Lending Club Case Study

Problem Statement

I am working for the Lending club a finance company which specializes in lending various types of loans to urban customers. When the company receives a loan application, the company have to make a decision for loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision:

   1. If the applicant is likely to repay the loan, then not approving   the loan results in a loss of business to the company
   2. If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company.

Objective

The company wants to understand the driving factors 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.

Approach

Data Understanding

Understanding the data completely to plan the analysis process.

Data Cleaning

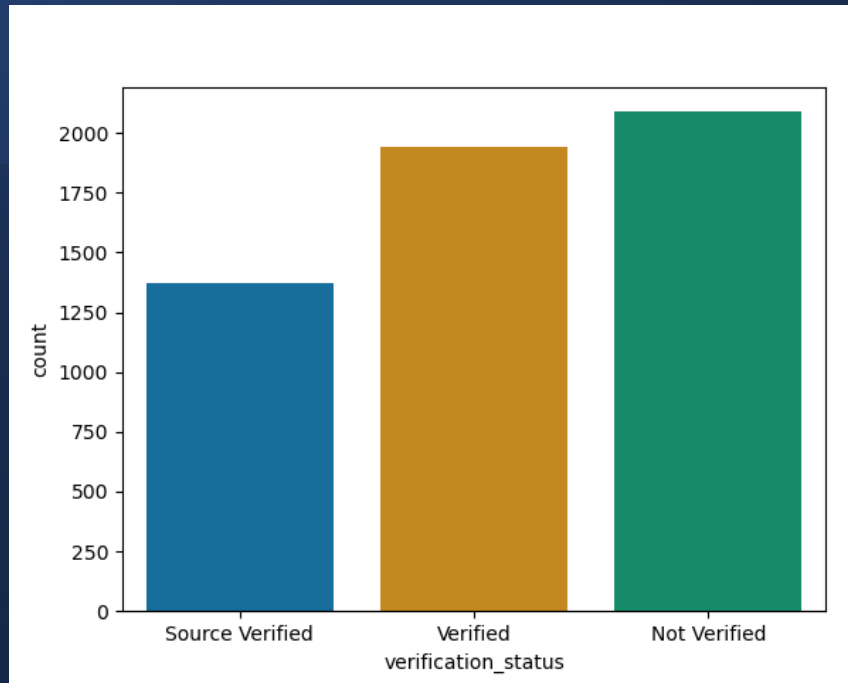Cleaning the data to get the most accurate results from our analysis.

Data Visualization

Doing Univariate and Bivariate analysis to discover the patterns between variables.
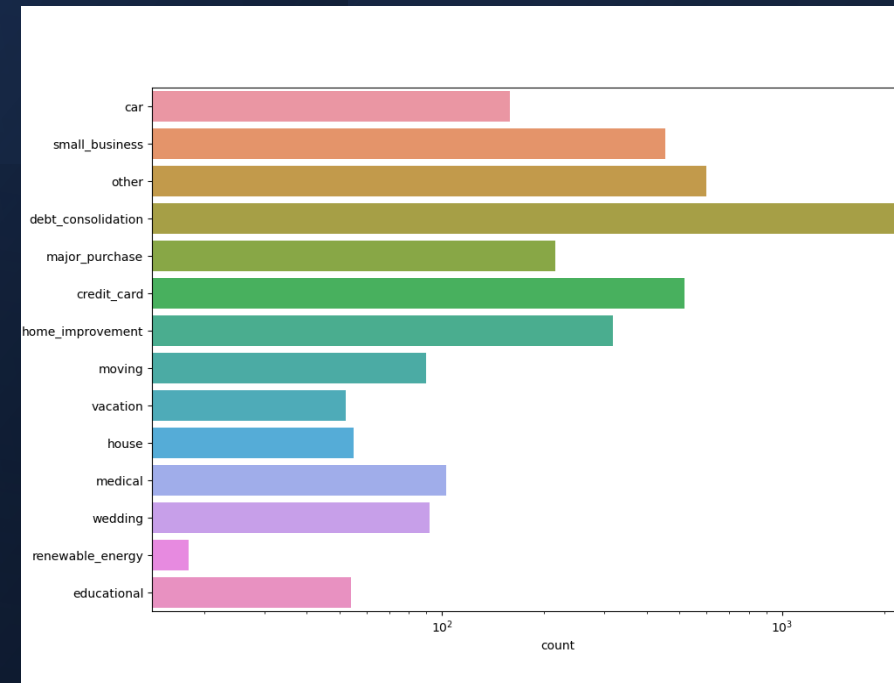
Observations
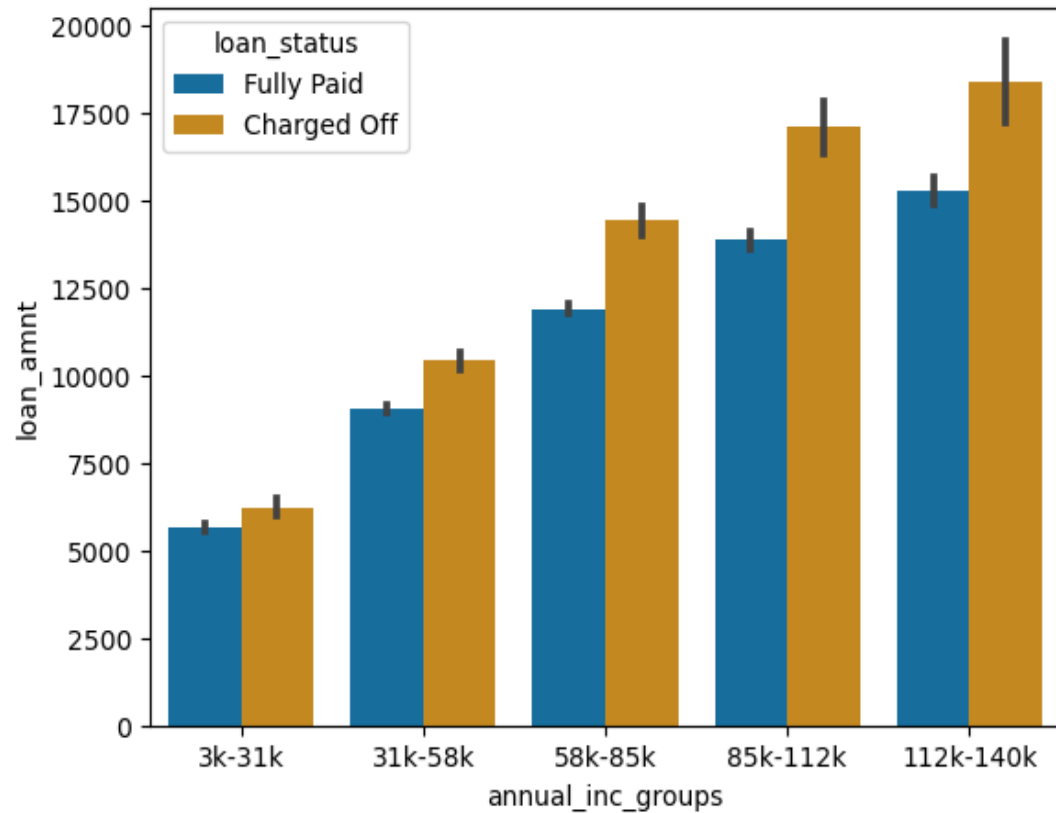
Univariate Analysis

Verified status

Purpose of Loan
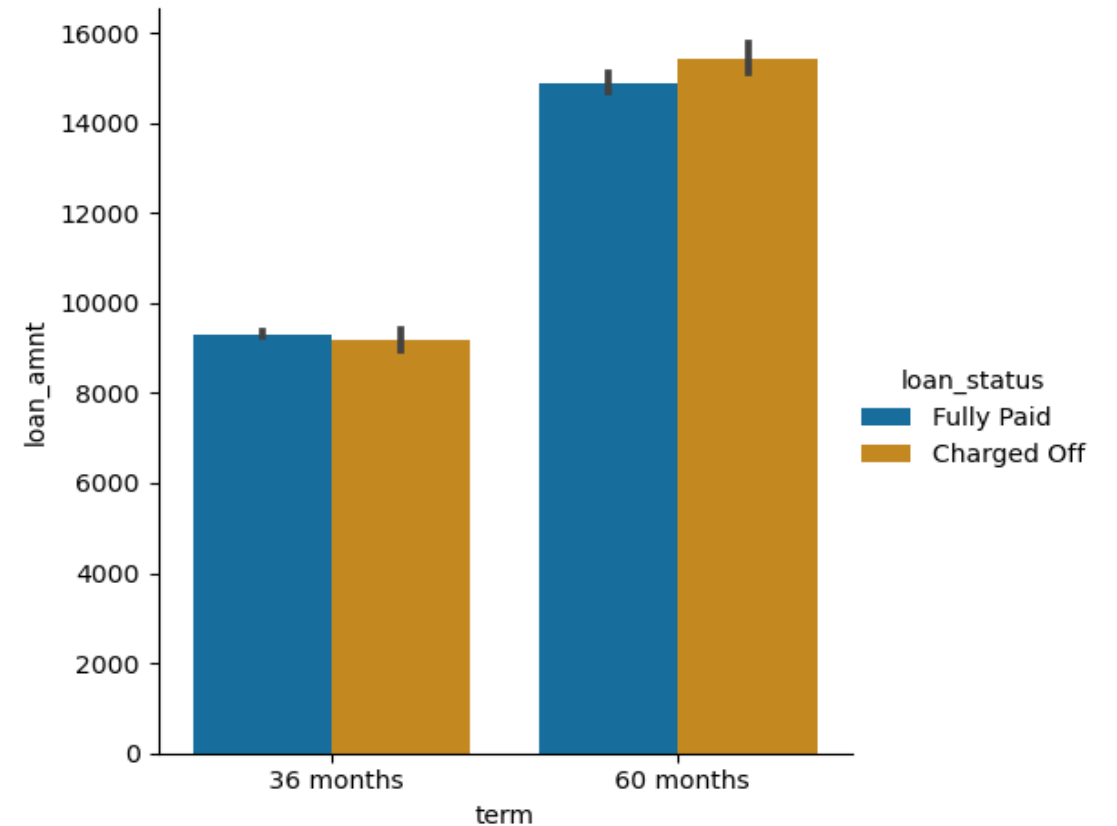
Univariate Analysis Observations

There is 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 verification 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

Annual_inc_groups vs loan_amnt

Term vs loan_amnt

Bivariate Analysis Observations

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%

Recommendations

Avoid giving loans to the applicant

- who would use the loan to pay debts.

- Who has DTI ratio between 12-18.

- Who is not verified.

- Who has grade G and interest rate above 20%.

- Who is taking loan for 'home improvement' and have income of 60k -70k.

- Who receive interest at the rate of 21-24% and have an income of 70k-80k.

- Who has grade F and loan amount is between 15k-20k.