Lending Club Case Study

Project group   :   Balaji S and Raghavendra R

Group ID   :   74992

Problem Statement

- The objective of this case study is to analyze risky loan applicants from the given data set using EDA

- A finance company which provides loans to customers, gets the customer loan application information to analyze and provide the validation of the loan approval

- Need to identify the right applicant to approve the loan application by checking the previous status.

Analysis

- We have used few of exploratory data analysis using below techniques to conclude the result obtained form the data
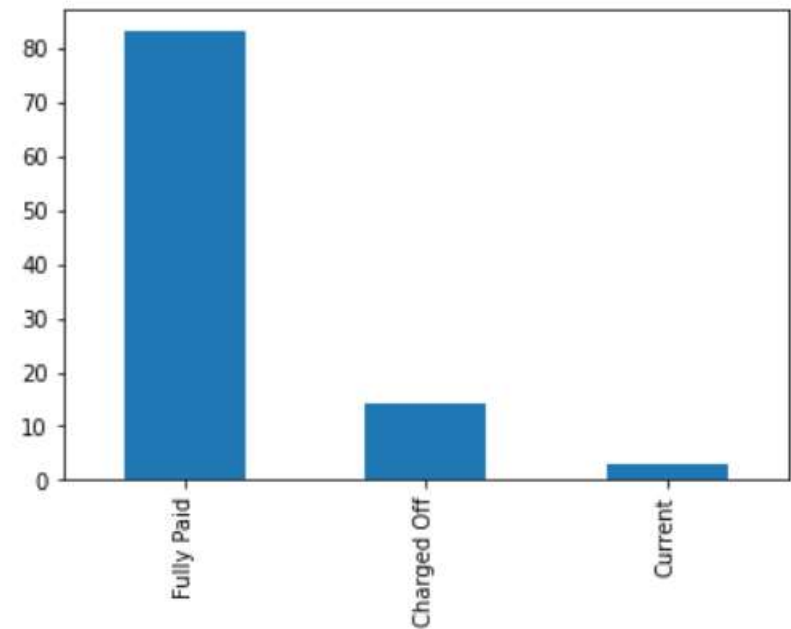
  - Univariate analysis
  - Bivariate analysis

Univariate analysis

- Loan status variable

    Fully paid % = 82.96
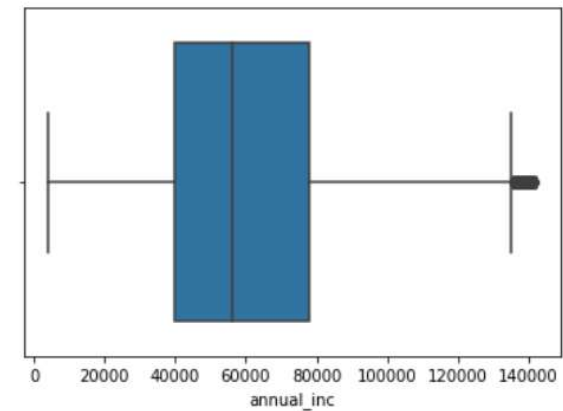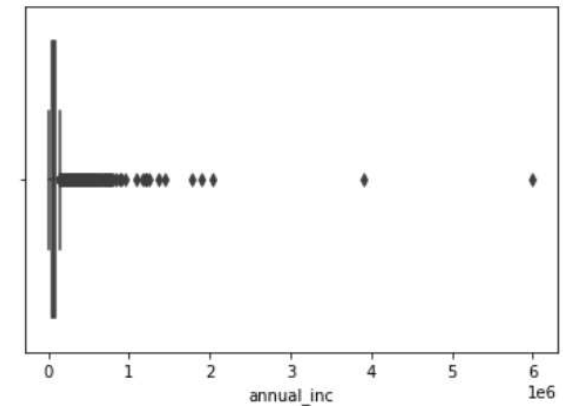
    Charged off % = 14.17

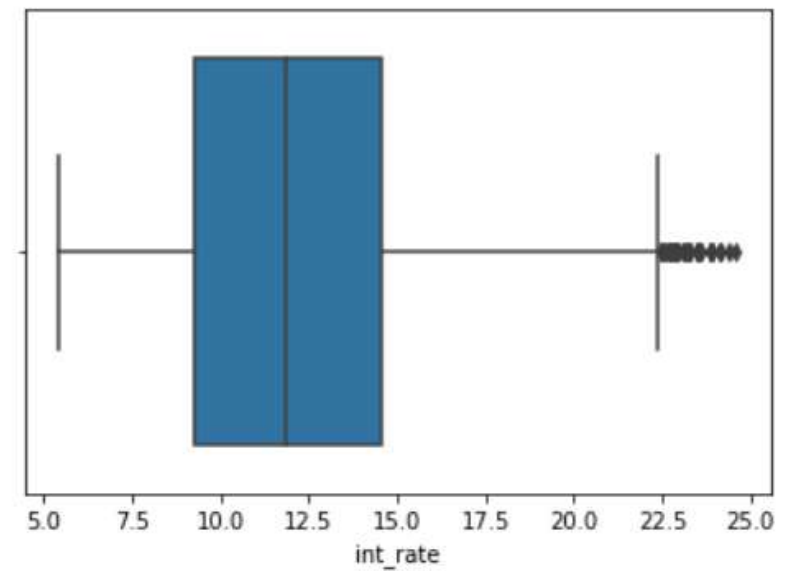   Current(running) = 2.87

Univariate analysis

- Outlier detection in annual income

  Maximum salary is 6000000.00 and
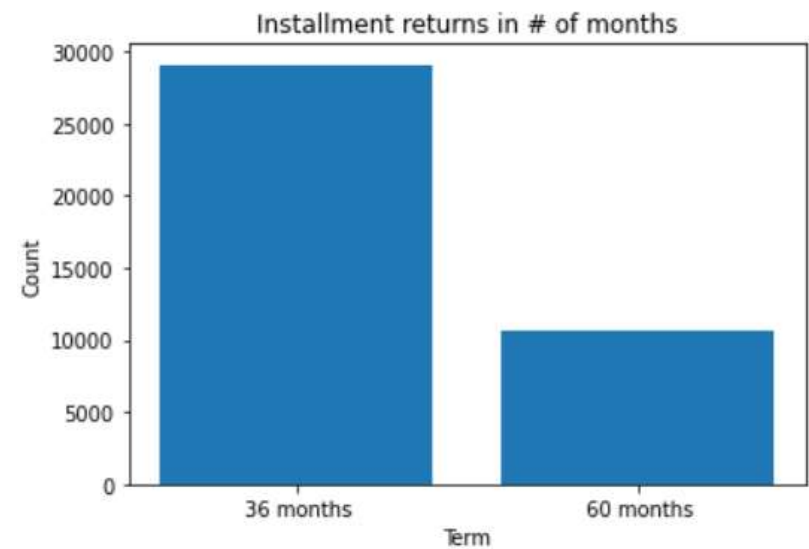 hence remove it from the dataset

Univariate analysis

- Interest rate

    Significantly increases   between
    15% to 25% after 75th quantile

Bivariate analysis

- Loan repayment term

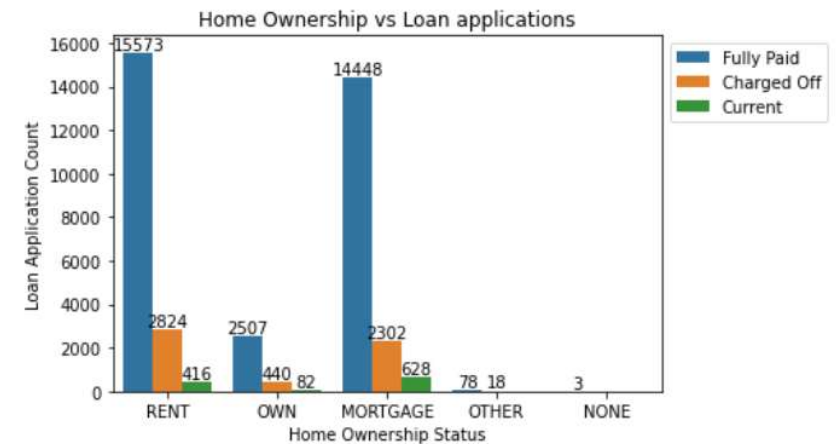   36 months:- 29096
    60 months:- 10621

Bivariate analysis

- House ownership vs Loan application

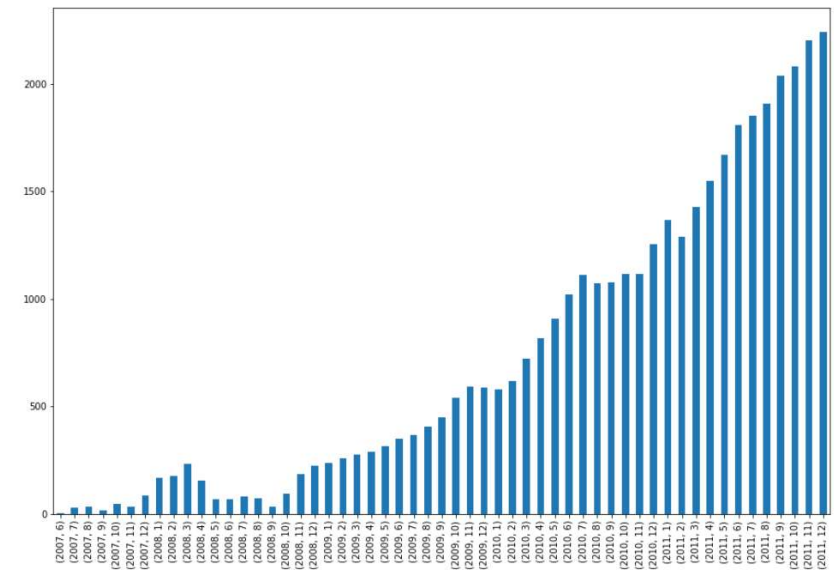 Applications received from rented house has higher charge off

  Applicants who have mortgaged property has higher charge off

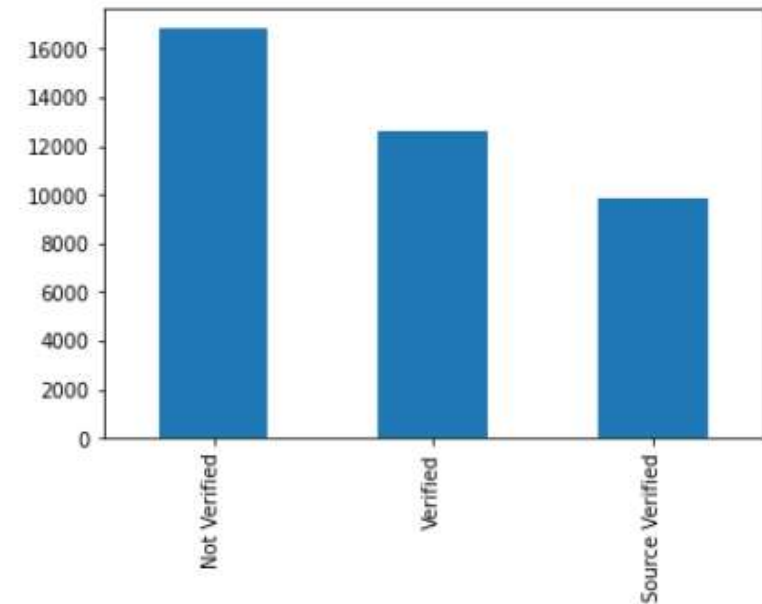# Bivariate analysis

- Loan disbursal YoY

  Year 2018-4 to 2018-10 has
  declining loan applications

Bivariate analysis

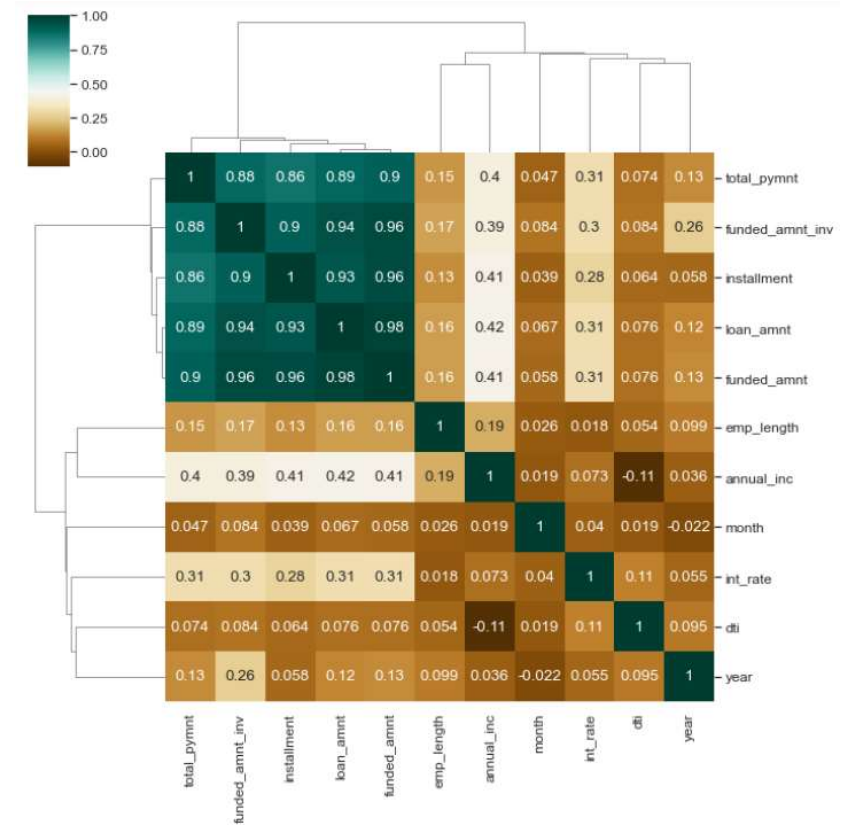- Income verification status vs Loan payment

  Income source verified and Not verified has high risk of Charge off

Bivariate analysis

- Heatmap with Dendogram to show closeness among numerical variables
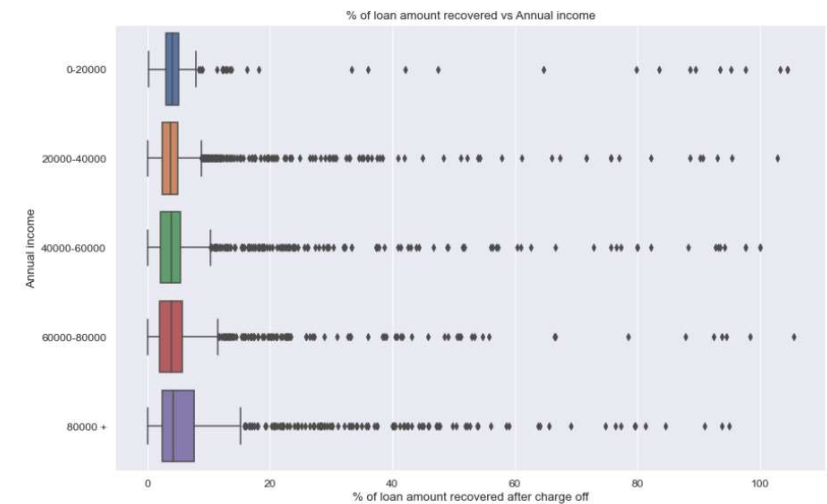
    Annual income with dti is negatively correlated.

Bivariate analysis

- Recovery vs Annual income

    There is slight increase in
    recovery percentage with
    increase in annual income

Summary

- Applications received from rented house has higher charge off
- Applicants who have mortgaged property has higher charge off
- Year 2018-4 to 2018-10 has declining loan applications
- Income source verified and Not verified has high risk of Charge off
- Annual income with dti is negatively correlated
- Positive correlation between annual income and employment years
- There is slight increase in recovery percentage with increase in annual income
- Annual income 0-20000 has higher charge off