# Lending Club Case Study Submission

Name: Jithin J & Senthilkumar L

#### **Abstract**

- This case study is to help a consumer finance company to take decision as to whether a applicant will default on the loan provided to him/her based on the various consumer attributes and loan attributes
- Based on the attributes the company can take two decisions, either to accept the loan or reject the loan.
- If the company accepts for the loan, then there are 3 scenarios which can be encountered, the loan can be either be fully paid or charged off or the applicant is still paying the loan.
- Our main concern is to find the factors leading to charged off so that company can avoid financial loss due to the loan being defaulted

# Problem Solving Methodology

Data Collection from the Loan Dataset

Cleaning the data in the dataset by removing the columns having all missing

Removing the columns having only a single value Bi Variate analysis is done mainly helped to compare each of the variable with loan status.

Univariate analysis is done know its distribution

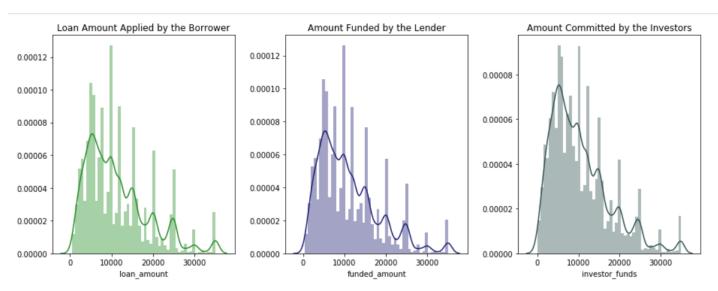
Performing Univariate and Bivariate analysis of different variables

The bivariate analysis helped us to know which

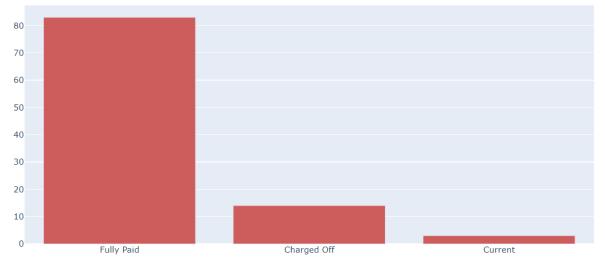
factors have a high impact

## **Univariate Analysis**

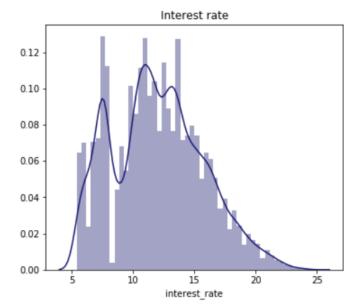
- Following are the findings of Univariate Analysis:
- Most of the loan amount applied was in the range of 5000 to 20000
- The amount applied, lend amount and amount invested by investors also has a similar distribution



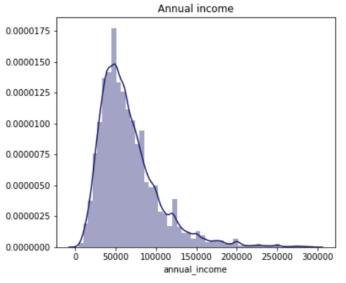
We can see that 14% of the loans are defaulted and 83% of the loans were fully paid.



Most of the interest rate was in the range of 9% to 16%

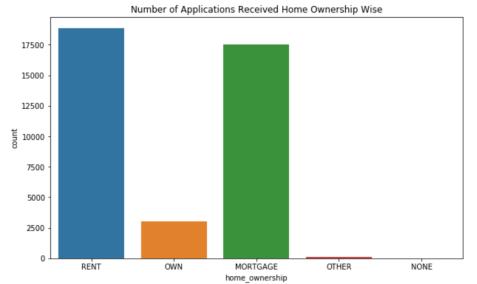


Annual income mostly lies in the range of 40000 to 82000.

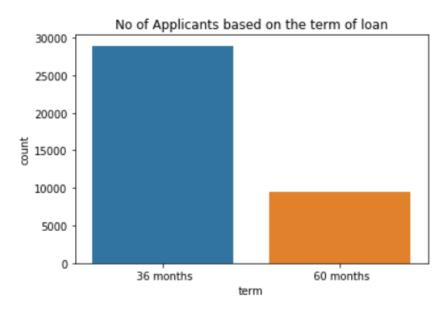


The maximum number of applications were received where the house was

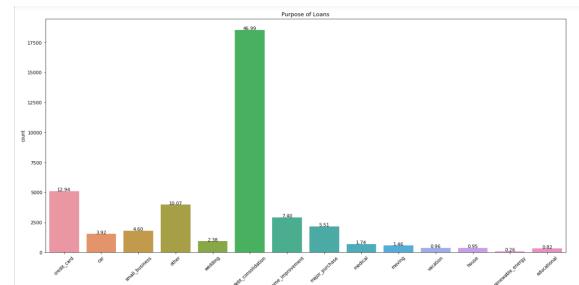
rented.



Most of the loans in our dataset have a term of 36 Months

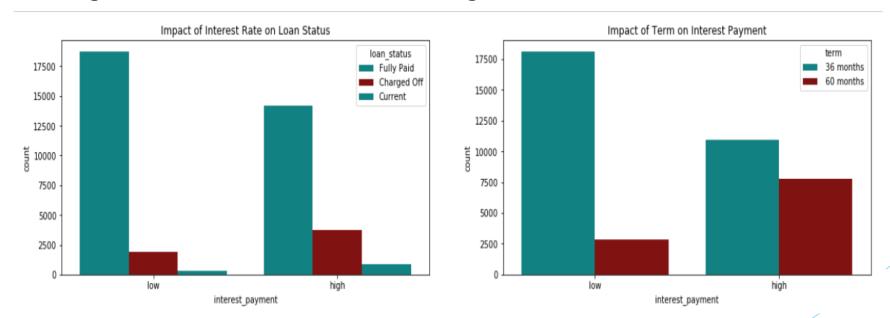


We can see from the below the loan was majorly taken for debt consolidation followed by credit card and home improvement

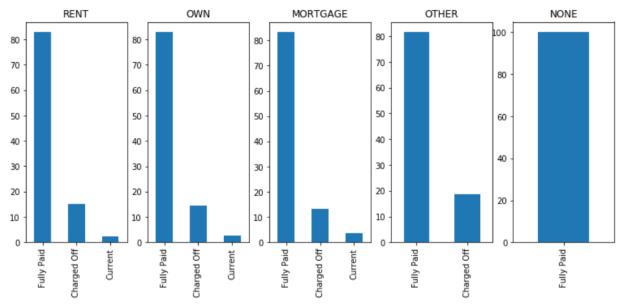


### Bivariate Analysis

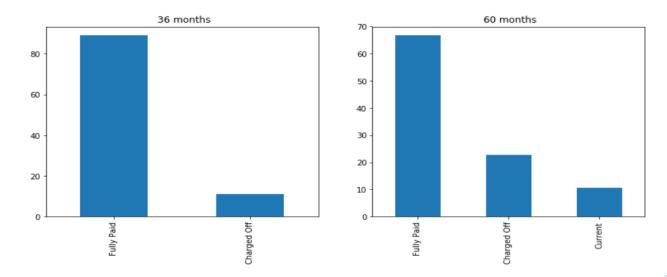
- We observed the following on performing the bivariate analysis
- From the below plot we can observe that we have a higher rate of loans being charged off when the interest rate is high also when the term is 60 months.



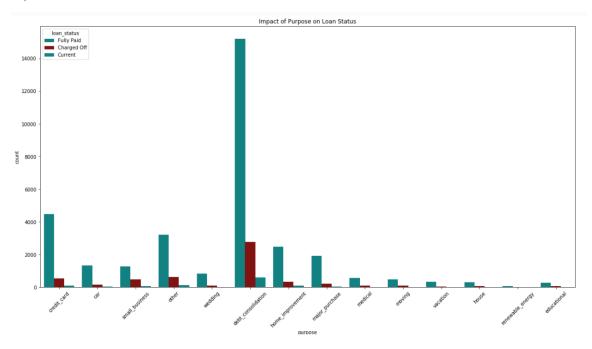
Most of the charged off loans are where the house is rented



► The charged off count increases when the term is longer

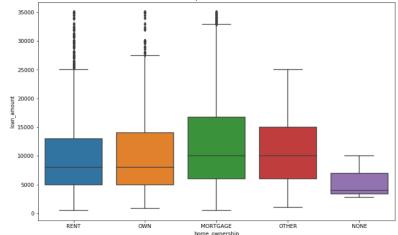


We can see that highest number of charged off loans were the ones taken for the purpose of debt consolidation.



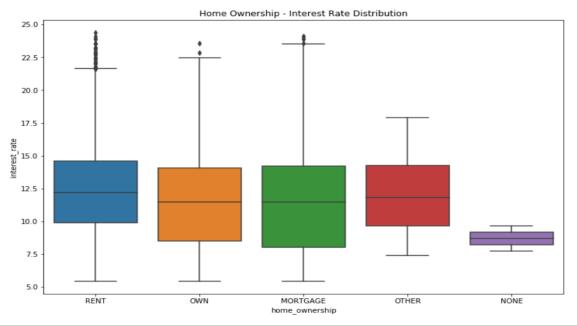
From the below plot we can observe that loan amount has a highest spread

for the mortgage type

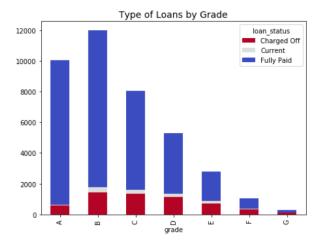


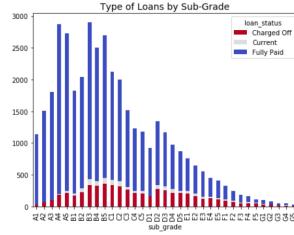
From the above we can see the highest interest rate of loan is for the rented

house

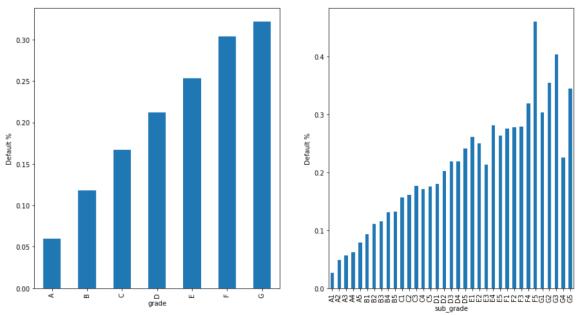


Most of the charged off loans are in the grades B and C followed by D

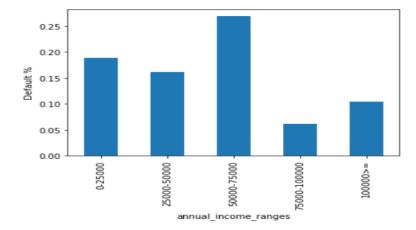




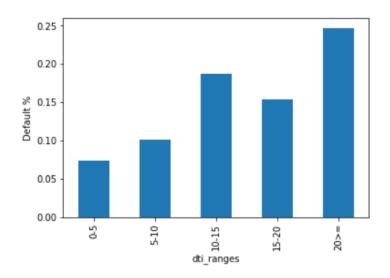
From the figure, it is clear that as grade increases so as default percentage



From the figure, it is clear that people with low income range are tend to pay their loans.

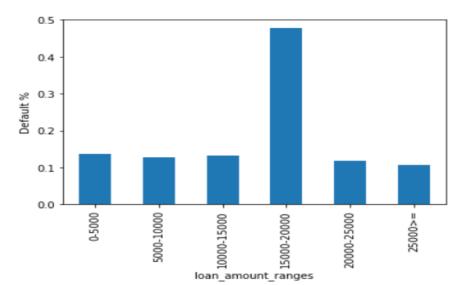


▶ It is clear that as dti range increases so as default percentage

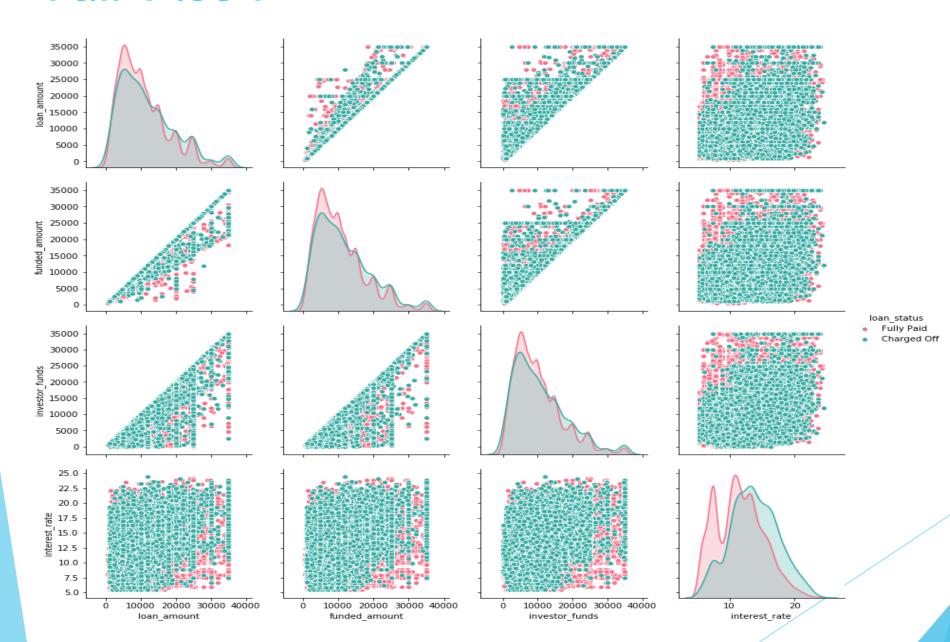


From the figure, it is clear that as loan amount increases so as default

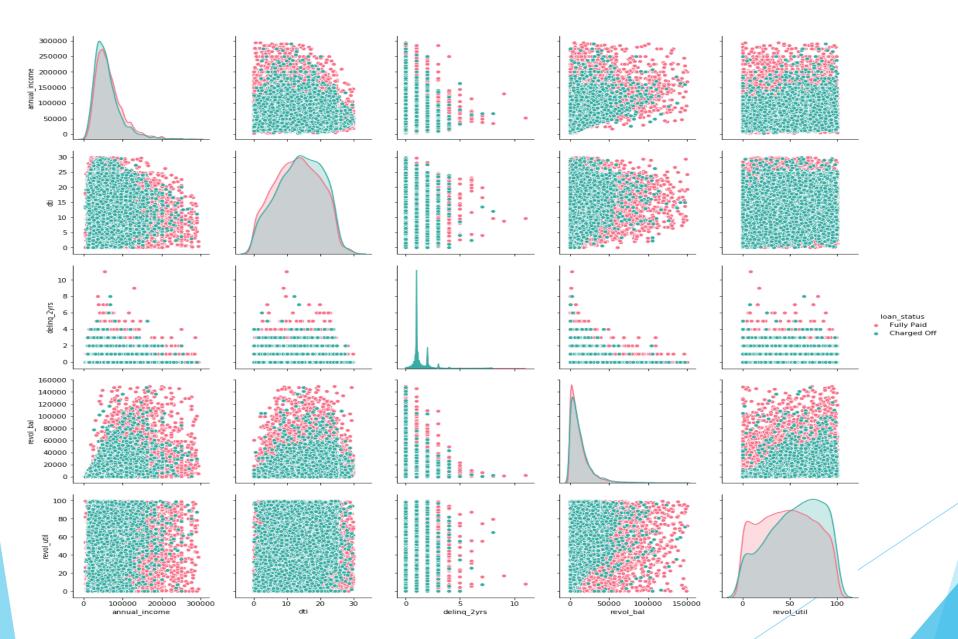
percentage



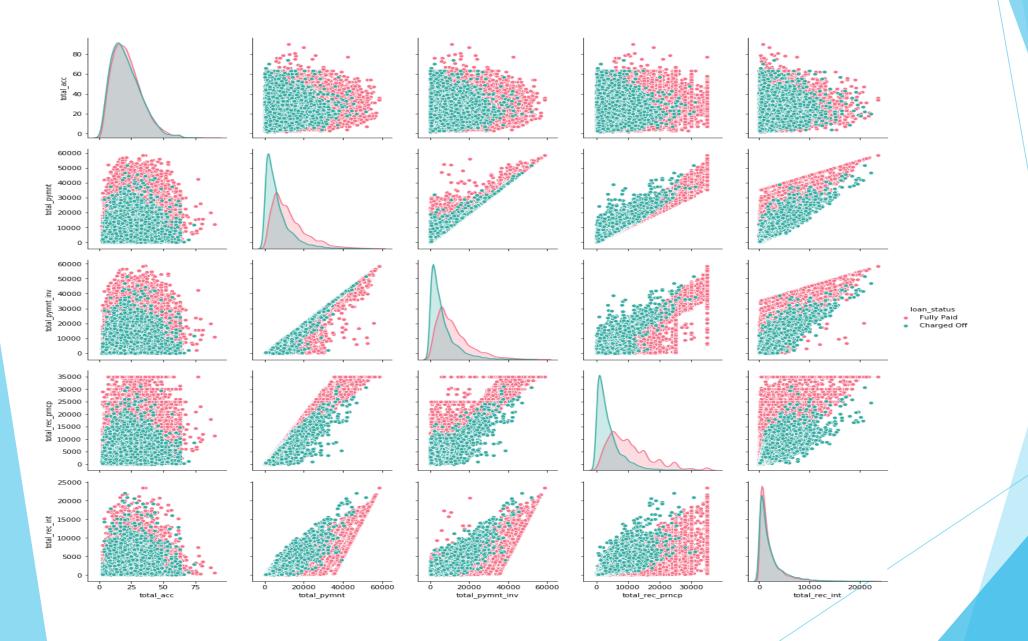
### Pair Plot 1



### Pair Plot 2



### Pair Plot 3



## **Correlation Plot**

loan_amount -	1	0.98	0.94	0.3	0.93	0.42	-0.061	0.068	-0.034	0.011	0.18	-0.049	0.31	0.064	0.25	0.88	0.85	0.84	0.73	0.048	0.14	0.077	0.47	-0.035	0.052
funded_amount -	0.98	1	0.96	0.3	0.96	0.41	-0.058	0.067	-0.034	0.011	0.17	-0.049		0.068	0.25	0.9	0.86	0.86	0.74	0.05	0.14	0.079	0.48	-0.036	0.042
investor_funds <sup>-</sup>	0.94	0.96	1	0.3	0.91	0.39	-0.039	0.075	-0.04	-0.0043	0.16	-0.05	0.28	0.073	0.24	0.87	0.91	0.84	0.73	0.03	0.13	0.064	0.47	-0.04	0.068
interest_rate -	0.3	0.3	0.3	1	0.28	0.069	-0.21	0.11	0.16	0.13	0.0035	0.1	0.091	0.47	-0.05	0.29	0.28	0.17	0.52	0.095	0.13	0.071	0.17	0.084	0.026
installment -	0.93	0.96	0.91	0.28	1	0.42	-0.028	0.057	-0.022	0.0091	0.17	-0.044	0.31	0.094	0.23	0.86	0.82	0.85	0.64	0.059	0.12	0.078	0.41	-0.032	0.03
annual_income -	0.42	0.41	0.39	0.069	0.42	1	0.063	-0.12	0.029	0.039	0.26	-0.022	0.42	0.039	0.37	0.4	0.38	0.4	0.28	0.014	0.039	0.024	0.22	-0.021	0.015
loan_status -	-0.061	-0.058	-0.039	-0.21	-0.028	0.063	1	-0.045	-0.021	-0.073	0.0094	-0.051	-0.0071	-0.1	0.022	0.24	0.23	0.34	-0.014	-0.17	-0.34	-0.21	0.22	-0.048	-0.025
dti -	0.068	0.067	0.075	0.11	0.057	-0.12	-0.045	1	-0.032	0.0047	0.29	-0.0053	0.24	0.28	0.23	0.065	0.071	0.042	0.11	-0.011	0.027	0.013	0.011	0.0057	0.015
delinq_2yrs -	-0.034	-0.034	-0.04	0.16	-0.022	0.029	-0.021	-0.032	1	0.0081	0.012	0.0082	-0.057	-0.043	0.066	-0.025	-0.032	-0.041	0.022	0.029	0.012	0.013	-0.015	0.0039	-0.011
inq_last_6mths -	0.011	0.011	-0.0043	0.13	0.0091	0.039	-0.073	0.0047	0.0081	1	0.093	0.024	-0.022	-0.069	0.11	-0.01	-0.021	-0.024	0.026	0.031	0.019	0.013	0.026	0.016	0.013
open_acc -	0.18	0.17	0.16	0.0035	0.17	0.26	0.0094	0.29	0.012	0.093	1	0.00076	0.29	-0.09	0.69	0.16	0.15	0.16	0.12	-0.018	0.017	0.0061	0.082	0.0066	0.0006
pub_rec -	-0.049	-0.049	-0.05	0.1	-0.044	-0.022	-0.051	-0.0053	0.0082	0.024	0.00076	1	-0.061	0.059	-0.023	-0.052	-0.052	-0.063	-0.0032	-0.0019	-0.0057	-0.0056	-0.033	0.85	-0.023
revol_bal -	0.31	0.3	0.28	0.091	0.31	0.42	-0.0071	0.24	-0.057	-0.022	0.29	-0.061	1	0.3	0.31	0.29	0.27	0.27	0.24	0.0072	0.045	0.025	0.12	-0.048	0.016
revol_util -	0.064	0.068	0.073	0.47	0.094	0.039	-0.1	0.28	-0.043	-0.069	-0.09	0.059	0.3	1	-0.071	0.074	0.078	0.022	0.19	0.039	0.052	0.029	-0.016	0.061	0.047
total_acc -	0.25	0.25	0.24	-0.05	0.23	0.37	0.022	0.23	0.066	0.11	0.69	-0.023	0.31	-0.071	1	0.22	0.22	0.23	0.15	-0.025	0.024	0.011	0.17	-0.01	0.0025
total_pymnt -	0.88	0.9	0.87	0.29	0.86	0.4	0.24	0.065	-0.025	-0.01	0.16	-0.052	0.29	0.074	0.22	1	0.97	0.97	0.82	0.015	0.03	0.028	0.51	-0.042	0.026
total_pymnt_inv -	0.85	0.86	0.91	0.28	0.82	0.38	0.23	0.071	-0.032	-0.021	0.15	-0.052	0.27	0.078	0.22	0.97	1	0.94	0.81	-1.4e-05	0.023	0.018	0.5	-0.045	0.049
total_rec_prncp -	0.84	0.86	0.84	0.17	0.85	0.4	0.34	0.042	-0.041	-0.024	0.16	-0.063	0.27	0.022	0.23	0.97	0.94	1	0.68	-0.019	-0.095	-0.059	0.57	-0.052	0.023
total_rec_int -	0.73	0.74	0.73	0.52	0.64	0.28	-0.014	0.11	0.022	0.026	0.12	-0.0032	0.24	0.19	0.15	0.82	0.81	0.68	1	0.079	0.093	0.042	0.24	-0.00039	0.022
total_rec_late_fee -	0.048	0.05	0.03	0.095	0.059	0.014	-0.17	-0.011	0.029	0.031	-0.018	-0.0019	0.0072	0.039	-0.025	0.015	-1.4e-05	-0.019	0.079	1	0.099	0.092	-0.062	-0.0056	-0.0051
recoveries -	0.14	0.14	0.13	0.13	0.12	0.039	-0.34	0.027	0.012	0.019	0.017	-0.0057	0.045	0.052	0.024	0.03	0.023	-0.095	0.093	0.099	1	0.79	-0.072	-0.0048	0.023
collection_recovery_fee	0.077	0.079	0.064	0.071	0.078	0.024	-0.21	0.013	0.013	0.013	0.0061	-0.0056	0.025	0.029	0.011	0.028	0.018	-0.059	0.042	0.092	0.79	1	-0.043	-0.0053	0.00068
last_pymnt_amnt -	0.47	0.48	0.47	0.17	0.41	0.22	0.22	0.011	-0.015	0.026	0.082	-0.033	0.12	-0.016	0.17	0.51	0.5	0.57	0.24	-0.062	-0.072	-0.043	1	-0.023	0.055
pub_rec_bankruptcies -	-0.035	-0.036	-0.04	0.084	-0.032	-0.021	-0.048	0.0057	0.0039	0.016	0.0066	0.85	-0.048	0.061	-0.01	-0.042	-0.045	-0.052	-0.00039	-0.0056	-0.0048	-0.0053	-0.023	1	-0.022
month -	0.052	0.042	0.068	0.026	0.03	0.015	-0.025	0.015	-0.011	0.013	0.0006	-0.023	0.016	0.047	0.0025	0.026	0.049	0.023	0.022	-0.0051	0.023	0.00068	0.055	-0.022	1
	loan_amount -	funded_amount -	investor funds -	interest_rate -	installment -	annual_income -	loan_status -	dfi -	delinq_2yrs -	ing_last_6mths -	open_acc -	- bnp rec	- led_bal	revol_util -	total_acc -	total_pymnt -	total_pymnt_inv -	total_rec_pmcp -	total_rec_int -	total_rec_late_fee -	recoveries -	ection_recovery_fee -	last_pymnt_amnt -	ub_rec_bankruptcies -	month -

- We can observe the following from our correlation plot:
- ► The loan status has a high negative correlation with the interest rate. We can see that as the interest rate increases, there are chances for applicants to default on the loan.
- Also the loan status shows a positive correlation with total received principle. This means that the loan will not be defaulted if the principal amount is being received on date.
- The interest rate has a positive correlation with total received interest.
- The total received interest depends highly on the loan amount and the funded amount.

#### Conclusion

- We can come to the below conclusion from our analysis.
- We have a higher rate of loans being charged off when the interest rate is high also when the term is 60 months. Hence it will be better for the company to approve loans with a lower interest rate and lesser term
- Its better to approve loans to people where the house is owned under them or mortgaged for the loan.
- Lesser debt to income ratio is better for Loan being paid on time.
- Low income applicants tend to pay the loan but it depends on the dti which means loan amount / income amount. Low income people tend to apply for lower amount.
- Grade also has an impact on default, lower grade ABC has lower default ratio that higher grade EFG
- Purpose may seem to make little difference but small business has higher probability to go default. So small business with lower DTI option are better for the loan.
- Hence the company needs to keep the above observations in mind while approving for a loan.