

## Lending Club Case Study

Submitted by:

Saheli Das

Harikrushn Raiyani

# Business understanding



#### About Company

Lending Club is a peer to peer lending company based in the United States, in which investors provide funds for potential borrowers and investors earn a profit depending on the risk they take (the borrowers credit score).

#### • The Problem

Lending Club wants to understand the driving factors behind loan default, i.e. the driver variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.

### Problem solving methodology



Data Univariate **Bivariate Conclusion Data Cleaning Understanding Analysis** Analysis Drop the features which Analysis to find out if The dataset contains Analysis the patterns Insights and observations have more than 50% or

complete loan data for all loans issued through the 2007-2011, including the current loan status (Current, Charged-off, Fully Paid) and latest payment

information. A Data Dictionary has been provided to get knowledge of all the columns and their domain specific uses

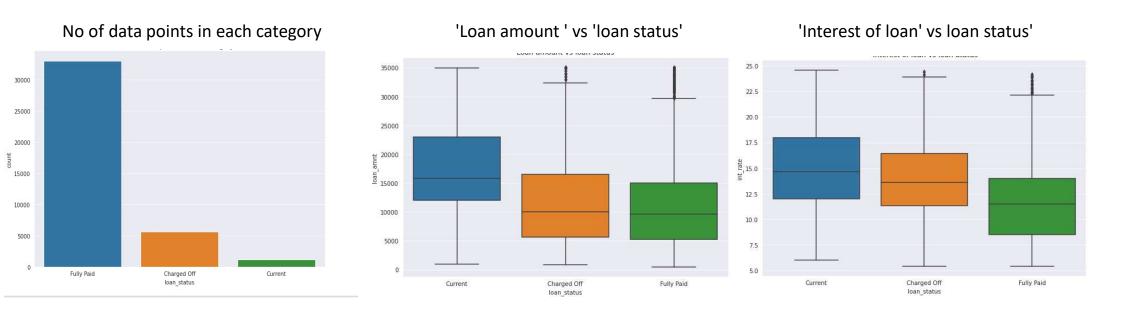
more Null values.

Remove the columns which contain single categorical value and also irrelevant to the analysis

found in univariate data include central tendency median. (mean. mode) and the distribution of the continuous data.

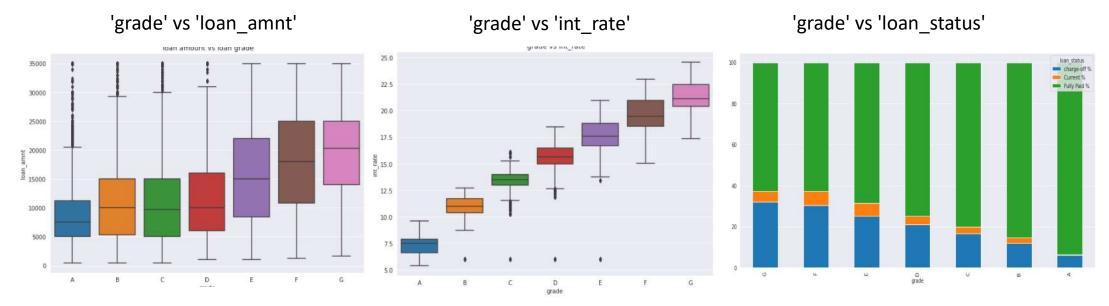
there is a relationship between two different variables.





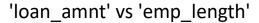
- Most of the loans which have been sanctioned are fully paid. only 14% loans are defaulters.
- The high loan amount can cause of high interest which will increase the chances of charged off loan.

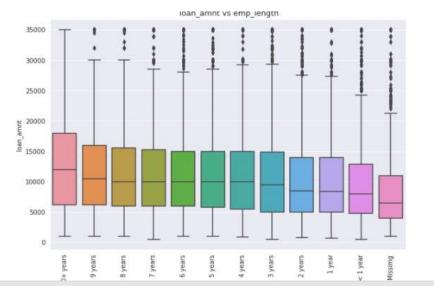




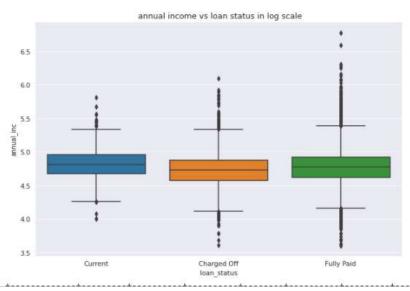
- Grade A loans are maximum fully recovered almost 93.63% and have very less chances to be charged off.
- The chances of fully recovered is decreased with the grade from A to G
- The loan amount of Grade G and F loans are comparatively higher than the others and also the interest rate is high.
- The recovery rate of Grade G and F loans are also less than others near about 62% and have very high chances to be charged off. So that we can conclude Grade G and F loans are very high risky loans





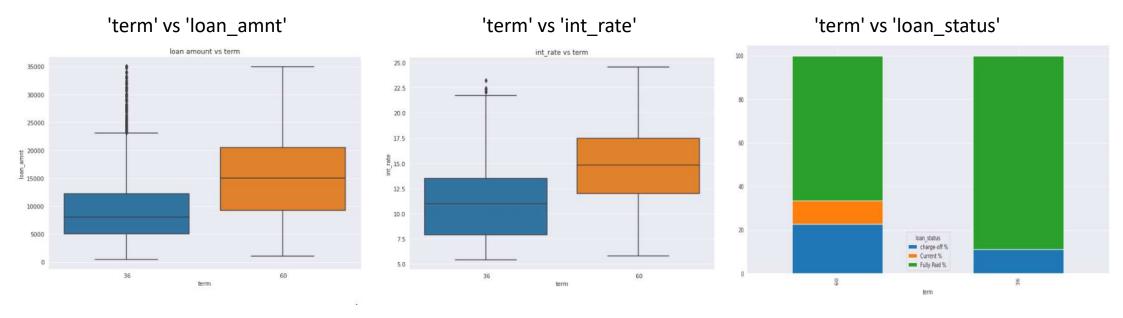


#### 'annual income' vs 'loan status' in log scal



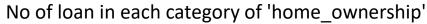
- The persons, who are having more than 10 years of working experience, are taking more loans and also the loan amount is high .It has been observed the percentage of charged off loans is high for the persons, who are having more than 10 years of working experience.
- From the above analysis we can conclude, in case of charged off loan borrower's avg annual income is little less which is not much significant.

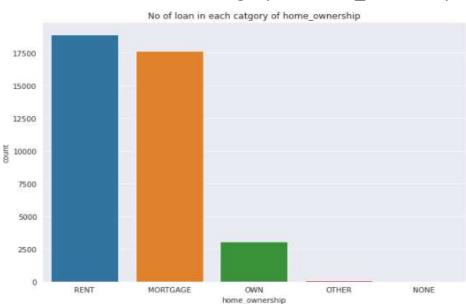




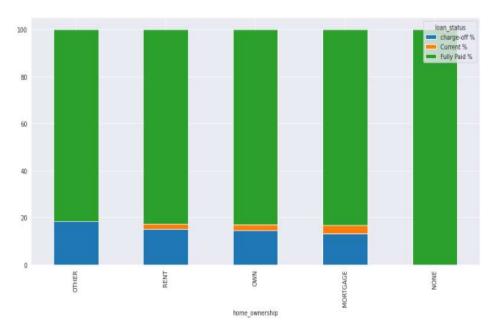
- The avg loan amount sanction for the term of 36 month is less than the avg loan amount sanction for the term of 60 month.
- If the loan amount is large and it has been taken for long duration then the installment paid by per month is also high. In such case there is a high risk to be defaulter





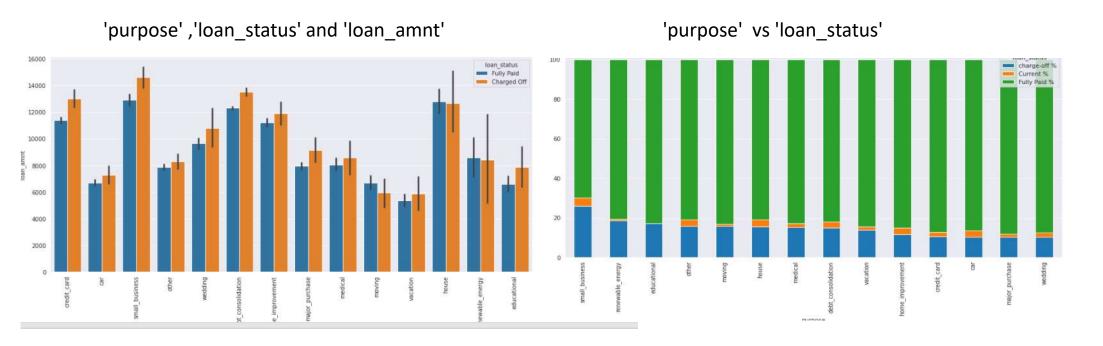


#### 'home\_ownership' vs 'loan\_status'



• The persons who are staying in rented house are taking maximum loan and they have a high possibility to become defaulter.

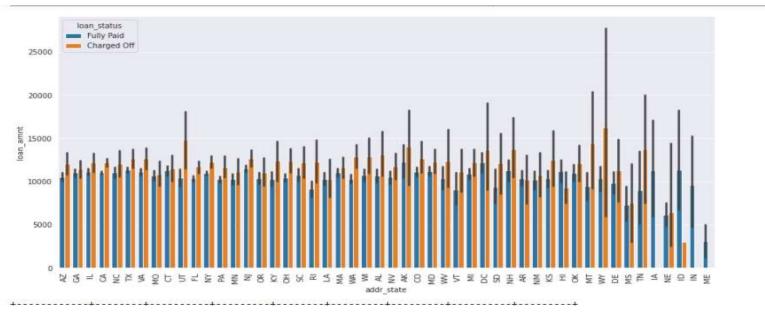




The loan of high amount are mostly taken by small business and the chances to be charged off is also high



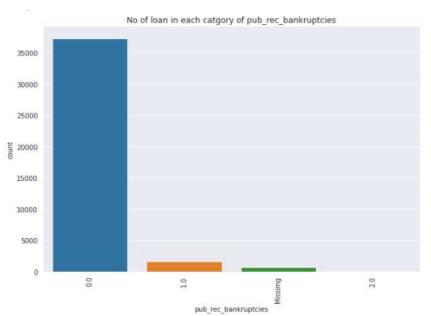




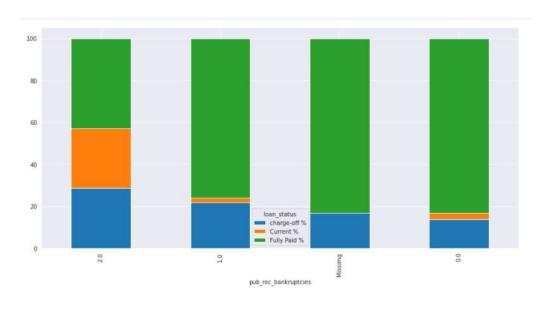
- •NE ,NV,AK are the state where the percentage of charged off is high.
- •States with higher Charge Off rates have very low number of loans.
- •The loans have been sanctioned for AK is the highest average loan amount and also the chances to be charged off is high.







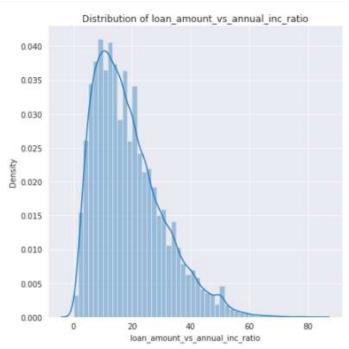
'pub\_rec\_bankruptcies' vs 'loan\_status'



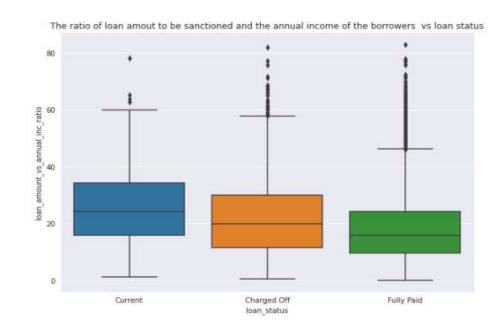
- •Most of the cases number of number of public record bankruptcies is 0.
- •From the above analysis we can conclude ,if the number of public record bankruptcies is high, the possibility to be defaulter is high



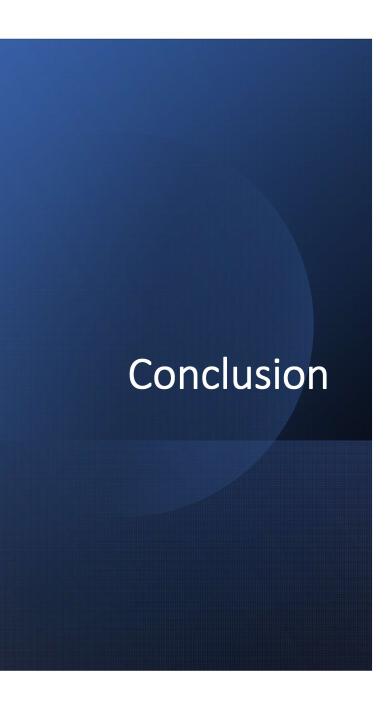
Distribution of loan\_amount\_vs\_annual\_inc\_ratio



The ratio of loan amount to be sanctioned and the annual income of the borrowers vs loan status



• If the loan amount to be sanctioned is more than 20% of the annual income of the borrowers ,then there will be a chance to be defaulters





- The high loan amount can cause of high interest which will increase the chances of charged off loan.
- The loan amount sanctioned for the low graded loans (G and F) are higher than the others with highest interest and they have a very high chance to be charged off.
- The persons, who are having more than 10 years of working experience, are taking more loans and also the loan amount is high.
- If the loan amount is large and it has been taken for long duration then the installment paid by per month is also high. In such case there is a high risk to be defaulter.
- The persons who are staying in rented house are taking maximum loan and they have a high possibility to become defaulter.
- The loan of high amount are mostly taken by small business and the chances to be charged off is also high
- The highest average loan amount that was charged off have been sanctioned for the state WY.NE ,NV,AK are the state where the percentage of charged off is high. States with higher Charge Off rates have very low number of loans.
- The number of public record bankruptcies is high, the possibility to be defaulter is also high.
- If the loan amount to be sanctioned is more than 20% of the annual income of the borrowers ,then there will be a chance to be defaulters