# Lending Club Case Study

By:

Manu Sijaria

Aryan Baruah

### Overview

- About Company: Lending Club is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface.
- Problem statement: Analyze data to identify
  - The driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default.
  - Consumer and loan attributes that influence the tendency of default.

## Methodology



### Data Analysis and Cleaning

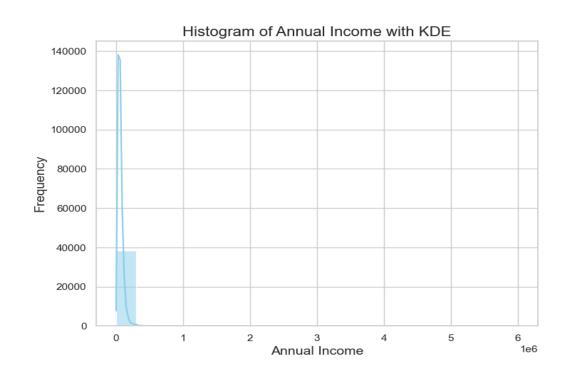
- Identifying and dealing with missing values in the dataset. Removing the columns with all Null values.
- Check for duplicate rows and remove them.
- Removing columns with unique values after analysis.
- Removing fields which are not useful in terms of the problem statement
- Remove fields not adding valuable meaning to the problem statement wrt loan domain.
- Outliers can significantly impact the results of statistical analysis.
  Identifying and deciding whether to remove, transform, or keep outliers is is essential.

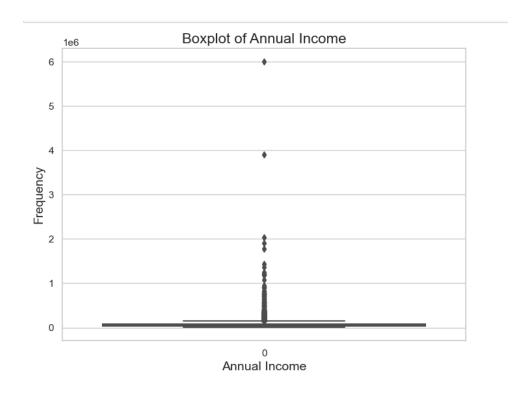
### Data Analysis and Cleaning

- Correcting Inconsistencies by identifying and rectifying inconsistencies in the data.
- Data Transformation by analysing and convert data into proper tyes eg int, float date etc
- If categorical variables have too many levels, they might be consolidated into broader categories for better analysis and interpretation.

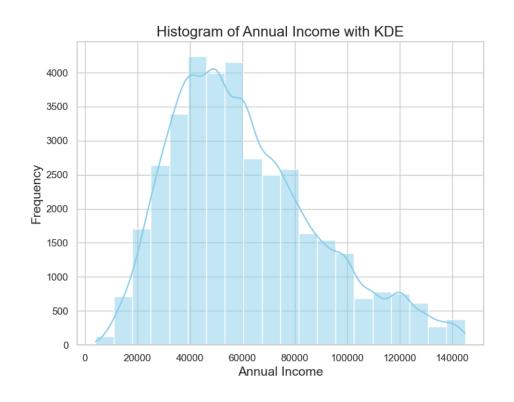
## Univariate Analysis

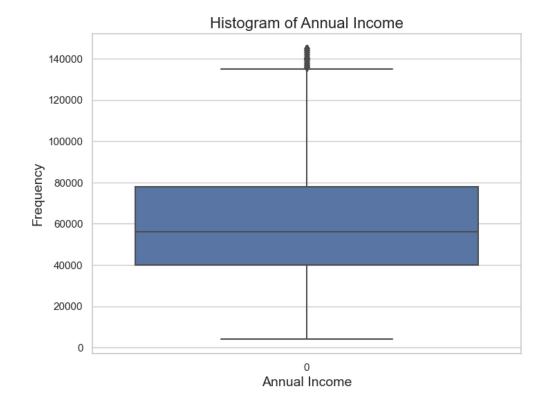
Observations: While plotting the annual income, outliers are observed.



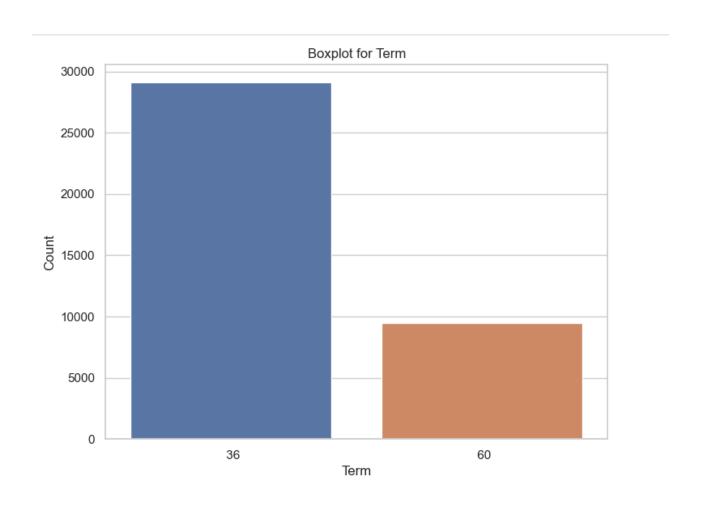


Observation: Inter Quartile Range is used to remove the outliers. The distribution is normal.



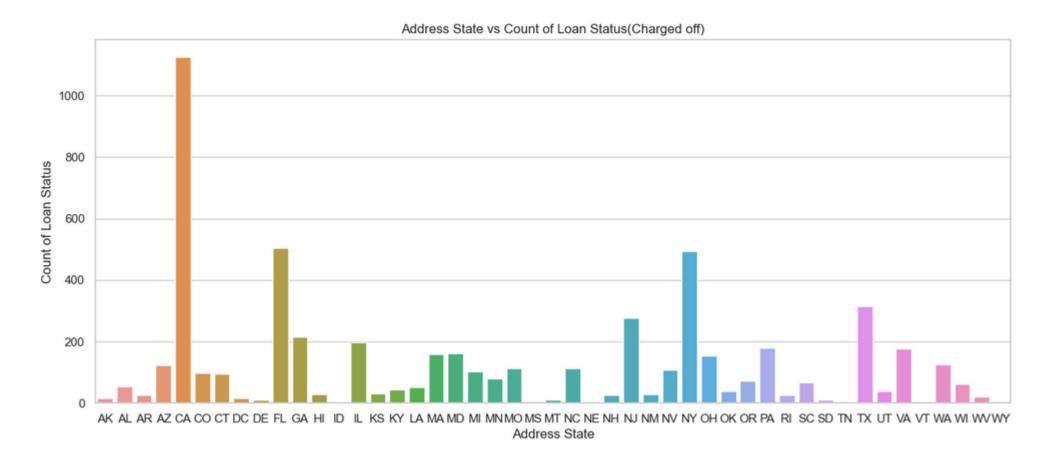


### Observations: Applicants take more loans for 36 months

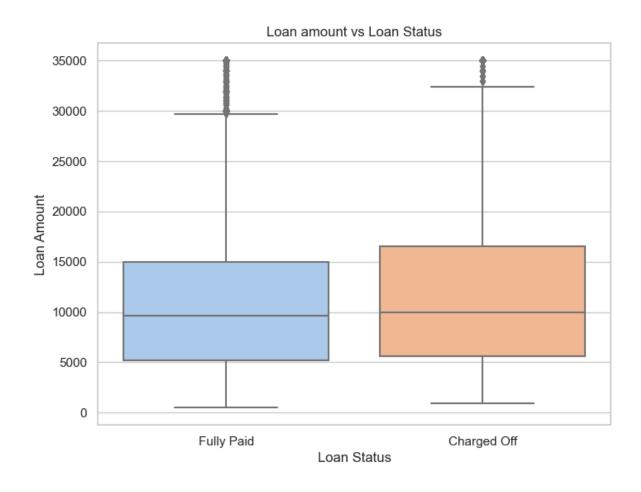


### Bivariate Analysis

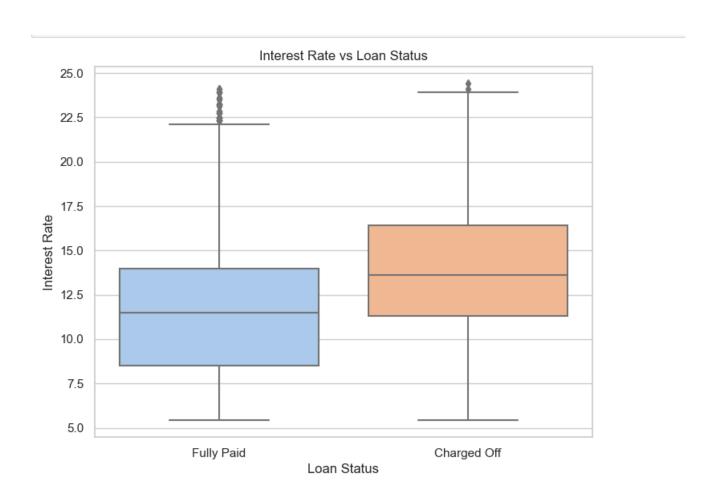
Observation: Maximum number of defaulters are from the state of CA



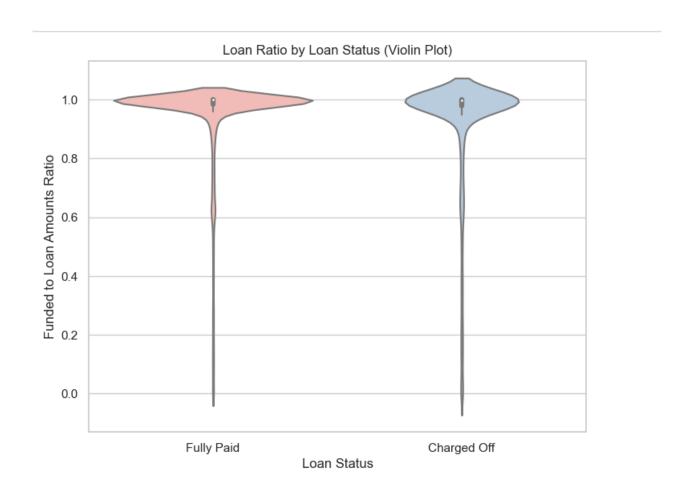
Observation: Charged off loans have higher loan amounts than fully paid loans.



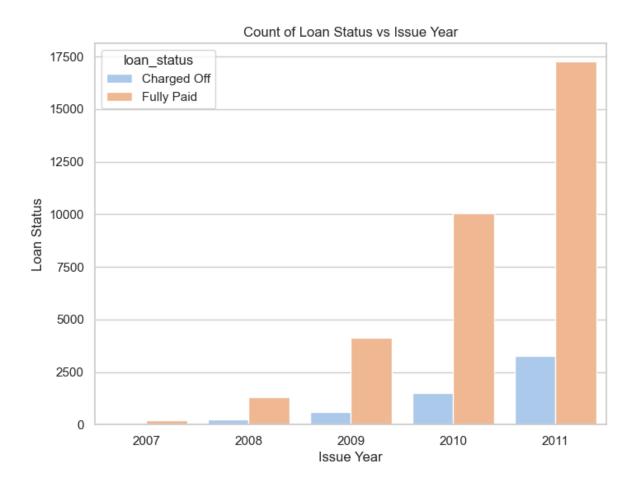
Observation: Higher rate of interest causes more defaulters.



Observations: There is a higher probability of loan getting fully paid.



Observations: Even though loan applicants are increasing with the years, there is a much higher chance of the loan getting paid off.



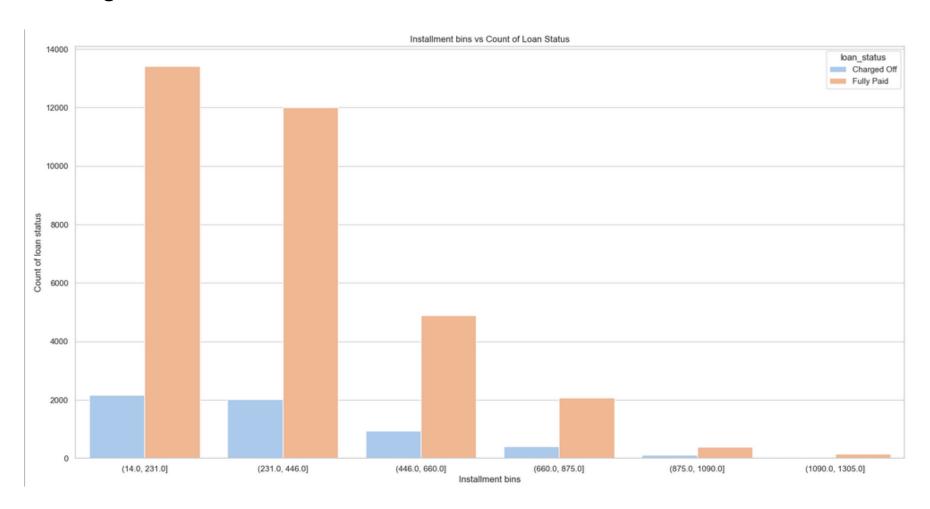
Observations: Lower DTI has significantly higher loan status of fully paid.



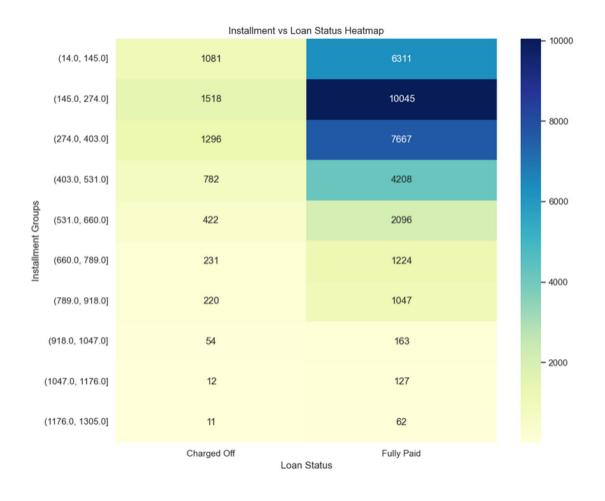
#### Observation:

If the installment is less, there is higher fully paid loans.

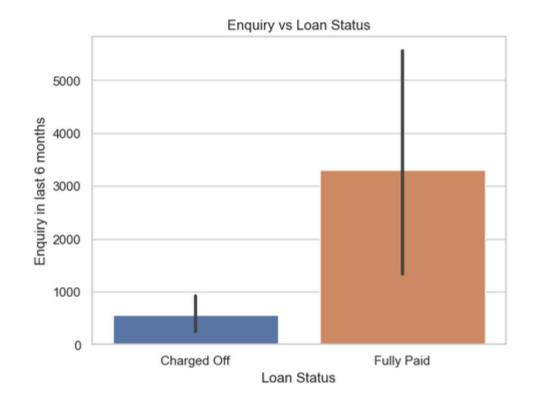
In each installment bin, it has been observed that fully paid loan status has a greater number of records wrt charged off status.



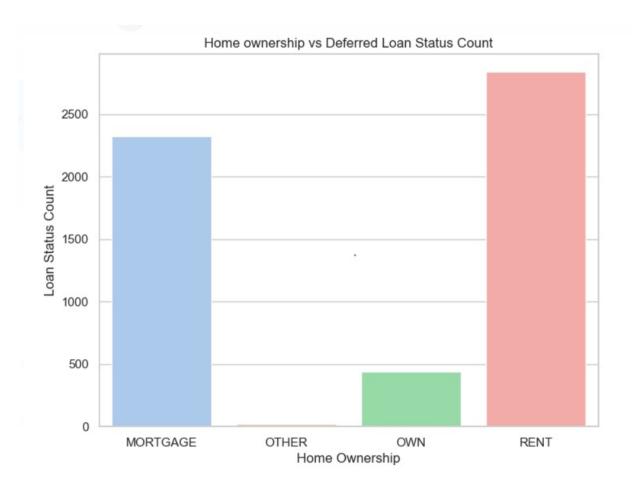
Observations: Less number of installment has more number of fully paid loan status records.



Observations: In the last 6 months it can be seen that People who fully paid the loan call more often to the bank then the people who charged off

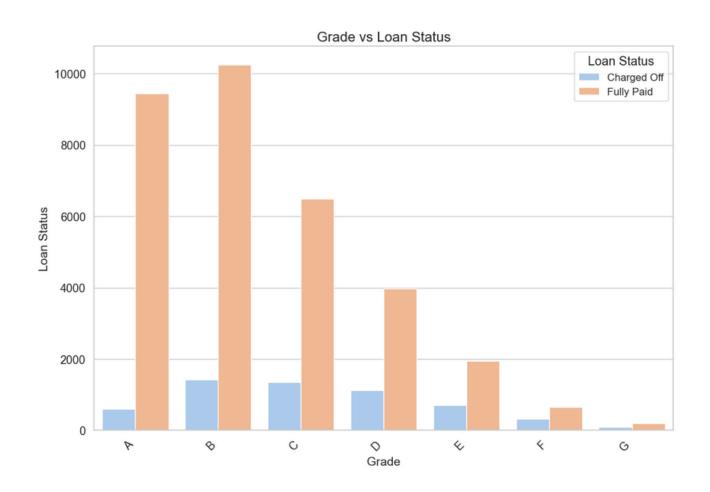


Observations: Applicants who rent would defer loans more.



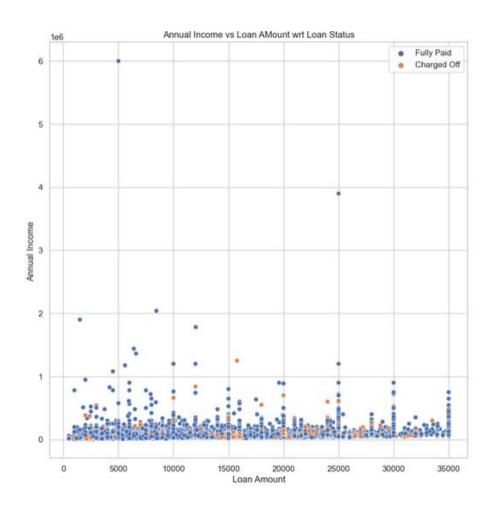
Observations:

Applicants with grade A, B,C are most likely to pay off the loan. Applicants with grade G is most likely to default the loan.



### Multivariate Analysis

Observations: Applicants with more annual income has paid off their loans more.



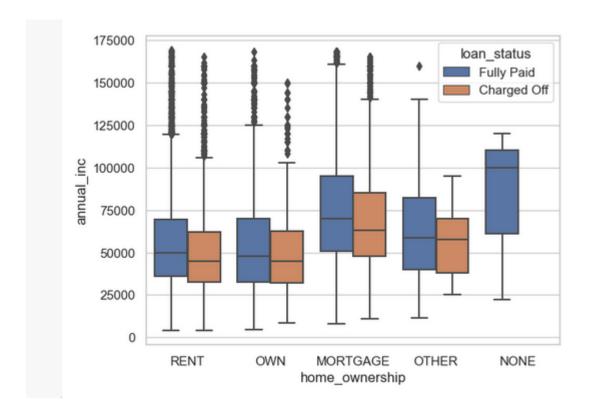
#### Observations:

There is a very high correlation between number of installment and total payments. Also observed in Funded amount and total payment.

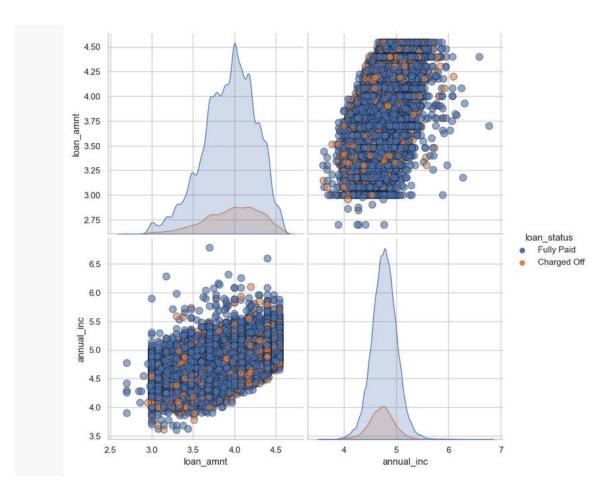
DTI and Annual income has negative correlation.

|                        |          |             |                |          |             |            |        |             | Corre           | alatio                | n Hea                  | itmap  | ,       |             |                 |               |                   |            |           |                        |     |
|------------------------|----------|-------------|----------------|----------|-------------|------------|--------|-------------|-----------------|-----------------------|------------------------|--------|---------|-------------|-----------------|---------------|-------------------|------------|-----------|------------------------|-----|
| loan_amnt              | 1        | 0.98        | 0.94           | 0/3      | 0.93        | 0.27       | 0.062  | -0.032      | 0.013           | 0.015                 | 0.0007                 | 7-0.05 | 0.31    | 0.88        | 0.85            | 0.85          | 0.47              | 0.052      | 0.11      | -0.11                  | 1   |
| funded_amnt            | 0.98     | 21          | 0.96           | 0.3      | 0.96        |            | 0.062  | -0.032      | 0.013           | 0.016                 | -0.011                 | -0.051 |         | 0.9         | 0.86            | 0.86          | 0.48              | 0.042      | 0.12      | -0.038                 |     |
| funded_amnt_inv        | 0.94     | 0.96        | 1              | 0.3      | 0.91        |            | 0.071  | -0.038      | 0.0028          | 0.073                 | 0.44                   | -0.051 |         | 0.87        | 0.91            | 0.84          | 0.47              | 0.068      | 0.25      | 0.19                   | - 0 |
| int_rate               | 0.3      | 0.3         | 0.3            | 1        | 0.28        | 0.049      | 0.11   | 0.16        | 0.13            | -0.054                | 0.33                   | 0.099  | 0.094   | 0.29        | 0.28            | 0.17          | 0.17              | 0.025      | 0.039     | 0.02                   |     |
| installment            | 0.93     | 0.96        | 0.91           | 0.28     | 1           |            | 0.052  | -0.02       | 0.011           | 0.0002                | 7-0.064                | -0.046 | 0.31    | 0.86        | 0.82            | 0.85          | 0.41              | 0.029      | 0.053     | -0.05                  |     |
| annual_inc             | 0.27     | 0.26        | 0.25           | 0.049    | 0.27        | 1          | -0.12  | 0.022       | 0.035           | 0.007                 | -0.032                 | -0.018 |         | 0.26        | 0.25            | 0.26          | 0.14              | 0.009      | 0.0086    | 0.031                  | -0  |
| æ                      | 0.062    | 0.062       | 0.071          | 0.11     | 0.052       | -0.12      | 1      | -0.033      | 0.0022          | 0.069                 | 0 19                   | 0.0047 | 0.23    | 0.059       | 0.066           | 0.037         | 0.0086            | 0.014      | 0.093     | 0.033                  |     |
| delinq_2yrs            | 0.032    | -0.032      | -0.038         | 0.16     | -0.02       | 0.022      | -0.033 | (1)         | 0.0087          | -0.57                 | -0.037                 | 0.0076 | 0.054   | -0.023      | -0.03           | -0.038        | 0.013             | 0.011      | 0.0031    | -0.016                 | -0  |
| inq_last_6mths         | 0.013    | 0.013       | 0.0028         | 0.13     | 0.011       | 0.035      | 0.0022 | 0.0087      | 1               | 0.005                 | 0.029                  | 0.024  | 0.021   | 0.0082      | -0.019          | -0.021        | 0.028             | 0.013      | -0.059    | -0.046                 |     |
| mths_since_last_deling | 0.015    | 0.016       | 0.073          | -0.05-40 | 00027       | 0.0074     | 0.069  | -0.57       | 0.0051          | 1                     | 0.49                   | 0.059  | 0.014   | 0.013       | 0.06            | 0.015         | 0.0076            | 0.053      | 0.14      | 0.10                   |     |
| mths_since_last_record | 0.0007   | 70.011      | 0.44           | 0.33     | 0.064       | -0.032     |        | 0.037       | -0.029          | 0.49                  | 1                      | 0.78   | -0.1    | 0.0058      | 0.37            | 0.048         | 0.084             | 0.22       | 0.75      | 0.69                   | -0  |
| pub_rec                | 0.05     | 0.051       | -0.051         | 0.099    | 0.046       | -0.018     | 0.0047 | 0.0076      | 0.024           | 0.059                 | 0.78                   | 1      | 0.061   | -0.053      | -0.053          | -0.064        | 0.033             | -0.022     | -0.006    | 0.0062                 |     |
| revol_bal              | 0.31     |             | 0.29           | 0.094    | 0.51        |            |        | -0.054      | -0.021          | 0.014                 | -0.1                   | -0.061 | 10      | 0.29        |                 |               | 0.10              | 0.015      | 0.0093    | 0.067                  | -0  |
| total_pymnt            | 0.88     | 0.9         | 0.87           | 0.29     | 0.86        |            | 0.059  | -0.023      | 0.0082          | 0.013                 | 0.0058                 | 0.053  |         | 11          | 0.97            | 0.97          |                   | 0.026      | 0.11      | -0.002                 |     |
| total_pymnt_inv        | 0.85     | 0.86        | 0.91           | 0.28     | 0.82        |            | 0.066  | 0.03        | -0.019          | 0.06                  | 0.37                   | -0.053 |         | 0.97        | /4              | 0.94          | 0.5               | 0.049      | 0.22      | 0.18                   |     |
| total_rec_pmcp         | 0.85     | 0.86        | 0.84           | 0.17     | 0.85        |            | 0.037  | -0.038      | 0.021           | 0.015                 | -0.048                 | -0.064 |         | 0.97        | 0.94            | 1             | 0.57              | 0.023      | 0.095     | 0.0051                 |     |
| last_pymnt_amnt        | 0.47     | 0.48        | 0.47           | 0.17     | 0.41        | 0.14       | 0.0086 | -0.013      | 0.028           | 0.0076                | 0.084                  | -0.033 | 0.13    | 0.51        | 0.5             | 0.57          | -1                | 0.065      | 0.13      | 0.007                  |     |
| issue_month            | 0.052    | 0.042       | 0.068          | 0.025    | 0.029       | 0.009      | 0.014  | -0.011      | 0.013           | 0.063                 | 0.22                   | -0.022 | 0.015   | 0.026       | 0.049           | 0.023         | 0.055             | 1          | -0.036    | 0.051                  | - 1 |
| issue_year             | 0.11     | 0.12        | 0.25           | 0.039    | 0.053       | 0.0086     | 0.093  | 0.0031      | 0.059           | 0.14                  | 0.75                   | 0.006  | 0 0093  | 0.11        | 0.22            | 0.095         | 0:13              | -0.036     | 18        | 0.56                   |     |
| funded_loan_amts_ratio | -0.11    | 0.038       | 0.19           | 0.02     | -0.05       | -0.031     | 0.033  | -0.016      | 0.046           | 0 16                  | 0.69                   | 0.0062 | -0.067  | 0.002       | 0.18            | 0.005         | 0.007             | 0.051      | 0.56      | 1                      |     |
|                        | ban_amit | funded_armt | Anded_amnt_inv | nf_rate  | installment | annual_inc | 6      | deling 2)rs | ing last Griths | mths_since_last_deinq | mths since last record | ph de  | led_low | total_pymot | total pymed inv | ptal rec pmcp | last_pyrroit_annt | gene mouth | sene hear | tunded_loan_amts_ratio |     |

Observations: Applicants who rent and have less annual income & applications with higher annual income and mortgage are more likely to have fully paid loan status.



Observation: Applicants with more annual income have higher rate of paid off.



### Summary

- Applicant with more Annual Income have paid off their loans fully.
- Applicant with more number of funded amount has paid off their loan fully.
- Applicant with home mortgage and high annual income have paid off their loans fully.
- Applicant with home ownership and less income have max number of fully paid off loans.
- Applicants with grade A, B,C are most likely to pay off the loan.
- Applicants with grade G is most likely to default the loan.
- Less number of installment has more number of fully paid loan status records.
- In each installment bin, it has been observed that fully paid loan status has a greater number of records wrt charged off status.
- Maximum number of defaulters are from the state of CA