
Lending Club Case Study

Assignment for EDA IITB, Executive AI and ML

Contributor

— Ashish Bhavsar —

Brief understanding

Lending club case study is to give us feel for real life problem, and how to use this problem to get overview of data. By doing various data manipulation on the given data set which include Data evaluation, Data Cleanup, Remove null values, Data Imputation.

Also it help to understand,

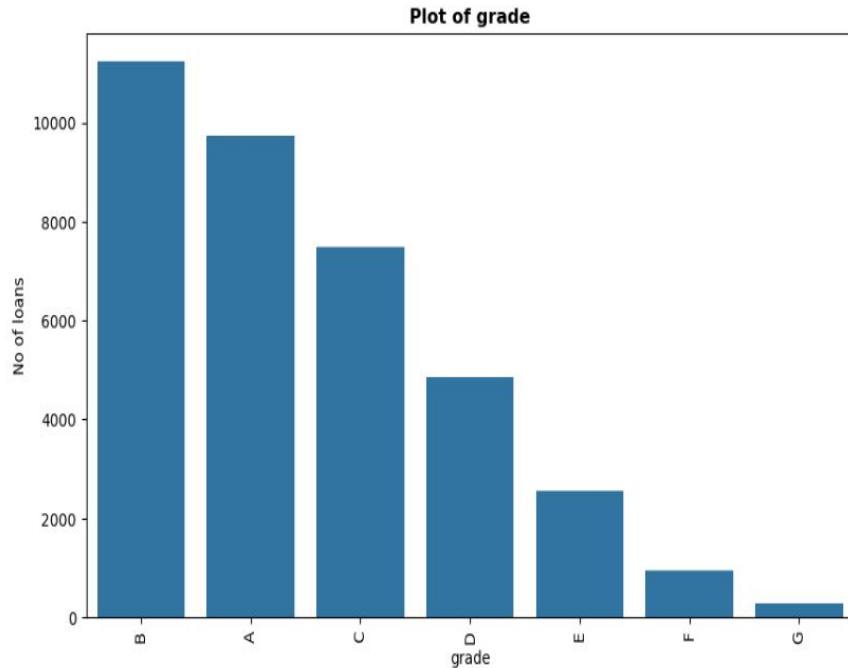
- 1.How can we use data set to get visualization about problem which will then help for further action to be taken.
- 2.Use of EDA on real time business problem
3. Business about lending money in banking and finance sector.
4. How to use data visualization to solve business critical decision as best clients to lend loans

EDA overview

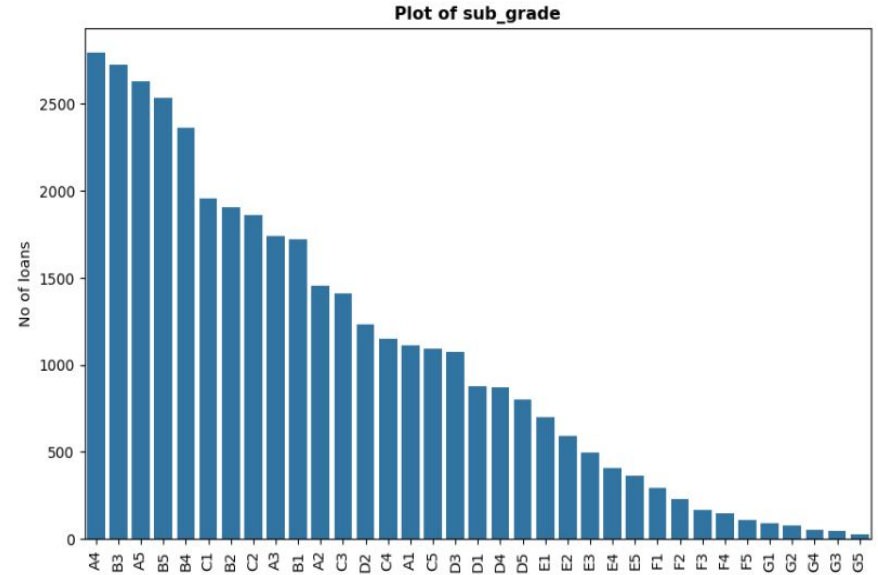
As explain in diagram below, step by step EDA



Grade and No of Loans

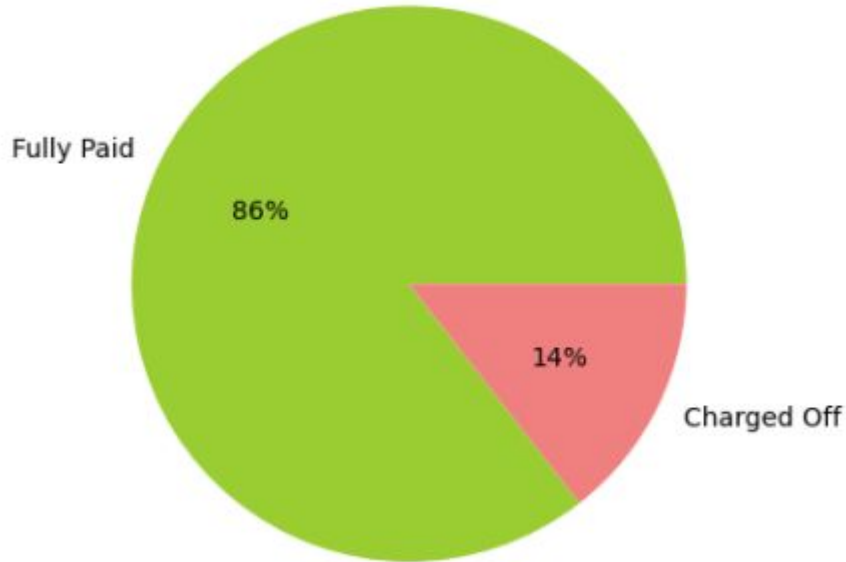


Plot of Grade shows Grade A and B are given more loans compared to other grades



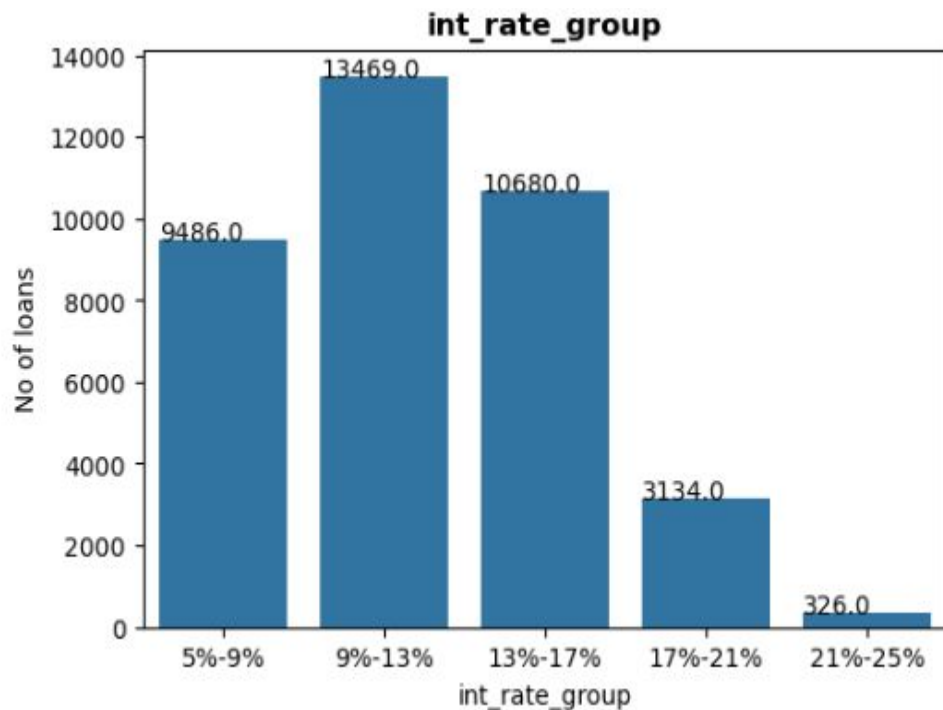
Plot of Sub Grade shows Grade A4, B3, A5, B5, B4 are given more loans compared to other grades

Loan status



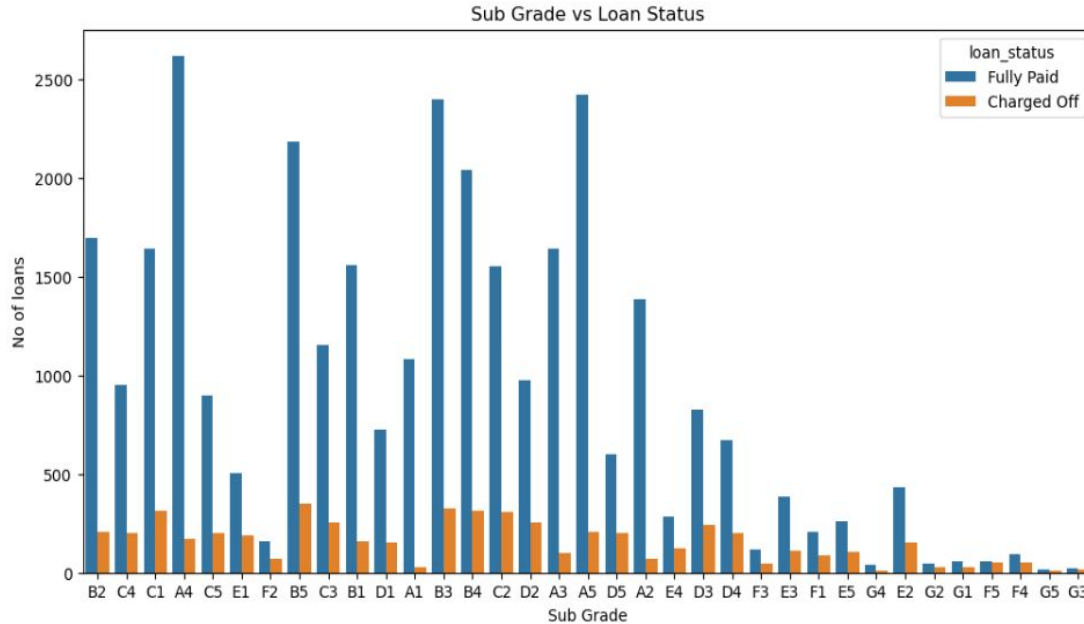
This shows that overall loans which was given till the time, 14% of them are charged off, which is nothing but loss making for business

Interest rate group and No of Loans



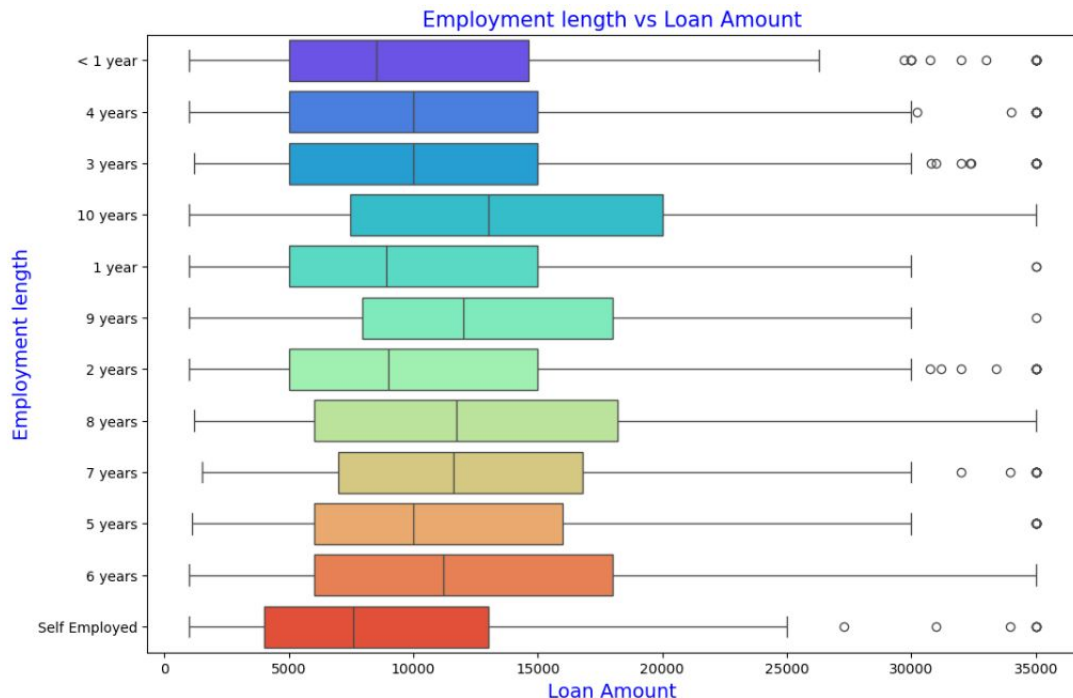
1. Interest rate range 9 to 13 is the range where maximum loans have been issued
2. 21 - 25% is the range where minimum loans have been issued

Sub grade and Loan status



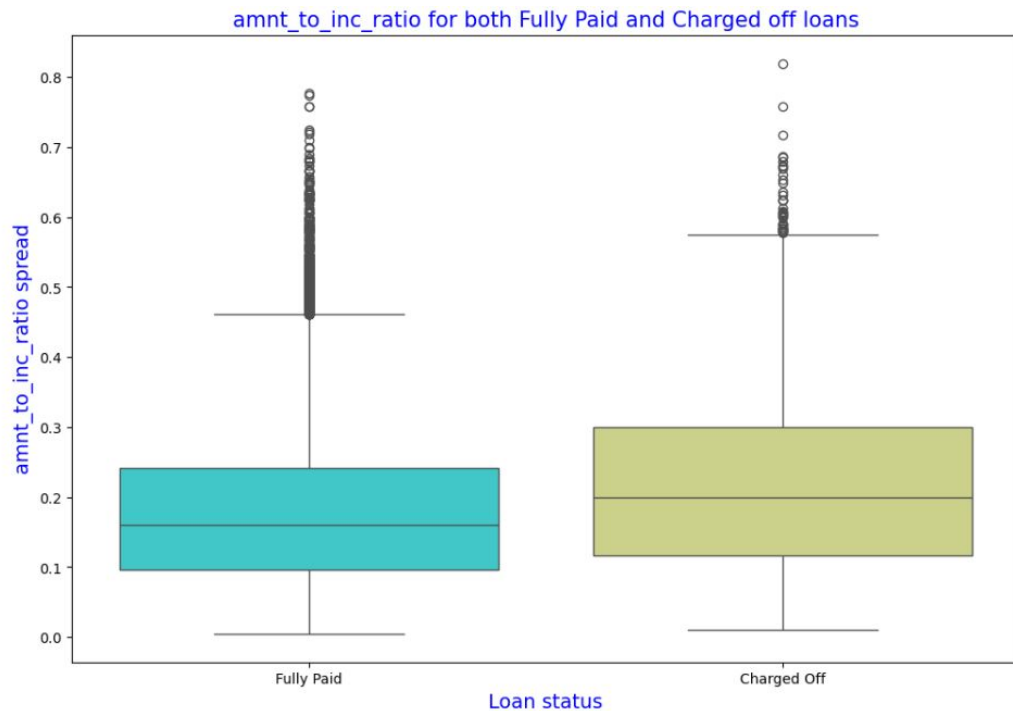
1. Based on the counts, Grade B3,B4,B5, C1,C2, D3, D4 top sub grades in Charged Off
2. Grade A4, B5, B3 and A5 are the top subgrade who has fully paid of loans. And those are the one who is having maximum number of loans

Employment length and Loan Amount



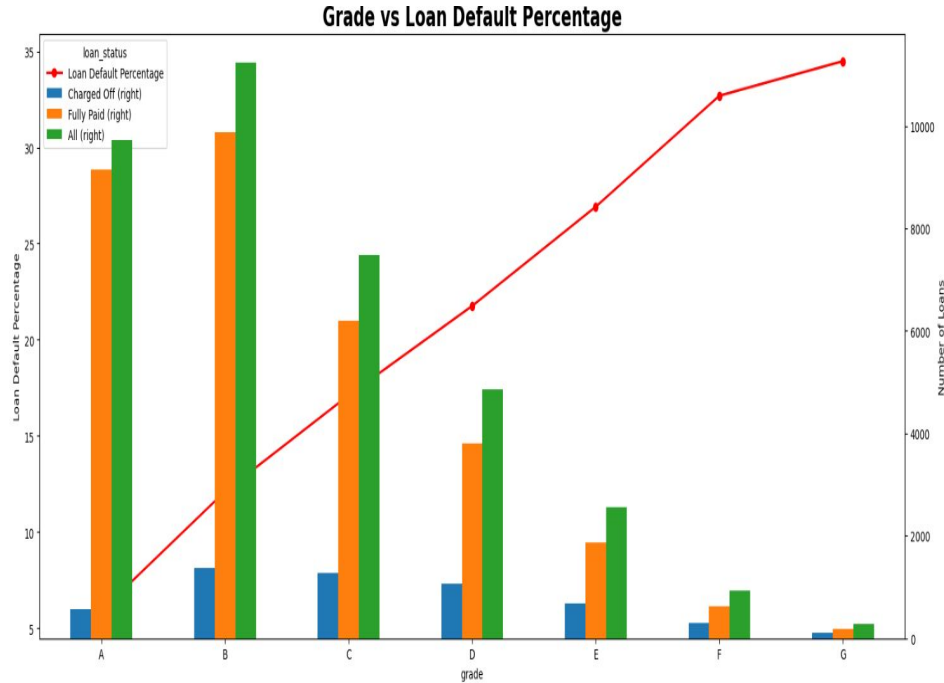
1. Employment length of 10 years got more loans and they are the maximum defaulters
2. Followed by year 6, 7, 8 and 9 where the medians are almost close

Loan status with Amount to interest ratio



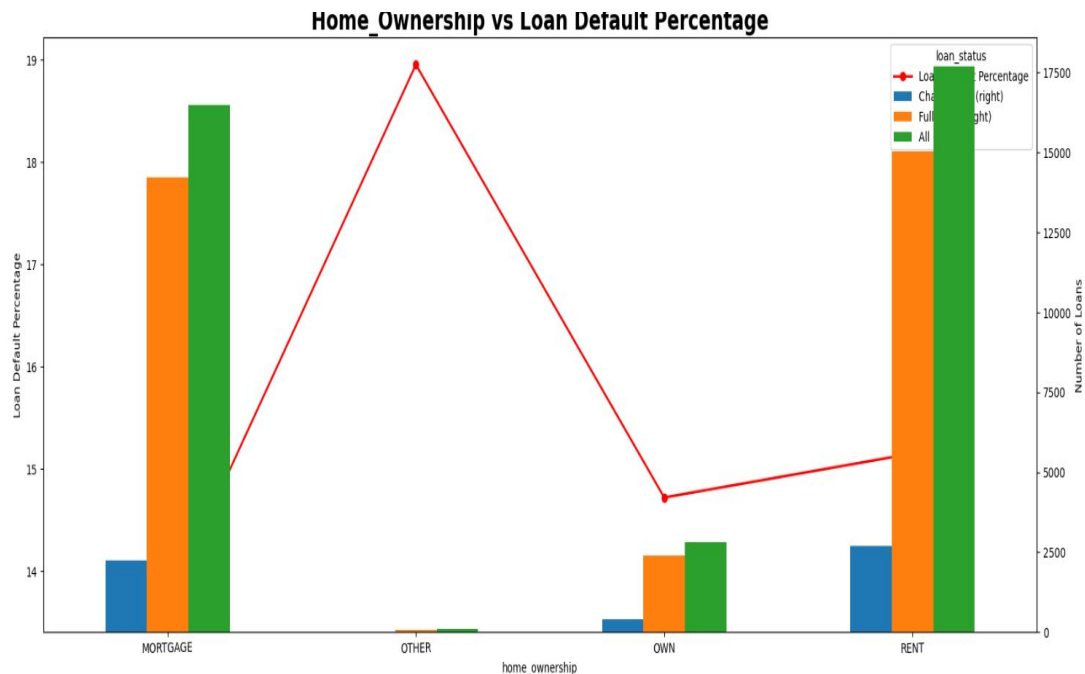
1. Amount to interest ratio is more in case of bad loans which are nothing but charged off loans
2. Overall business impact is more due to charged off loans due to no recovery.

Grade and Loan default percentage



Tendency to default the loan is increasing from Grade A to Grade G

Home ownership and Loan default percentage



Borrowers from Other category have highest tendency to default the loan.

Summary

1. Based on the counts, Grade B, C and D are top three in Charged Off
2. Based on the counts, Grade B3,B4,B5, C1,C2, D3, D4 top sub grades in Charged Off
3. 36 month loan default is more compared to 60 month
4. Maximum loans are issued for category 10 years and they are the maximum defaulters
5. Plot of loan issue year shows maximum loans were taken in the year 2011
6. Plot of the loan issue month shows maximum loans were given in the month of Oct, Nov, Dec.
7. Borrowers from states CA, FL, NY and NJ have shown maximum tendency to default.
8. People who are in Rent or Mortgage have failed to pay the loan compared with people in Own house
9. State AK, MT, KY, UT, TN are marginally above the 10k level and at 13k
10. The amount of Verified loans which are Charged Off is more compared to Not Verified
11. Employment length of 10 years got more loans and they are the maximum defaulters
12. Followed by year 6, 7, 8 and 9 where the medians are almost close looking at the box plot