

Leading Club Case study

Batch:C37 - AL/ML

Group Member:

1. Name: Brynold.T

Gmail: <u>brynold1997@gmail.com</u>

2. Name: Tausif Khan

Gmail: <u>Tausifkit@gmail.com</u>



Abstract

- We apply data analytic methods to predict loan status for the borrowers and lenders. Loans taken by the borrowers can be fully paid, default or charged off and these situations depend on hundreds of factors.
- In this paper we will primarily deal with few of those predictor variables and try to make prediction for loan status.

LOAN DATASET



Problem solving Steps



Data Data Data Analysis Analysis

Data Cleaning

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

Data Understanding

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

Univariate Analysis

Analysing each column, plotting the distributions of each column.

Segmented Univariate Analysis

Analysing the continuous data columns with respect to the categorical column

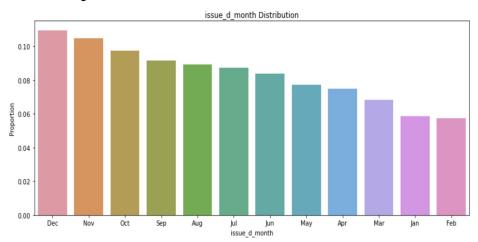
Bivariate Analysis

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

Recommendations

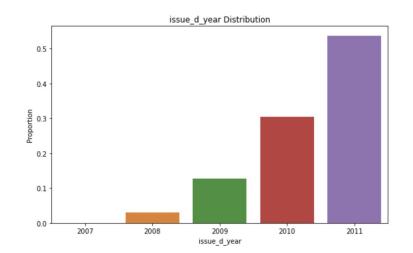
Analysing all plots and recommendations for reducing the loss of business by detecting columns best which contribute to loan defaulters.

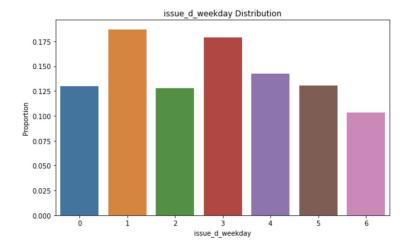
Analysis:

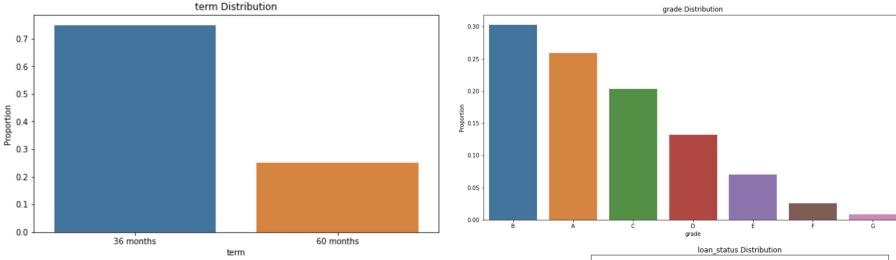




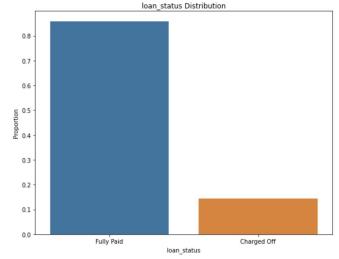
- The number of loan have been issues doubling every year. So lending club business extending every years.
- By the end of every year the loan issued are gradually increased (january to December) this could be because of some reason like year end, Christmas vacation etc.

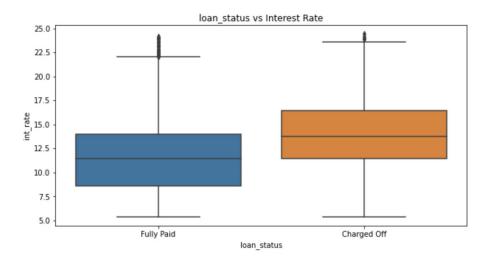


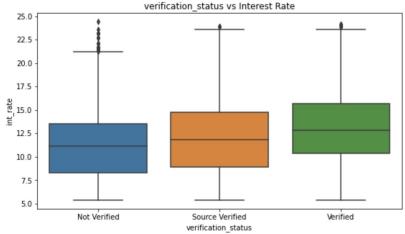




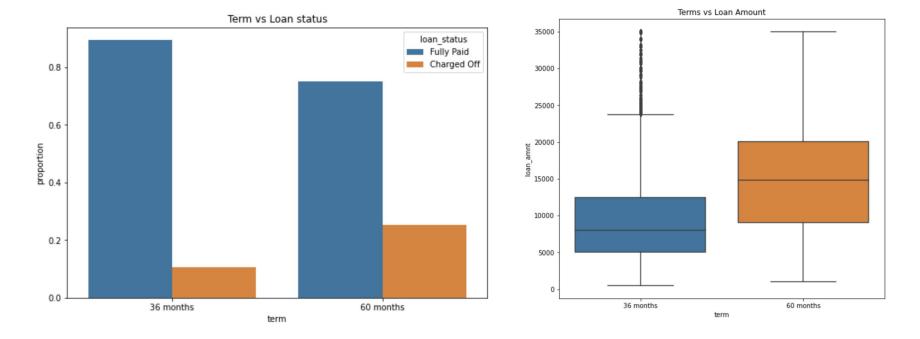
- Loan issued 36 months tenure more than 60 months. ie Around 75% of loans are issued for 36 months.
- Most borrowers fall under A and B grades than other grades mostly on B grades
- There are 85% of borrowers has paid the loan fully. where are 14% are defaulted the loan respectively.



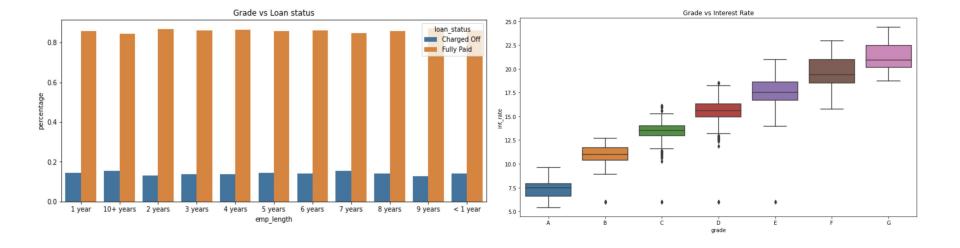




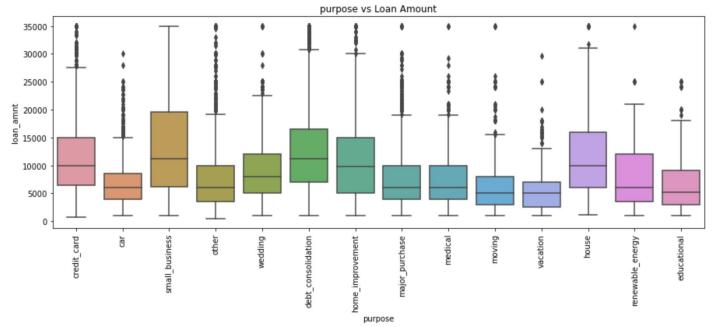
- Barrowers with Own and Mortgage got loans with less interest rates as the security will be there. And The Non-Verified Barrowers got less interest rates compared to Verified and Source Verified barrowers.
- This clearly tells us, greater the interest rate more the chance of Defaulting the loan.



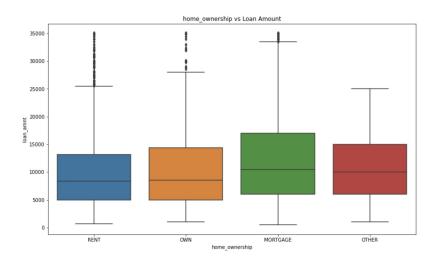
• Looks like there are more proportion of borrowers defaulted loan in 60 months term then 36 months. Also the Fully Paid rate is higher in 36 months tenure.

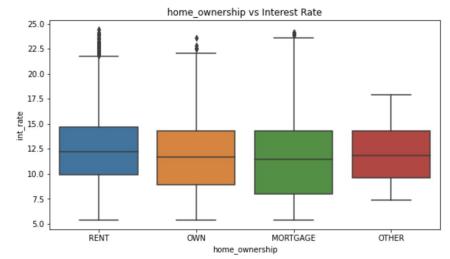


- There is not big changes or pattern observed defaulters across employment lengths.
- As grade decreases the interest rate gradually increases. and they are more and more prone to default the loan.

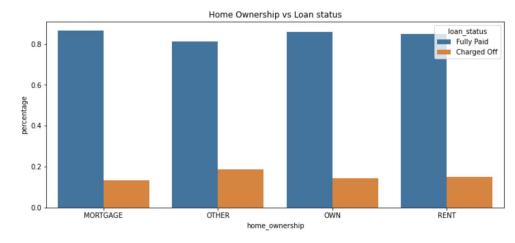


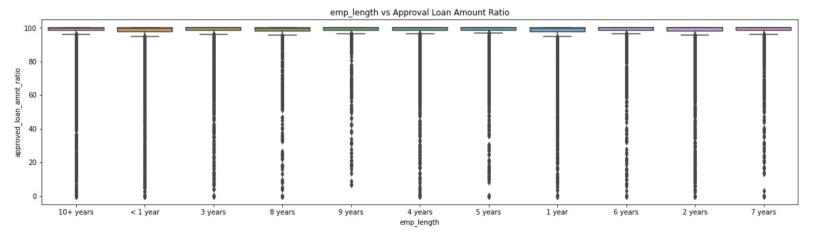
- More loan amount is from Small bussiness followed by debt consolidation.
- Least loan amount for picnic and vacations



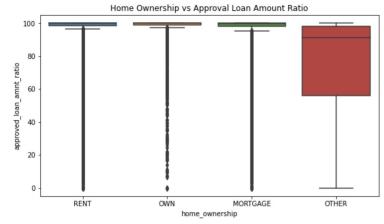


- There is bit high percentage of defaults are recorded in other home ownership category.
- People living in mortgage has high loan demand





- Borrowers with Other home ownership are having less approved ratio which mean they got less amount than request amount.
- There is not much relation between approval of loan amount ratio and employment length



									Correla	tion bet	ween (Columns										-10
id ·	- 1	0.12	0.14	0.2	0.046	0.082	0.0097	0.08	-0.0039	-0.038	0.019	-0.025	0.0024	0.088	0.032	0.11	0.17	0.14	-0.01	0.33		
loan_amnt -	0.12	1	0.98	0.95	0.3	0.93	0.27	0.06	-0.035	0.011	0.17	-0.047	0.32	0.061	0.26	0.88	0.86	0.48	-0.032	-0.12		
funded_amnt ·	0.14	0.98	1	0.97	0.3	0.96	0.27	0.06	-0.035	0.011	0.17	-0.048	0.31	0.065	0.25	0.9	0.88	0.48	-0.033	-0.035		
funded_amnt_inv	0.2	0.95	0.97	1	0.29	0.92	0.26	0.063	-0.039	0.00016	0.16	-0.052	0.29	0.068	0.24	0.89	0.91	0.47	-0.038	0.14		- 0.8
int_rate	0.046	0.3	0.3	0.29	1	0.28	0.048	0.11	0.16	0.14	0.0086	0.099	0.096	0.46	-0.046	0.29	0.28	0.17	0.083	0.0032		
installment -	0.082	0.93	0.96	0.92	0.28	1	0.27	0.05	-0.023	0.009	0.17	-0.043	0.31	0.091	0.23	0.86	0.83	0.41	-0.029	-0.037		
annual_inc	0.0097	0.27	0.27	0.26	0.048	0.27	1	-0.12	0.021	0.035	0.16	-0.014	0.28	0.016	0.24	0.26	0.25	0.14	-0.012	-0.03		- 0.6
dti -	0.08	0.06	0.06	0.063	0.11	0.05	-0.12	1	-0.035	0.00067	0.29	-0.0054	0.23	0.28	0.23	0.058	0.06	0.0077	0.0064	0.014		0.0
delinq_2yrs	-0.0039	-0.035	-0.035	-0.039	0.16	-0.023	0.021	-0.035	1	0.0084	0.011	0.011	-0.056	-0.044	0.067	-0.025	-0.03	-0.013	0.0058	-0.0085		
inq_last_6mths	-0.038	0.011	0.011	0.00016	0.14	0.009	0.035	0.00067	0.0084	1	0.094	0.023	-0.025	-0.07	0.12	-0.0088	-0.016	0.028	0.014	-0.041		
open_acc	0.019	0.17	0.17	0.16	0.0086	0.17	0.16	0.29	0.011	0.094	1	0.0034	0.29	-0.088	0.69	0.16	0.15	0.08	0.0091	-0.04		- 0.4
pub_rec		-0.047	-0.048	-0.052	0.099	-0.043		-0.0054	0.011		0.0034	1	-0.059	0.057	-0.02	-0.05	-0.053	-0.034	0.84	-0.017		
revol_bal	0.0024		0.31	0.29	0.096	0.31	0.28	0.23	-0.056	-0.025	0.29	-0.059	1	0.3	0.31	0.29	0.28	0.13	-0.047	-0.071		
revol_util		0.061	0.065	0.068	0.46	0.091	0.016	0.28	-0.044	-0.07	-0.088	0.057	0.3	1	-0.069	0.073	0.075	-0.018	0.06	0.026		
total_acc ·	0.032	0.26	0.25	0.24	-0.046	0.23	0.24	0.23	0.067	0.12	0.69	-0.02	0.31	-0.069	1	0.22	0.22	0.17	-0.0067			- 0.2
total_pymnt ·		0.88	0.9	0.89	0.29	0.86	0.26	0.058		-0.0088		-0.05	0.29	0.073	0.22	1	0.98	0.51	-0.039			
total_pymnt_inv		0.86	0.88	0.91	0.28	0.83	0.25	0.06	-0.03	-0.016	0.15	-0.053	0.28	0.075	0.22	0.98	1	0.5	-0.043			
last_pymnt_amnt -		0.48	0.48	0.47	0.17	0.41	0.14	0.0077	-0.013	0.028	0.08	-0.034	0.13	-0.018	0.17	0.51	0.5	1	-0.023	-0.0055		- 0.0
pub_rec_bankruptcies		-0.032	-0.033	-0.038	0.083	-0.029	-0.012	0.0064	0.0058	0.014	0.0091		-0.047	0.06	-0.0067	-0.039	-0.043	-0.023	1	-0.021		0.0
approved_loan_amnt_ratio	0.33	-0.12	-0.035	0.14	0.0032	-0.037	-0.03	0.014	-0.0085	-0.041	-0.04	-0.017	-0.071	0.026	-0.035	-0.0021	0.14	-0.0055	-0.021	1		
	- Þi	loan_amnt -	funded_amnt -	funded_amnt_inv -	int_rate -	installment -	annual inc -	- ifi	deling_2yrs -	ing_last_6mths -	open_acc -	- pan dud	- led_bal -	revol_util -	total_acc -	total_pymnt -	total_pymnt_inv -	last_pymnt_amnt -	pub_rec_bankruptcies -	pproved_loan_amnt_ratio -	-	

The public derogatory records column is highly (+ve) correlated with public bankruptcies records. Interest rates are high for people with high revol utilisation.

Conclusions:



- The public derogatory records column is highly (+ve) correlated with public bankruptcies records. Interest rates are high for people with high revol utilisation.
- Loan issued 36 months tenure more than 60 months. ie Around 75% of loans are issued for 36 months.
- Most borrowers fall under A and B grades than other grades mostly on B grades
- There are 85% of borrowers has paid the loan fully. where are 14% are defaulted the loan respectively
- Barrowers with Own and Mortgage got loans with less interest rates as the security will be there.
- The Non-Verified Barrowers got less interest rates compared to Verified and Source Verified barrowers.
- This clearly tells us, greater the interest rate more the chance of Defaulting the loan.
- Borrowers with Other home ownership are having less approved ratio which mean they got less amount than request amount.
- There is not much relation between approval of loan amount ratio and employment length