Lending Club Case Study

Here the objective is to help the investors to take the right decision in terms of which borrower to lend loan who will not be defaulting in future.

Case Study:

LendingClub

Problem statement

Lending Club specializes in lending various types of loans to urban customers. 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

---

Case Study:

LendingClub

Business Objective

If the company approves the loan, there are 3 possible scenarios described below:

● Fully paid: Applicant has fully paid the loan (the principal and the interest rate)
● Current: Applicant is in the process of paying the instalments, i.e. the tenure of the loan is not yet completed. These candidates are not labelled as 'defaulted'.
● Charged-off: Applicant has not paid the instalments in due time for a long period of time, i.e. he/she has defaulted on the loan

-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. 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'.

-Company wants to understand the driving factors 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.

Case Study:
LendingClub

Treatment of null values and outliers

Most frequent method seems to be the correct method to treat null values.

Refer the Null_check.ipynb file for in depth steps.

Box plot/Scatter plots can be used to check outliers, however there are various ways to replace them.

Ways to treat the outliers are discussed in the outliers.ipynb file for reference.

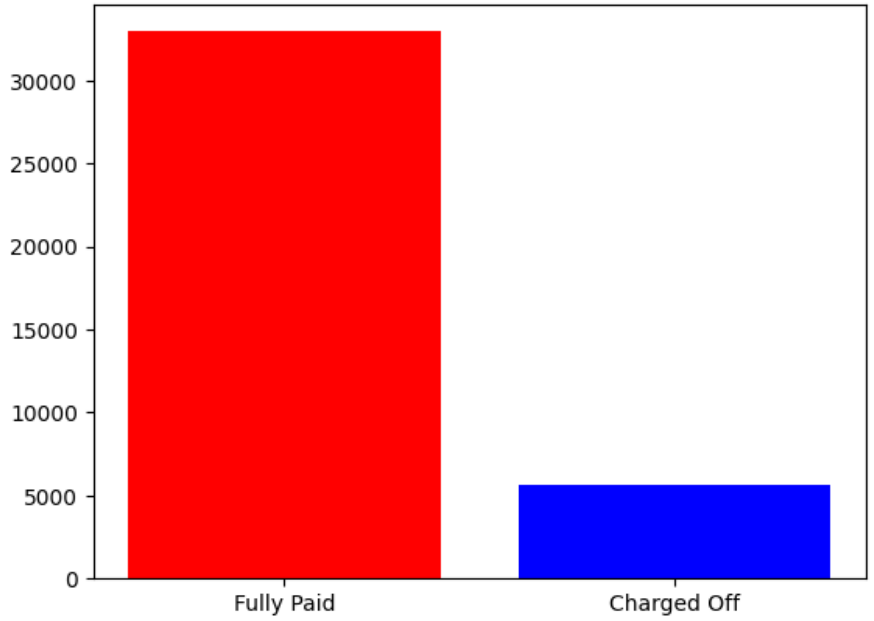
Understanding the dataset after data cleaning(removing null values)

Understanding pending features of the dataset:-

1. member_id :- Id for the borrower member
2. loan_amnt :- The amount of the loan applied for by the borrower
3. funded_amnt :- Total amount committed to that loan
4. funded_amnt_inv :- amount committed by investors
5. term :- number of payments on the loan whicn cvary between 36 to 60 months
6. int_rate :- Interest Rate on the loan
7. installment :- monthly payments for the loan taken
8. grade :- Grades are assigned from A to G depending on how risky the loan is under a category.
9. sub_grade :- After grades Sub Grades are assigned from 1 to 5 depending on risk category.
10. emp_title :- Job title assigned by the borrower
11. emp_length :- Employment of the borrowser in year
12. home_ownership :- Home ownership status of the borrower
13. annual_inc :- Annual income of the borrower
14. verification_status :- States if the income of the borrower was verified or not
15. issue_d :- loan issues month
16. loan_status :- Status of loan(defaulted or non defaulted)
17. pymnt_plan :- Payment plan for the loan taken
18. desc :- Loan description
19. purpose :- purpose of the oan taken by borrower
20. zip_code :- zip code of the borower
21. addr_state :- address of the borrower
22. dti :- ratio calculated using the borrower's total monthly debt payments on the total debt obligations, excluding mortgage and the requested LC loan, divided by the borrower's self-reported monthly income.

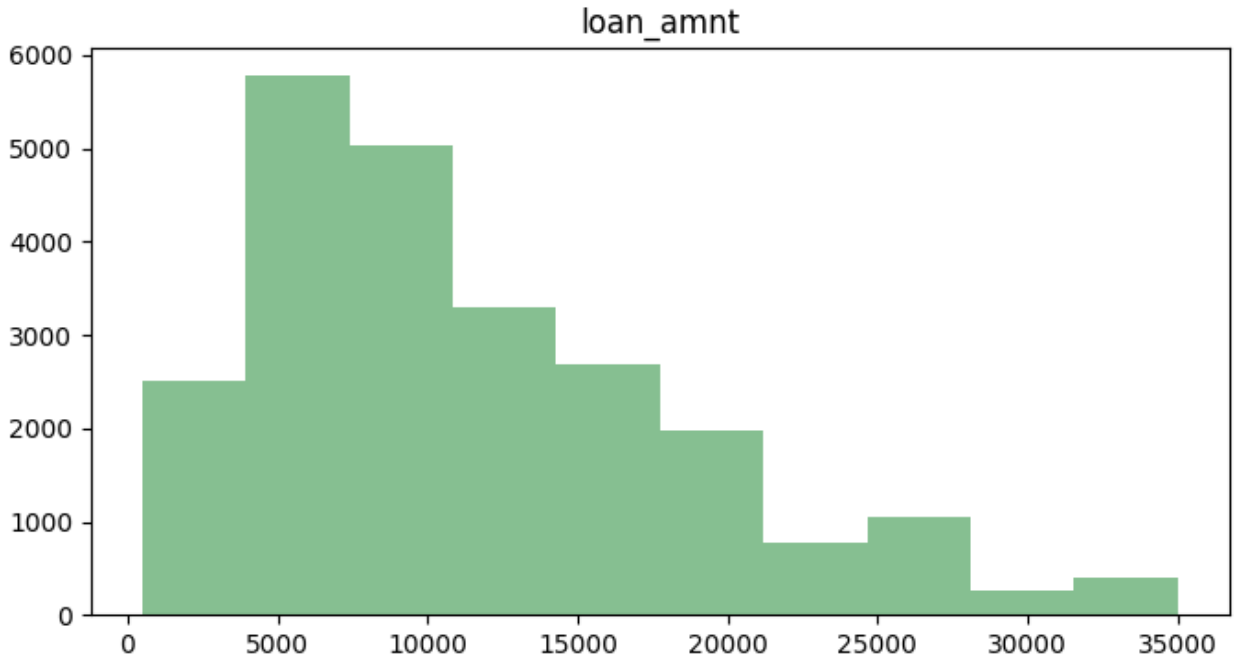
Here loan_status is the target column which will predict if the customer is defaulter.

Validation for balanced data



From the above diagram we understand that the data is not balanced hence we can use a smote technique to make it balanced

Histogram for the loan amount

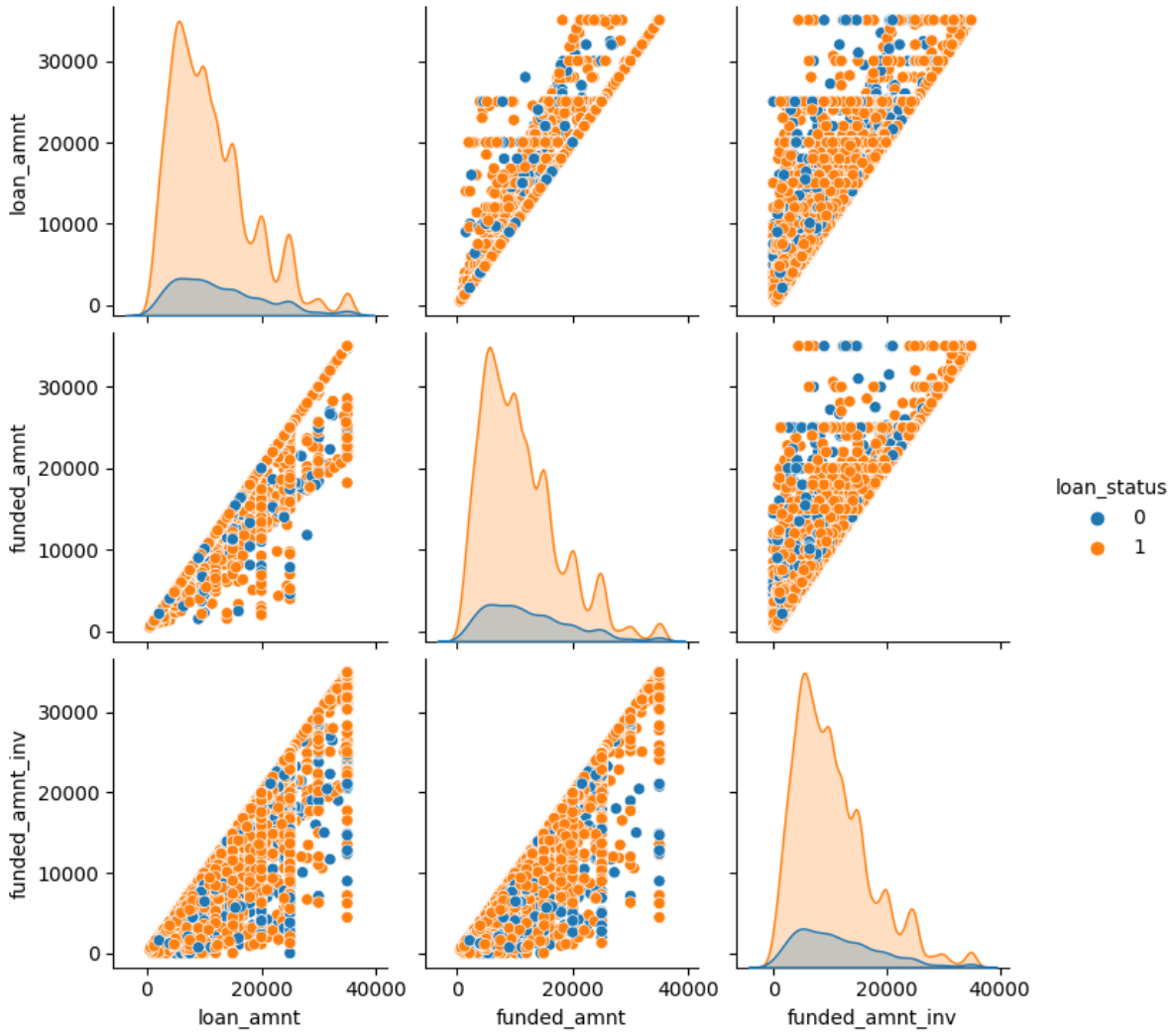


From the above diagram we understand that the most of the loan amount were most around range 5000 – 1000 from the dataset.

Comparing Loan_amt, funded_amt, funded_invest_amt vs Loan_status

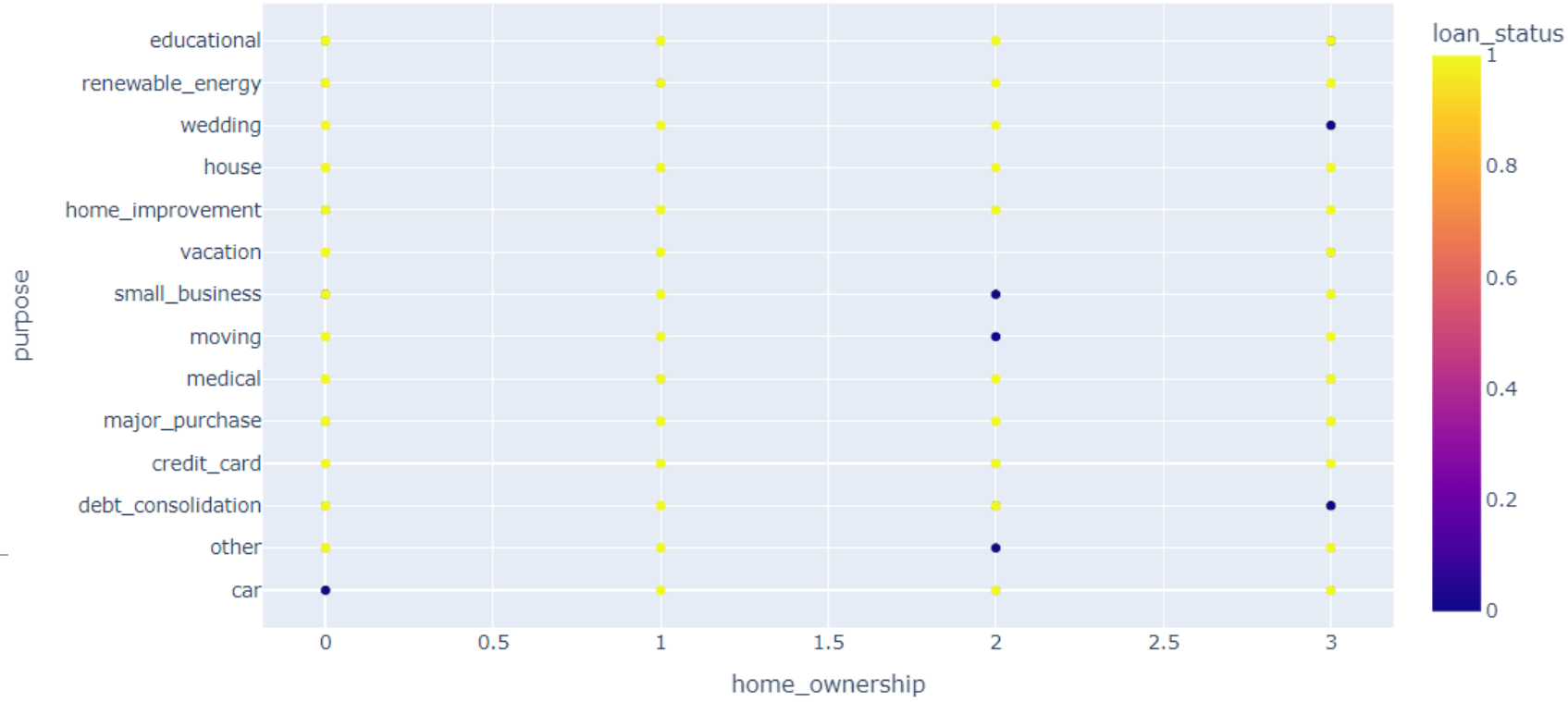
Insights :-

- We don’t see any specific relation between loan_amt vs loan_status as the defaulters are split across the whole region.
- We also don’t see any relation between funded_amt vs loan _status
- No relation from funded amt_inv vs loan_status
- Hence these columns will not help to take a decision whether to approve a loan for a borrower or not.



Comparison of home ownership vs loan_status

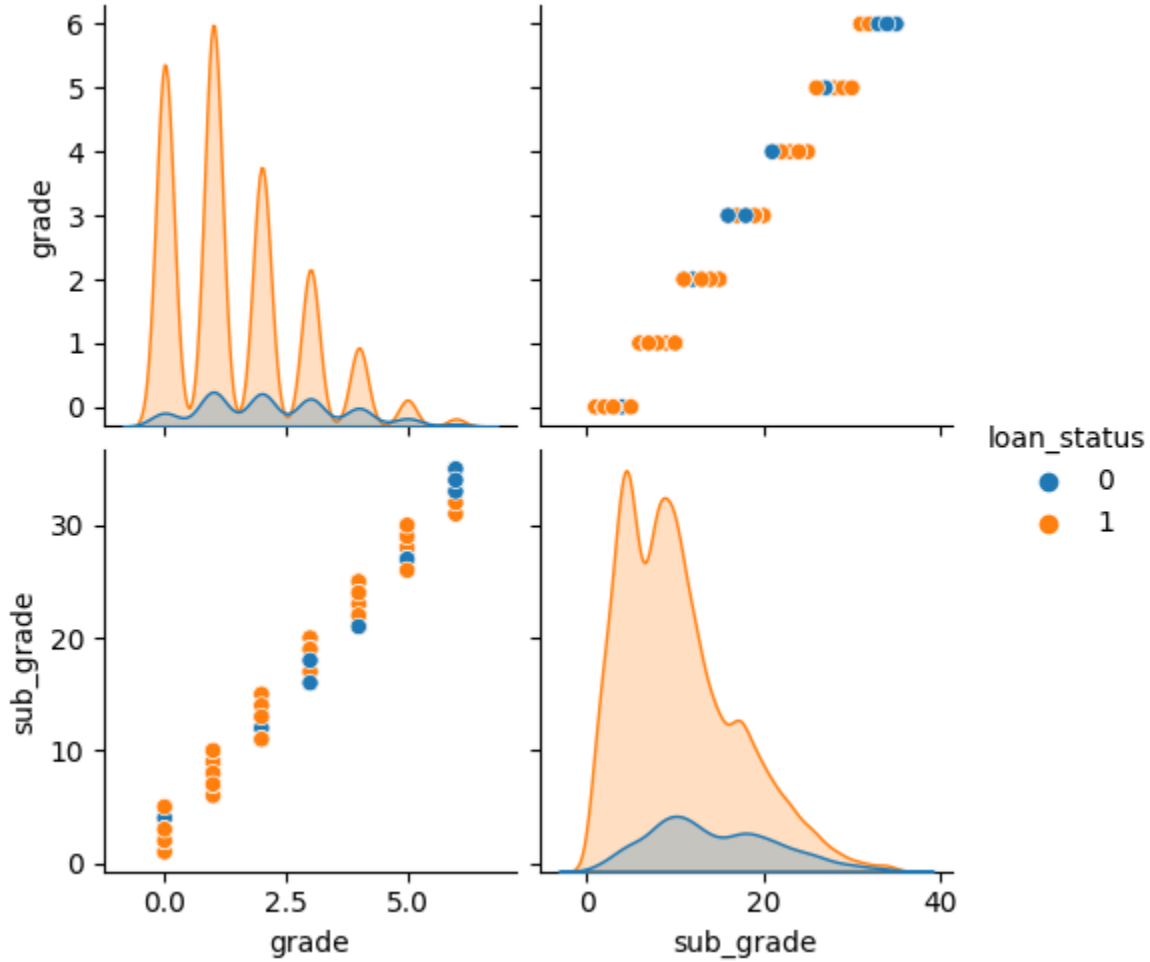
Borrowers falling under the others category and the Mortgaged their home ownership have higher chances of defaulting the loan vs people who have their own home will have less chances to default.



Comparison with grades, sub grades vs loan_status

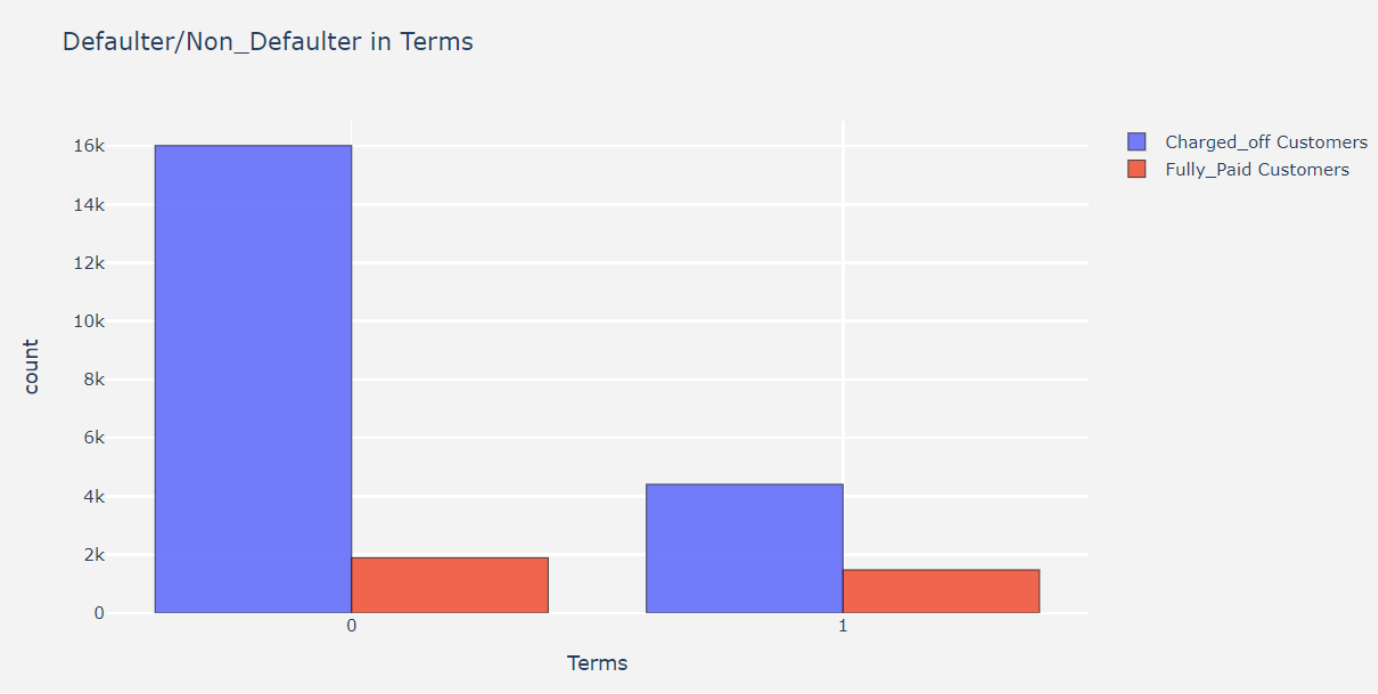
Grades from D and above have higher chances to default the loan payment in future.

Sub grades from 3 and above have higher chances of being defaulters.



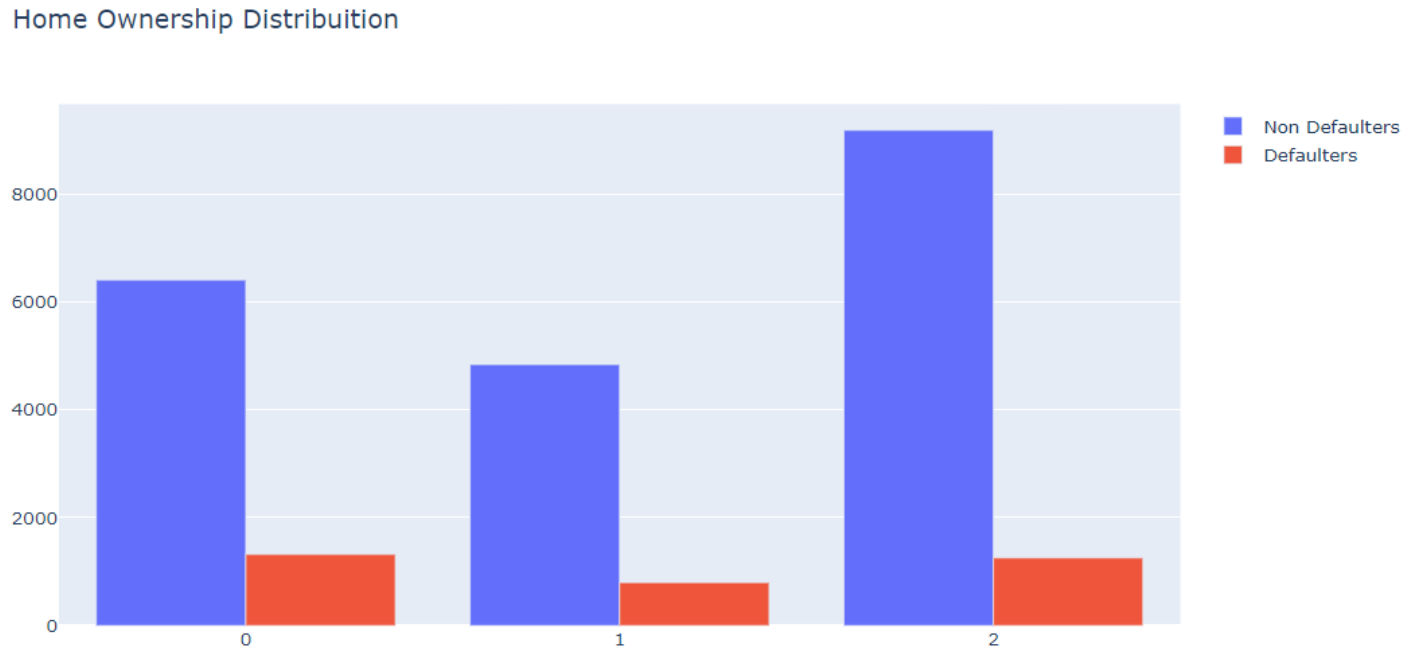
Comparison with loan term vs loan_status

We can clearly see that the borrowers with shorter duration of loan term have defaulted the payment 25% more than the ones who had opted for 60 months duration.



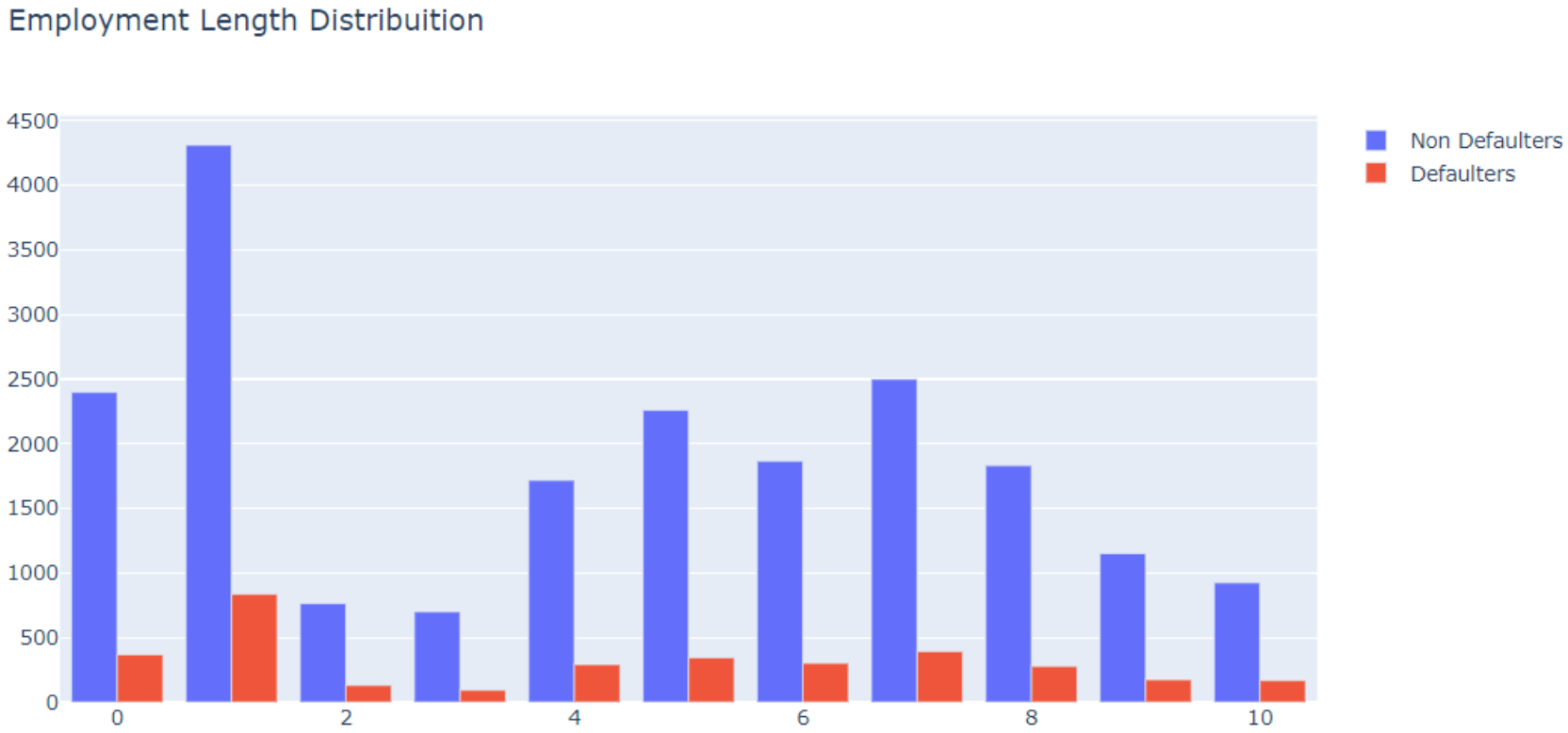
Comparison with verification status vs loan_status

Here we find the non verified and verified without a proper source have the highest chances of defaulting loan in future.



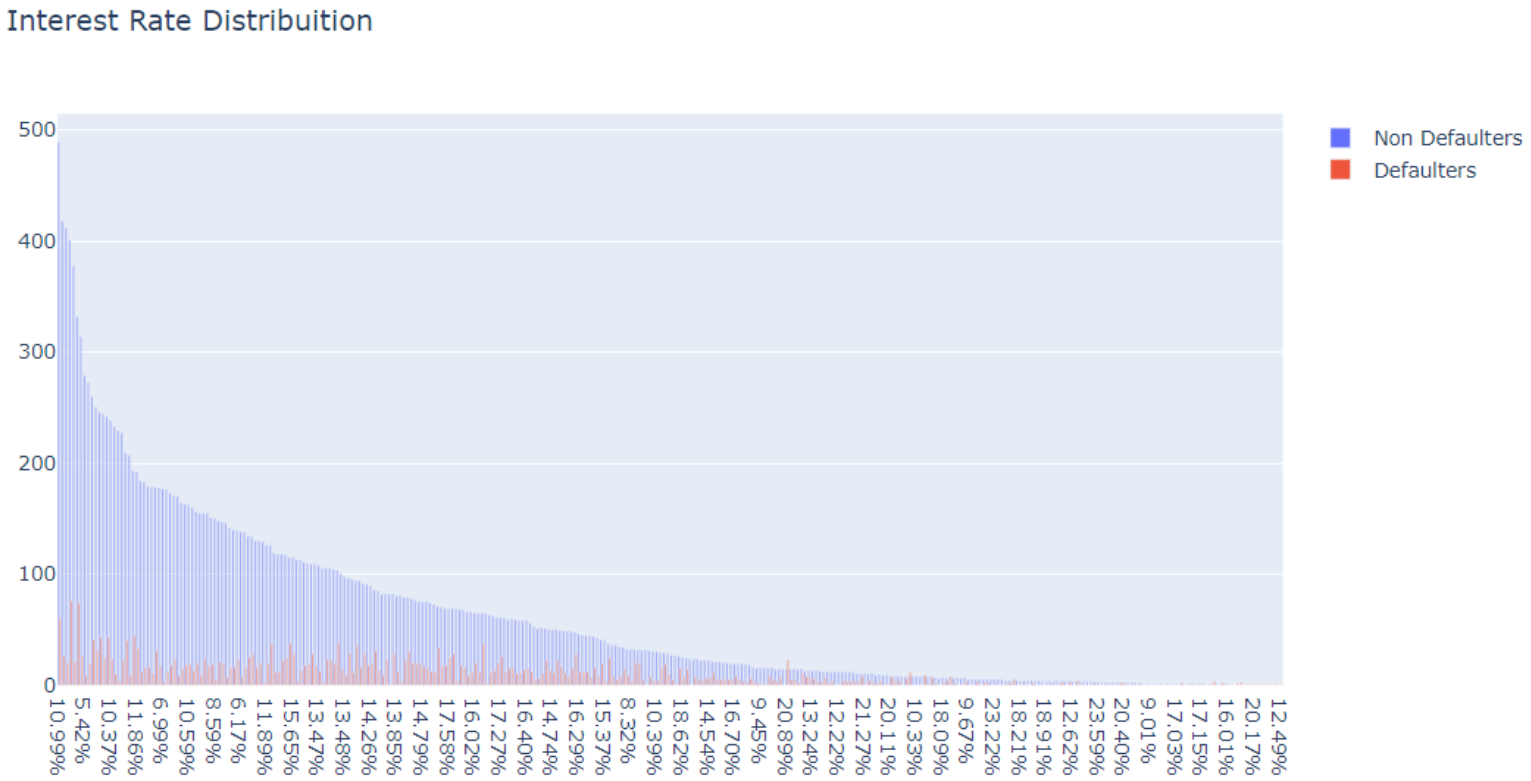
Comparison with employment length vs loan_status

Borrowers whose employment term is upto 1years has maximum chances to default, additionally also borrowers who have employment term of 4-8yrs still have chances to default.



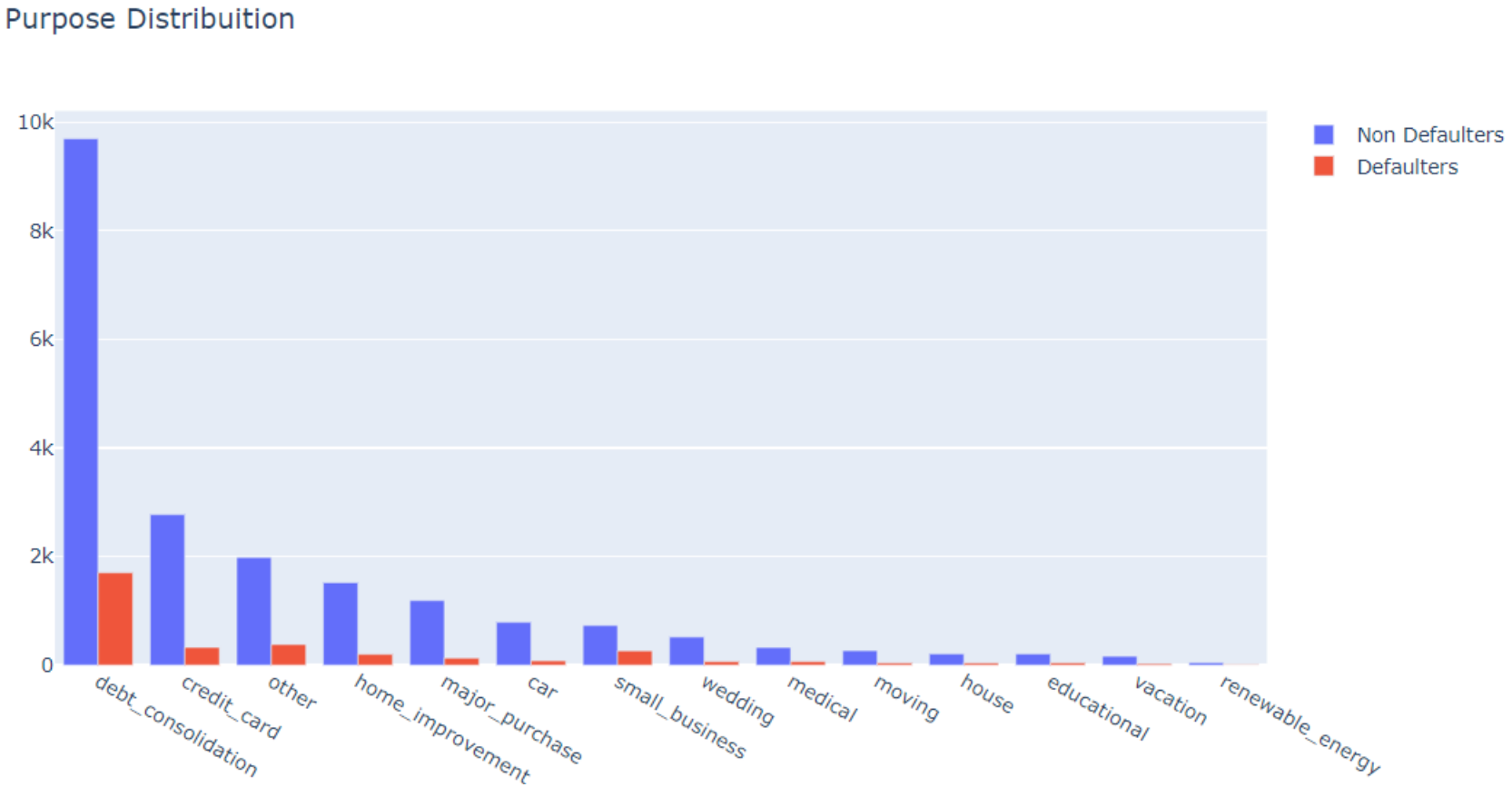
Comparison with interest rate vs loan_status

We really don't see any important relation between the interest rate and the defaulters hence this cannot be considered an important metric.



Comparison with purpose vs loan_status

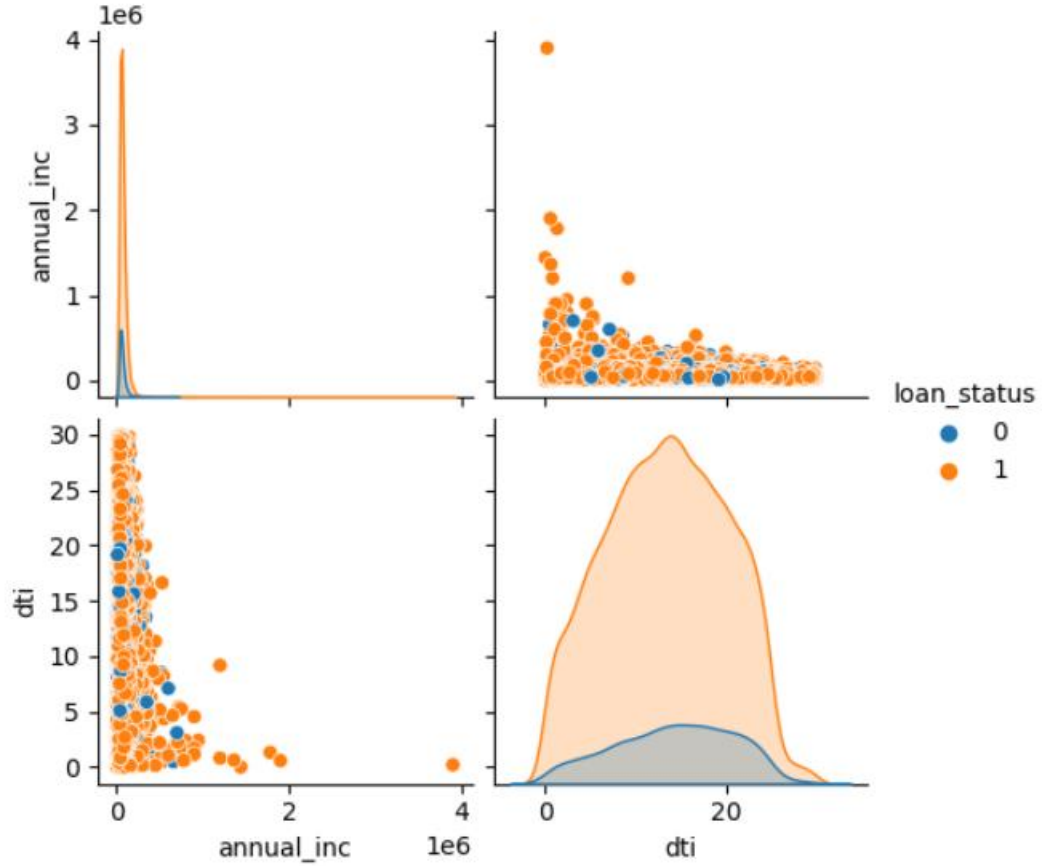
Borrowers falling under the category of debt consolidation, small business, credit card are most likely to default.



Comparison with Annual income, dti vs loan_status

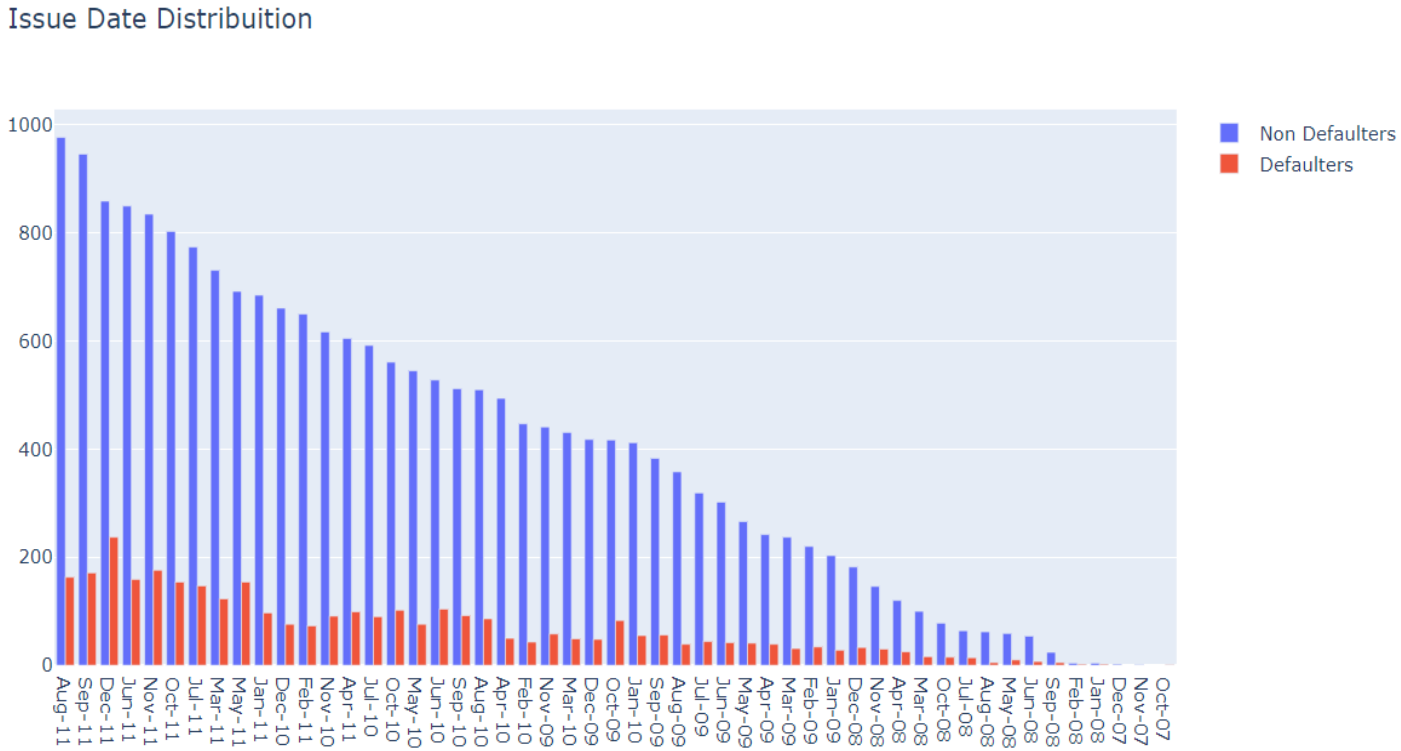
Borrowers who have salaries upto 1,500,000 are more likely to default loan payment.

But as compared to dti feature there is no significant relation which can help to predict the borrowers who can default loan payment in future.



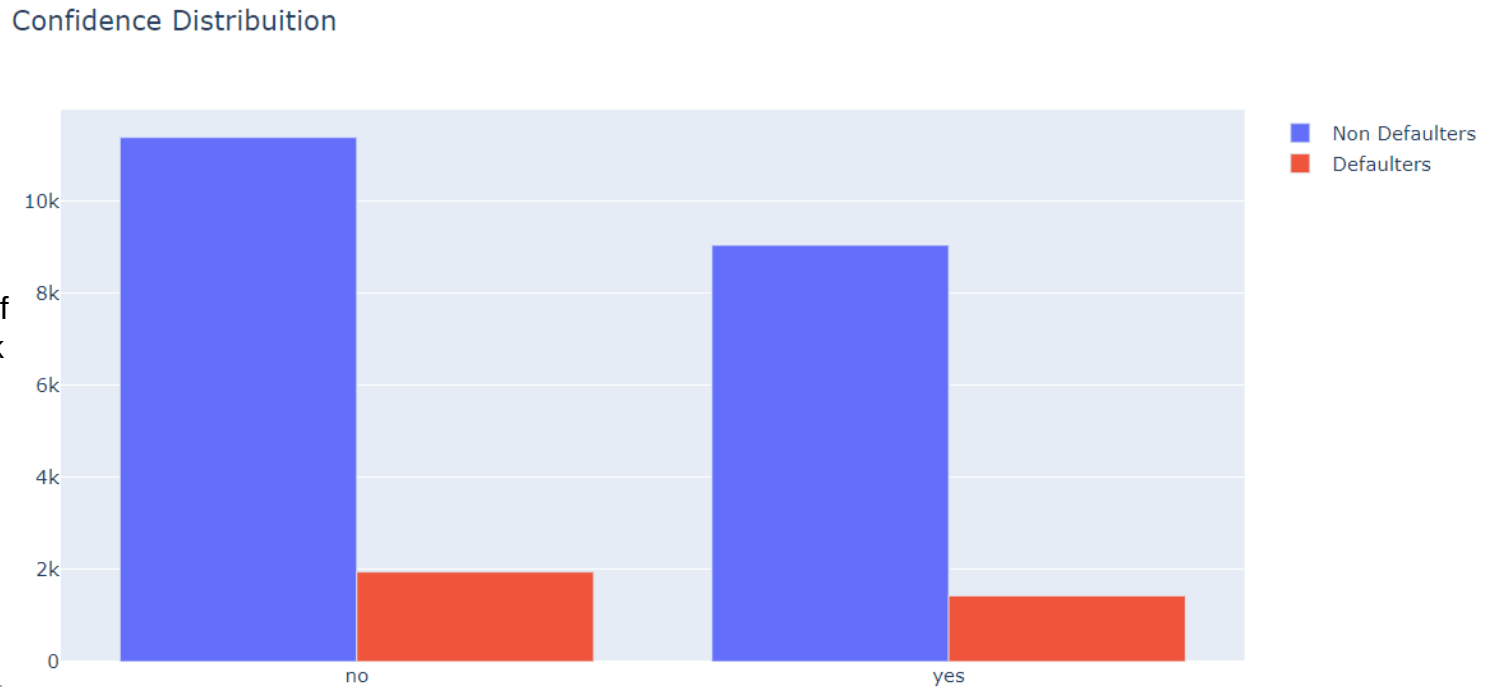
Comparison with loan issue date vs loan_status

We see that there are more defaulters for the years 2010 or before where there could be covid or recession which would have impacted so many defaulters. But the trend has settled down in the further years.



Comparison with confidence(investors) vs loan_status

Here we can see that the consumers on which the investors had more confidence have equal chances of defaulters, hence this is not a correct measure to look for approving the loan.



Confidence is calculated as the borrowers on which the investors had high confidence who would not default.

Recommendations

- ✓ Lending Club should be more cautious for approving loans falling under the group D and above, and also falling under sub group 3 and above.
- ✓ Since borrowers with lower term have higher chances of defaulting hence they need to be careful in approving their loan.
- ✓ With approving loan for the borrowers there can be a mechanism put up by Lending Club that the borrowers can also regularly make some long term investment which will be also help in case they fall in the position to default in future. Lending Club can also charge very low rate of interest out of the return earned by borrower making the investment as a another source of income.
- ✓ Borrowers need to be verified through a proper source or they would be high chances to default in loan EMI's.
- ✓ It is adviseable for the borrowers to consult a financial advisor so that they don't fall under the situation where they have to default the loan payment.
- ✓ Investors had wrong assumptions/conclusions with regards to the faith they had on borrowers who would not default in future.
- ✓ Lending club has to be more cautious for approving loans for borrowers who are in rent or falling under others category.
- ✓ Borrowers also need to be advised on having secondary source of income in case of financial distress.
- ✓ Lending club also have to be careful while approving loan for low income group people.