

GRAMENER CASE STUDY

EDA GROUP ASSIGNMENT

Group Members:

1. Surabhi Varshney
2. Sekhar Sahu
3. Prasasthy K.B
4. Anargha Biswas

Abstract

The consumer finance company specializes in different types of loans to the urban customer. The major problem they face is when the customer fails to replay the loan . 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

Business Objective :

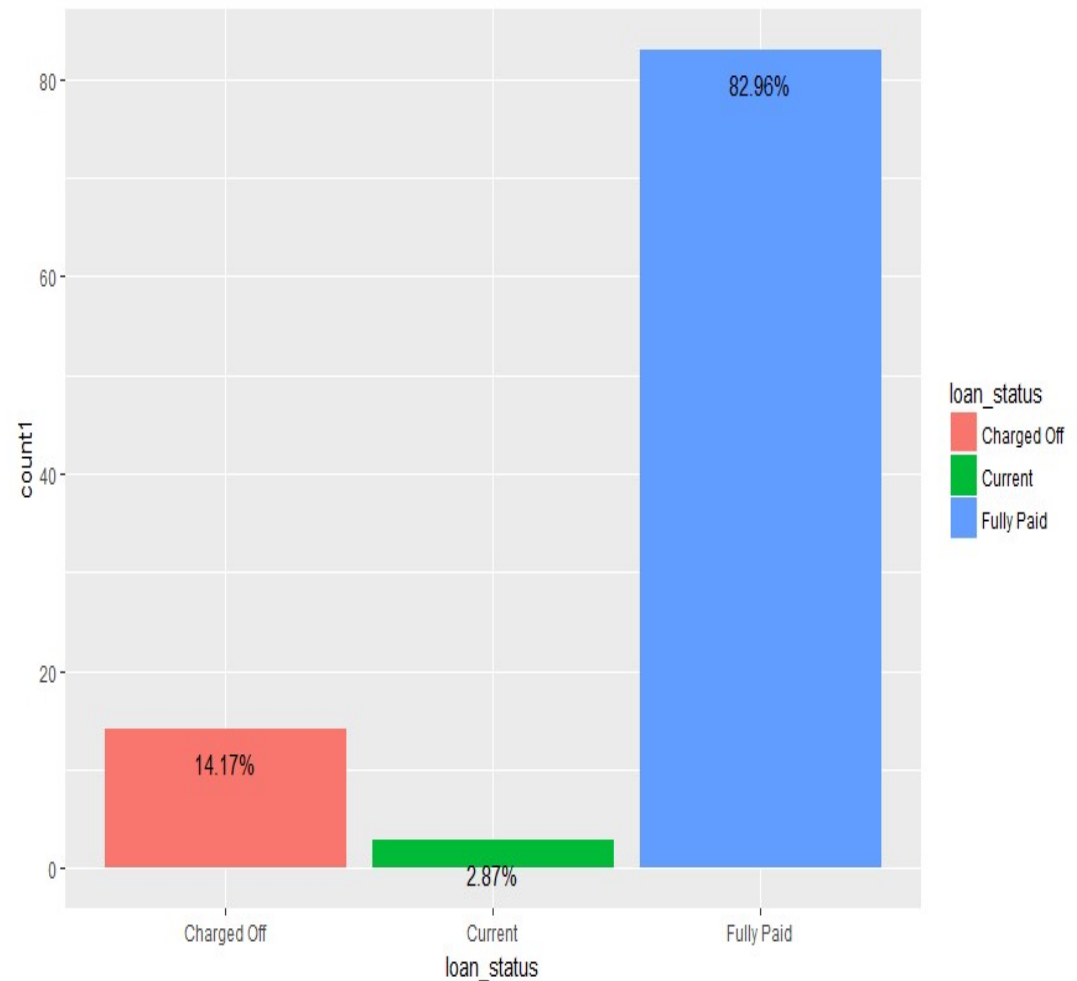
- The aim of this analysis is to try to find the customers who are likely to pay the loan and who are likely to default.
- We aim to find the parameters that significantly influences a customer's probability of being a defaulter thereby taking necessary actions.

A glimpse at the data..

- We have analyzed three states of loan, namely, fully paid, current and defaults.
- Fully paid are those for which customers have cleared the loan, current being the ones that are going on, and default being the ones that the customers failed to pay off.
- After cleaning the data we plot the loan status with different parameters and try to find the influences of those parameters on the status of loan.

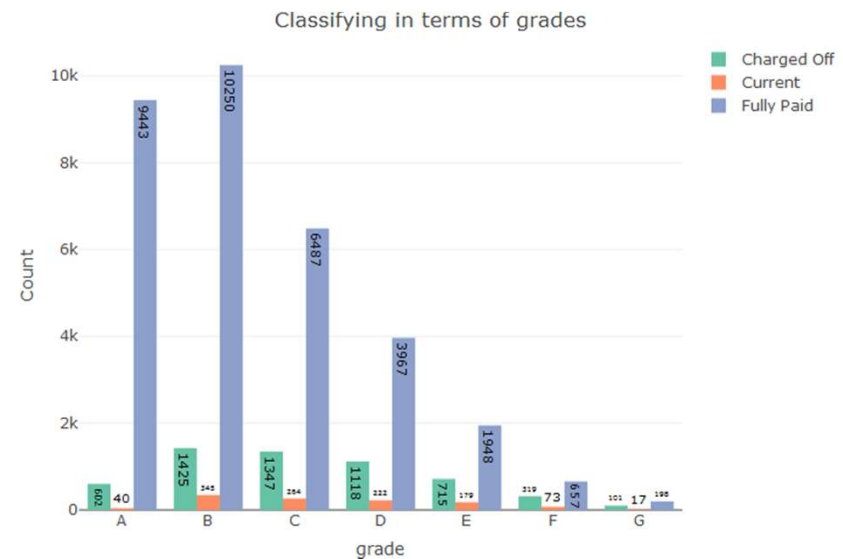
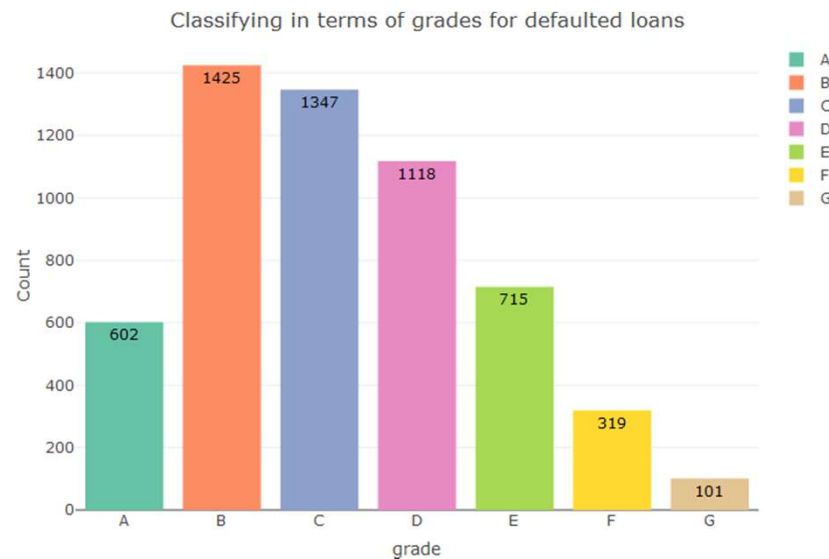
An overview of the data

There is a overall 14.17% of loan defaulters. Lets further try to find the different parameters which influences the defaulters.



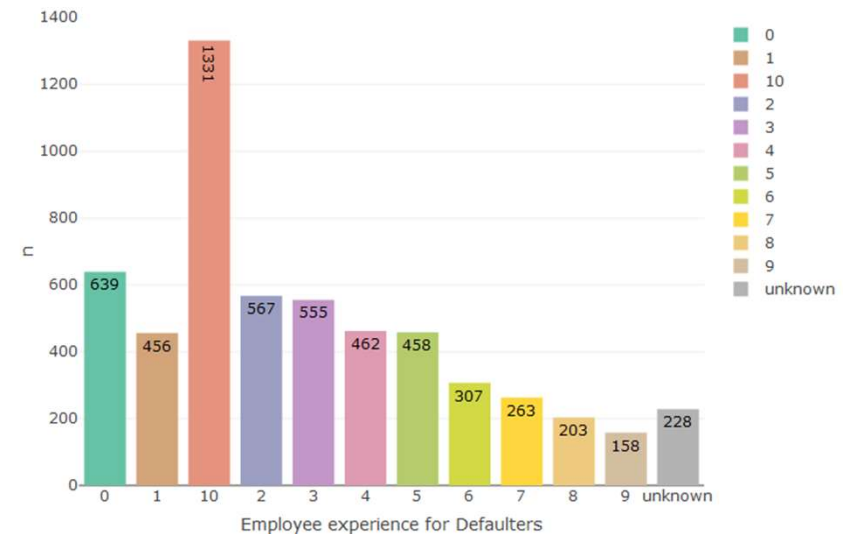
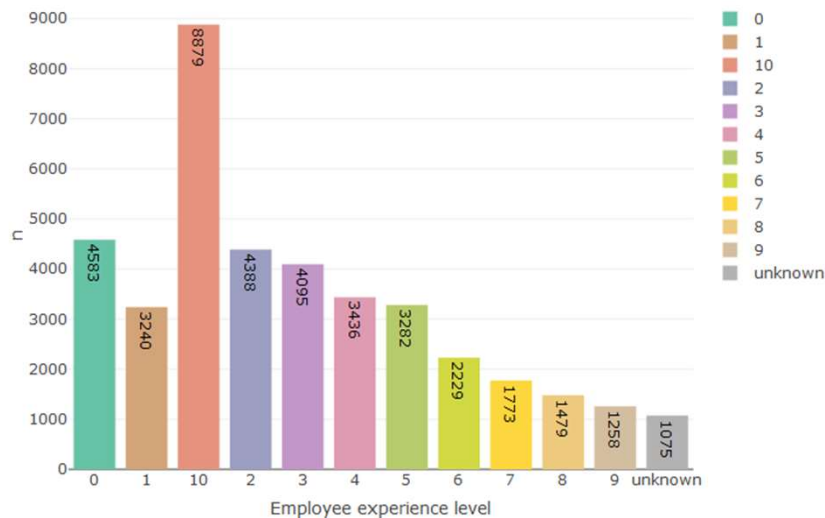
Default rate and grade of loans

Lets have a look at how the default changes with the different grades of loans. For grades B, C and D the defaulters are more. Loan grade A is comparatively safer than the other grades.



Employee work Experience

Customers with exp of more than 10 years and less than 1 year, tend to take more loans, and tend to default as well. The number of defaulters steadily decreases with increase in work exp.

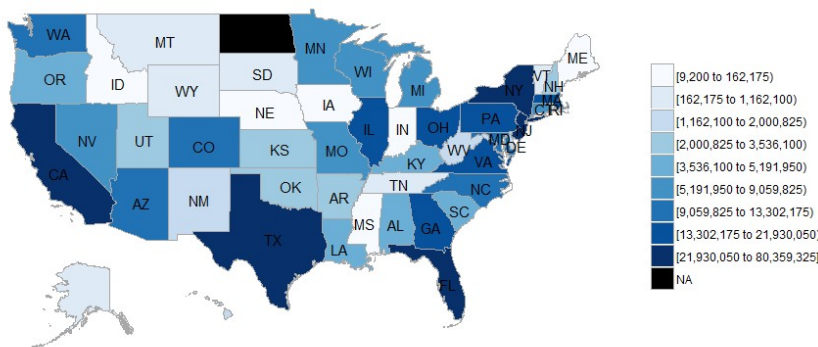


Influence of Geographical Location

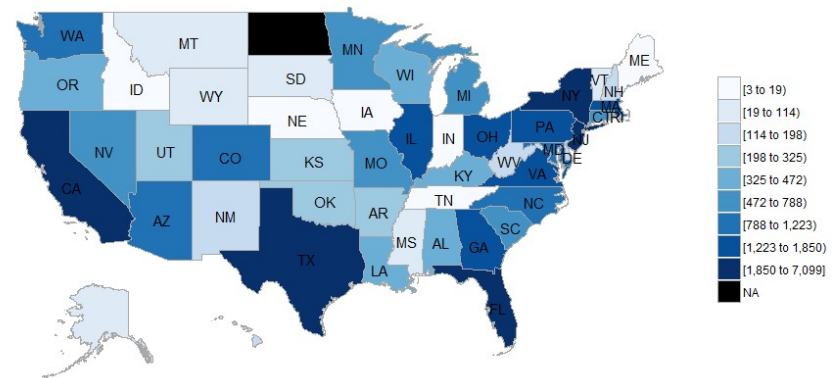
Loan Amount and volume varies from state to state

California, Texas, New York, Florida, and Illinois have the largest dollar amounts and volumes of loans.

Loan Amount by State

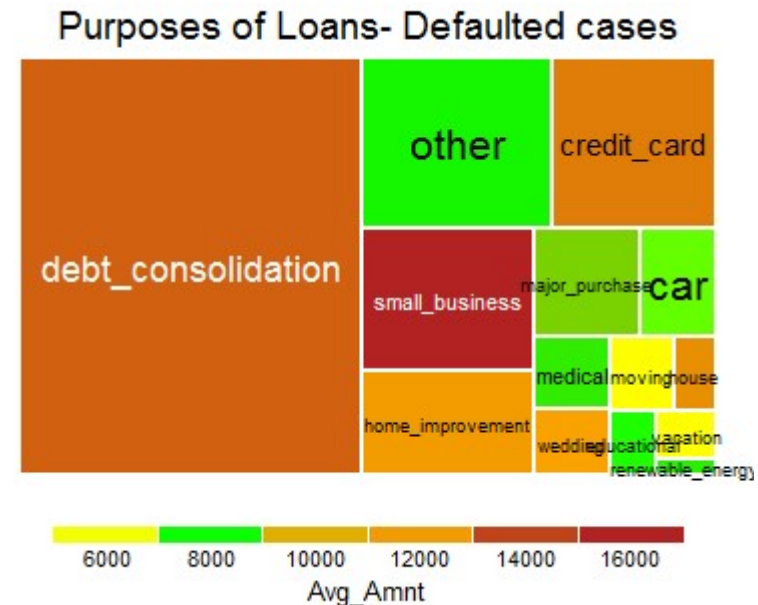
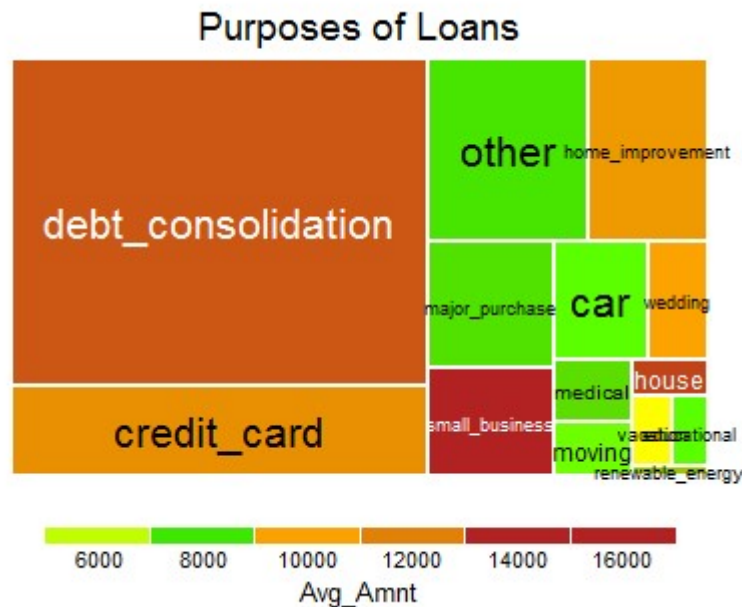


Loan Volume by State



Purpose of loans..

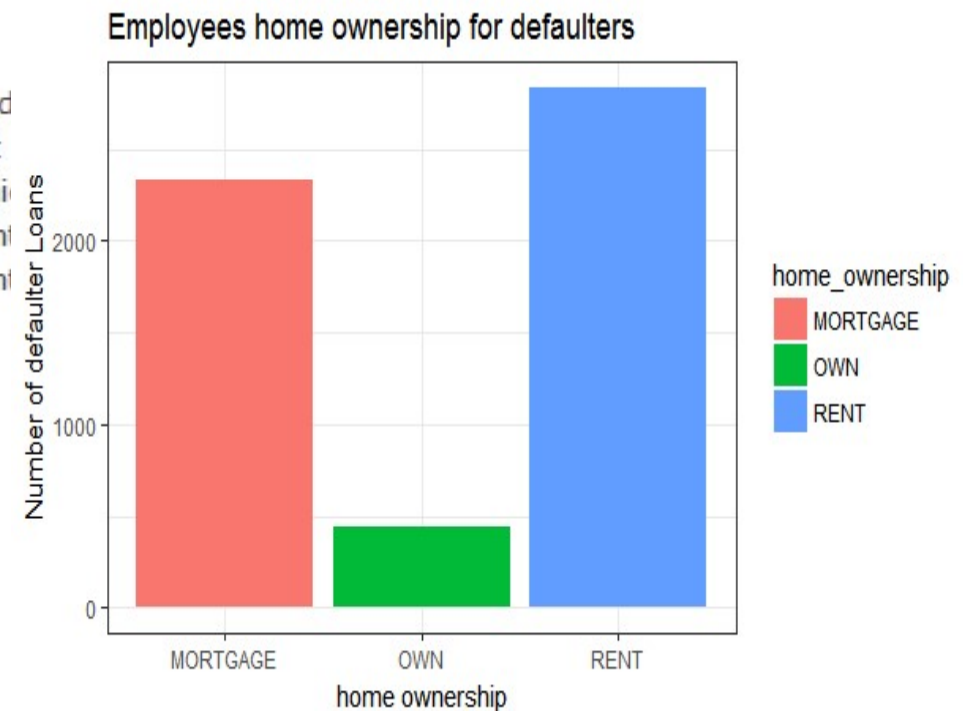
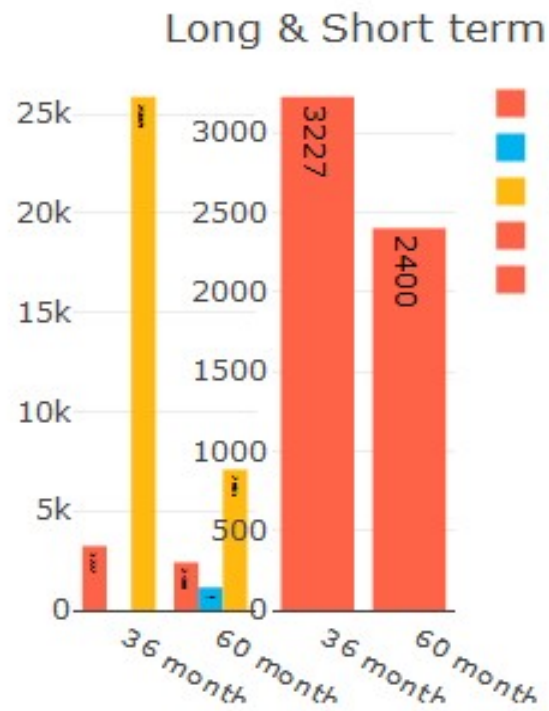
The purpose of loans have a major influence on the number of defaulters. debt consolidation and credit card are the most popular reasons for borrowing. Loans for debt consolidation, credit card, house, and small business usually have higher average amount than other purposes.



Influence of Short term and long term loans and home ownership

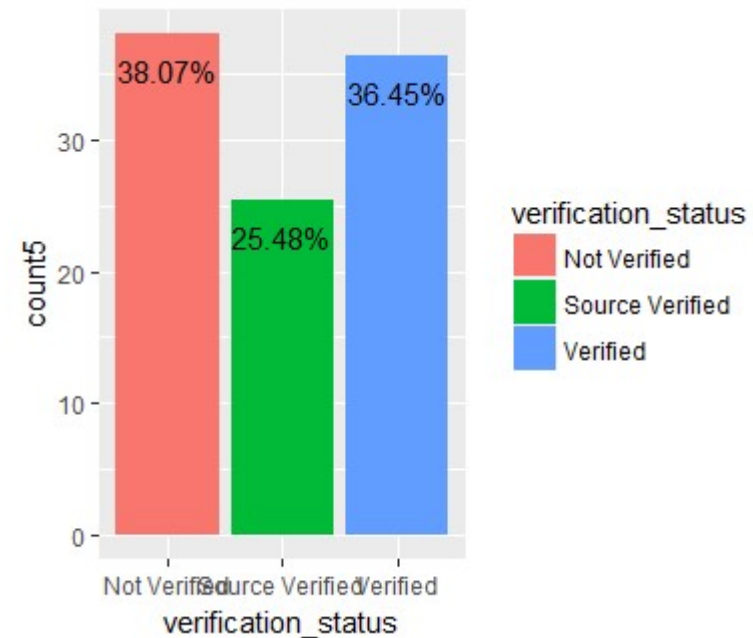
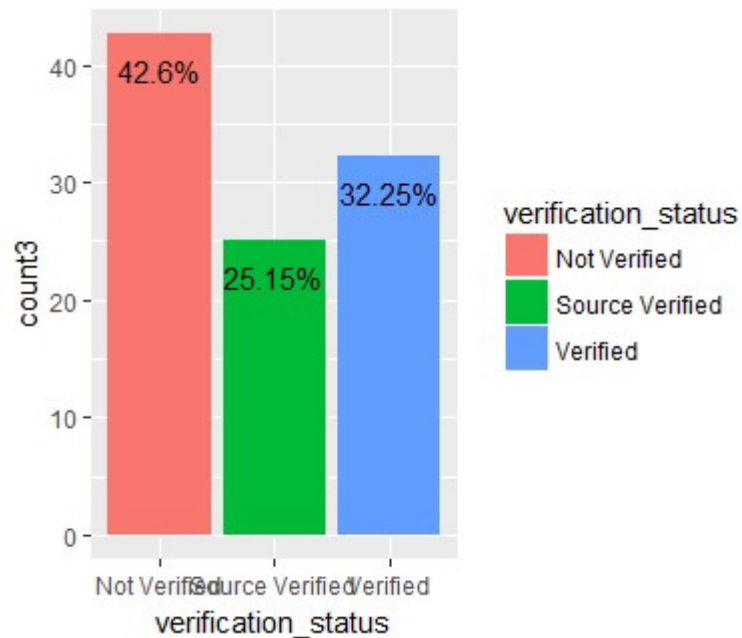
Short Term loans tend to default more than long term loans.

Customers with homes on rent and mortgage tend to default more.



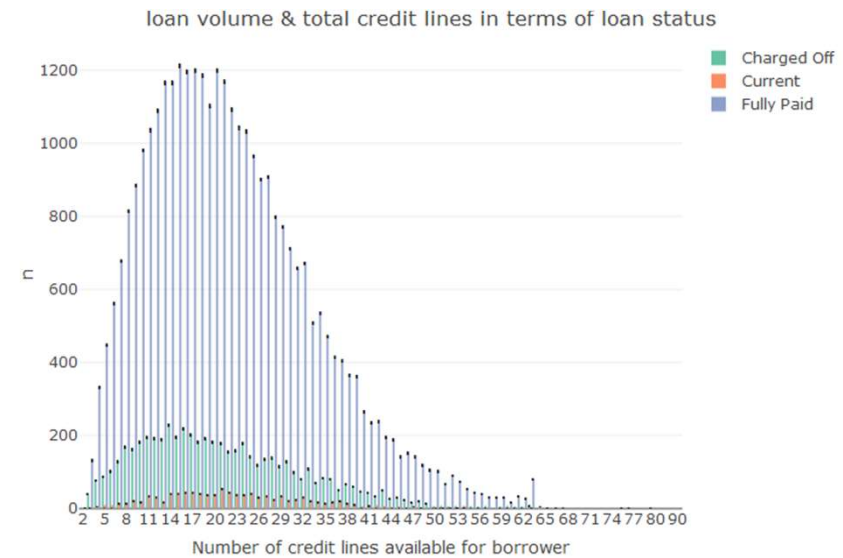
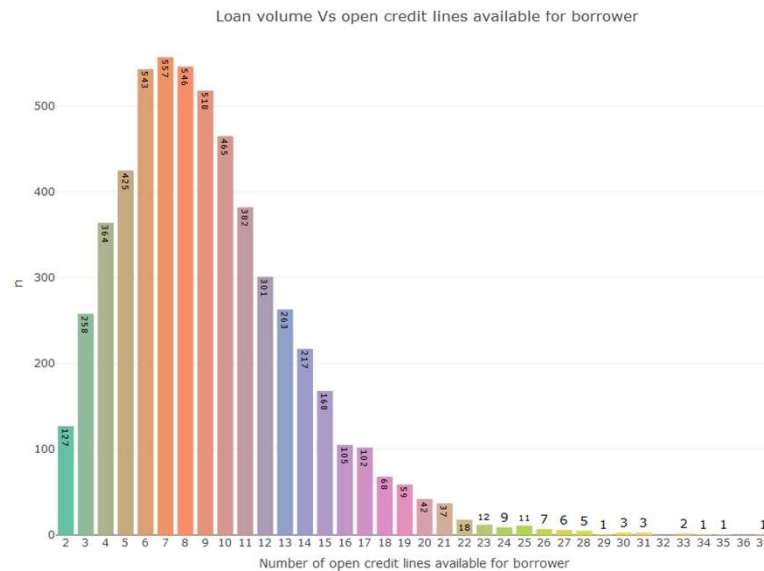
Verification status.. How risky the loan is?

42.6% of the loans provided are not verified, of which 38.07% of the unverified loans are likely to default. However, 36.45% of the verified loans have defaulted as well.



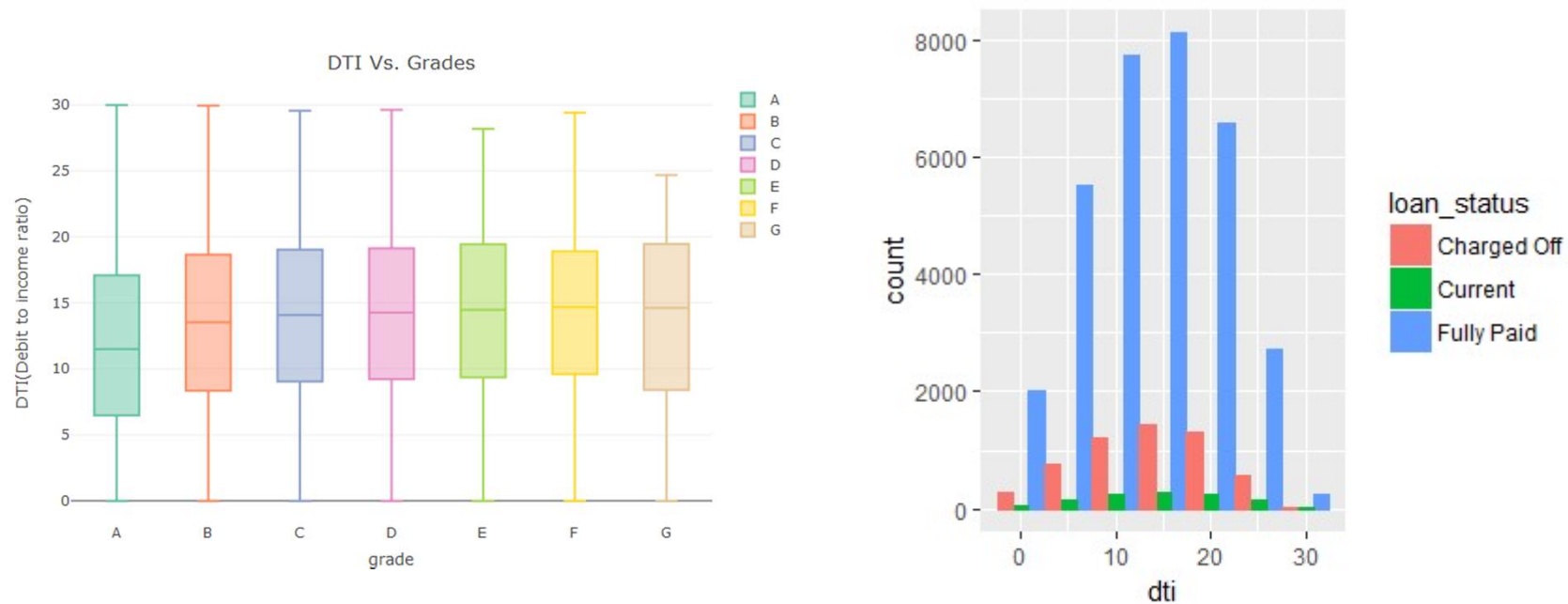
Credit line status

Loan volume vs the overall credit line available to the customers. Also, the loan status vs the total credit shows an interesting trend.



Influence of DTI on Loan Status

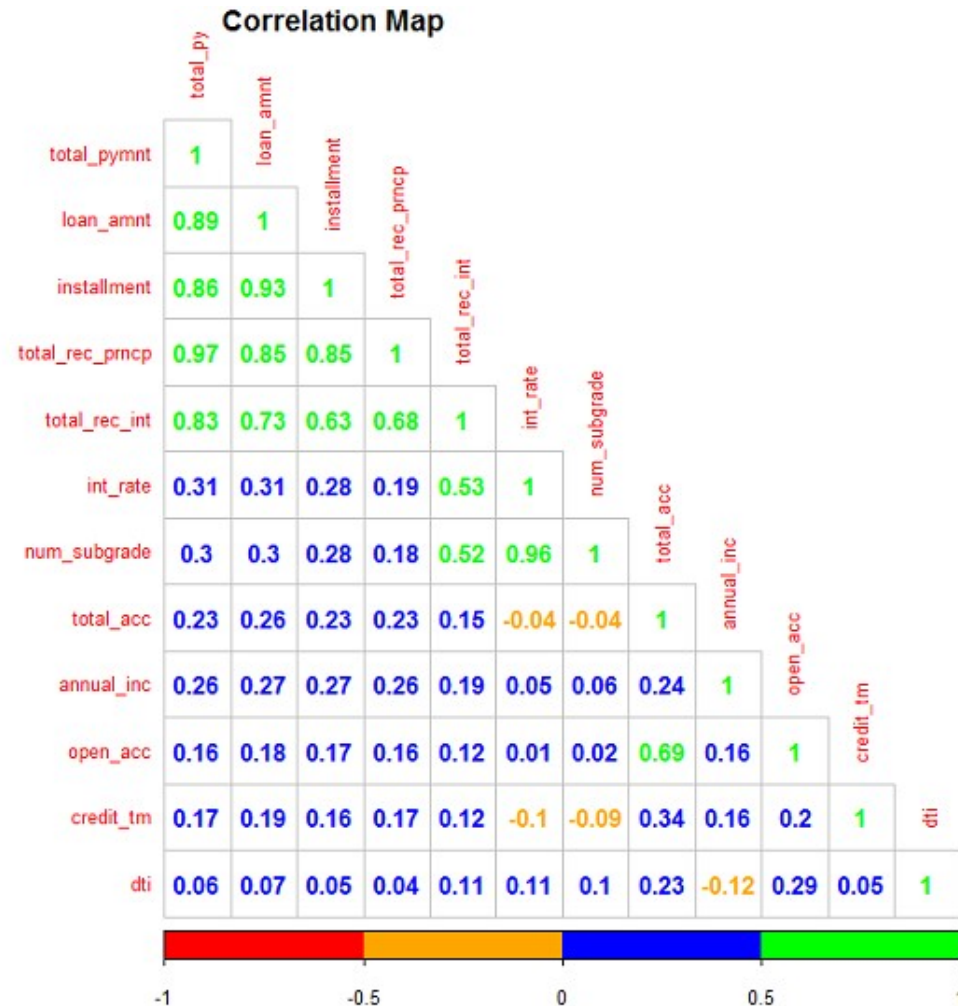
DTI doesn't seem to have much influence on the loan being default. For almost all grades of loan the pattern is similar. Also the loan being default doesn't seem to depend much on DTI value, against the conventional trend.



The Correlation Map

Here we have the correlation Map which shows us the dependencies of different variables.

- there is a high dependency between total principal received till date and Payments received to date for total amount funded till date.
- Among the variable that we have analyzed
 - high dependency can be seen between interest rate and sub grade of the applicant
 - Also, there is a negative dependency between dti and annual income of the loan applicant



Conclusion

- We have presented just a part of our analysis about some variables which are seemingly relevant, and have some visible pattern and influence on loan status.
- Full analysis is described in corresponding R file i.e effect of following variables on a loan customer ending up as a defaulter:
 - dti
 - earliest_cr_line
 - emp_length
 - annual_inc
 - grade
 - home_ownership
 - int_rate
 - issue_d
 - loan_amnt
 - loan_status
 - purpose
 - sub_grade
 - term
 - verification_status
 - addr_state
- There are other fields as well, which have significant influences on loan status, like recoveries and collection_recovery_fee but due to lack of available data, those fields are not included as part of this analysis.