

LENDING CLUB CASE STUDY

BY-VAIBHAV SINGH

CONTENT

- Problem Understanding
- Data Cleaning
- Univariate Analysis
- Bivariate Analysis
- Recommendations

PROBLEM UNDERSTANDING

- Lending Club 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.
- Like most other lending companies, lending loans to ‘risky’ applicants is the largest source of financial loss (called credit loss). 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'.
- If one is able to identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicants using EDA is the aim of this case study.

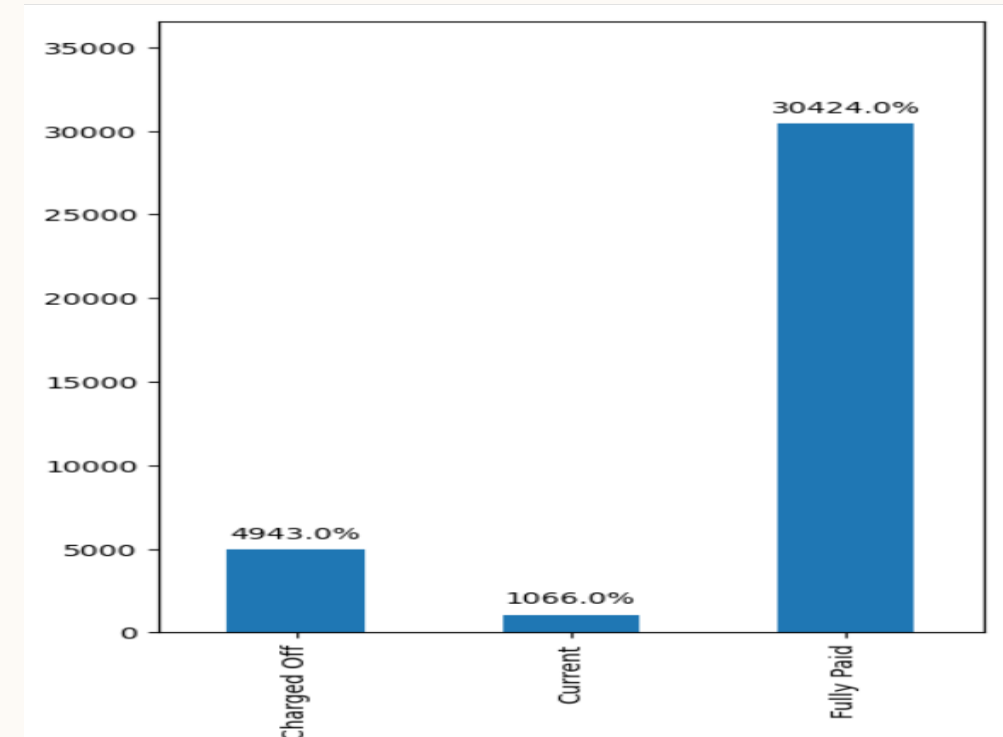
DATA CLEANING

- Removed all the columns with Null values.
- Removed all the rows with majority Null Values.
- Changed the datatype of the columns according to the value of the columns.

UNIVARIATE DATA ANALYSIS

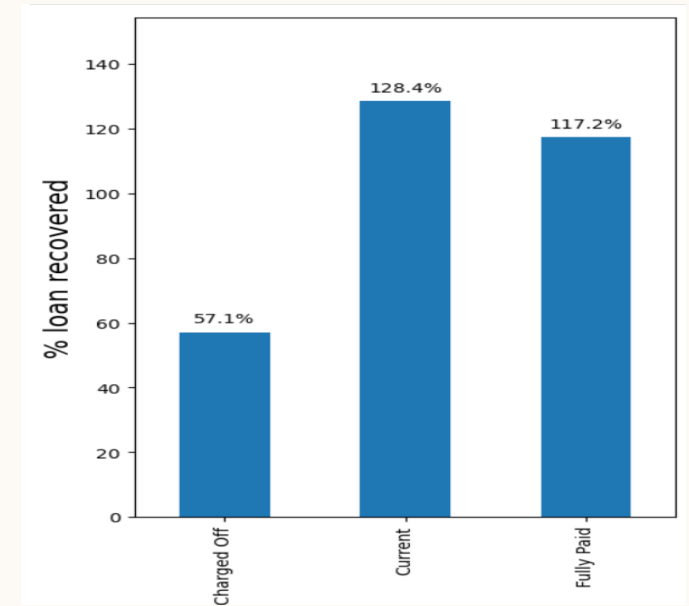
LOAN STATUS

- Funded amount lies between 0 to 35000.
- In graph majority of the people fully paid off the loan so it is better to provide them loans and reduce the business loss.



PERCENTAGE OF LOAN RECOVERED

- Lending Club recovers only 57% of the loan amount when loans are defaulted.
- On fully paid up loans, the company makes 17% profit.



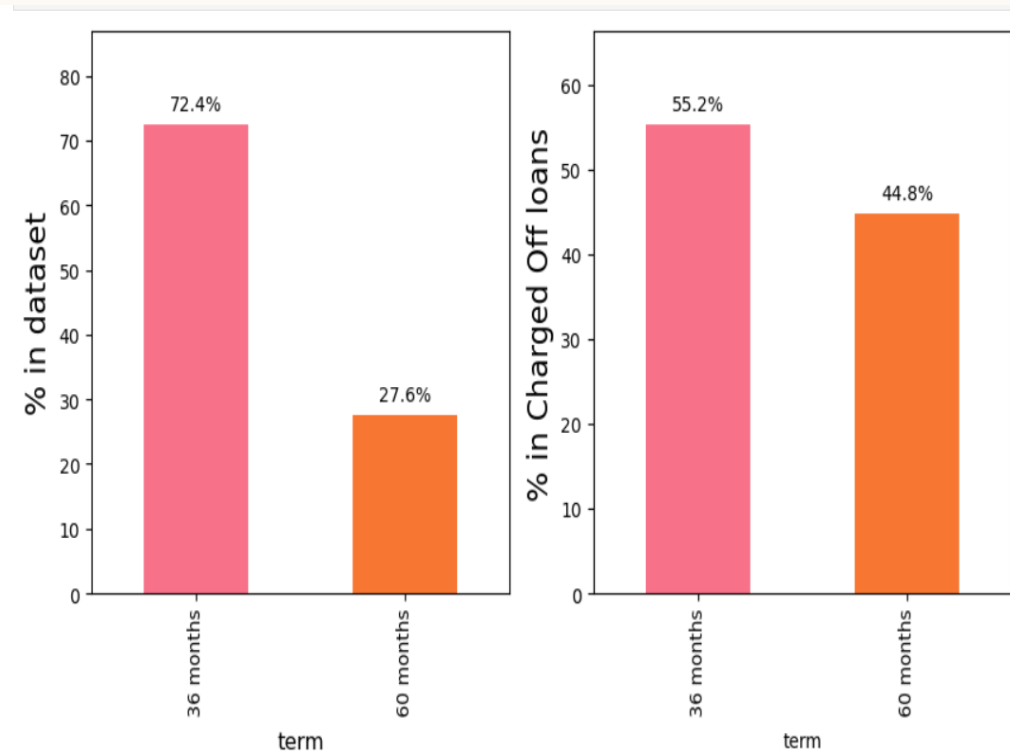
LOAN AMOUNT DISTRIBUTION

- Overall, the applied loan amount distribution is slightly right-skewed with mean greater than the median.
- Most of the loans granted are below the 15000(75% percentile).
- Charged off loans are shifted towards higher average loan amount request

	count	mean	std	min	25%	50%	75%	max
loan_status								
Charged Off	4943.0	12286.147077	8064.255961	1000.0	6000.0	10000.0	17000.0	35000.0
Current	1066.0	17100.703565	8571.093163	1000.0	12000.0	16000.0	22800.0	35000.0
Fully Paid	30424.0	10945.731166	7181.867451	500.0	5500.0	9600.0	15000.0	35000.0

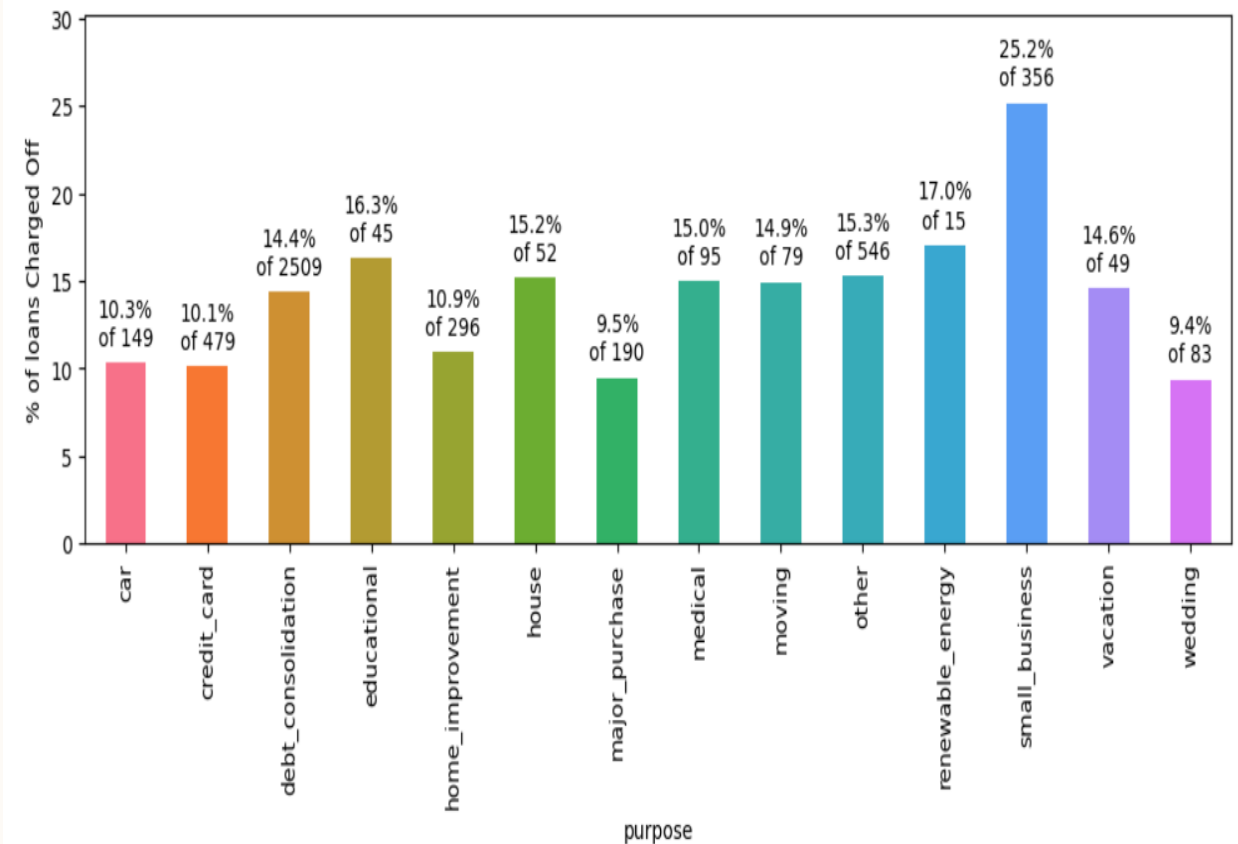
LOAN TERM

- Around 75% of the total loans are given for duration of 3 years, while just 25% of the loans are those given for 5 years.
- Among Charged Off loans, percentage of term 60 months rises to 45%. The higher term loans have a higher chance of default.



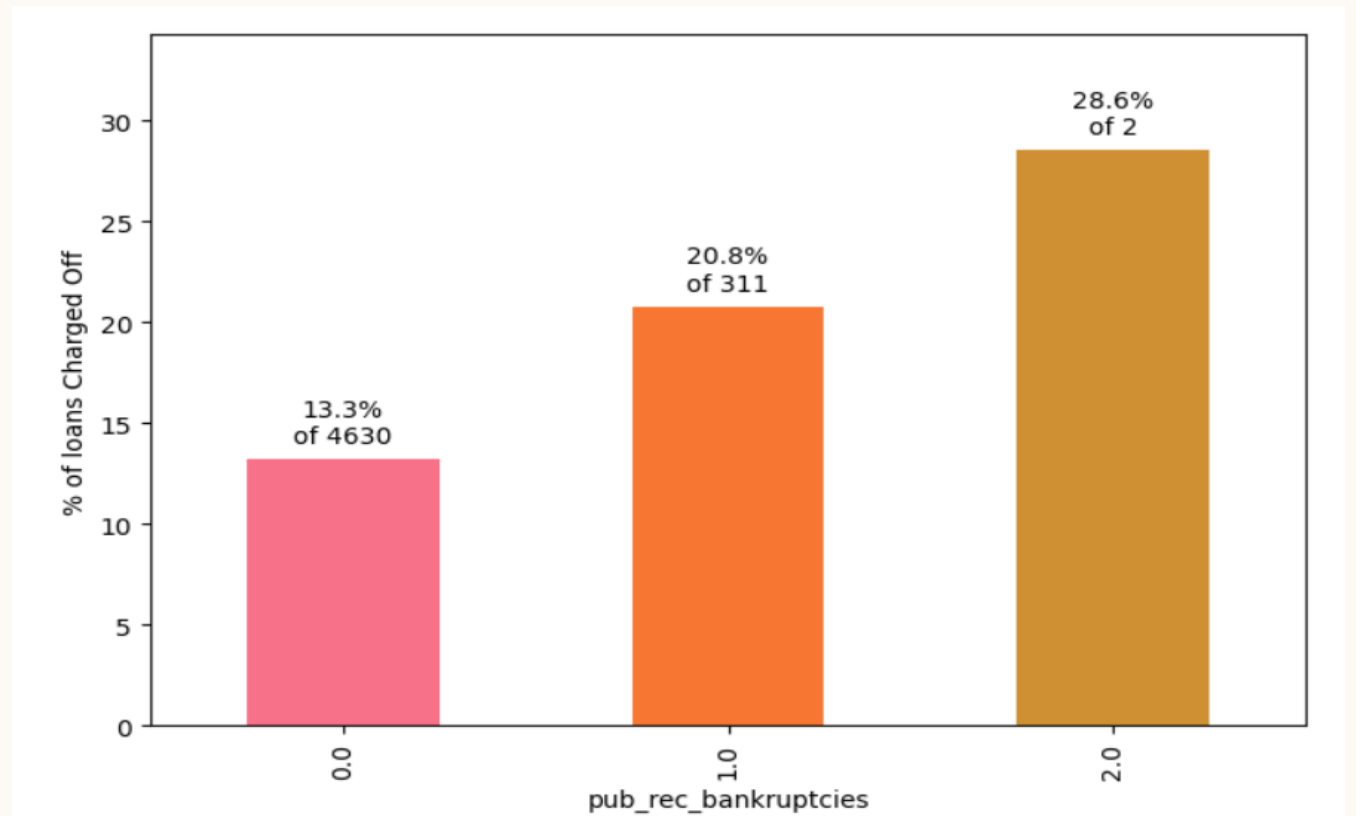
PURPOSE OF LOAN

- 26% of loans for small business are Charged Off. Making them the most risky purpose.
- Approximately ~49% of the loans are issued for the purpose of dept consolidation.
- 17% of the loans for renewable_energ are Charged Off, but the number is too less to be of significance.



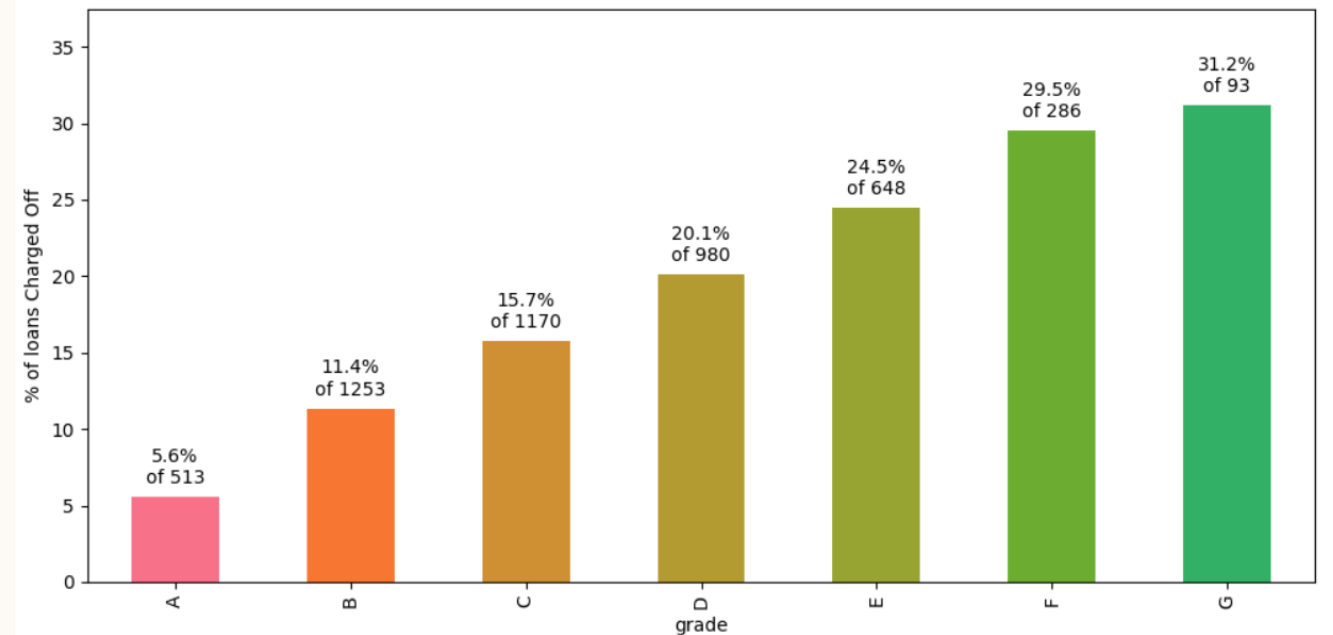
NUMBER OF PUBLIC RECORD BANKRUPTCIES

- The percentage of Charged Off loans is markedly higher when the borrower has a prior record of bankruptcy.



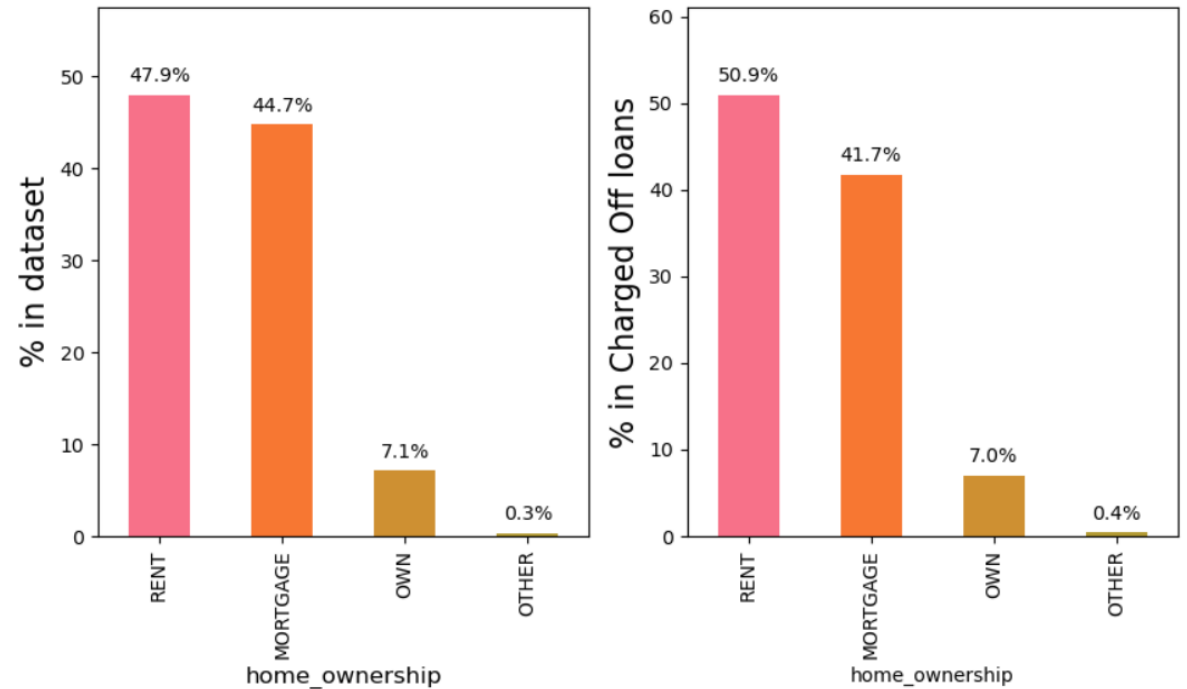
ANALYSIS OF GRADE ASSIGNED BY LC

- Nearly 30% of all loans in Grades F and G see a default.
- Grade E onwards are risky, and less numerous. Lending Club should either refuse loans or charge high interest rates.



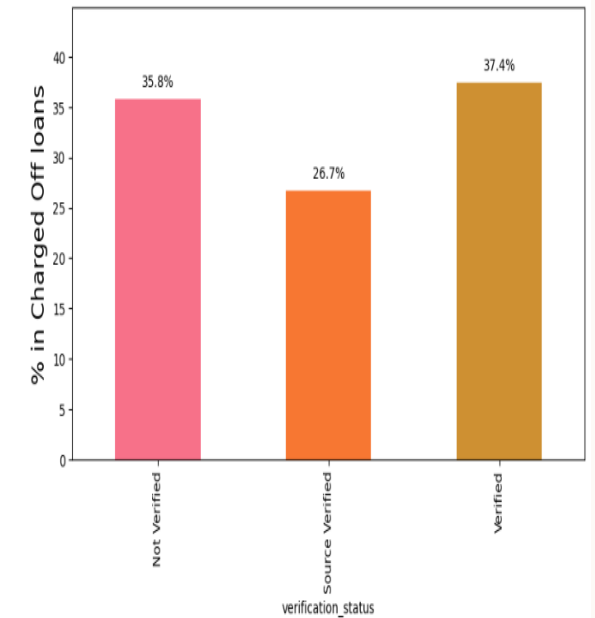
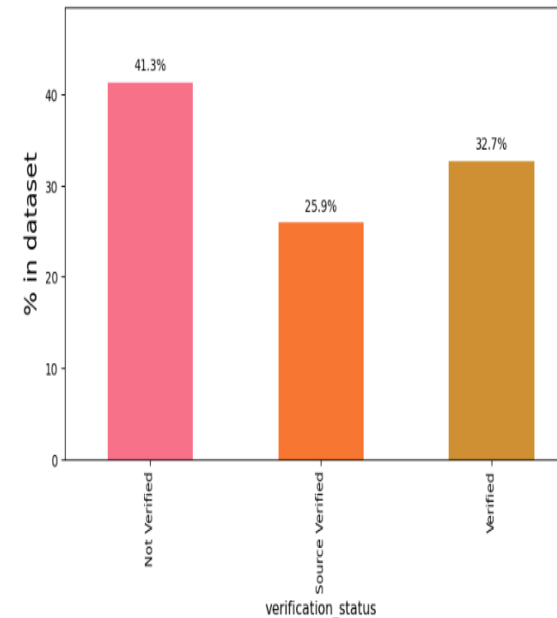
HOME OWNERSHIP DATA WITH LOAN COUNT

- There is no such impact of home_ownership to default on loan



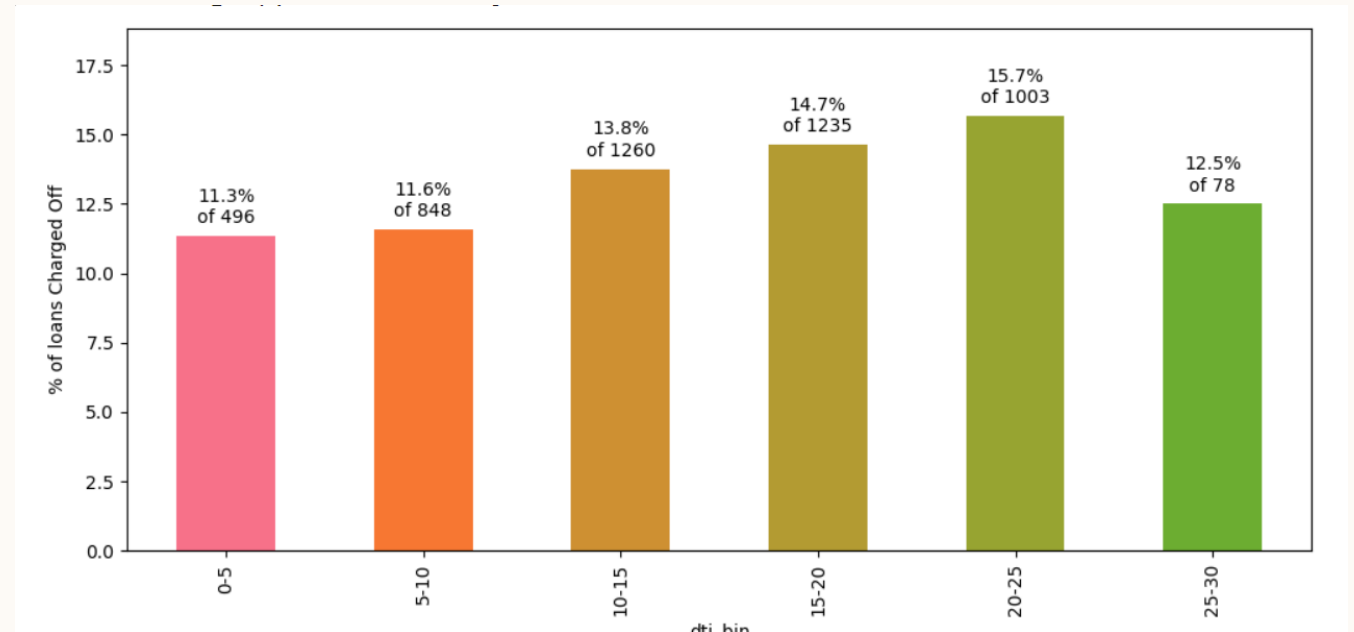
INCOME VERIFICATION STATUS

- Overall, it appears that there are more 'Non Verified' loans than the 'Source Verified' and 'Verified' loans. LC should take steps to verify these loans income so as to reduce the Charge Off percentage.
- Verified loans show more charged Off percentage as compared to Not verified loans. This is unexpected and will be investigated later in bivariate analysis



DEPT PAYMENT TO INCOME RATIO

- When the dti or dept payment to income ratio is higher than 20, higher percentage of loans are Charged Off
- Higher the dti higher the chances of loan being Charged Off

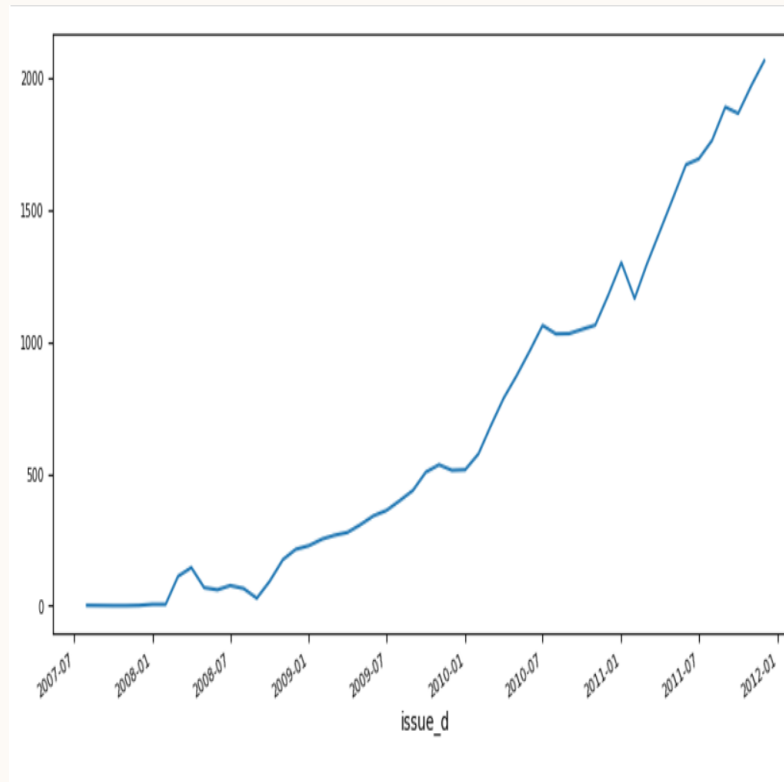


BIVARIATE ANALYSIS

15

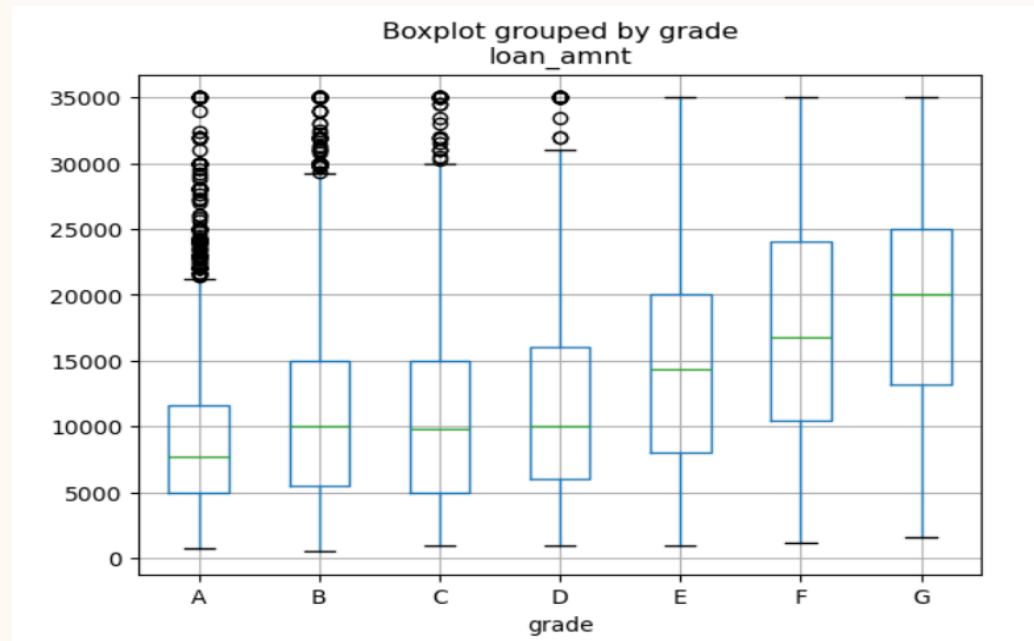
LOAN AMOUNT OVER TIME

- Loan amount over time increases.



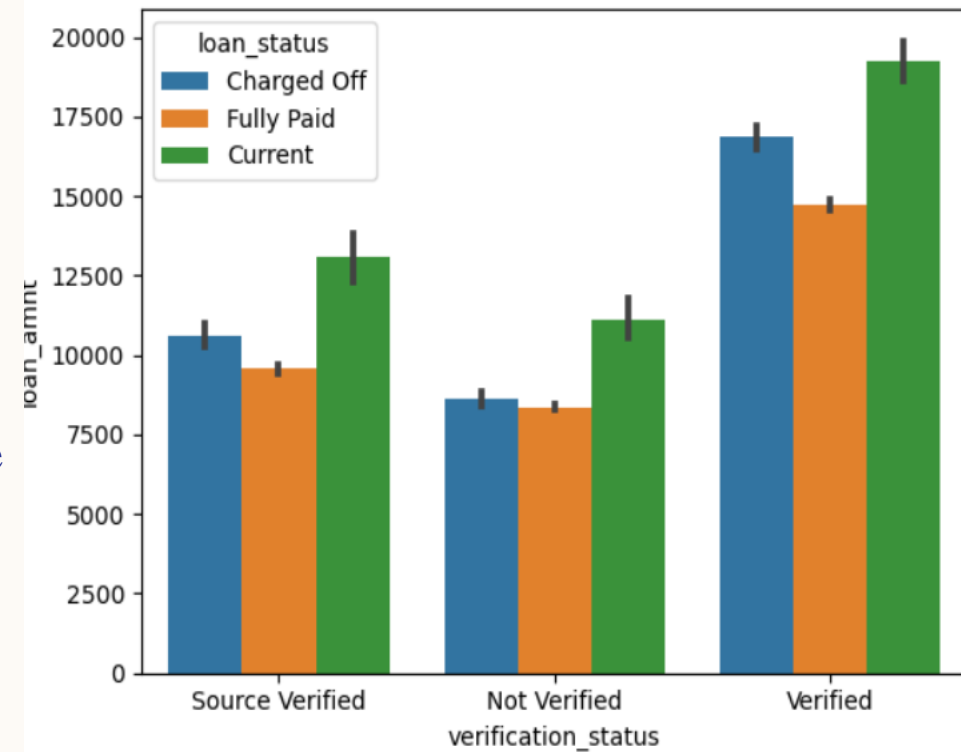
LOAN AMOUNT VS GRADE

- Larger loans generally appear to be given a lower grade, with the median loan amount for a grade G loan being almost 10000 higher than that of a grade A, B, or C loan.



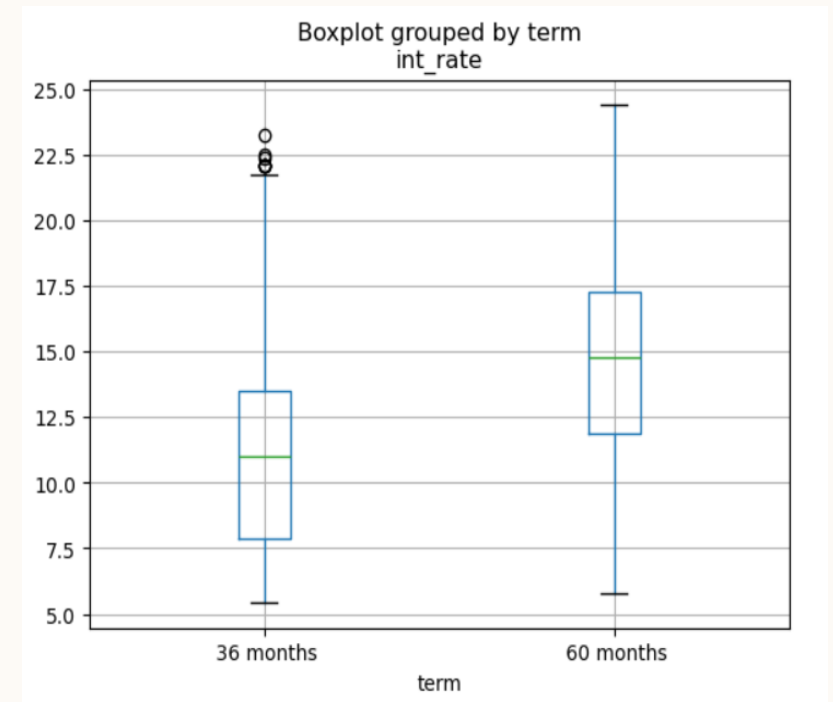
LOAN AMOUNT VS VERIFICATION STATUS VS LOAN STATUS

- Higher loan amounts are Verified more often.
- We already know that larger loans are less in number, but see a higher charge off rate.
- This, combined with previous observation, explains why verified loans see a higher rate of default. It's not the verified status per se, it's the fact that higher loan amounts are riskier and are also verified more often by Lending Club.



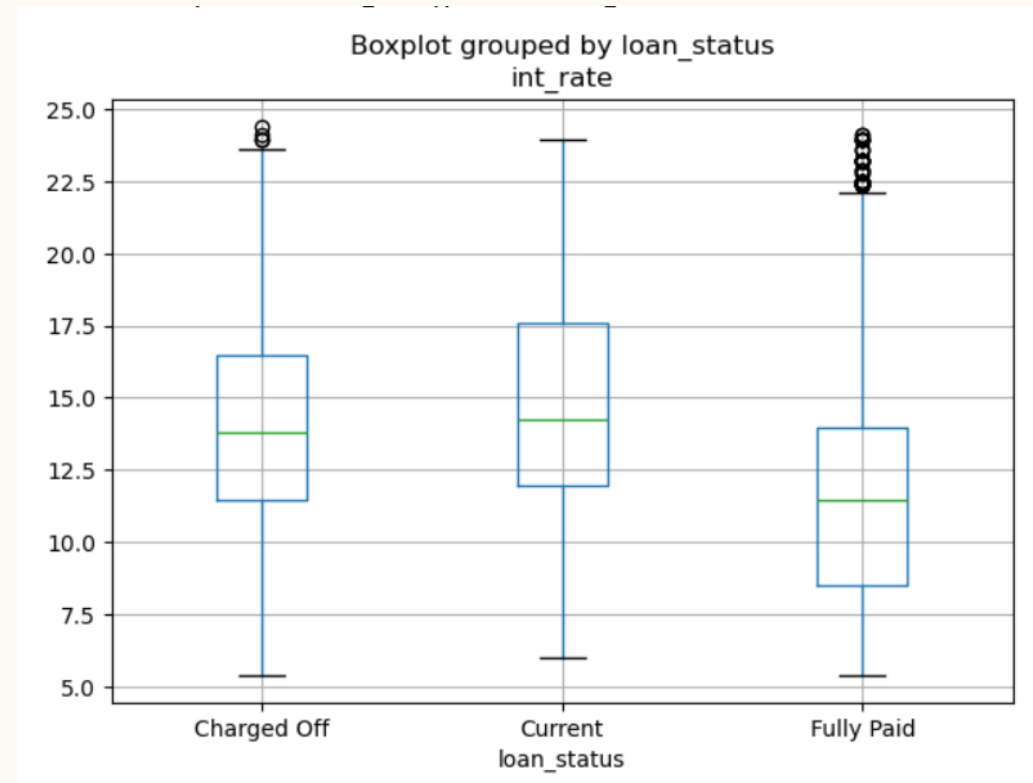
INTEREST RATE VS TERM

- Interest rates are based on term. Larger amounts were seen to be given for higher term. The rate of interest associated with them is also high.



INTEREST RATE VS LOAN STATUS

- Loans at a higher interest rate are more likely to be Charged Off.



RECOMMENDATIONS

- Lending club should check borrowers grade in order to reduce the risk. Grade A candidates are more likely to pay full loan.
- Grade E onwards are risky, and less numerous. Lending Club should either refuse loans or charge high interest rates.
- Lending club should reduce the loans for 60 months tenure, they are prone to loan defaulters.
- Lower the dti ratio, lower the chance of loan default. It is likely that low dti ratio borrower are more likely to pay full loan.
- Lending Club should take steps to seriously verify income status so as to reduce the Charge Off percentage.
- While giving loans for small business purpose, LC should check all the parameters as they prone to loan defaulters.



THANK YOU!