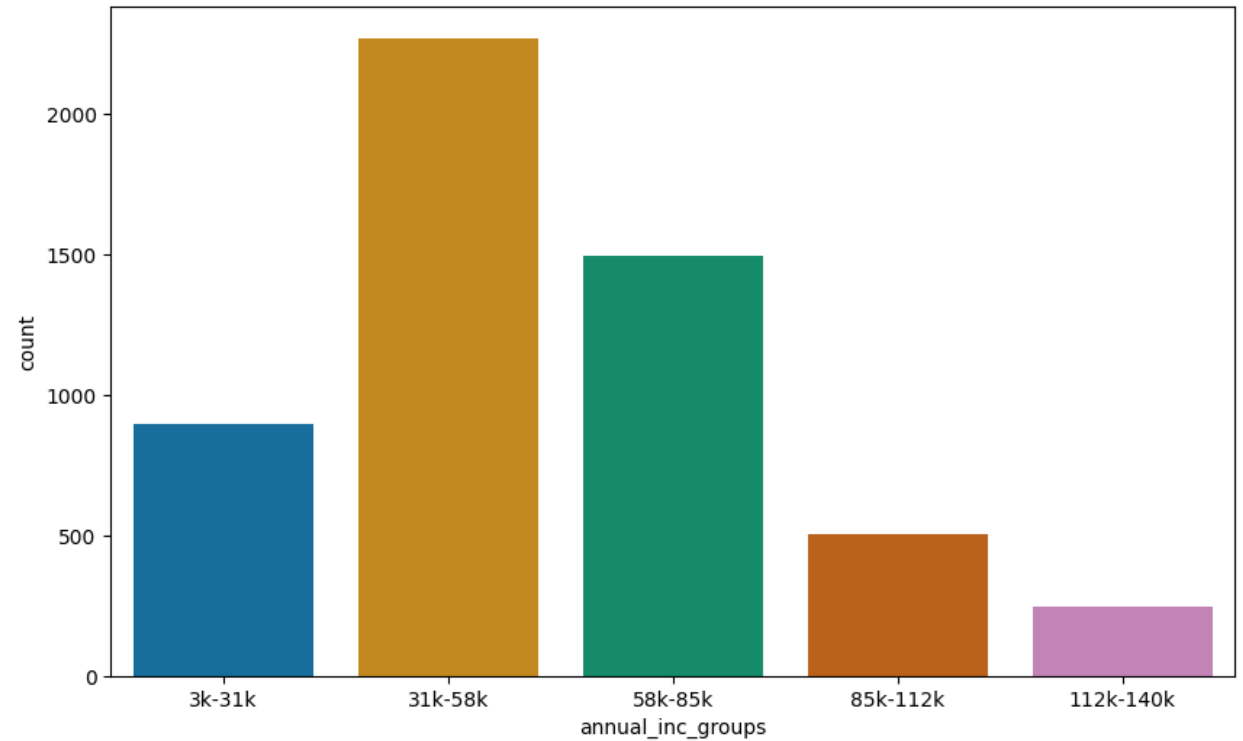
Analysis Result

Analysis :- By annual income

Findings/Observation:

Applicants having income in range of 31-59K are most likely to default loan.
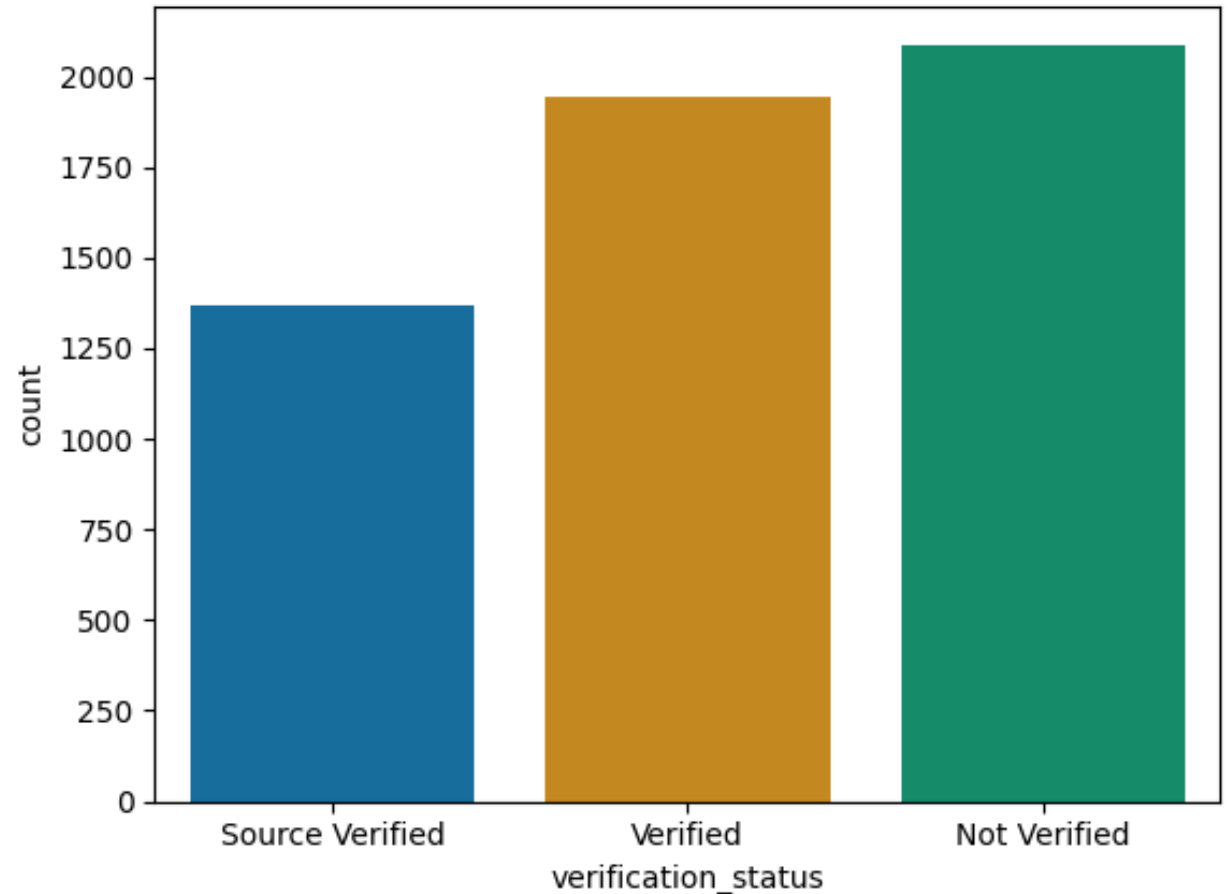
Analysis Result

Analysis :- By verification status

Findings/Observation:

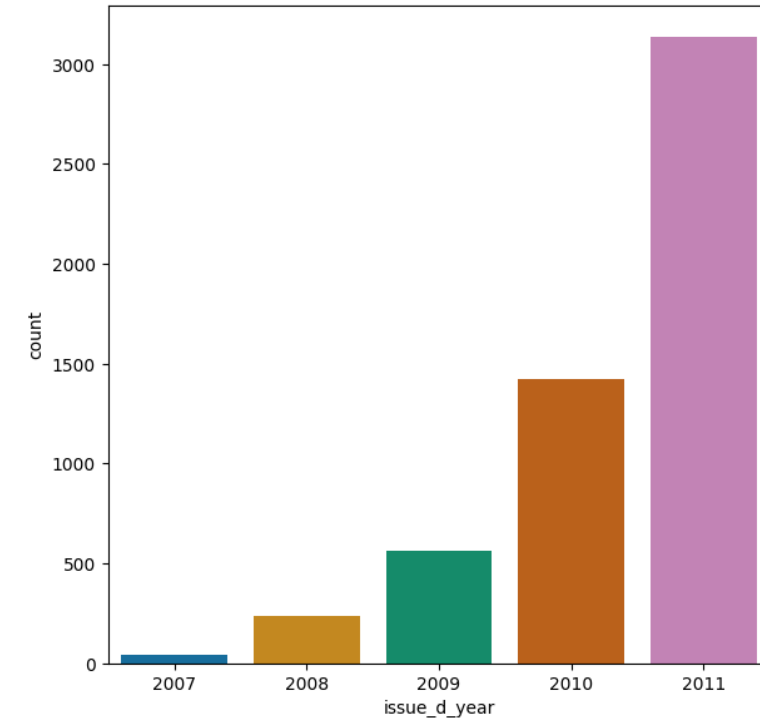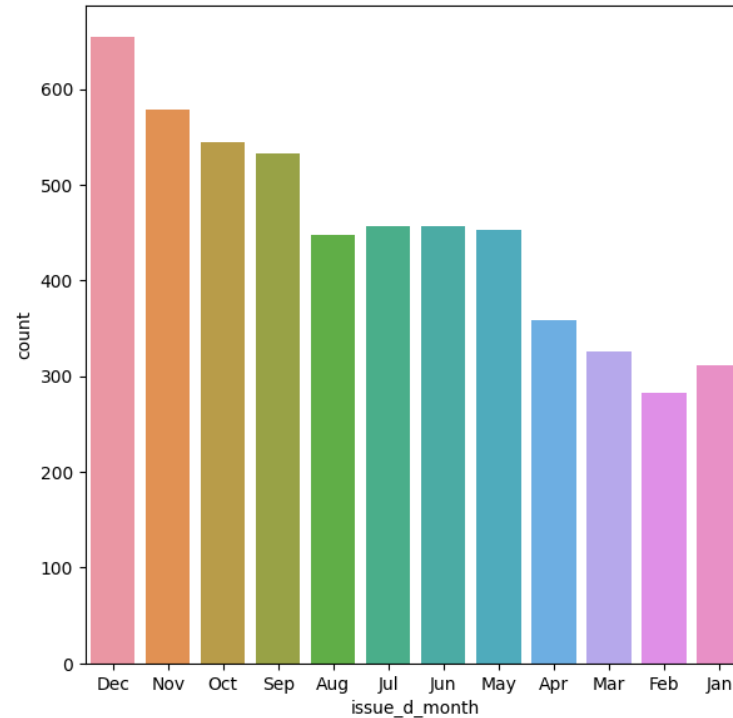Applicants who are not verified have high chance of defaulting loan.

Analysis Result

Analysis :- By issue_d

Findings/Observation:

Maximum number of defaults occurred when loan issued at the end of the year.

Maximum number of defaults occurred when loan issued in year 2011

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

➢Applicants having house_ownership as 'RENT'

➢Applicants who use the loan to clear other debts

➢Applicants who receive interest at the rate of 13-17%

➢Applicants who have an income of range 31201 - 58402

➢Applicants who have 20-37 open_acc

➢Applicants with employement length of 10

➢When funded amount by investor is between 5000-10000

➢Loan amount is between 5429 - 10357

➢Dti is between 12-18

➢When monthly installments are between 145-274

➢Term of 36 months

➢When the loan status is Not verified

➢When the no of enquiries in last 6 months is 0

➢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 Results

Presenting observations and finding after preforming bivariate analysis on cleaned data.
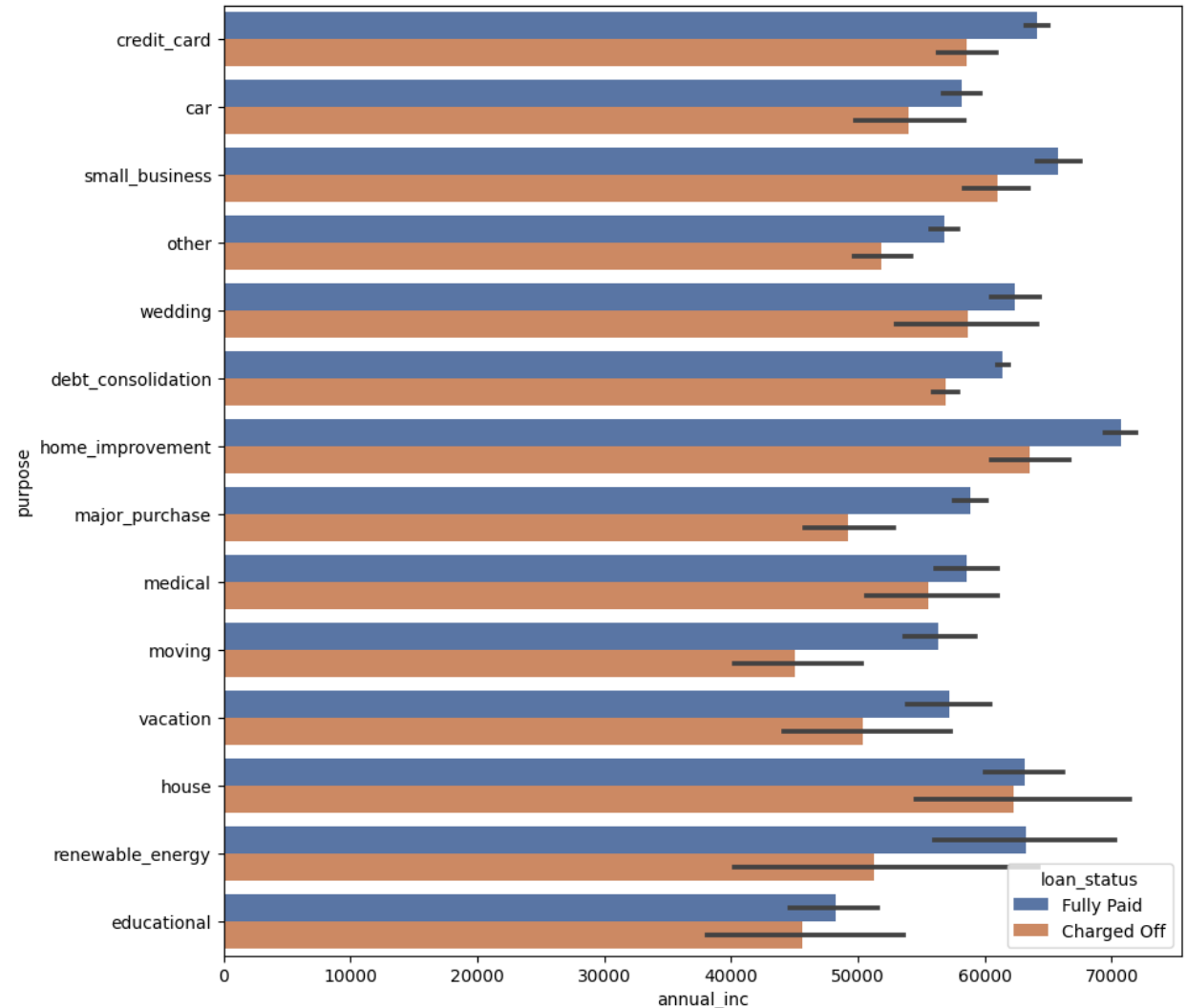
Analysis Result

Analysis :- By Annual income vs loan purpose

Findings/Observation:

Even if number of loans applied and defaulted are the highest in number for "debt_consolation", the annual income of those who applied isn't the highest.

Applicants with higher salary mostly applied loans for "home_improvment", "house", "renewable_energy" and "small_businesses"
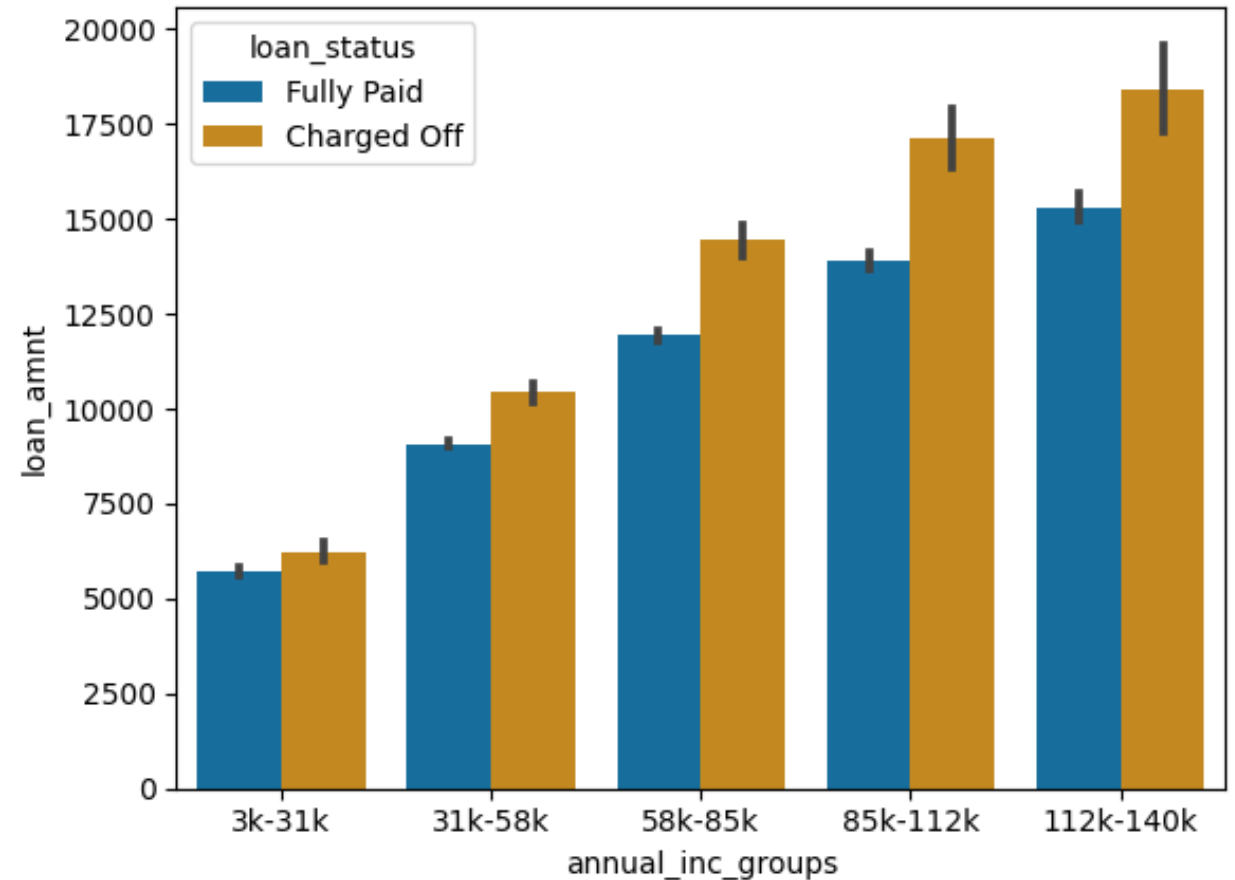
Analysis Result

Analysis :- By Annual income vs loan amount

Findings/Observation:

Across all the income groups, the loan_amount is higher for people who defaulted.

Analysis Result

Analysis :- By Annual income vs loan amount

Findings/Observation:

Across all the income groups, the loan_amount is higher for people who defaulted.
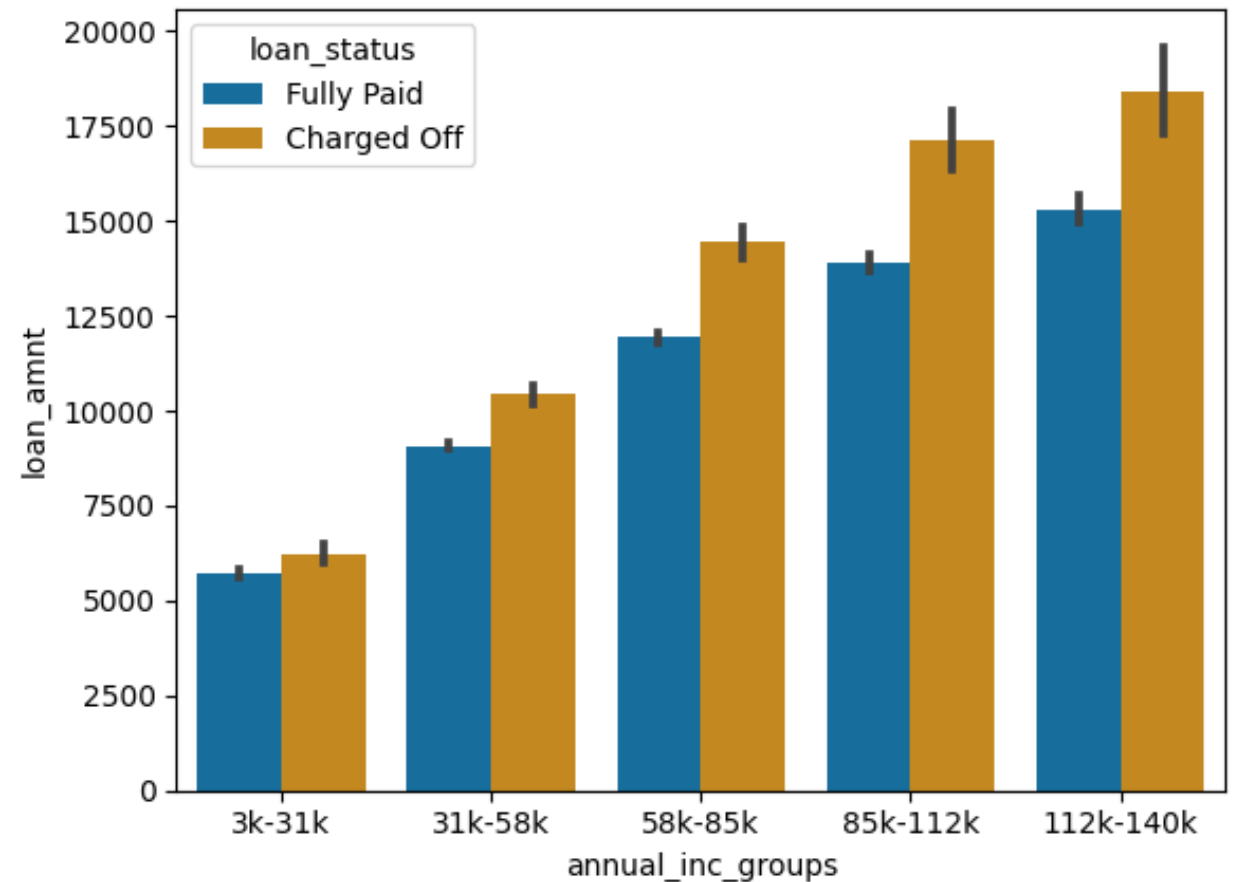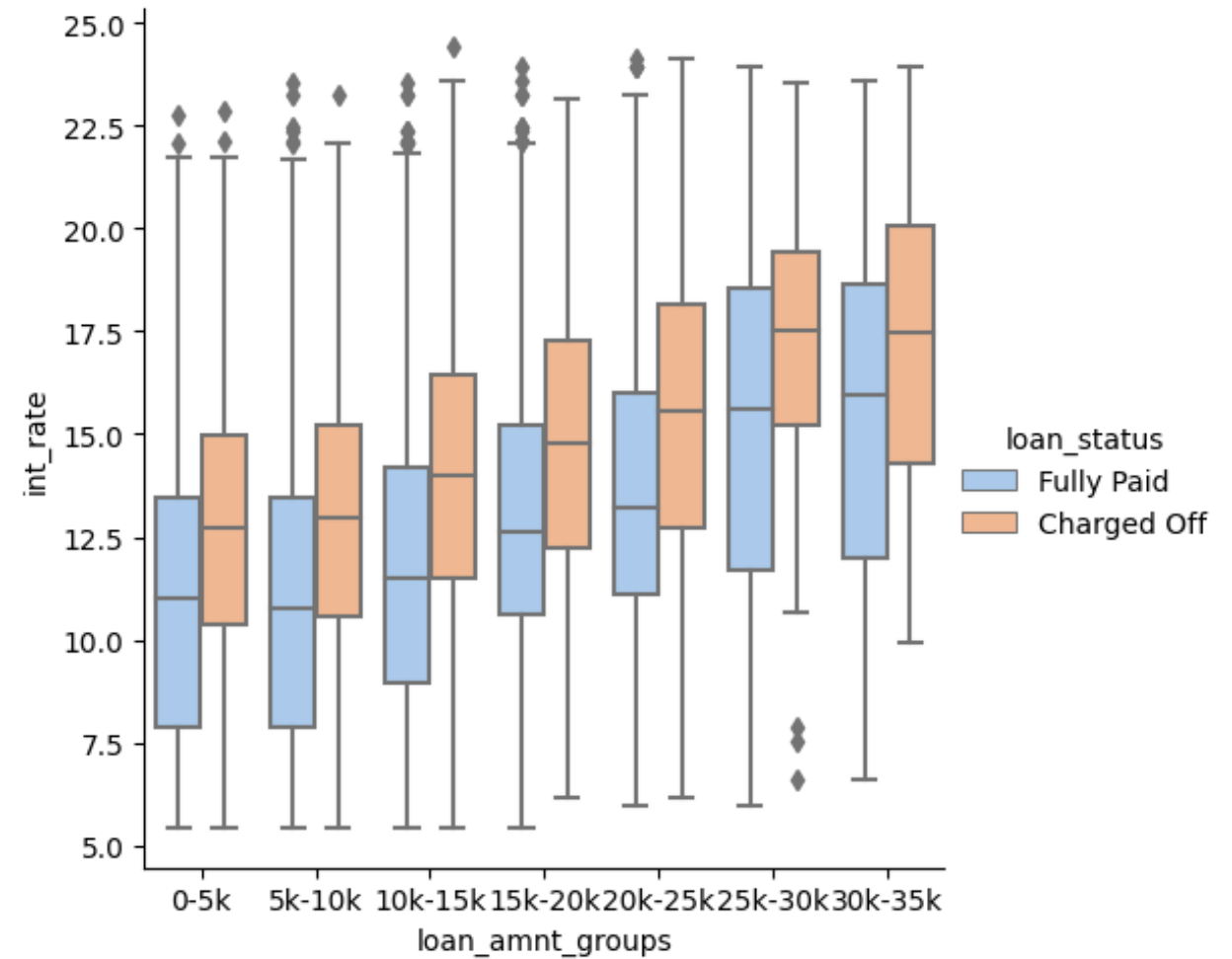
Analysis Result

Analysis :- By interest rate vs loan amount

Findings/Observation:

The interest rate for charged off loans is pretty high than that of fully paid loans in all the loan_amount groups.
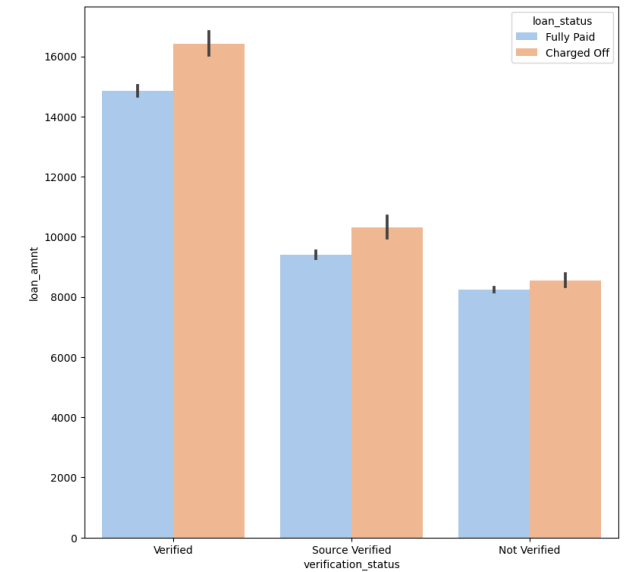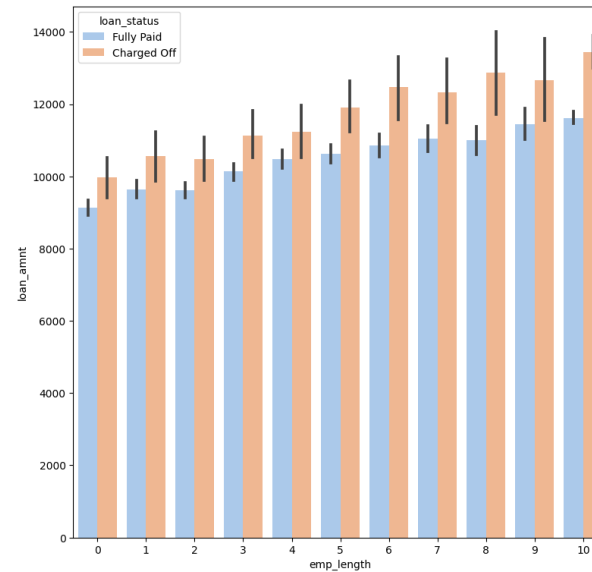
Analysis Result

Analysis :- By employee length vs loan amount and verified status vs loan amount

Findings/Observation:

Probability of defaulting loan is high when employment length is 10yrs and loan amount is 12k-14k and the loan is verified and loan amount is above 16k

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

❖ Applicants taking loan for 'home improvement' and have income of 60k -70k

❖ Applicants whose home ownership is 'MORTGAGE and have income of 60-70k

❖ Applicants who receive interest at the rate of 21-24% and have an income of 70k-80k

❖ Applicants who have taken a loan in the range 30k - 35k and are charged interest rate of 15-17.5 %

❖ Applicants who have taken a loan for small business and the loan amount is greater than 14k

❖ Applicants 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%

Multivariate Analysis

Findings/Observation:

There is a positive correlation between some entities and a negative correlation between some. Chances of default are strongly correlated with term and interest rate