



# Lending Club Case Study

Group Members: Manisha Madane Prateek Meher





### **Abstract**

- Lending Club is the world's largest online loan marketplace, offering personal loans, business loans, and medical procedure financing.
- Through a quick online interface, borrowers can easily access lower interest rate loans.
- The goal of the analysis is to use information about previous loan applicants to determine whether or not they 'defaulted.'





### Problem solving methodology

Data Data Analysis Analysis Analysis

#### Data Cleanin

g

Removing the null valued columns, unnecessary variables and checking the null value percentage and removing the respective rows.

#### Data Understandin

g

Working with the Data Dictionary and getting knowledge of all the columns and their domain specific uses

#### Univariat

е

#### Analysis

Analysing each column, plotting the distributions of each column.

#### Segmente

d

#### Univariate Analysis

Analysing the continuous data columns with respect to the categorical column

#### **Bivariat**

E

#### **Analysi**

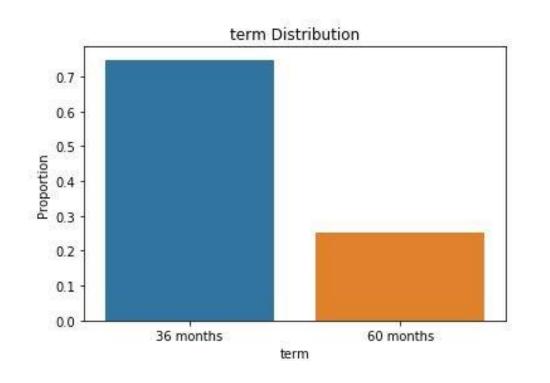
Analysing the two variable behaviour like term and loan status with respect to loan amount.

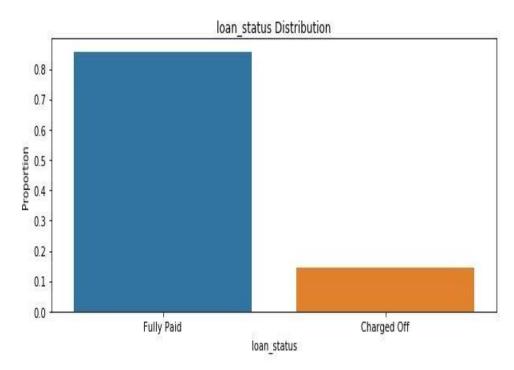
### Recommendation

Analysing all plots and recommendations for reducing the loss of business by detecting columns best which contribute to loan defaulters.





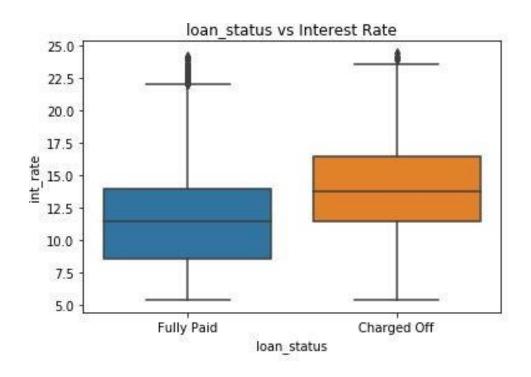




- There are only two loan terms 36 and 60 months. Around 75% borrowers took loans with 36 months term.
- The charged off borrowers are around 15% and fully paid is around 85% in the given data set.



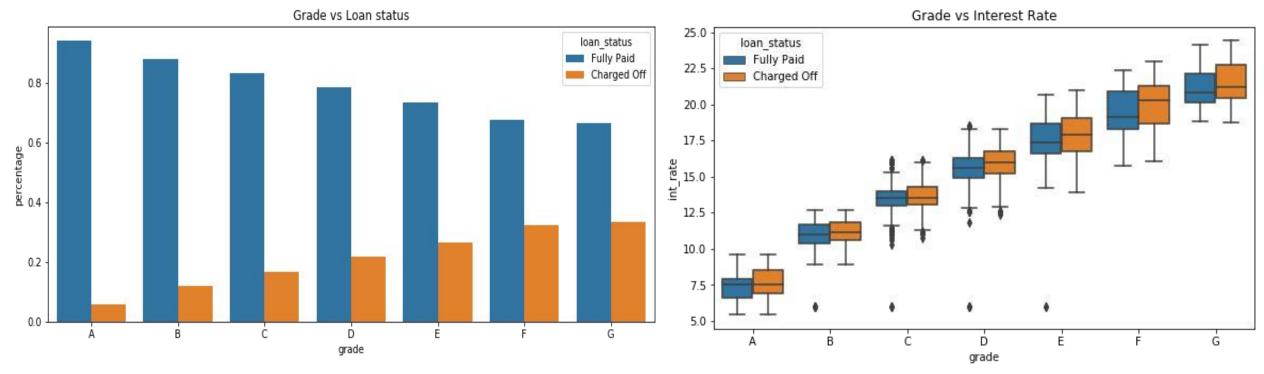




When the loan interest rate is high there is high chance of loan getting defaulted.



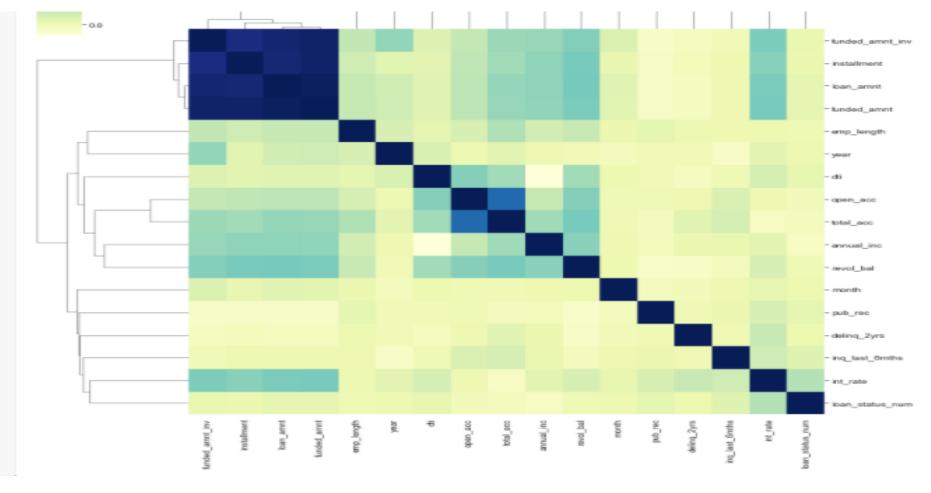




- Grades are very good category to tell the borrower probability of defaulting the loan.
- The Lower grades(E,F,G) have higher chances of defaulting the loan than Higher ones(A,B)
- Also the Lower grades are getting loans for higher interest rates which might be the cause for loan default.







we know darker the value higher the correlation, we can clearly see loan\_amnt, funded\_amnt, funded\_amnt\_inv and instalment have huge correlation These fields are proportional to each other. Next, the public records related fields pub\_rec & pub\_rec\_bankrupcies and number of accounts related fields open\_acc & total\_acc are correlated.



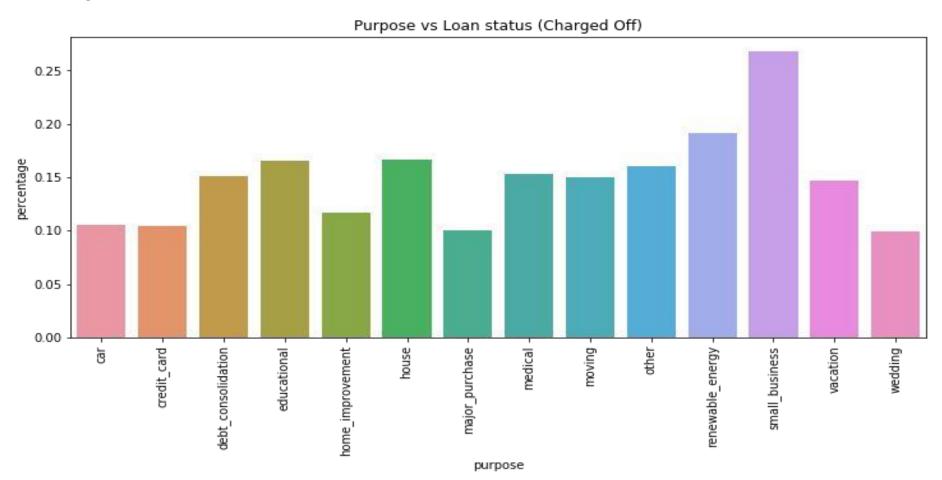




The loans which are given for CA, FL and TX state borrower's are defaulted more than other states.



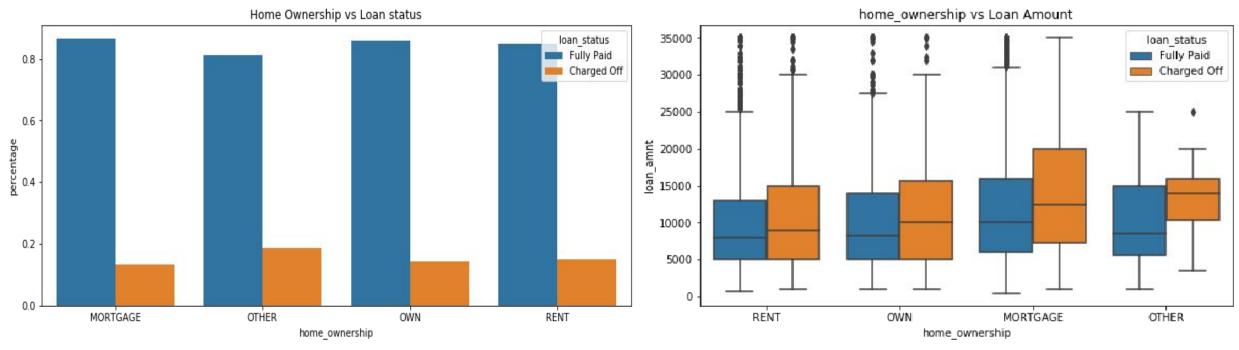




Borrower's who took loans for small business purpose have defaulted more.







- There is around 20% chance of loan default in each home ownership category.
- From the 2nd plot we can see the people with higher loan amounts in **mortgage** home ownership has high default rate than others.





### **Conclusions**

- Lending club should reduce the high interest loans for 60 months tenure, they are prone to loan default.
- Grades are good metric for detecting defaulters. Lending club should examine more information from borrowers before issuing loans to Low grade (G to A).
- Lending Club should control their number of loan issues to borrowers who are from CA, FL and NY to make profits.
- Small business loans are defaulted more. Lending club should stop/reduce issuing the loans to them.
- Borrowers with mortgage home ownership are taking higher loans and defaulting the approved loans. Lending club should stop giving loans to this category when loan amount requested is more than 12000.
- People with more number of public derogatory records are having more chance of filing a bankruptcy. Lending club should make sure there are no public derogatory records for borrower.