LENDING CLUB CASE STUDY

ML/AI/C69

NOVEMBER 20, 2024

PRESENTED BY
Shweta Sandeep Vaidya & Rajat Kashyap

Contents

Background

Lending club is the largest pear to pear market place connecting borrowers with the lenders. Borrowers apply through and online platform and where they are assigned and internal score. Lenders decide 1) Whether to lend and 2) The terms of loans such as interest rate, monthly installment tenure, etc.

Some popular products are debt consolidation loan, house loan, car loan, etc.

Business Objective

To identify variable which are strong indicator of default and potential used the insights in approval/ rejection decision making.

Technology Used

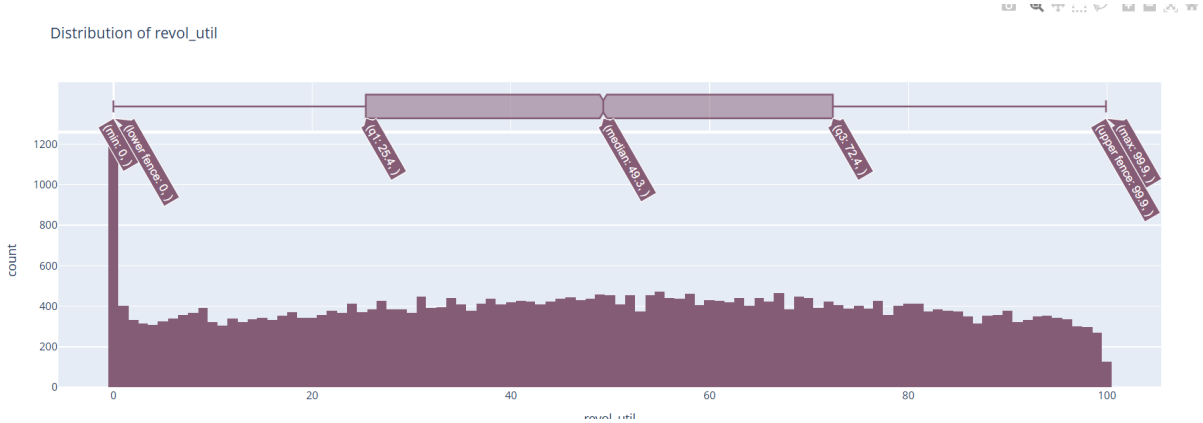
Python, Jupyter

Library used

Pandas, Numpy, Matplotlib, Seaborn, plotly.express

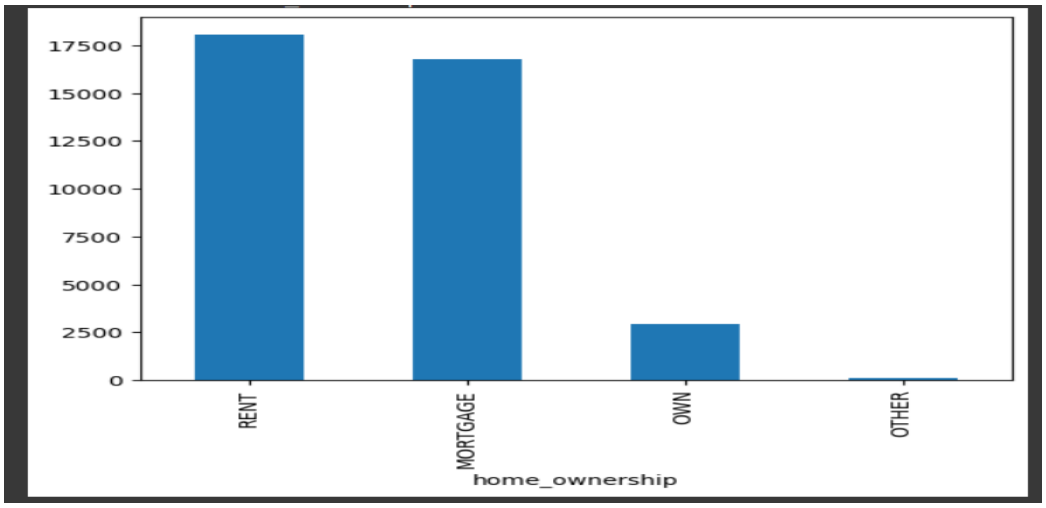
Conclusion

Univariant Analysis

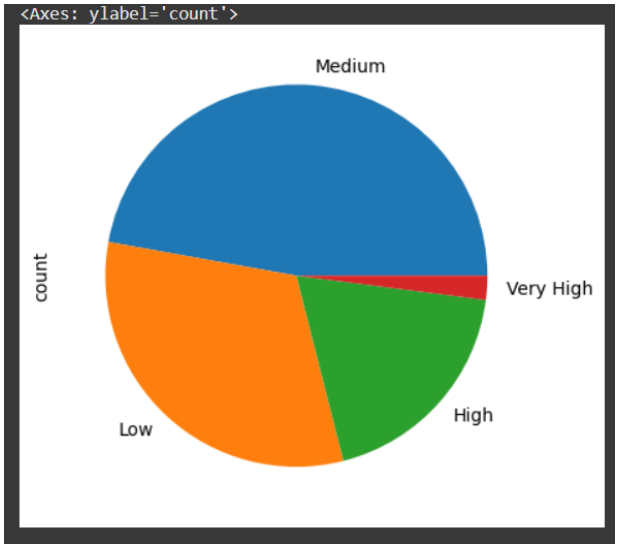


revol_util :- Revolving line utilization rate, or the amount of credit the borrower is using relative to all available revolving credit.

In this visualization we get know about the Lower Limit, Upper Limit, q1, q2, q3, and we can also see the distribution the data points.



home_ownership : We can see majority of the loan taken by people who are on "RENT" and "MORTGAGE"



int_rate_bin : are distrubuted uneven. There are "LOW" int rate also available, which we can consider in bivariant analysis

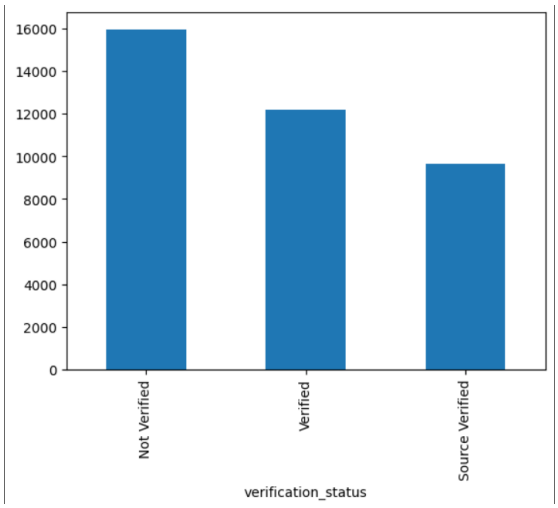
Very Low = interest rate (<=5)

Low = interest rate (5-10)

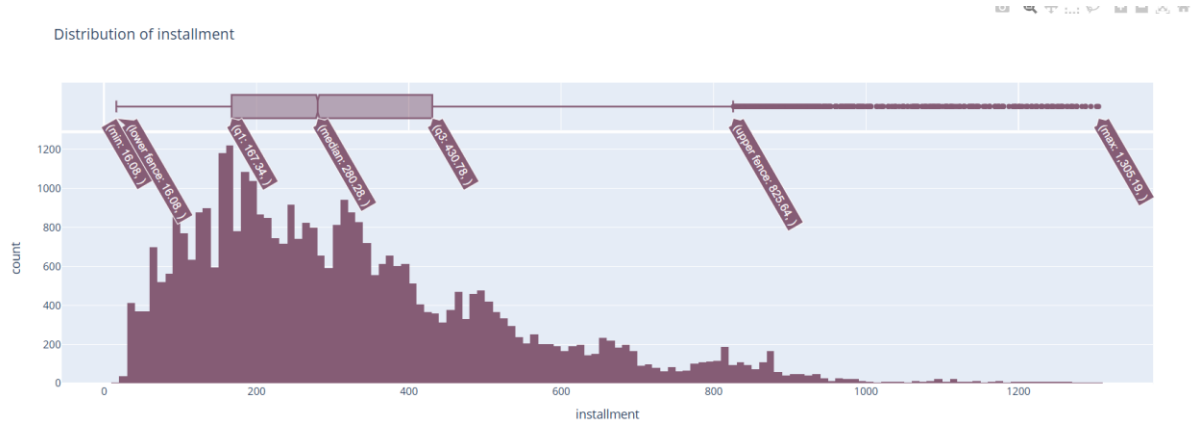
Medium = interest rate (10-15)

High = interest rate (15-20)

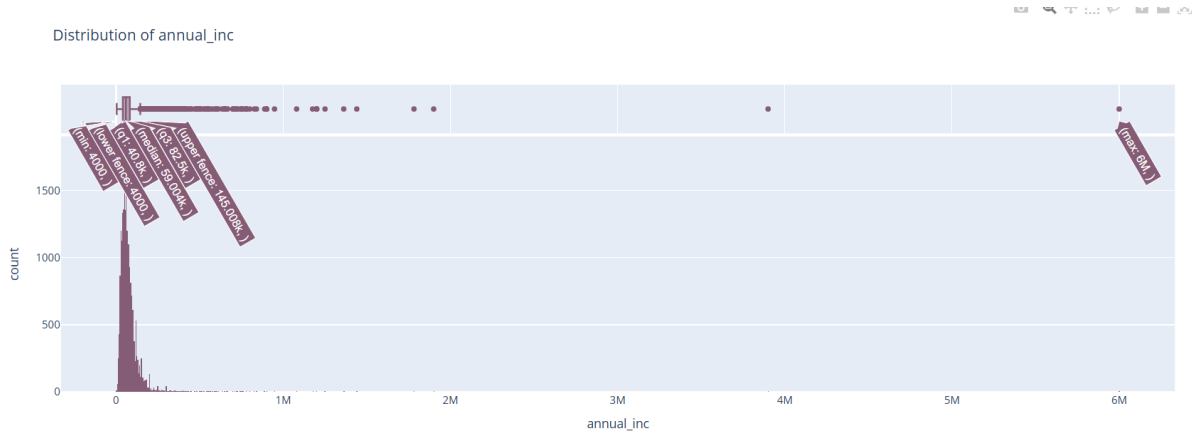
Very High = interest rate (>20)



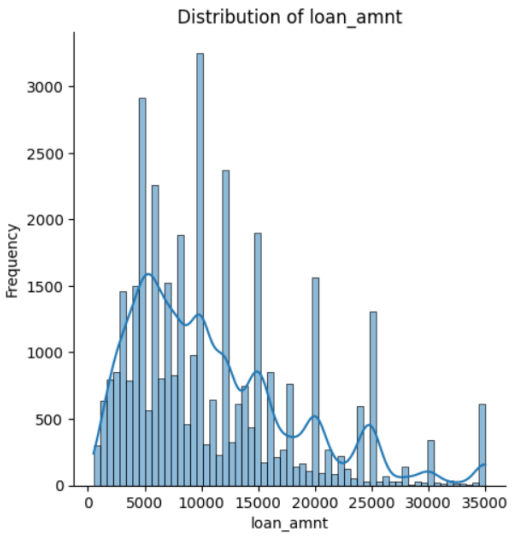
Here, we can see that the not verified status applicants are more than verified and source verified. So, the chances of fraudulent is more.



This is the Distribution of Installment in histogram. Here we can see outliers and the skewness this particular column.

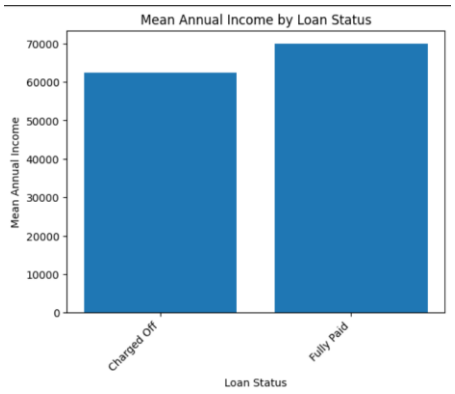


This is the Distribution of annual_inc in histogram plot. Here we can see the outliers and skewness this data.

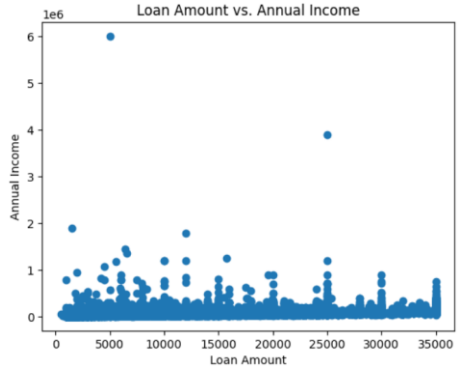


Frequency of loan_amnt. This shows that the 70% of the data comes under 0 to 15000 loan amount.

Bivariant Analysis



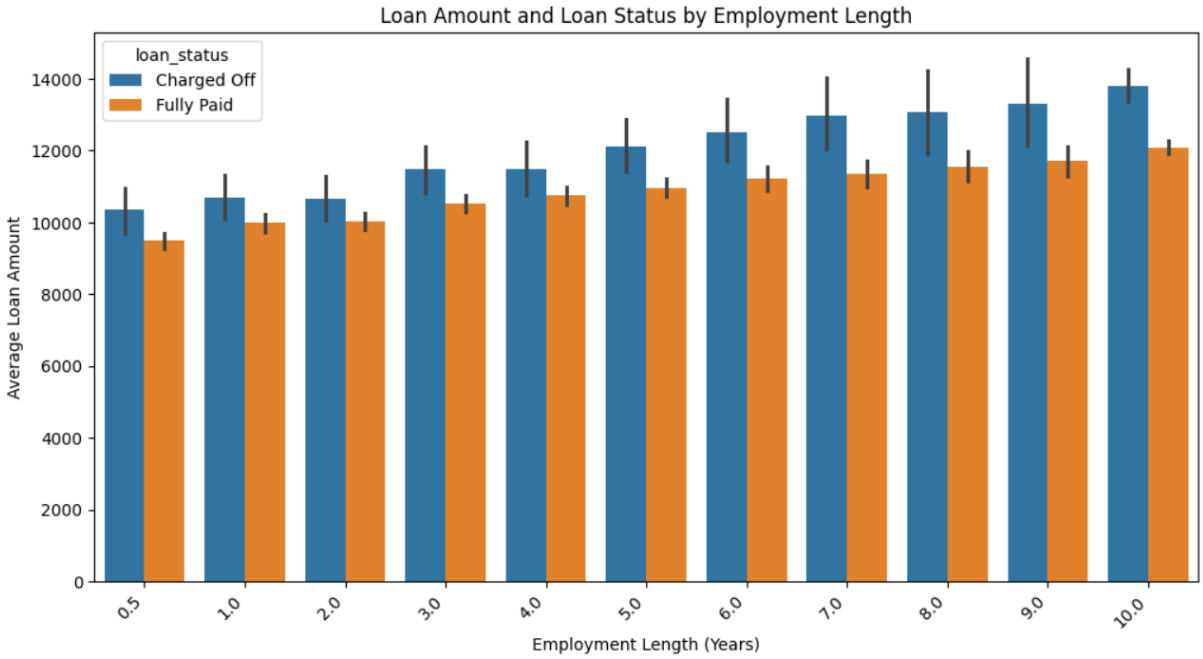
This is the Mean annual income vs loan status bar plot. This plot helps us to know that, applicants whos salary is more than than 60000 comes under fully paid rest comes under Charged Off.



Loan Amount vs Annual Income – This shows that maximum numbers of applicants who's salary range is comes under 0 to 0.3 has the highest loan amount.

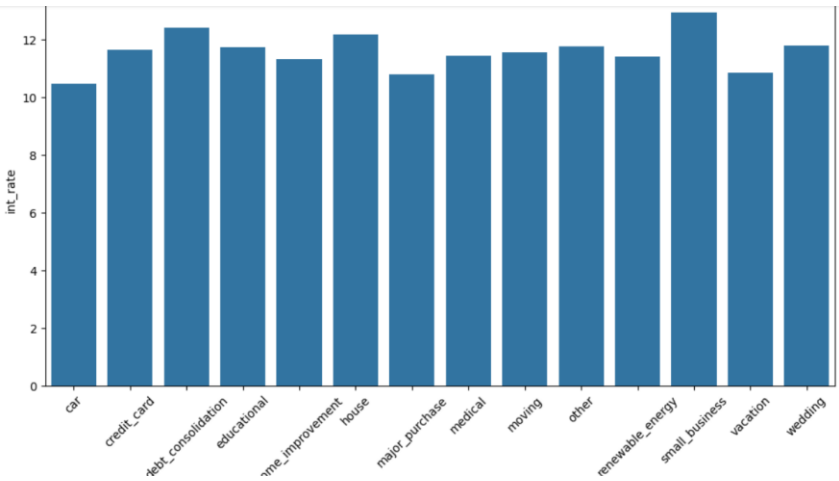
grade	term	loan_status	
A	36	Charged Off	7732.142857
		Fully Paid	8647.700860
	60	Charged Off	9694.594595
		Fully Paid	10202.513966
B	36	Charged Off	9221.102863
		Fully Paid	10073.986652
	60	Charged Off	14507.876712
		Fully Paid	14723.936450
C	36	Charged Off	8852.278481
		Fully Paid	9278.348402
	60	Charged Off	14570.732932
		Fully Paid	15389.048673
D	36	Charged Off	10004.748603
		Fully Paid	11022.979013
	60	Charged Off	14575.560748
		Fully Paid	14318.142097
E	36	Charged Off	11997.685185
		Fully Paid	12980.997001
	60	Charged Off	17032.209738
		Fully Paid	17063.117038
F	36	Charged Off	16808.163265
		Fully Paid	13476.027397
	60	Charged Off	19926.431298
		Fully Paid	18701.195219
G	36	Charged Off	15566.666667
		Fully Paid	18074.242424
	60	Charged Off	19440.064103
		Fully Paid	21697.852761

Here we can see that, Grade A has the lowest loan amount whereas Grade G has the highest loan amount in both the terms.

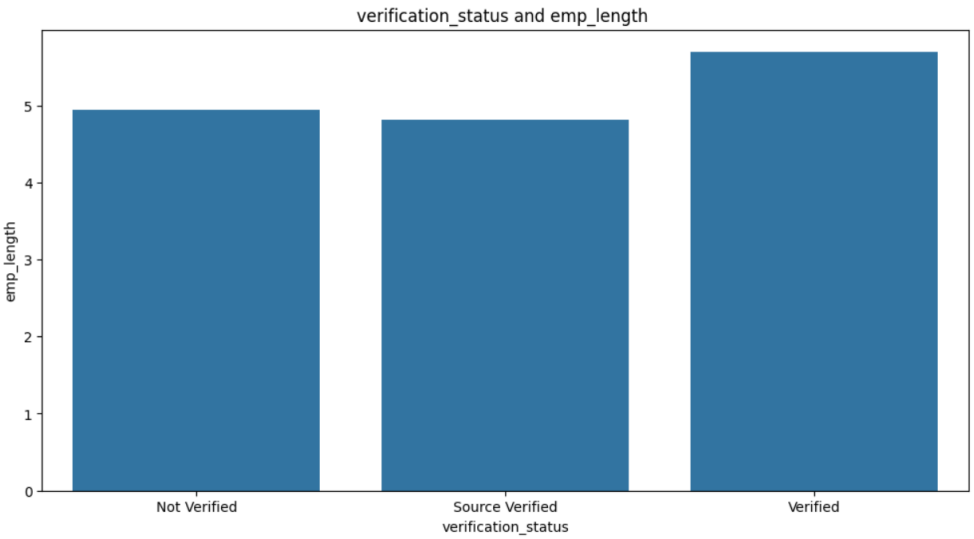


Bar plot of loan amount and loan status by Employment Length. It shows that Employee with more experience has the highest loan amount and with more charged off whereas lesser experience has the less loan amount.

	purpose	int_rate
0	car	10.467072
1	credit_card	11.642253
2	debt_consolidation	12.413020
3	educational	11.749243
4	home_improvement	11.311841
5	house	12.189611
6	major_purchase	10.794487
7	medical	11.445544
8	moving	11.558342
9	other	11.761037
10	renewable_energy	11.425446
11	small_business	12.937908
12	vacation	10.865014
13	wedding	11.801698



Purpose and interest rate bar plot. Indicates that small business have more Interest rate, whereas car loan has the substantial lowest interest rate.



Verification status VS Mean of Employee length. By looking into this chart, we can predict that the not verified applicant can be defaulter in future and by considering average employee length.