LENDING CLUB CASE STUDY

Final Submission

Case Study Group:

- Surajit Khan (Group Facilitator)
- Shamanth B C

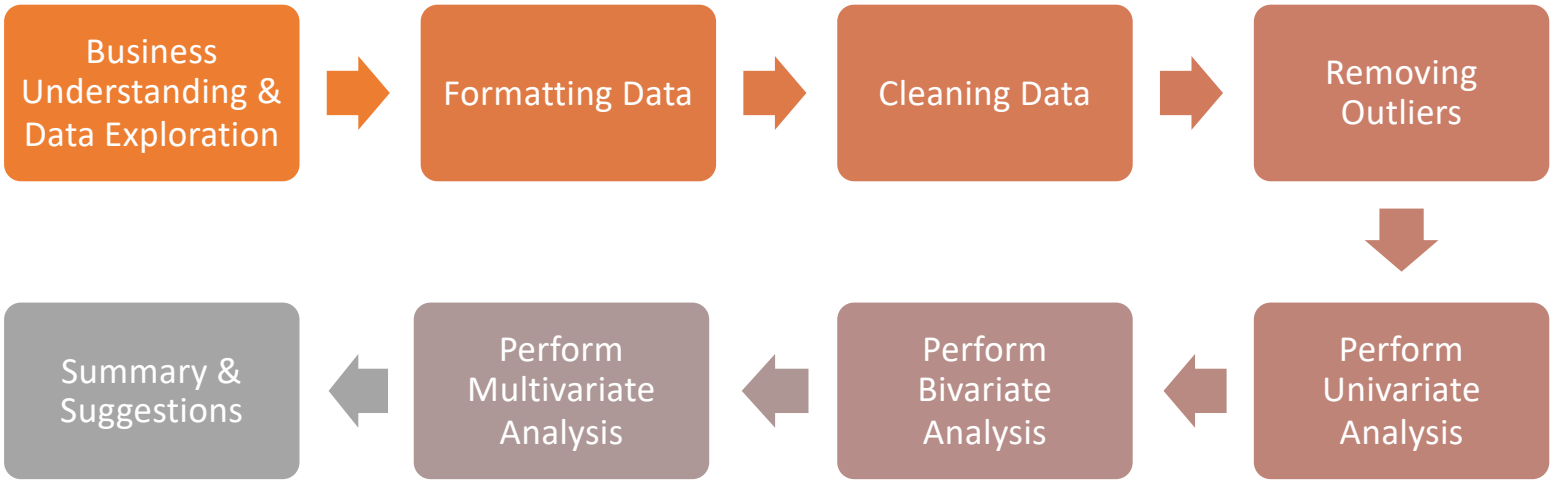
Problem Statement :

You work for a consumer finance company which specializes in lending various types of loans to urban customers. When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision:

- If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company.
- If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company
- The data given below contains the information about past loan applicants and whether they 'defaulted' or not.
- The aim is to identify patterns which indicate if a person is likely to default, which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc.

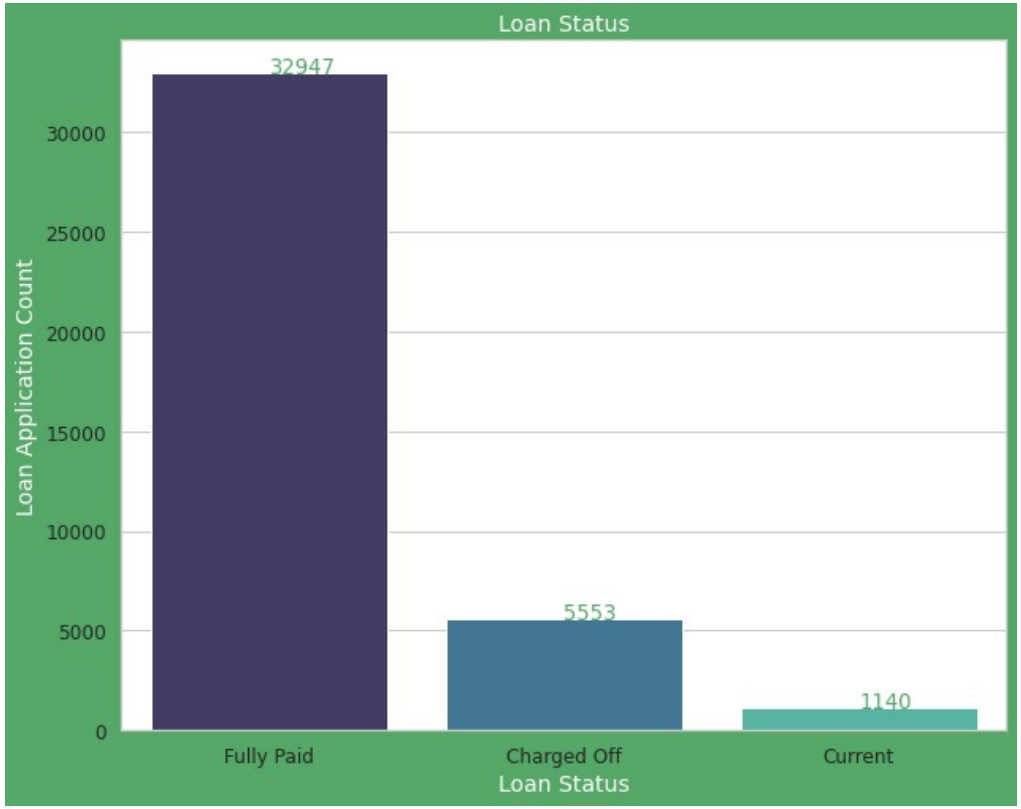
Overall Approach of the Analysis :

- Like most other lending companies, lending loans to 'risky' applicants is the largest source of financial loss (called credit loss).
- If one is able to identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss. Identification of such applicants using EDA is the aim of this case study.



Data preparation & cleaning

- All the columns having only or more than 30 percent of null values and not useful for analysis have been deleted(ex: tot_hi_cred_lim, total_bal_ex_mort, next_pymnt_d etc)
- All columns having same data or not necessary have been deleted (ex: Application type = individual)
- Unnecessary string value attached to int or float type data have been removed for ease of analysis (ex: term, int rate etc)
- Data types have been changed accordingly for analysis(int rate, term etc)
- Columns have few null values have been imputed with necessary methods such as Median/Mode/mean or filled as missing (ex: emp_title, emp_length_years)
- Date manipulation is done wherever necessary(ex: issue_d)
- Deleted columns with same correlation(ex: out_prncp & out_prncp_inv)
- Deleted few rows of last pymnt d and last credit pull d in case of null entry as total number of rows are only few



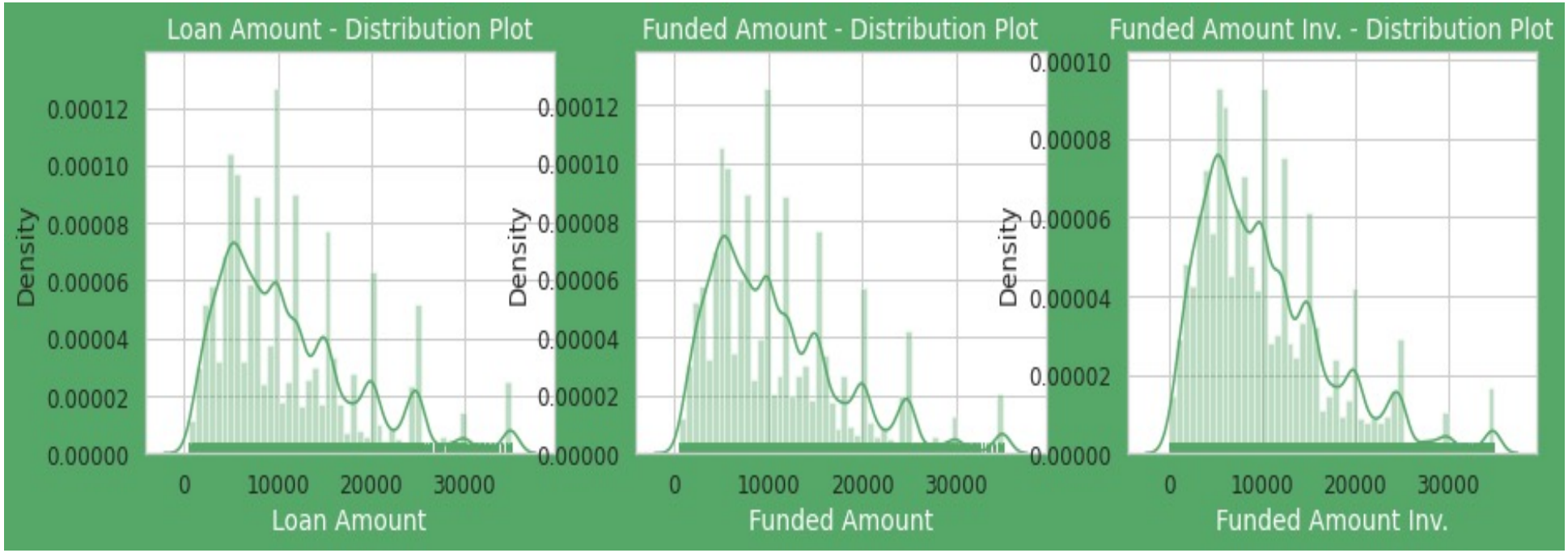
Plot and Data shows that :

- 14% loans were charged off out of total loan issued.
- 83% loans were fully paid out of total loan issued.

- Fully Paid 83.12%
- Charged Off 14 %
- Current 2.87 %

Lets analyze the loan status those who are charged off due to various other factors

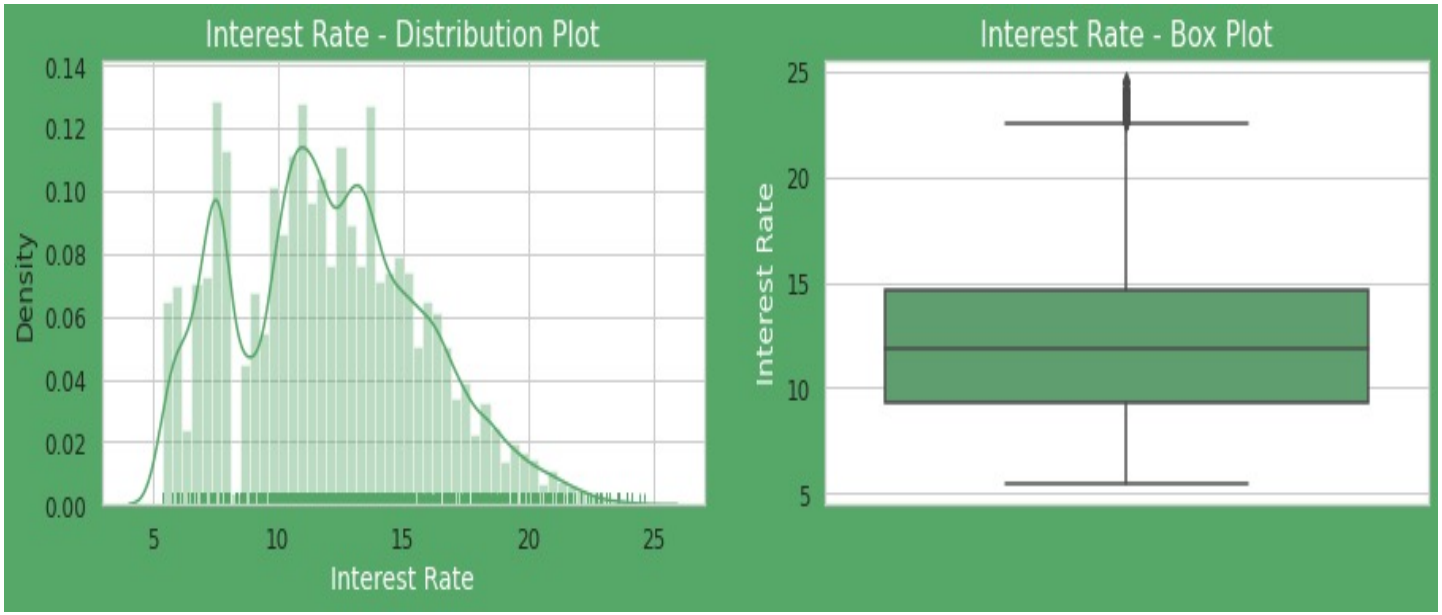
Distribution plot:



Observation:

Distribution of amounts for all three types look very similar.

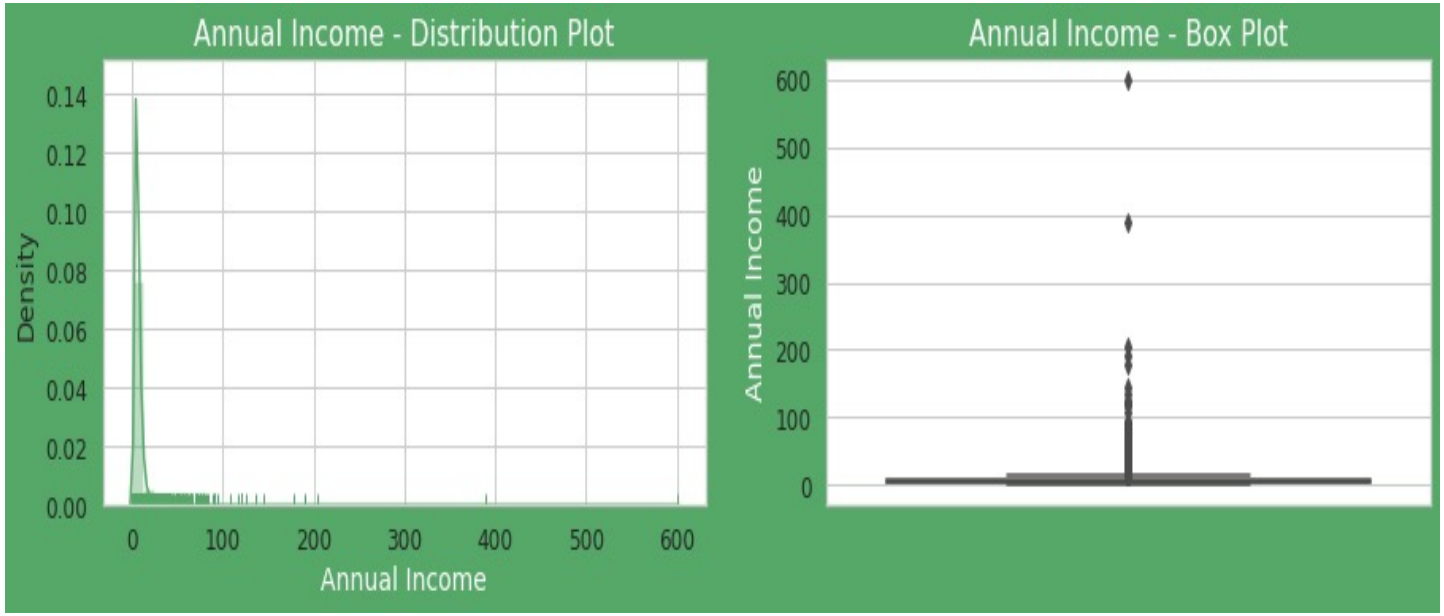
Distribution plot:



Observation

most of the interest rate is 9-15

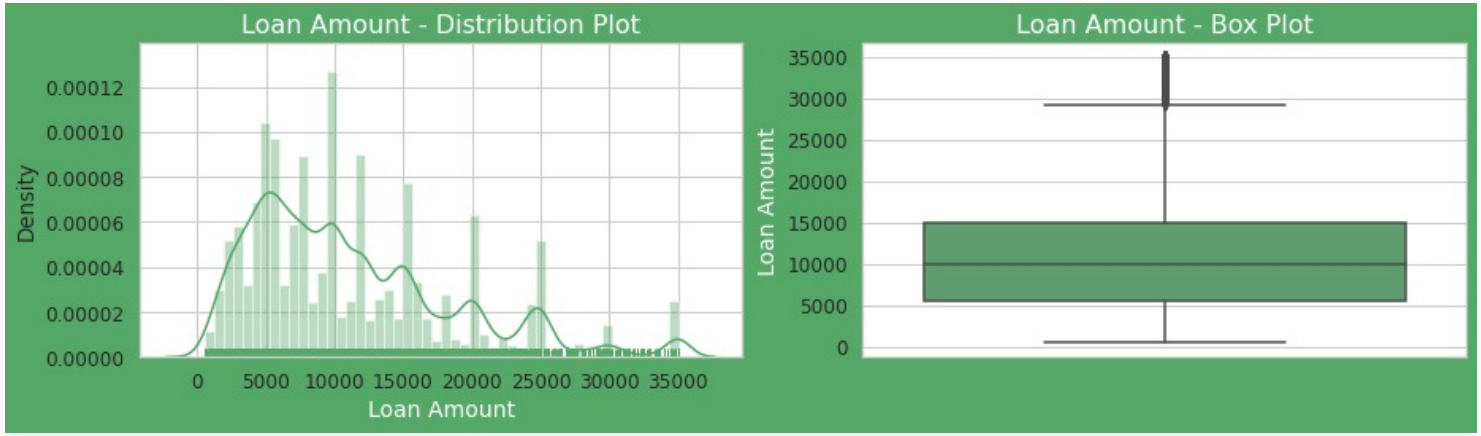
Distribution plot:



Observation:

Annual income is 4-8L range

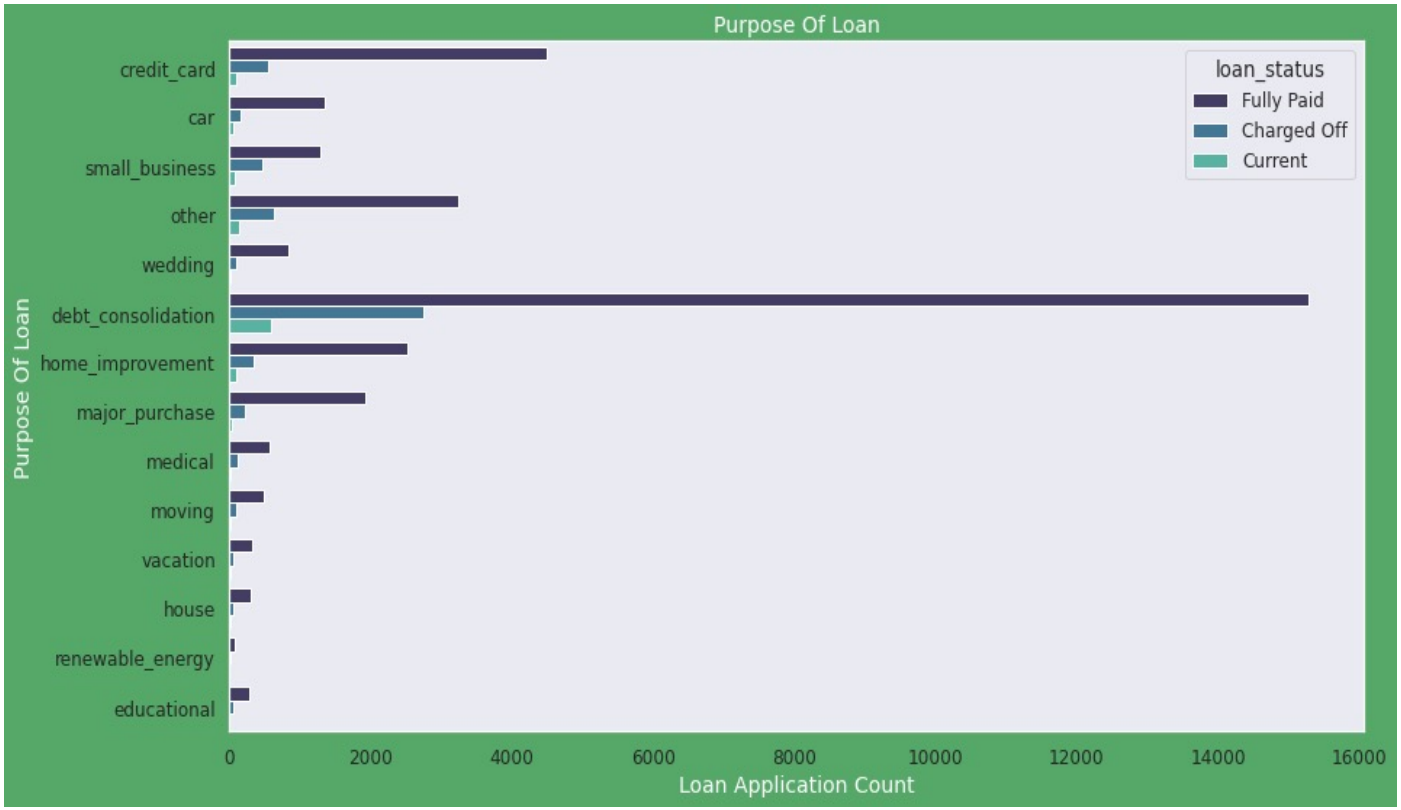
Distribution plot:



Observations :

Most of the Loan amounts are in range of 5000 - 15000

Purpose of Loans:

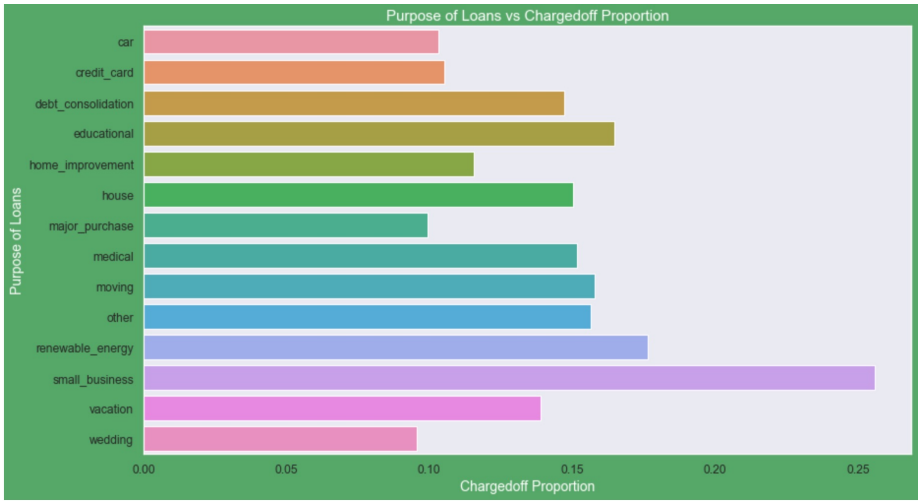


Purpose of Loans:

- Most of the loans were taken for the purpose of debt consolidation & paying credit card bill.
- Number of charged off count also high too for these loans.

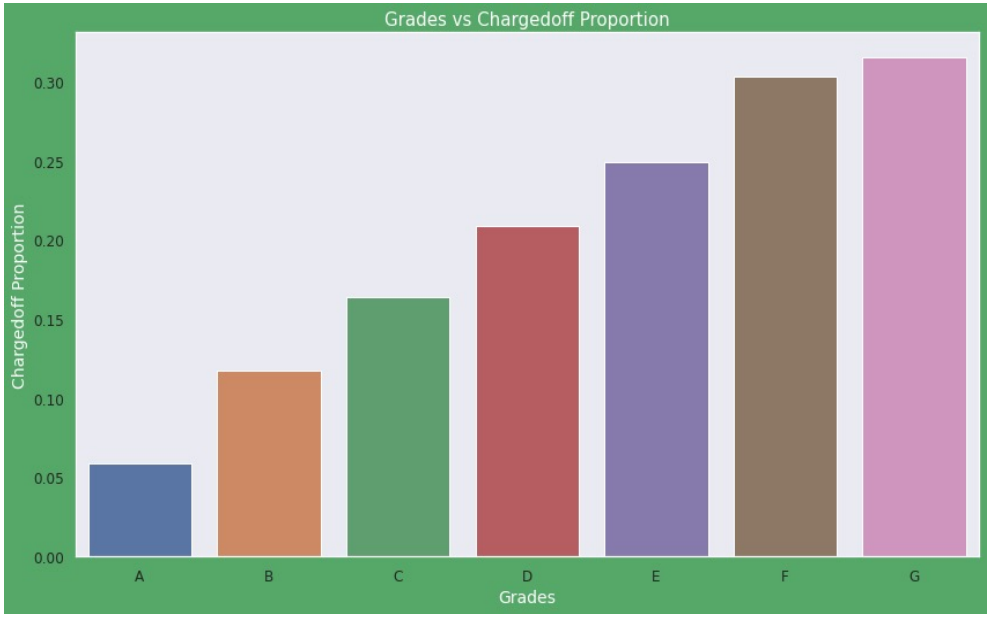
Loan Purpose Percentage:

- **debt_consolidation 46.95%**
- **credit_card 12.93%**
- **other 10.04%**
- **home_improvement 7.95 %**
- **major_purchase 5.5%**
- **small_business 4.5 %**



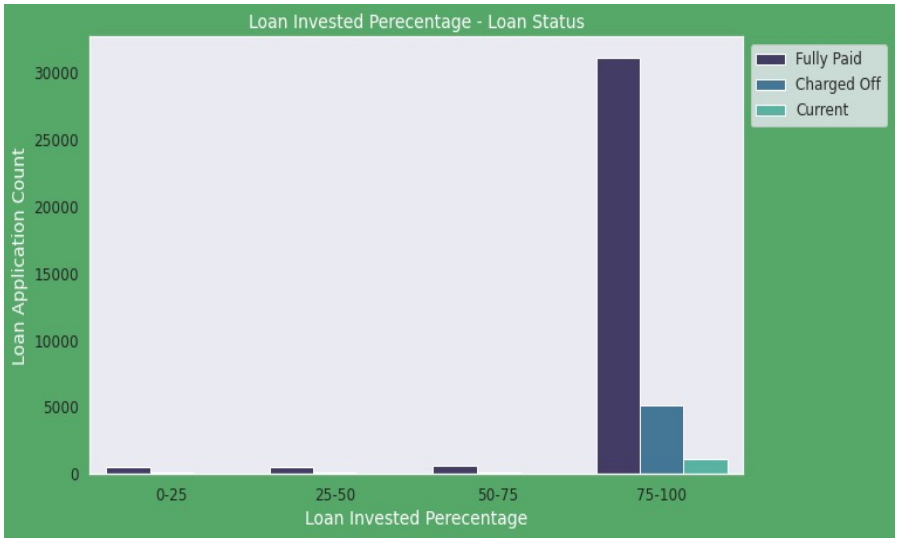
Purpose of Loans v/s Loan charged off Proportion :

- Small Business applicants have high chances of getting charged off.
- Renewable energy & Educational purpose have changed off proportion high as compare to other categories.



Grades v/s Loan charged off Proportion:

- Grade "A" has very less chances of charged off.
- Grade "F" and "G" have very high chances of charged off.
- Chances of charged of is increasing with grade moving from "A" towards "G"

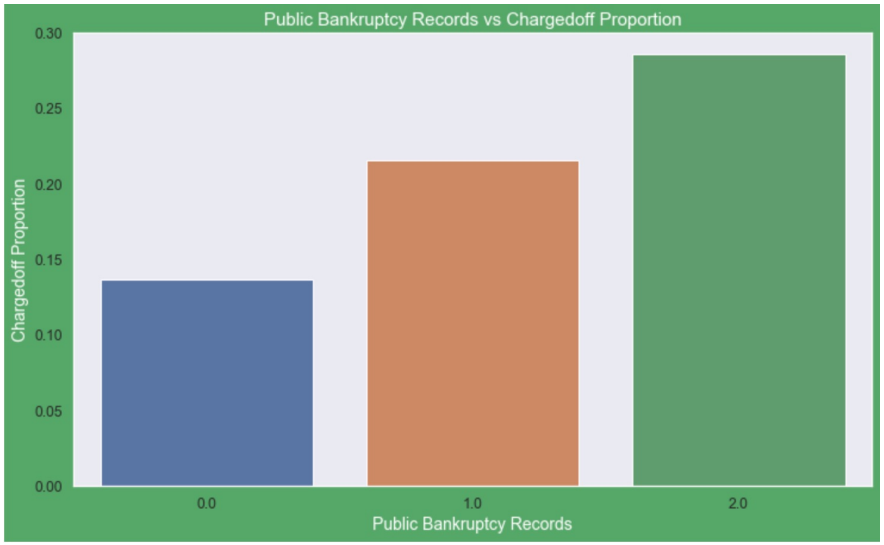


Loan invested percentage v/s Loan application count :

- Loan invested percentage of 75 to 100 are more likely to get charged off

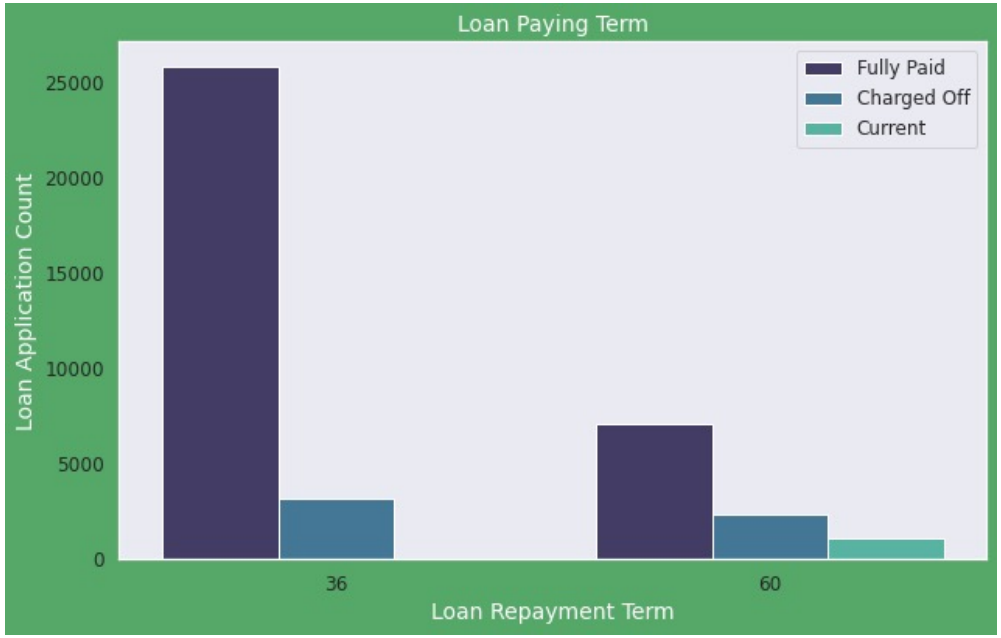
Public Bankruptcy Records v/s Loan charged off Proportion :

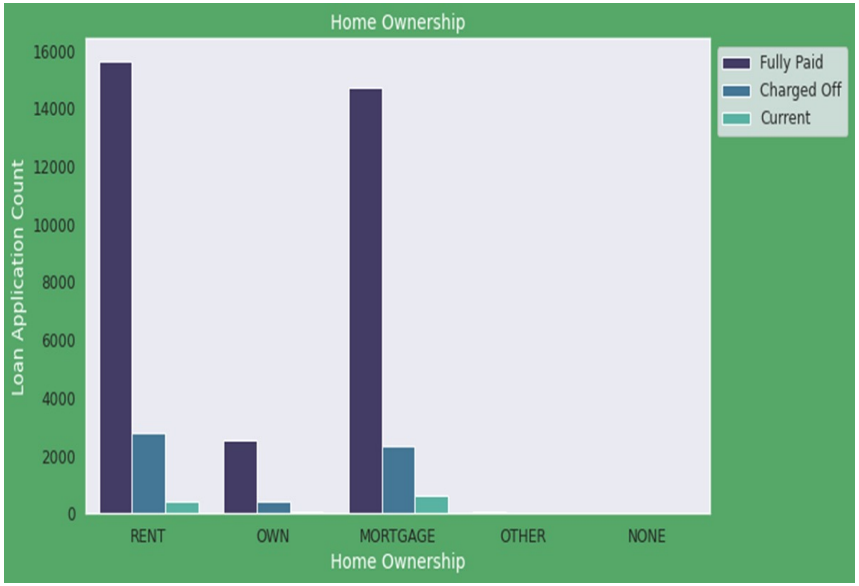
- Those with public bankruptcy record of 1 or 2 have higher chance of getting charged off
- Those with public bankruptcy record of 3 and 4 can be ignored as they are few in number



Loan application count v/s Home ownership:

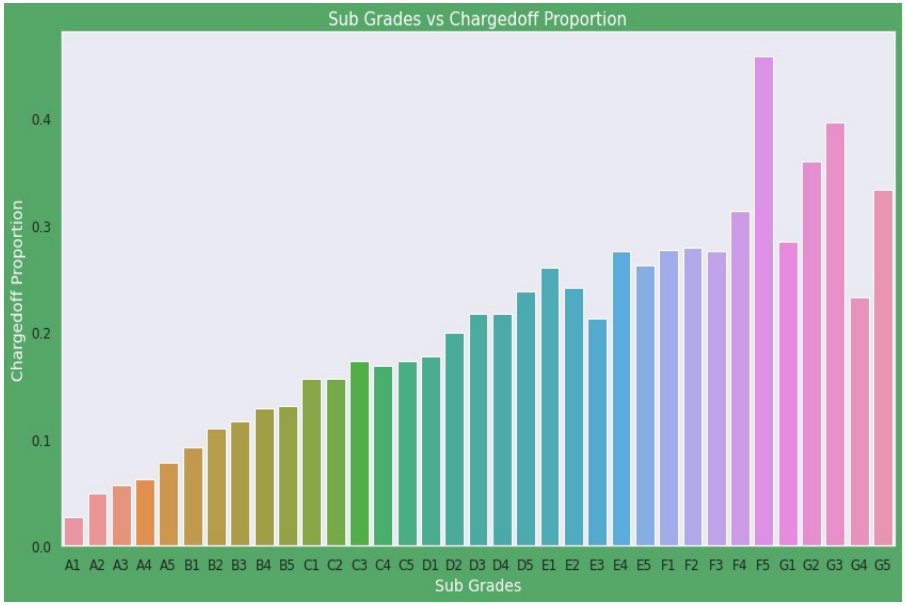
- People living in rented house or mortgaged their home have a higher chance of getting charged off than people owning a house





Loan repayment term v/s Loan application count

- People taking loan to repay in 60 months are more likely to get charged off than people who opted for 36 months

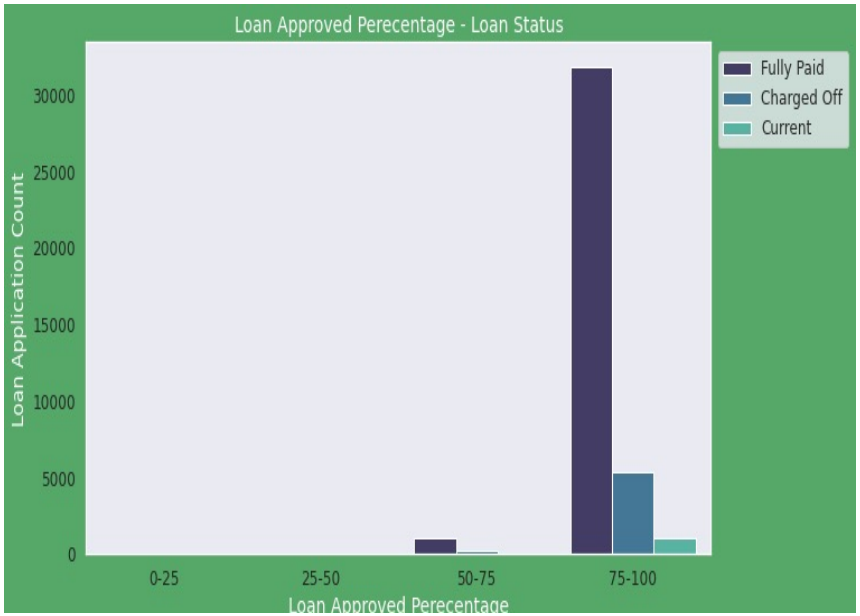


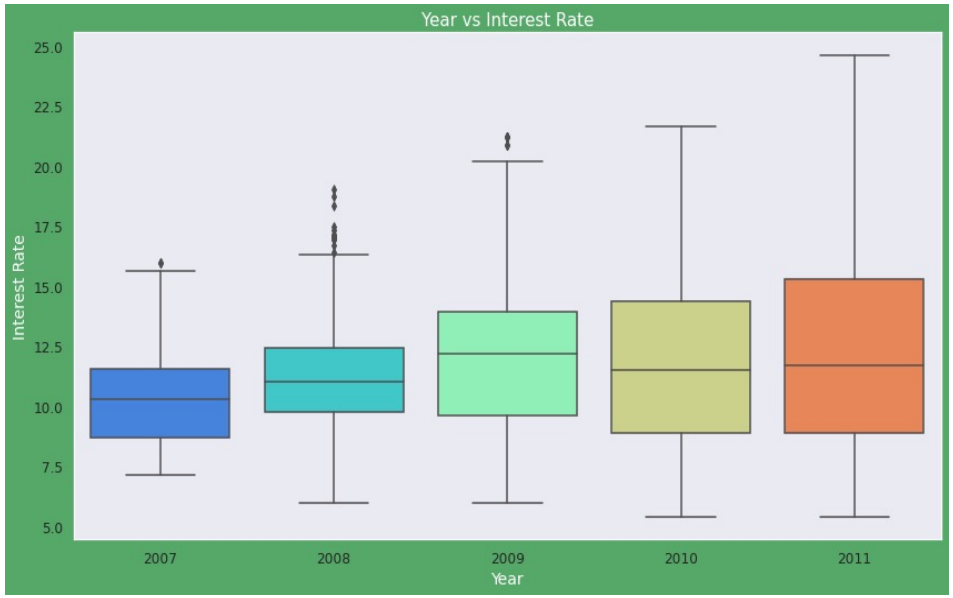
Sub grades v/s Chargeoff proportion

- .Those who are 4 years of work experience have high chances of getting charged off.
- This seems random and can be ignored

Loan approved percentage v/s Loan application count:

- 75 to 100 percentage group are more likely to get charged off



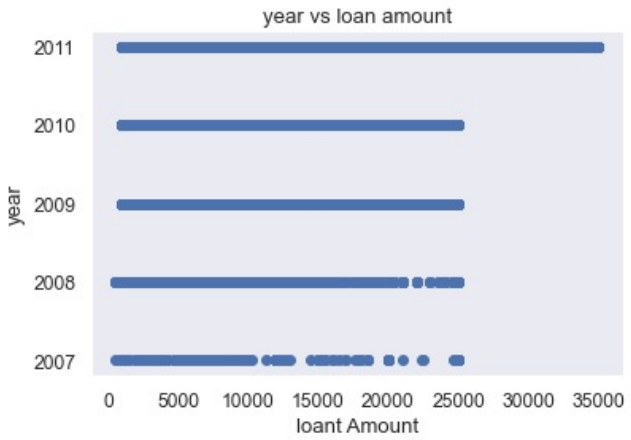


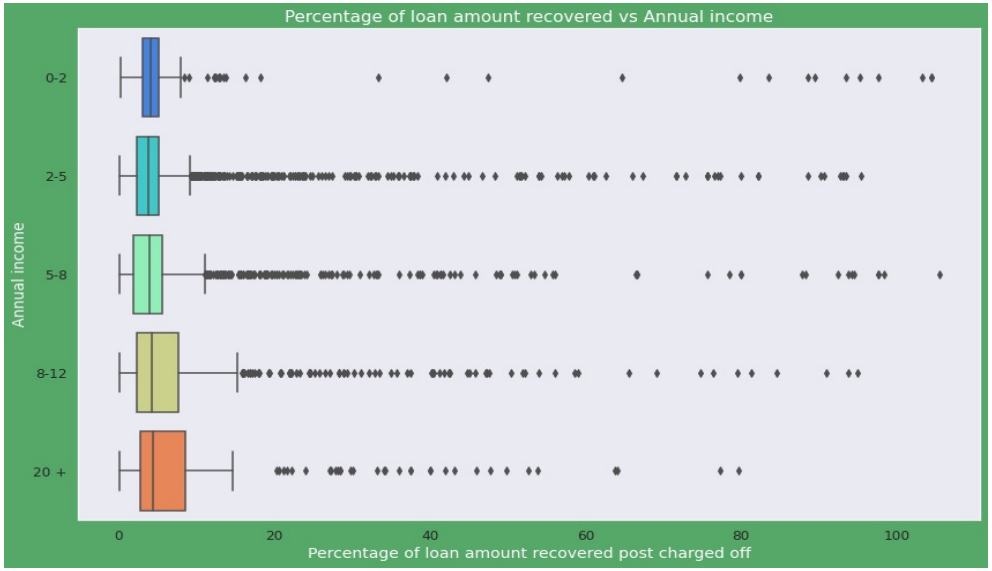
Year v/s Interest Rate:

- The interest rate is increasing along with the Year but after 2009 it is stable.

Year vs Loan amount:

- The loan amount is increasing with year





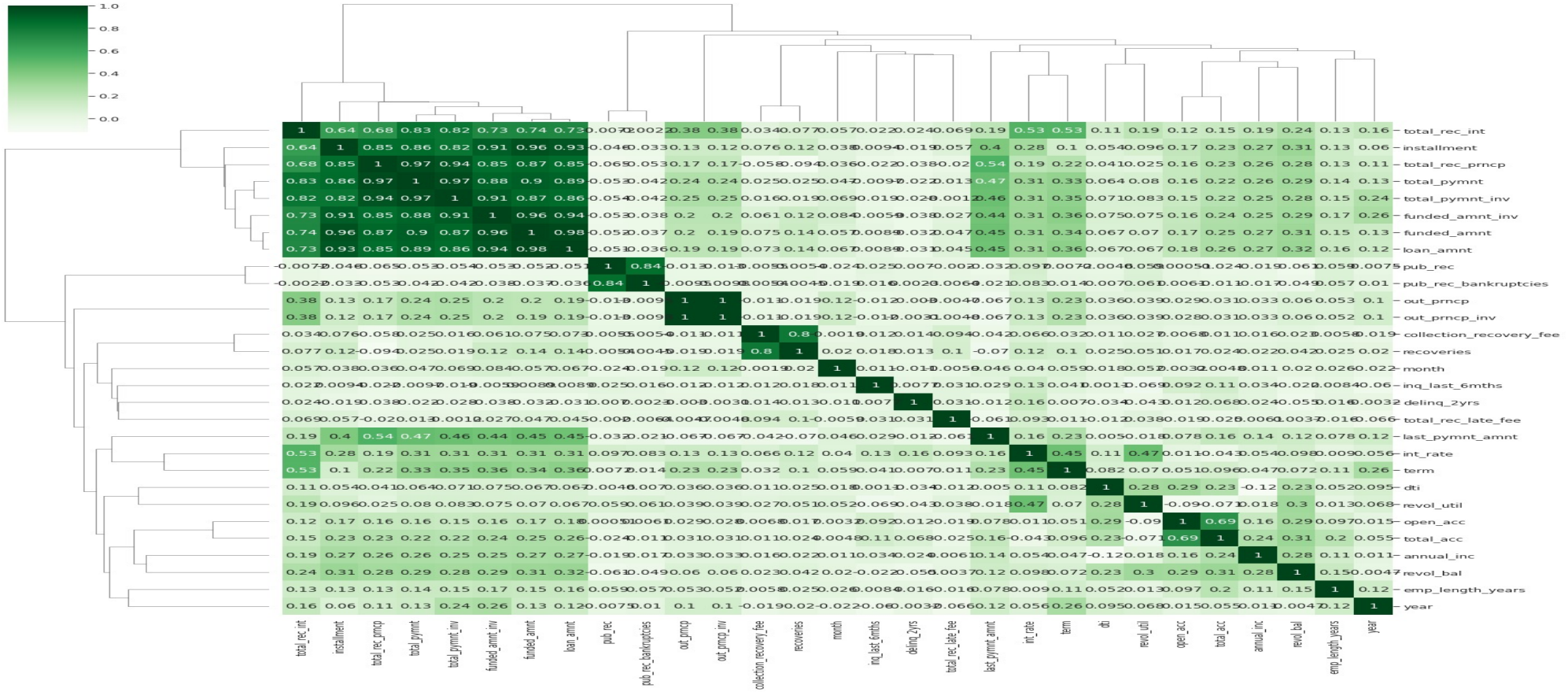
Loan Amount Recovered v/s Annual Income:

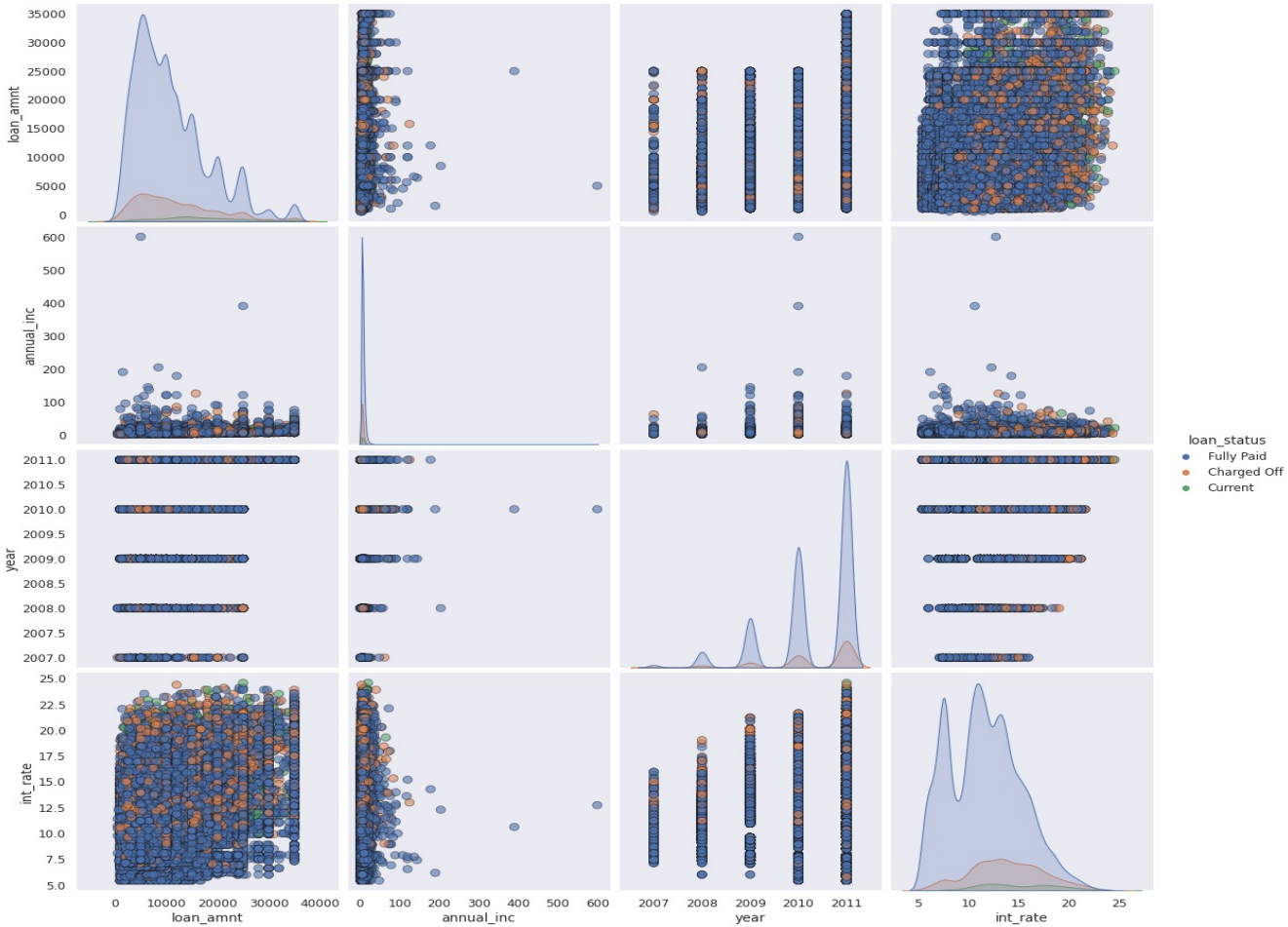
- As the annual income decreases, the charge off is getting higher

Grade v/s Applicant's Annual Income:

- From this we can conclude that the ones getting charged for delayed or no repayment of loan have lower annual incomes than the ones who has fully paid for each and every grade (i.e. at same interest range).

Bivariante Analysis -Correlation Matrix- Quantitative Variables





Observing Loan Amount, Annual Income, Year and Interest Rate :

Observing Loan Amount, Annual Income, Year when loan was taken, Interest Rate at which loan was taken to each other :

- Higher the interest rate , Higher charged off ratio
- Higher the annual income, Higher the loan amount slightly.
- Increase in number of charged off with increase in year.
- Interest rate is increasing with loan amount increase

Conclusion:

- Small Business Applicants have high chances of getting charged off.
- Charged off proportion increases with grades moving from “A” towards “G”.
- Charged off proportion increases as Interest Rate Increases.
- Higher the public bankruptcy record greater the charged-off proportion.
- We see more charge-off when 75-100% loans approved and 75-100% amount is invested.