



Lending Club Case Study

By Tejas Hankare & Madhur Gupta



Problem Statement :

- Like most other lending companies, 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.
- Expected Solution:
 - 1. Identify these risky loan applicants
 - 2. Understand the 'Driver Factors' behind 'Loan Default'



Abstract

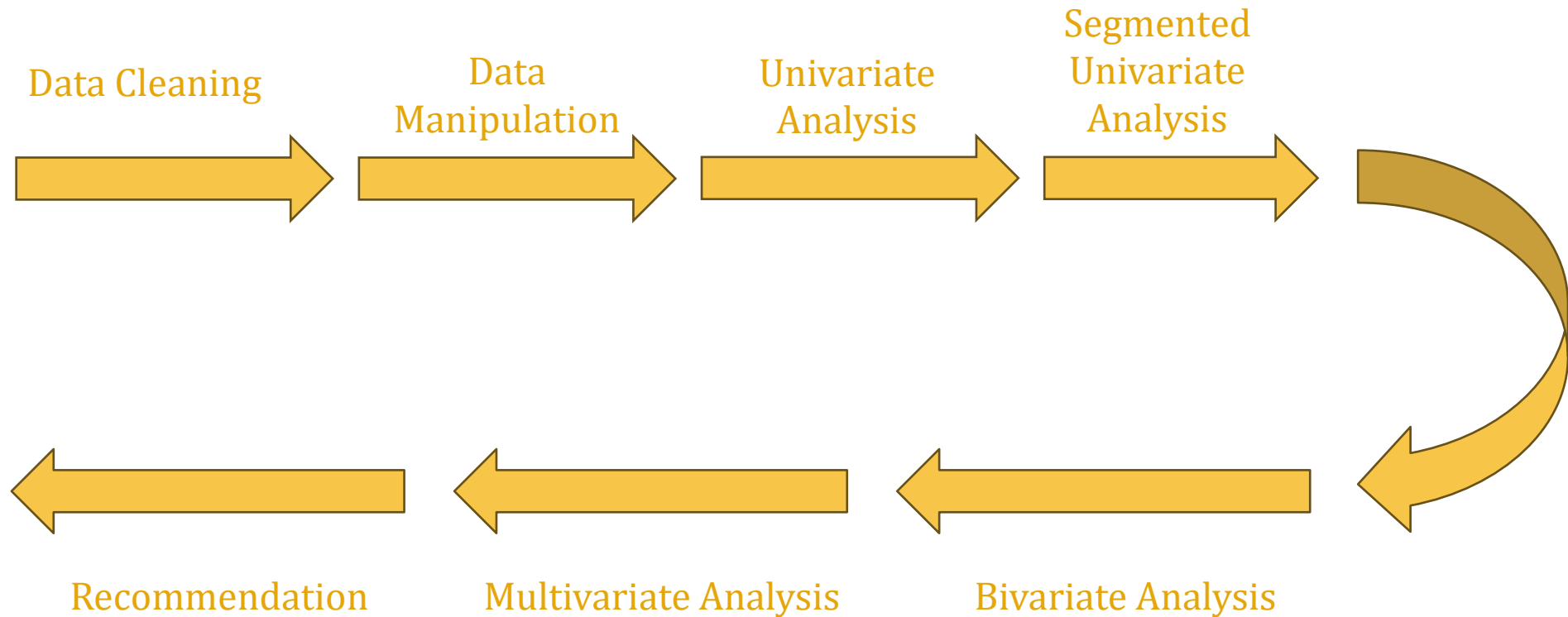
Lending Club is a marketplace for personal loans that matches borrowers who are seeking a loan with investors looking to lend money and make a return.

When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision:

- If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company.
- 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.

So, 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 utilize this knowledge for its portfolio and risk assessment.

Problem Solving Methodology



A decorative vertical bar on the left side of the slide, featuring a gold color and a pattern of various currency symbols (dollar, euro, yen, pound, etc.) in a 3D, embossed style.

Data Cleaning

- Removed all the columns with Null values.
- Removed all the rows with majority Null values.
- Removed Columns which are not defined in the Column Definition dataset

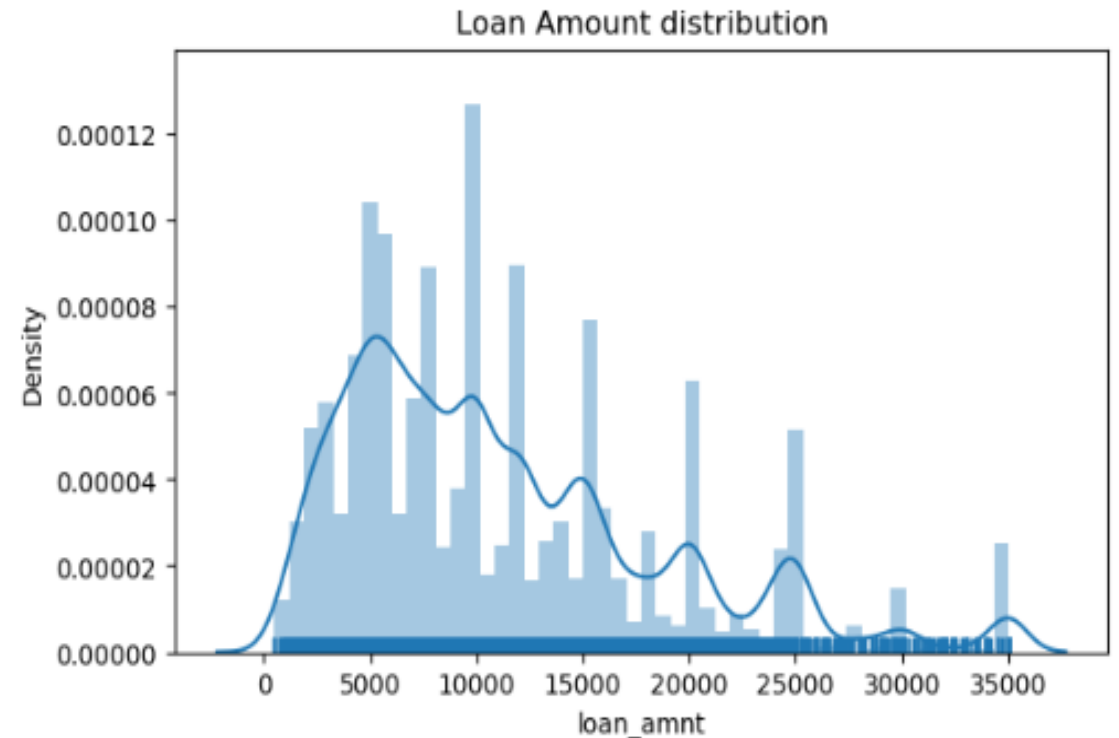
A decorative vertical bar on the left side of the slide, featuring a golden-yellow background with various 3D currency symbols like the dollar sign (\$), yen sign (¥), and euro sign (€) in a repeating pattern.

Data Manipulation

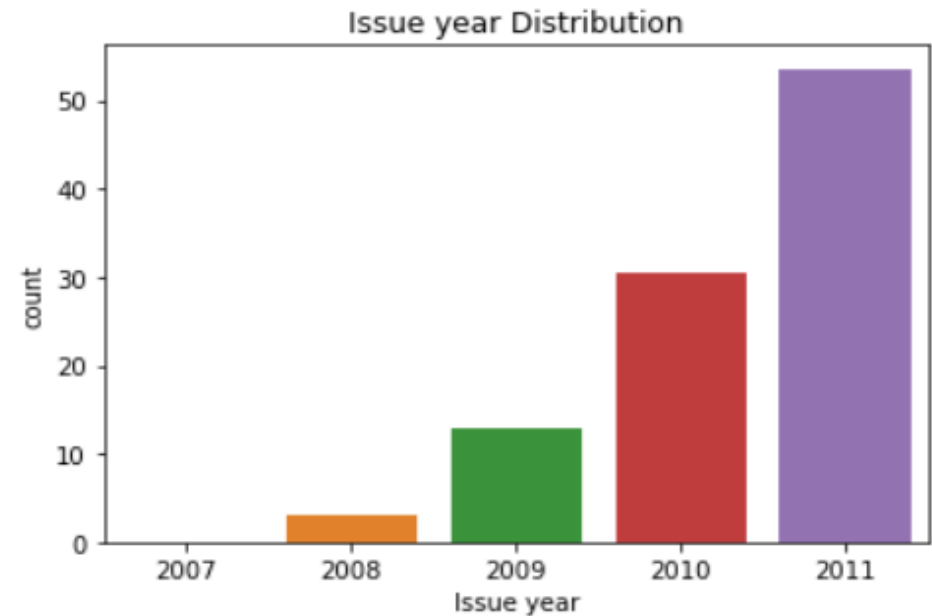
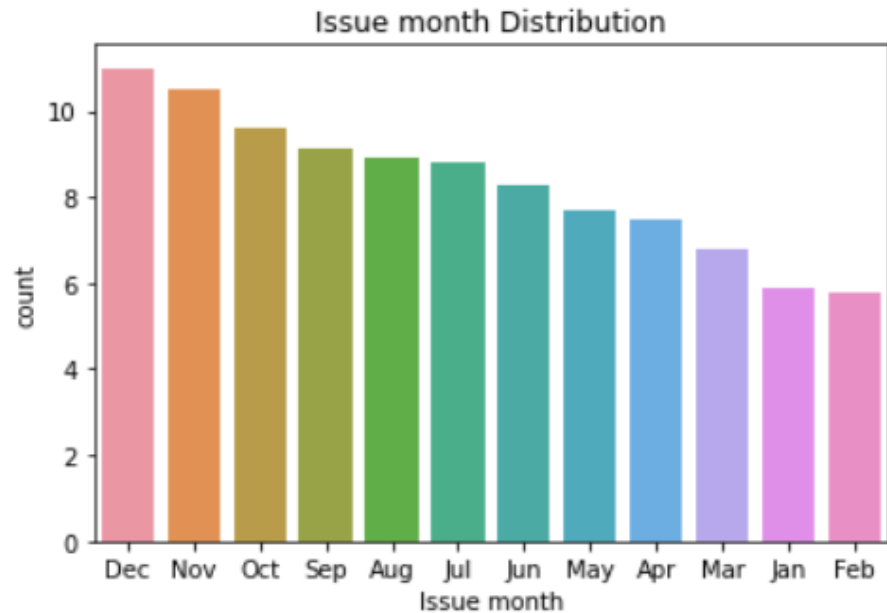
- Data manipulation refers to the process of changing or transforming data to make it more suitable for analysis or visualization.
- Standardizing data formats, such as converting date columns to a consistent format.
- Handling missing values by either imputing them or removing rows with missing data.

Loan Amount Distribution

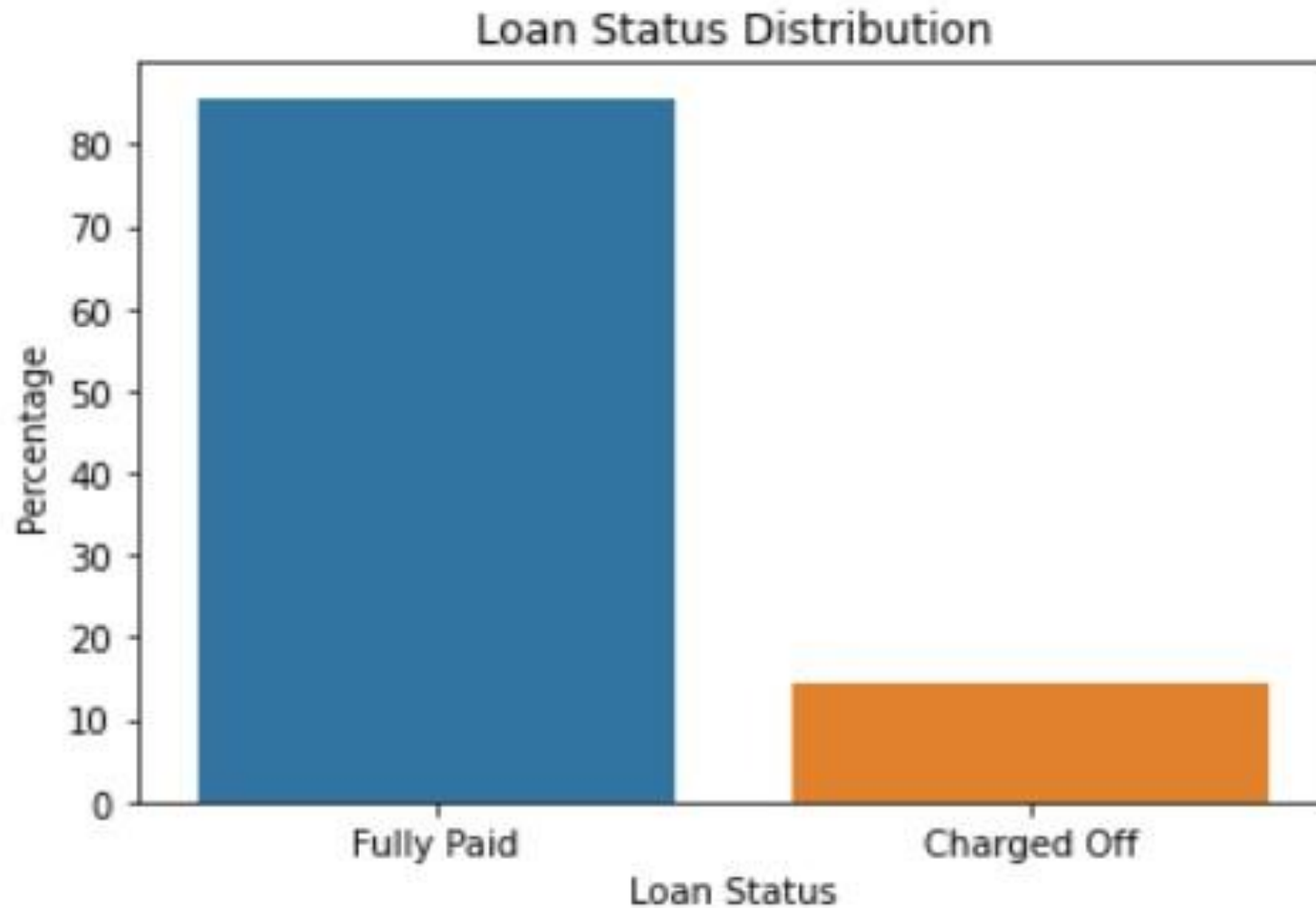
- Loan applied lies between 500 to 35000
- By this graph we got to understand that most of the loan borrowers need money under 15000
- Majority of borrowers asked for rounded number that is 5000 , 10,000 as observed from the Distribution graph



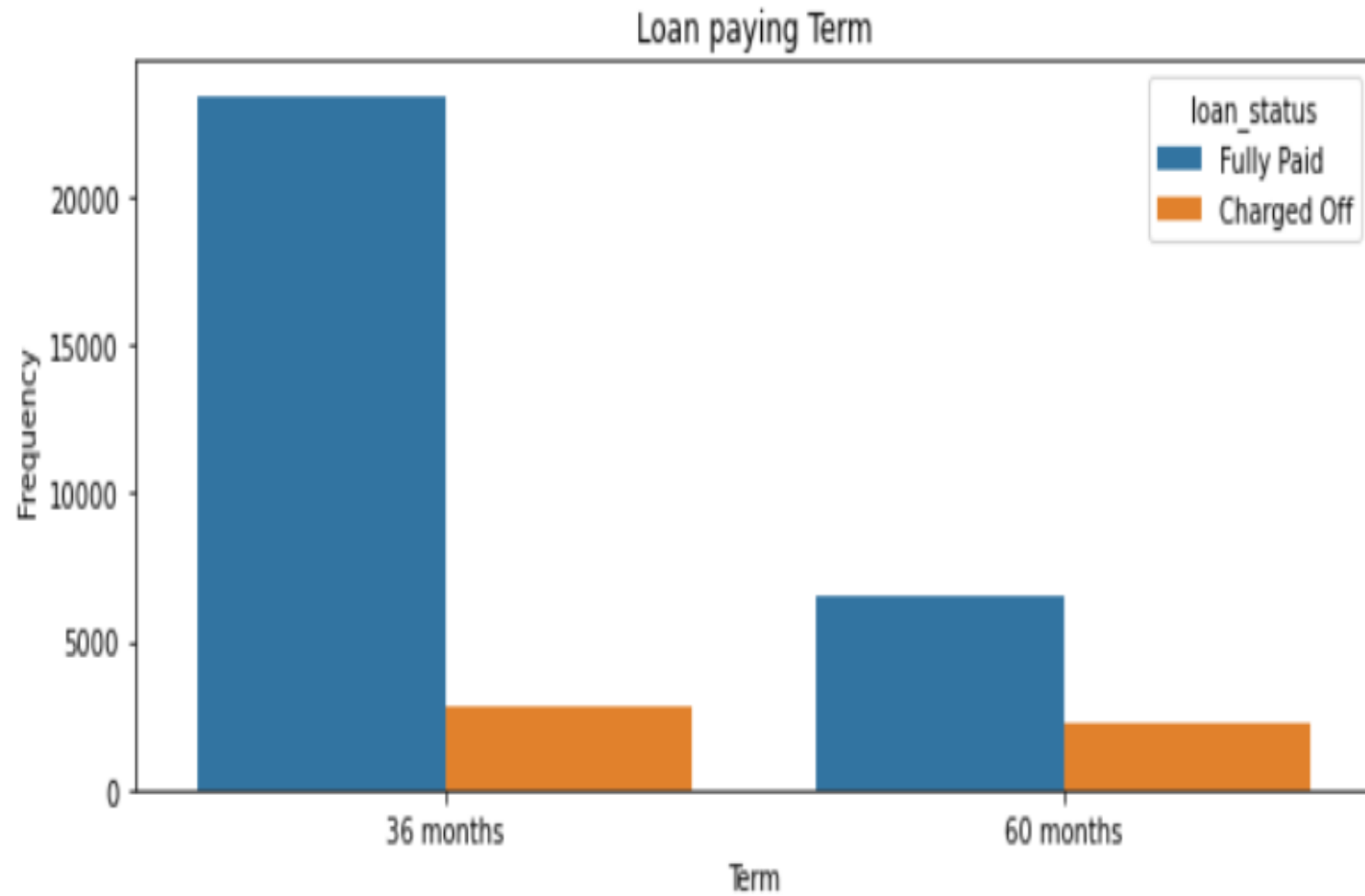
Analysis



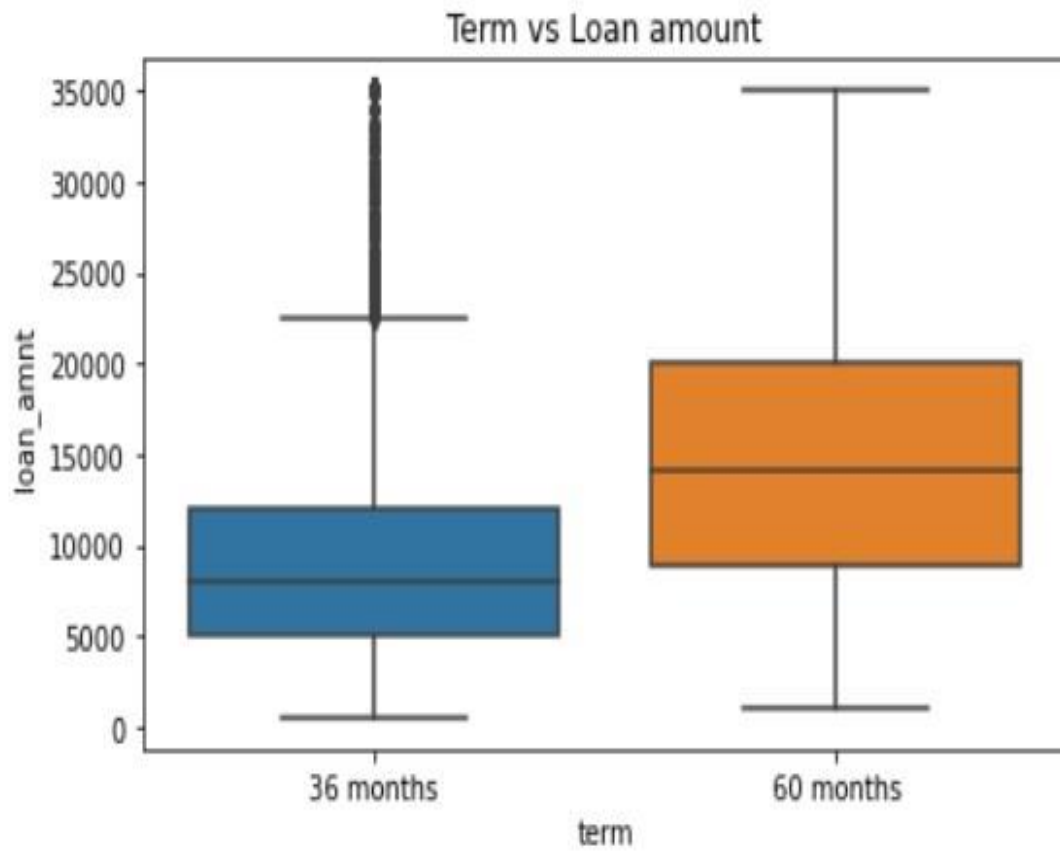
- Lending club has really expanded year by year, every year the number of loans has doubled.
- Issue of loan is also increasing every month from January to December and in the final quarter of year there are more loans issued due to vacation.



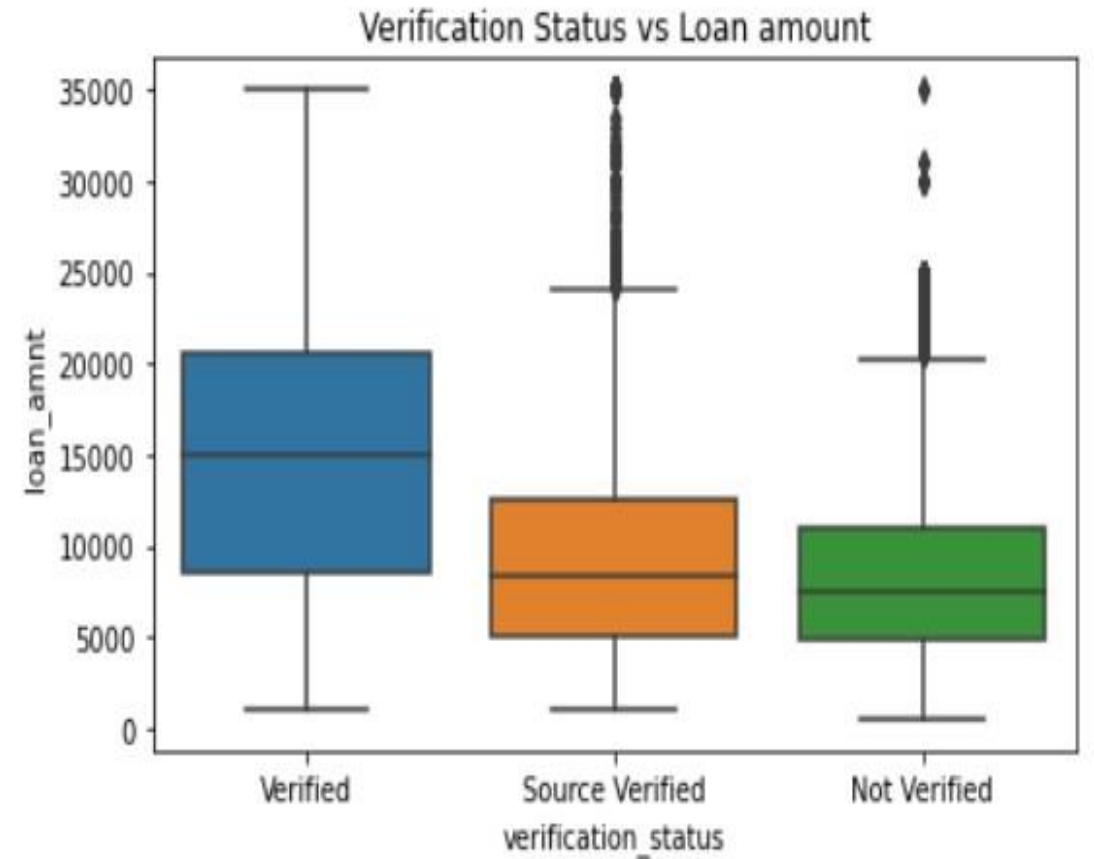
85% of borrower have fully paid their loan, Where 'as 15% of the borrowers are charged off.



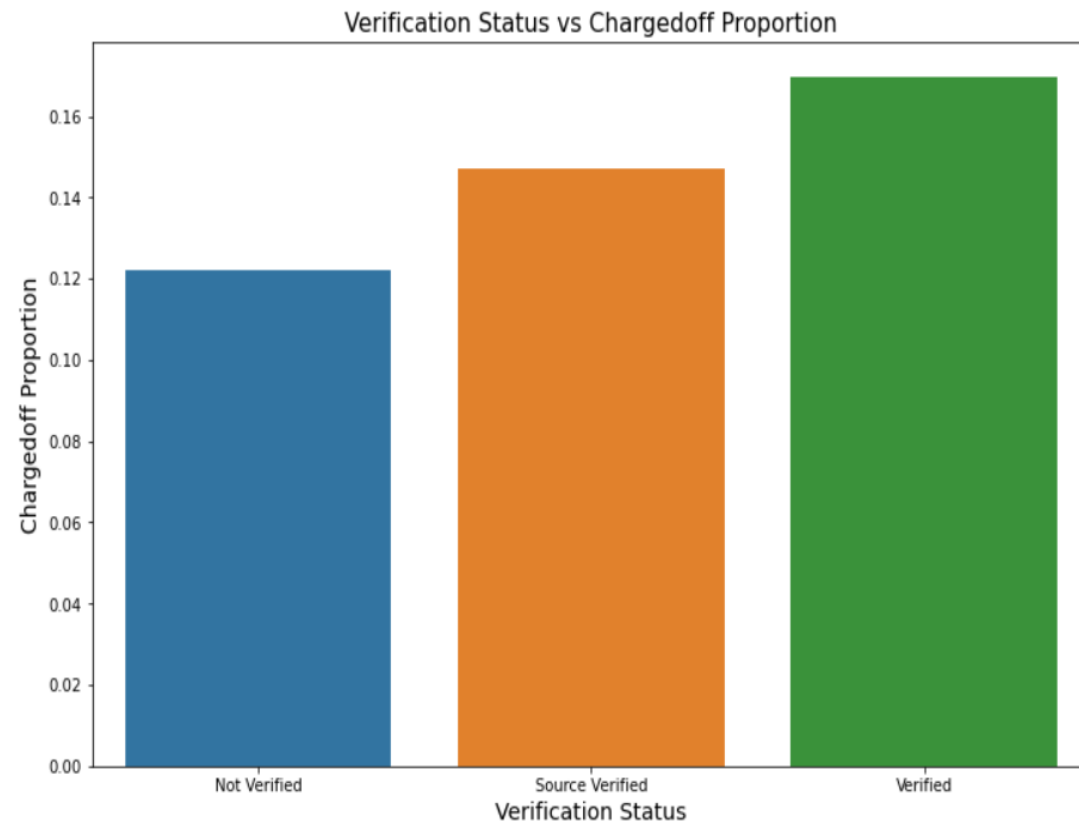
There are 2 loan term and most of the borrower took 36 months tenure. But the ratio of charged off is high in 60 months tenure.



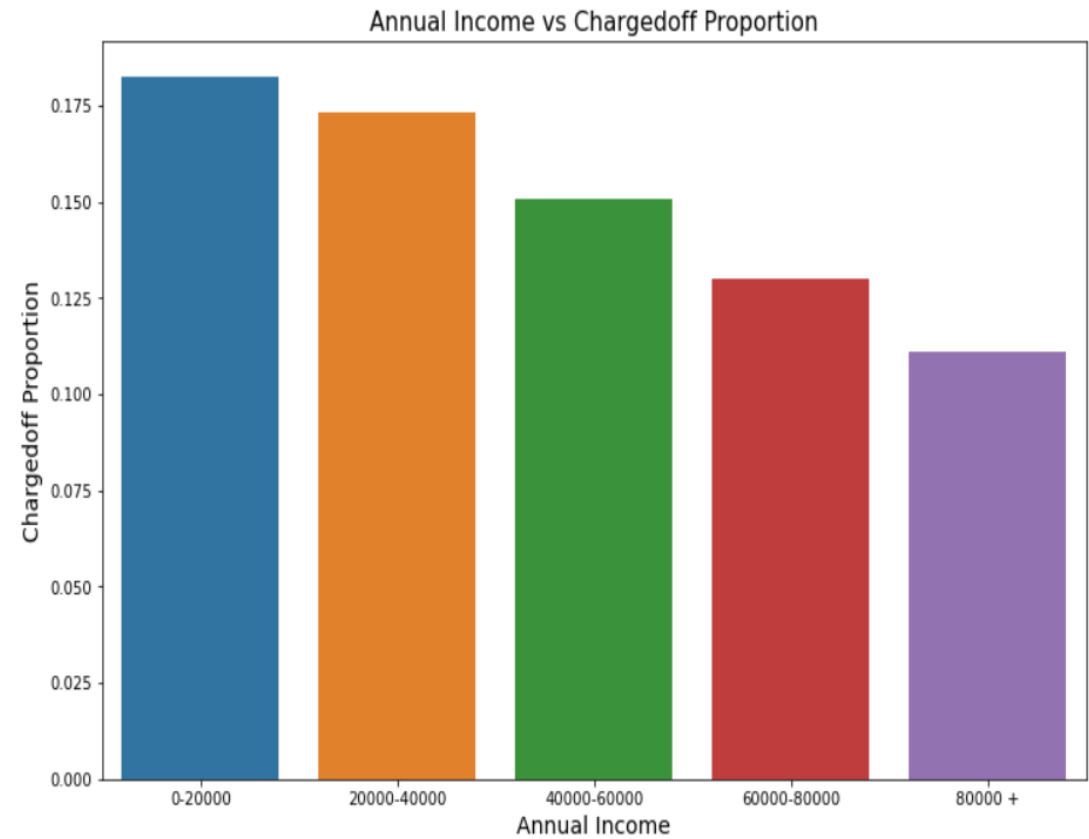
The borrower who takes higher loan amount tends to choose the loan term 60 months.



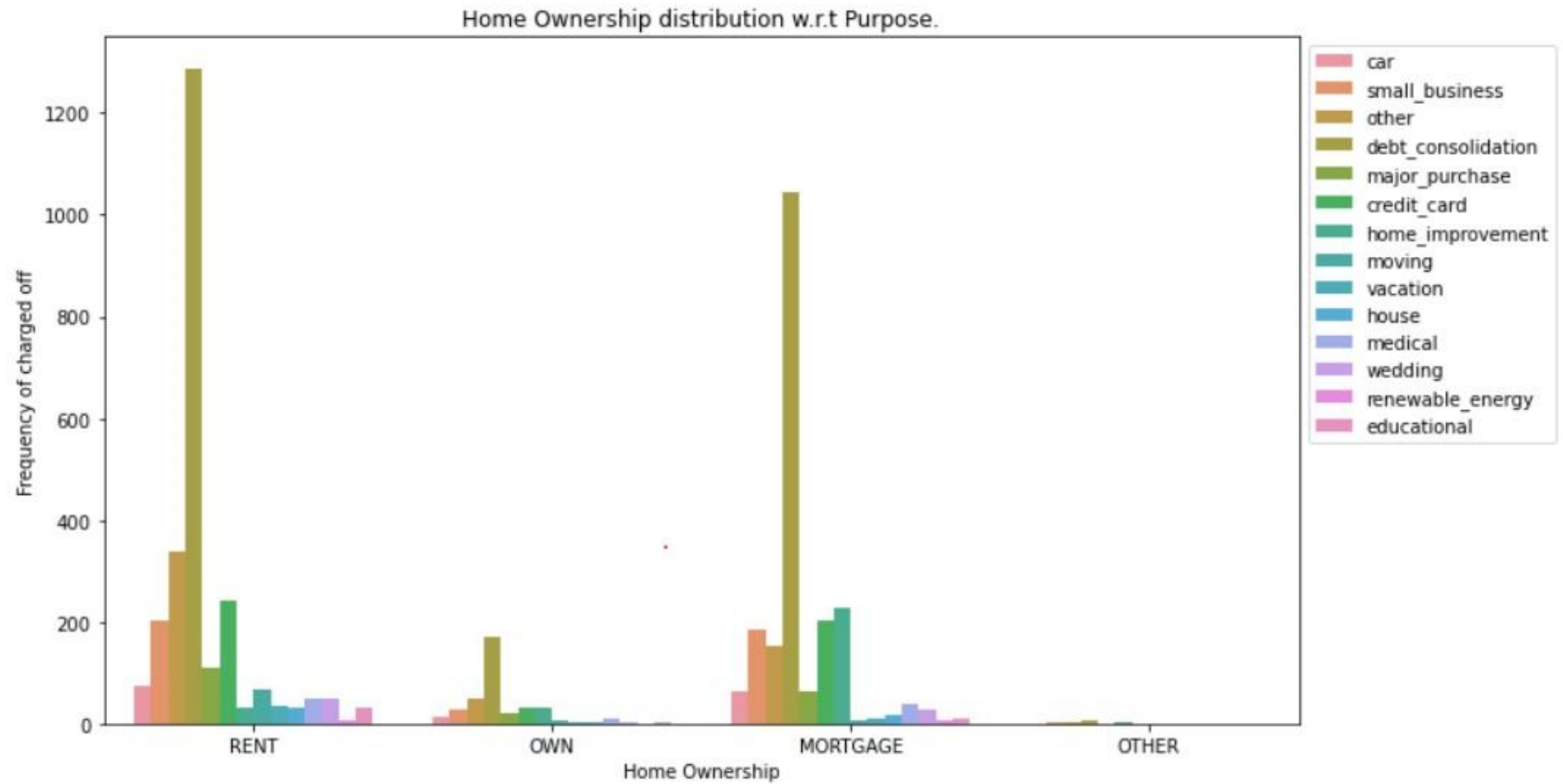
Mostly verified borrowers are getting higher Loan amount because of security reasons.



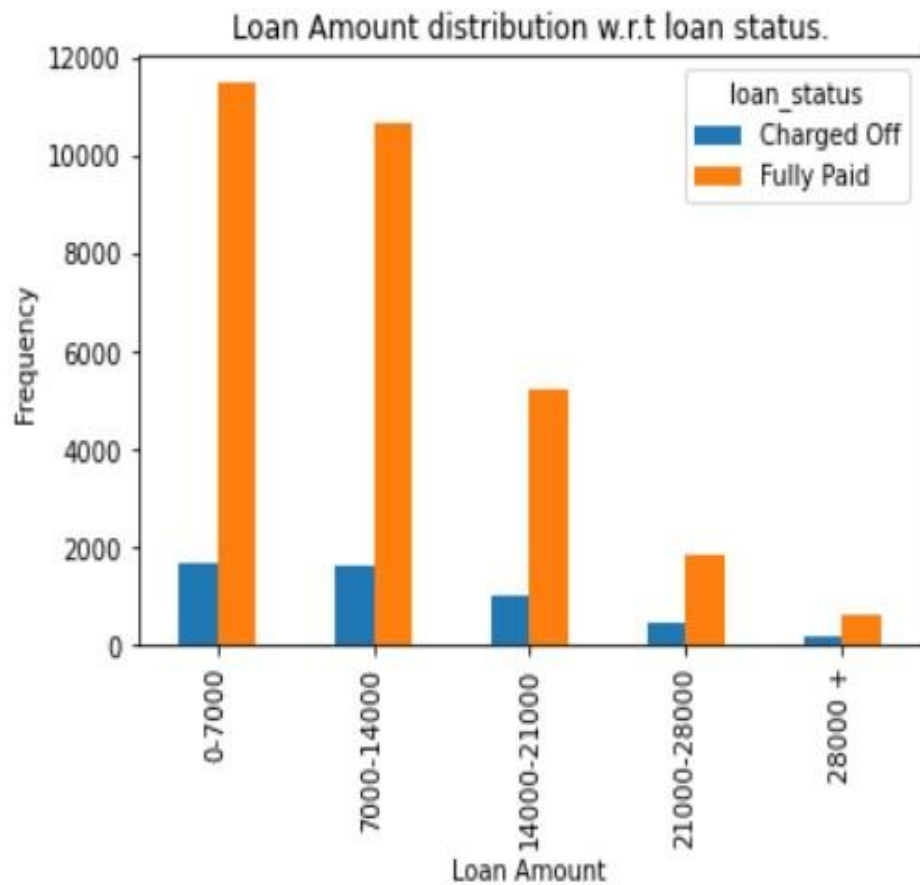
Most of the verified borrower's are getting charged off due to high loan amount and 60 months tenure.



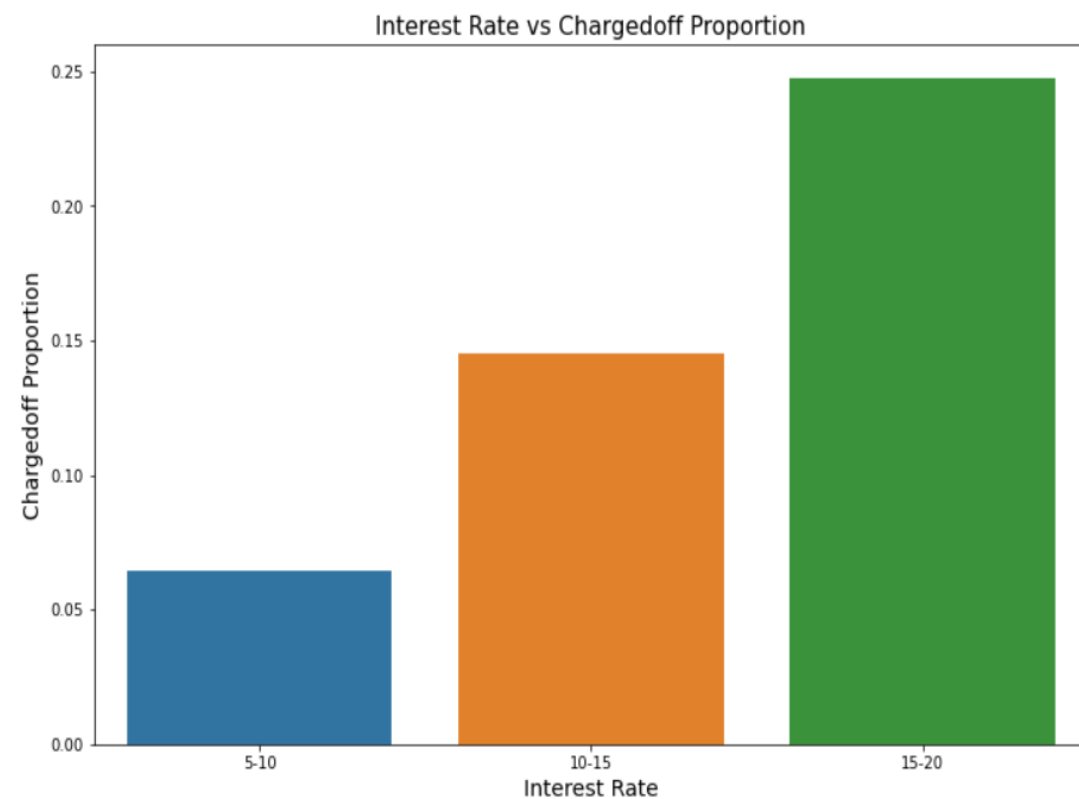
As annual income is increasing, charged off portion is decreasing.
So, we can say that lower the annual income Most is the chance of getting defaulted



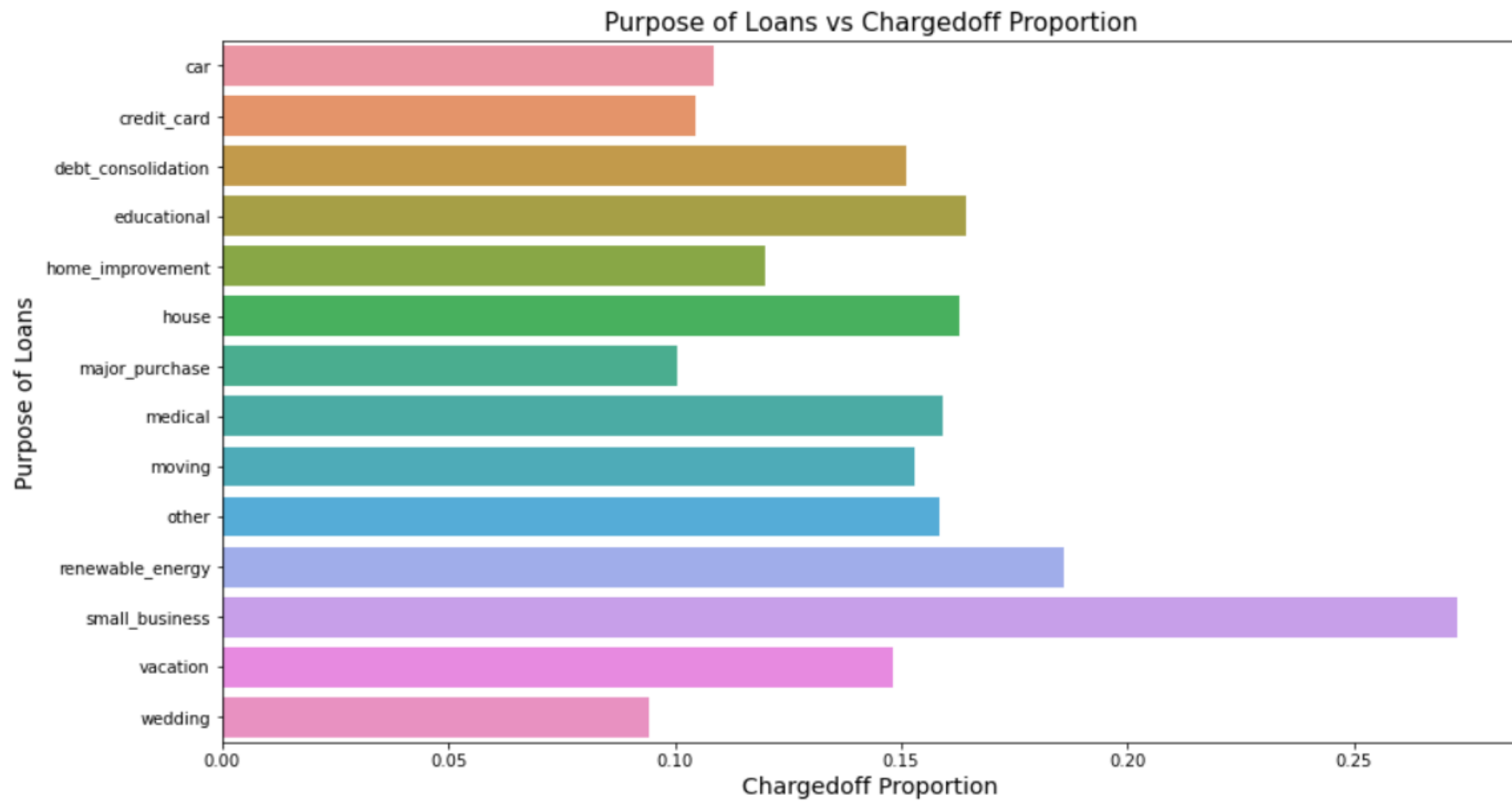
According to this, we can say that more loan is taken by the borrower whose home ownership is rent and they get charged off when they take loan for the purpose of debt consolidation.



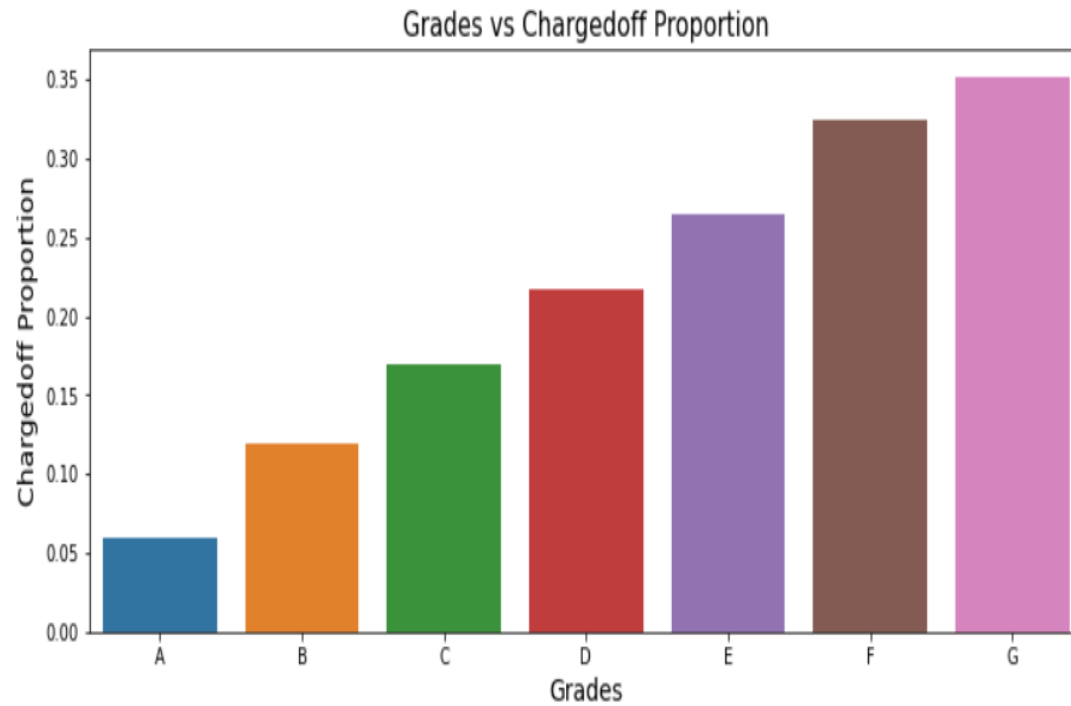
The borrower who took loan in the range 0 to 14k are mostly getting defaulted.



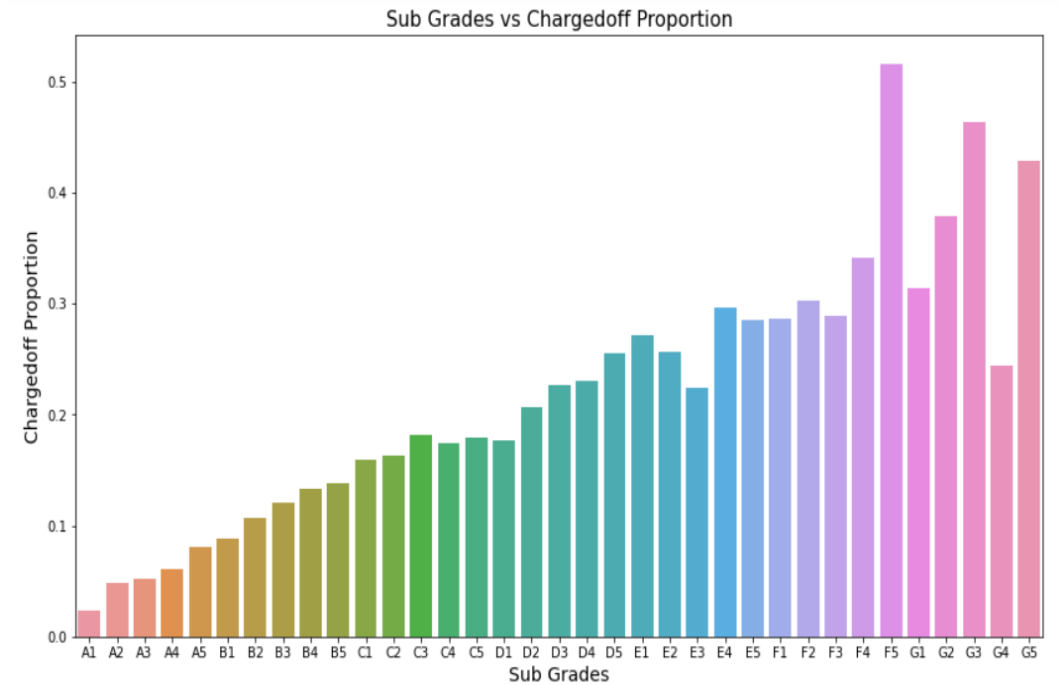
Loan with 15% to 20% interest rate are mostly getting defaulted.



The borrower who took loan for the purpose of small business are mostly getting charged off.



As grades are decreasing, charged off proportion is increasing.
So, we can say that lower the grade more is the chance of getting defaulted due to high loan amount and high interest.



In subgrade, mostly F5, G3 and G5 are getting charged off.



RECOMMENDATIONS

From above analysis we can say, that there is more probability of defaulting when

- Borrowers are taking loan for the term '60 months'.
- Borrower's whose loan status is 'Verified' as they taken high amount of loan with 60 months tenure.
- Borrower's who are having home ownership as 'Rent' and they take loan for the purpose of debt consolidation.
- Borrower's whose annual income is low i.e. (0-20000).
- Borrower's who takes loan amount in the range 0 to 14000.
- Borrower's who receive interest at the rate of 15-20%.
- Borrower who takes loan for the purpose of small business.
- Borrowers with lower Grade i.e. $F < G$.
- Borrower's whose subgrade is F5,G3,G5.



Thank You!

Presentation By Tejas Hankare &
Madhur Gupta