LENDING CLUB CASE STUDY

VARUN M R

OBJECTIVE

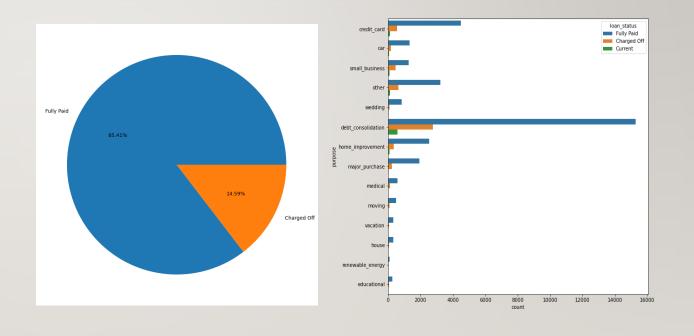
- The client is a financial institution focused on lending using a fast online process system.
- Problem Statement is to predict the risk of defaulting customers using already available customer loan data.
- Exploratory Data Analysis is used to clean, manipulate data followed by Univariate and Bivariate Analysis to find attributes contributing to default of loans.

DATA SET AND PREPARATION OF DATA.

- Data set contains basic characters of a loan like amount, term, interest rate, customer income, status of previous defaults of the customer, and final status of the loan taken, loan taken year etc.
- Data cleaning was done for single value columns, all null value columns, irrelevant columns.
- String and time variables changed into categorical and numerical for better analysis
- New variables created for better analysis of the data.

LOAN STATUS & PURPOSE OF LOAN.

- 83% loans are fully paid. Nearly 14% of the loans are charged off.
- People took majority of the loans(47%) for debt consolidation. Second main purpose of loans are Credit card payments (13%).



Correlation between Different Variables

- Loan amount, investor amount, funding amount are strongly correlated.
- Annual income with DTI(Debt-toincome ratio) is negatively correlated.
- Positive correlation between annual income and employment years showing income increases with work years.

loan_amnt -	1	0.98	0.92	0.24	0.92	0.14	0.37	0.071	-0.046	0.0026	0.025	-0.02	0.16	-0.048	0.29	0.059	0.24	0.87	0.82	0.83	0.7	0.039	0.13	0.066	0.43	-0.032	0.054	0.04
funded_amnt -	0.98		0.94	0.25	0.95	0.14		0.071						-0.049		0.064		0.89	0.84	0.85		0.041	0.13				0.066	
funded_amnt_inv	0.92	0.94		0.24	0.89	0.15		0.079	-0.053	-0.015	0.094	0.46	0.15	-0.049		0.07		0.86				0.016	0.12	0.051		-0.037		0.056
int_rate -	0.24	0.25	0.24	1	0.22	-0.0016	0.023	0.13	0.16	0.14	-0.054	0.32	-0.0061	0.1	0.061	0.46	-0.07	0.23	0.23	0.12		0.087	0.12	0.059	0.13	0.088	0.011	0.014
installment -	0.92	0.95	0.89	0.22	1	0.12	0.37	0.062	-0.033	0.0017	0.0061	-0.09	0.16	-0.045		0.09		0.85	0.79	0.84		0.051	0.1	0.062		-0.031	-0.0066	0.014
emp_length -	0.14	0.14	0.15	-0.0016	0.12	1	0.19	0.045	0.016	0.007	0.045	0.18	0.097	0.052	0.15	0.018		0.13	0.14	0.12	0.11	-0.019	0.021	0.0036	0.083	0.048	0.088	0.018
annual_inc	0.37			0.023	0.37	0.19	1	-0.11	0.025	0.035	-0.0064	-0.0031		-0.02		0.034						0.0062	0.028	0.015		-0.018	0.013	0.011
đti -	0.071	0.071	0.079	0.13	0.062	0.045	-0.11	1	-0.026	0.0076	0.064	0.19		-0.00045				0.065	0.072	0.041	0.12	-0.011	0.031	0.017	0.0084	0.01	0.076	0.016
delinq_2yrs -	-0.046	-0.047	-0.053	0.16	-0.033	0.016	0.025	-0.026	1	0.0075	-0.56	-0.028	0.013	0.01	-0.06	-0.041	0.067	-0.038	-0.044	-0.051	0.011	0.028	0.0094	0.011	-0.019	0.0053	-0.0027	-0.0096
inq_last_6mths -	0.0026	0.0025	-0.015	0.14	0.0017	0.007	0.035	0.0076	0.0075	1	-0.0028	-0.023	0.092	0.024	-0.02	-0.065	0.11	-0.02	-0.032	-0.033	0.02	0.032	0.02	0.013	0.022	0.015	-0.063	0.012
mths_since_last_delinq -	0.025	0.026	0.094	-0.054	0.0061	0.045	-0.0064	0.064	-0.56	-0.0028	1	0.49	0.022	0.057	0.022	0.11	0.014	0.022	0.078	0.023	0.016	-0.03	-0.0074	-0.016	0.017	0.025	0.15	0.053
mths_since_last_record	-0.02	-0.033	0.46	0.32	-0.09	0.18	-0.0031	0.19	-0.028	-0.023	0.49	1	0.017	0.78	-0.1		0.1	-0.025	0.4	-0.064	0.11	-0.097	-0.0086	-0.033	0.079	0.83	0.76	0.21
open_acc	0.16	0.16	0.15	-0.0061	0.16	0.097	0.26	0.29	0.013	0.092	0.022	0.017	1	0.00022	0.29	-0.083	0.68	0.15	0.14	0.15	0.11	-0.022	0.0072	0.0031	0.071	0.0067	-0.00079	-0.0013
pub_rec ·	-0.048	-0.049	-0.049	0.1	-0.045	0.052	-0.02	-0.00045	0.01	0.024	0.057	0.78	0.00022	1	-0.058	0.059	-0.02	-0.051	-0.051	-0.062	0.0025	-0.0011	-0.0045	-0.0041	-0.027	0.85	-0.0016	-0.022
revol_bal -	0.29			0.061		0.15	0.41		-0.06	-0.02	0.022	-0.1	0.29	-0.058	1	0.31	0.3	0.25				0.0029	0.038	0.016	0.098	-0.045	-0.031	0.013
revol_util -	0.059	0.064	0.07	0.46	0.09	0.018	0.034		-0.041	-0.065	0.11		-0.083	0.059		1	-0.069	0.07	0.074	0.019		0.034	0.046	0.018	-0.023	0.061	0.065	0.044
total_acc -	0.24	0.23	0.22	-0.07	0.21	0.2			0.067	0.11	0.014	0.1	0.68	-0.02		-0.069	1	0.2	0.2	0.21	0.12	-0.029	0.017	0.0074	0.15	-0.0063	0.039	-0.00043
total_pymnt -		0.89	0.86	0.23	0.85	0.13			-0.038			-0.025		-0.051		0.07		1	0.96	0.98		0.00037				-0.04		0.016
total_pymnt_inv -		0.84		0.23	0.79	0.14				-0.032		0.4		-0.051		0.074		0.96		0.93				0.0059		-0.043		0.04
total_rec_prncp -		0.85	0.82	0.12	0.84	0.12				-0.033				-0.062		0.019		0.98	0.93			-0.03		-0.066		-0.05	0.043	
total_rec_int - total_rec_late_fee -		0.71	0.69	0.5	0.6	0.11	0.23	0.12	0.011	0.02	0.016	0.11		-0.0025	0.2	0.2	0.12	0.79	0.77	0.65	0.064	0.064	0.084	0.03	0.16	-0.0054	0.089	
total_rec_late_fee -		0.041	0.016	0.087	0.051	0.019		0.031						-0.0011				0.00037			0.064	0.11	0.11	0.092		-0.0054		
collection recovery fee																						0.092	0.81	1		-0.0042		
last_pymnt_amnt -			0.43	0.13	0.37	0.083								-0.027				0.47	0.46	0.53	0.16	-0.067	-0.077	-0.044	1	-0.017		
pub_rec_bankruptcies -												0.83	0.0067			0.061						-0.0054			-0.017		0.0046	
issue_year		0.066	0.22					0.076				0.76	-0.00079	-0.0016	-0.031	0.065	0.039	0.059		0.043	0.089	-0.074	0.019	-0.021	0.11	0.0046	1	-0.042
issue_month ·	0.04	0.027	0.056	0.014	0.014	0.018	0.011	0.016	-0.0096	0.012	0.053	0.21	-0.0013	-0.022	0.013	0.044	-0.00043	0.016	0.04	0.014	0.014	-0.0087	0.016	-0.0014	0.044	-0.021	-0.042	1
	loan_amnt -	funded_amnt -	funded_amnt_inv -	int_rate -	installment -	emp_length -	annual_inc -	offi -	deling_2yrs -	ing_last_6mths -	ths_since_last_deling -	hs_since_last_record -	oben_acc -	- par qnd	revol_bal -	revol_util -	total_acc -	total_pymnt -	total_pymnt_inv -	total_rec_prncp -	total_rec_int -	total_rec_late_fee -	recoveries -	lection_recovery_fee -	last_pymnt_amnt -	ub_rec_bankruptcies -	issue_year -	issue_month -

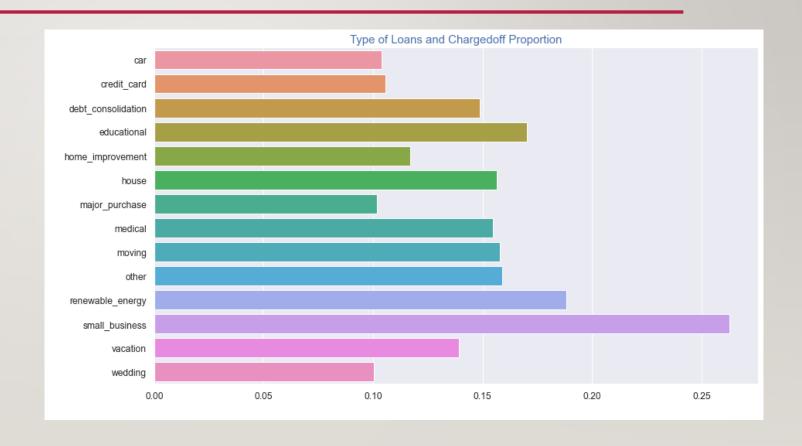
LOAN TERM VS LOAN TIME.

The plot shows proportion of charged off loans increases with increase in loan term.



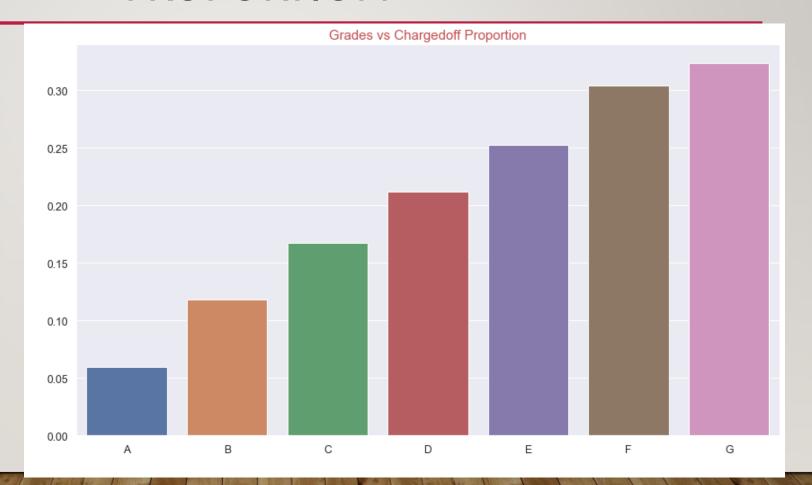
TYPE OF LOAN GIVEN AND DEFAULT.

- Small business, renewable energy, educational loans shown highest chargedoff proportion.
- Wedding, vacation, home improvement loans seen less default. May be also due to small amounts and lower loan term.



GRADE OF THE LOAN VS CHARGED OFF PROPORTION

 Change of the Grades from A to G shows steady increase in charged off loans.



RECOMMENDATIONS

- Income range 80000+ has less chances of charged off but 0-20000 has high chances of charged off.
- Debt Income Ratio, Grade of the Loan, Annual Income, Previous Record of Bankruptcy are strong indicator loan default.
- Small scale Industries, Educational Loans need thorough risk assessment.
- Grades F, G have higher risk of default compared to others.