

LENDING CLUB CASE STUDY

Name: Tuhin Kumar Dutta

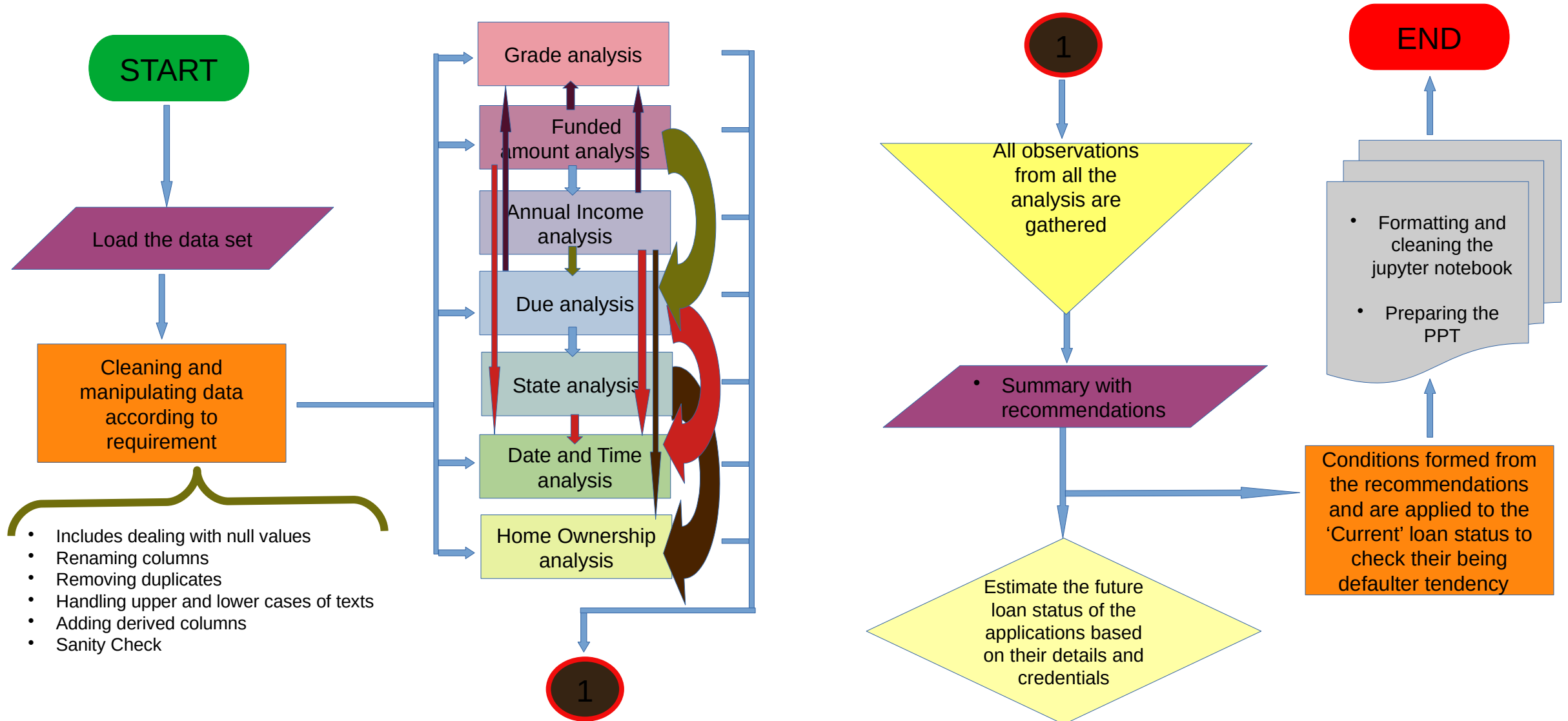
Abstract

Major objective is to present a general trend of loan applicants and formulate an efficient way to the Lending Club to let them sanction loan and thus assisting them in gaining optimum profits. Data provided by Lending Club is in form of a data set, which contains all details of applicants, loan characteristics and customer behavior variables.

In order to achieve the goal, the programming language, Python has extensively been used as the prime tool during the entire process of analysis along with an IDE called Jupyter Notebook and other some external python libraries like numpy, pandas, matplotlib and seaborn.

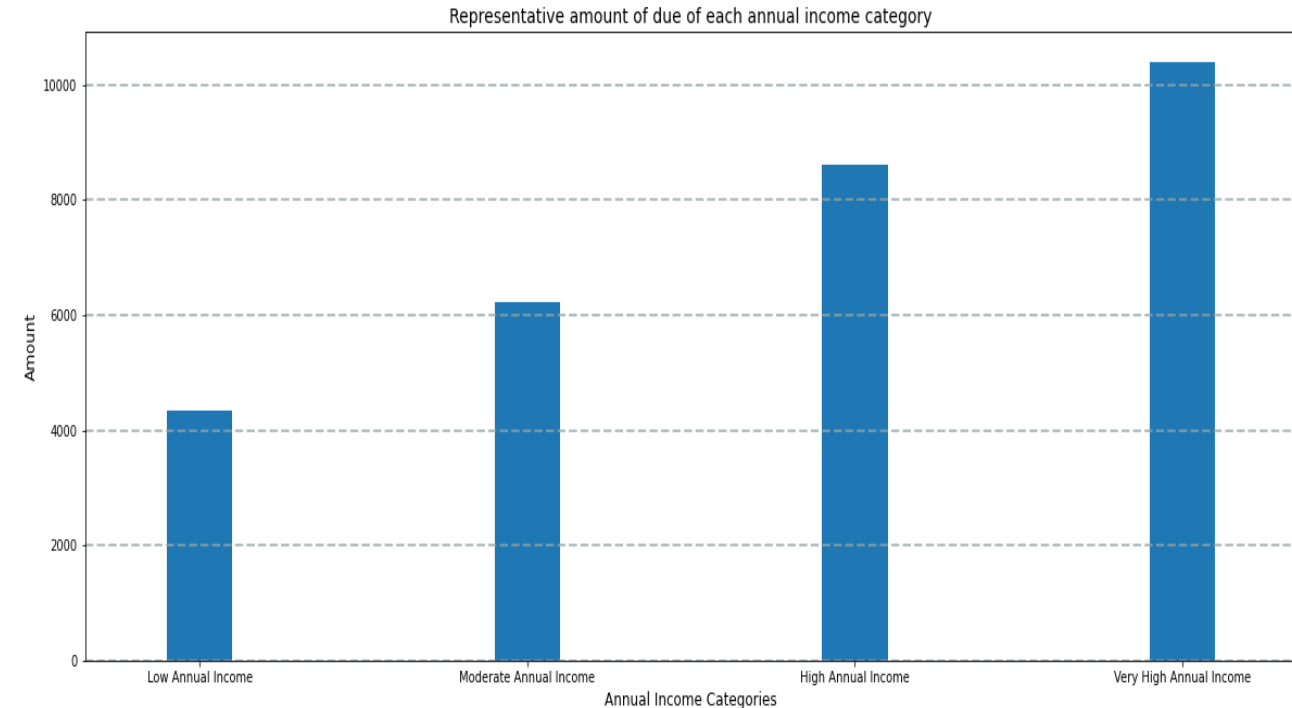
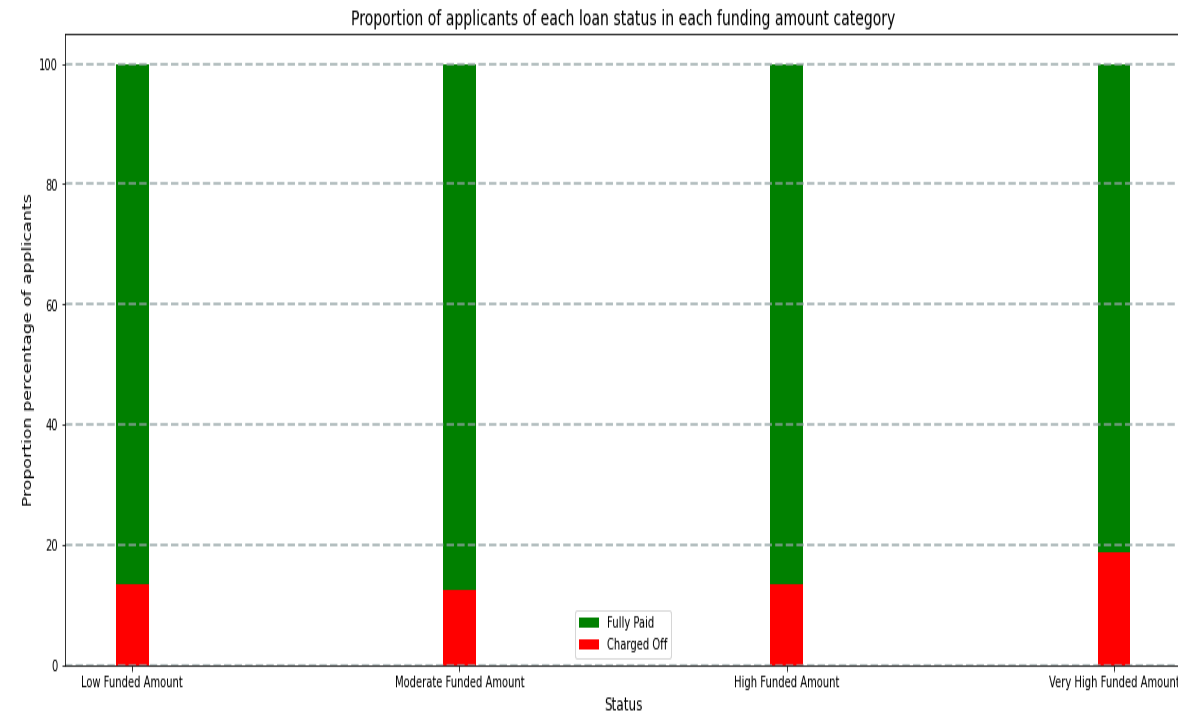
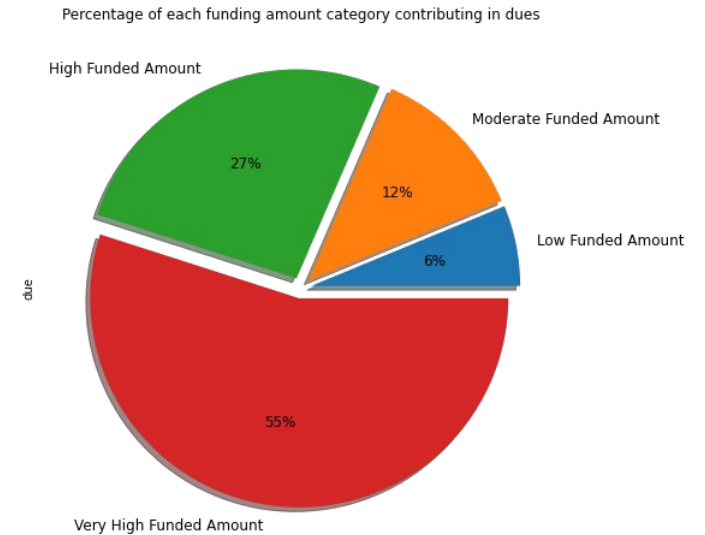
The approach begins with loading and selecting the relevant features of the data set which would contain all the necessary information. This data set then undergoes various processes to attain a clean and sane state. Various analysis based on funded amount, annual income, dues, state, etc. at various stages are performed to figure out the overall tendency and pattern of funding and applicants' loan status. Bar plot, stacked bar plot, line plot, box plot, strip plot, pie chart and heat-map have been used as part of data visualization which ultimately helped in presenting the gist and concluding that funding to some applicants could fetch profit whereas others may cause to face loss.

Problem solving methodology



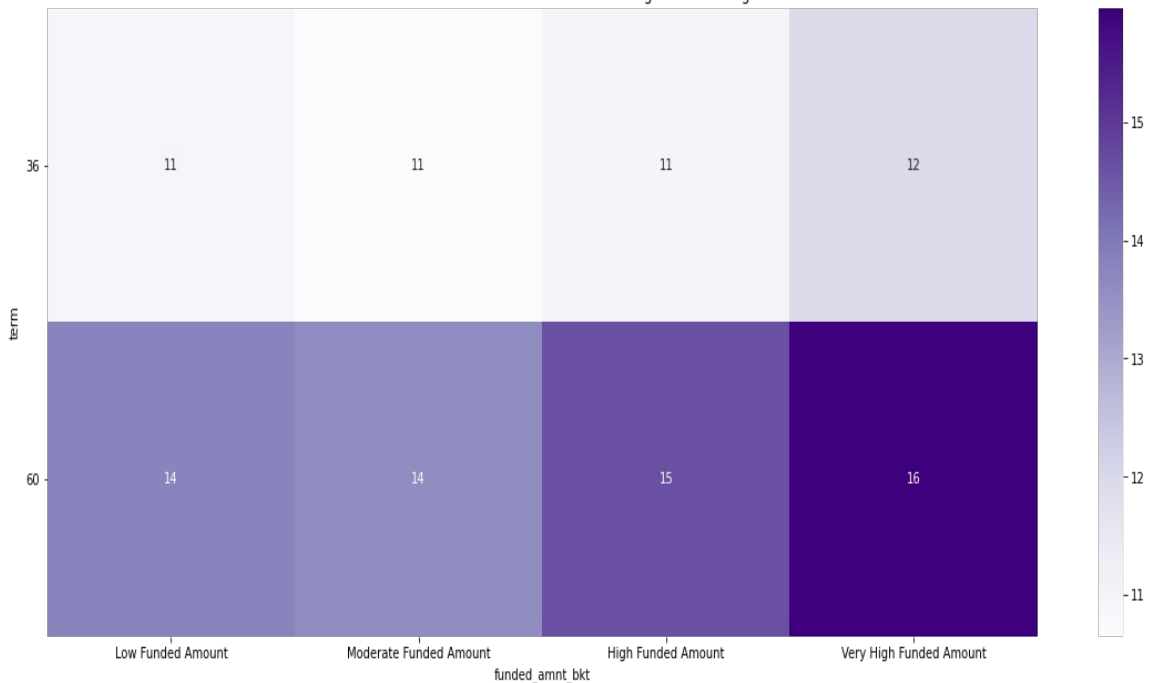
Funding Amount Analysis

- Around 20% of very high loan receivers tend to be defaulters.
- Very High funded amounts category contributed to around 55% of the dues.
- A Very High annual income applicant causes monetary loss around as twice as a low annual income applicant since their funding amounts are quite high.

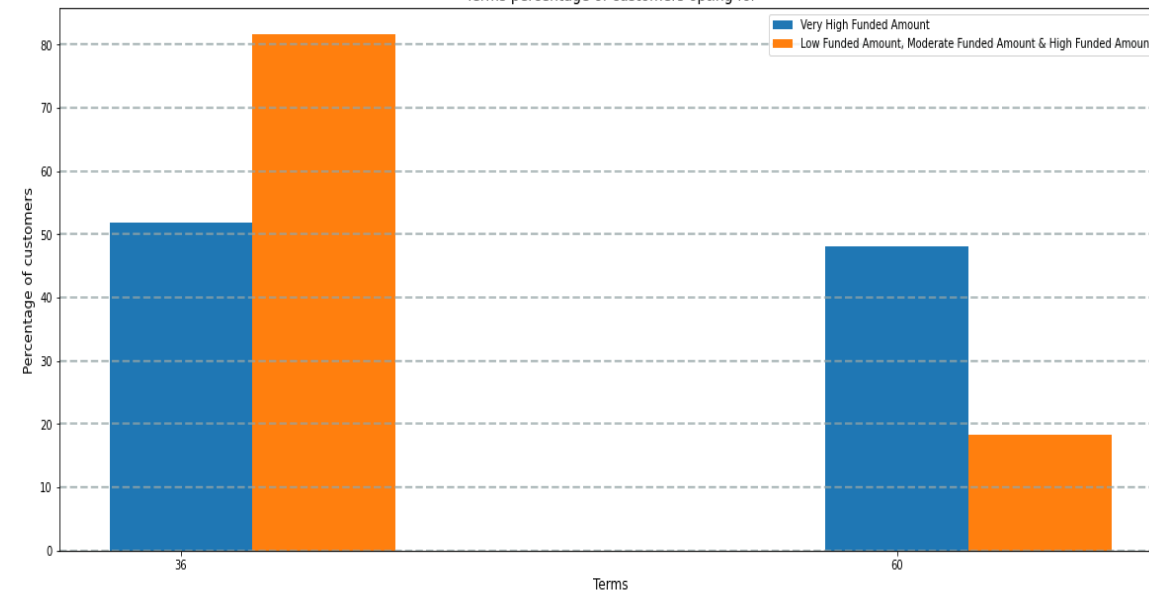


- Very high funded amounts are least in count but maximum in amount
- Around 80% low, moderate and highly funded customers tend to pay more payments than the installment amount but around 75% very highly funded customers can pay more, which is 5 percent less.
- Customers with very high funding amounts go for a plan of 60 months than the customers with lower amount of funding.
- Since the interest amount for 60 months term and very high funded amount is comparatively higher than other amounts' and term plans' interest rate, customers can not pay more than the fixed installment amount and finally may end up being charged off.

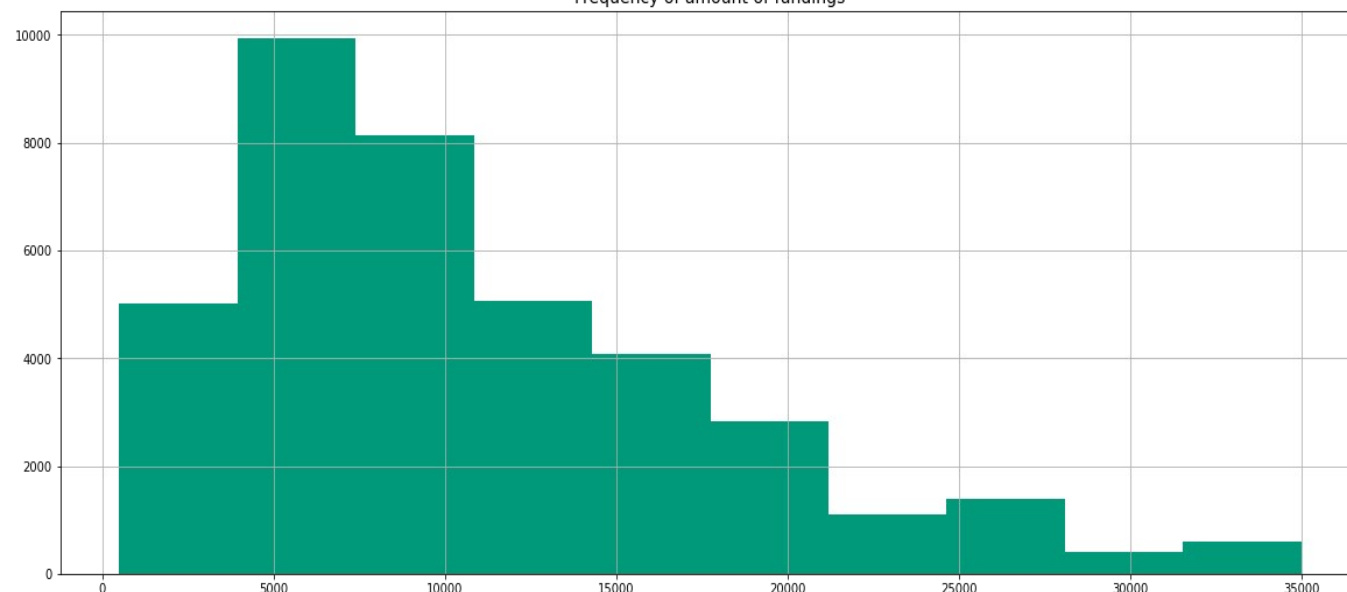
Distribution of interest rates with terms and funding amount categories



Terms percentage of customers opting for



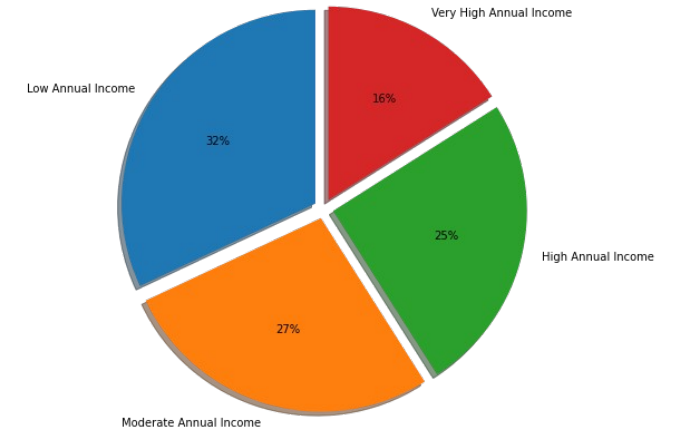
Frequency of amount of fundings



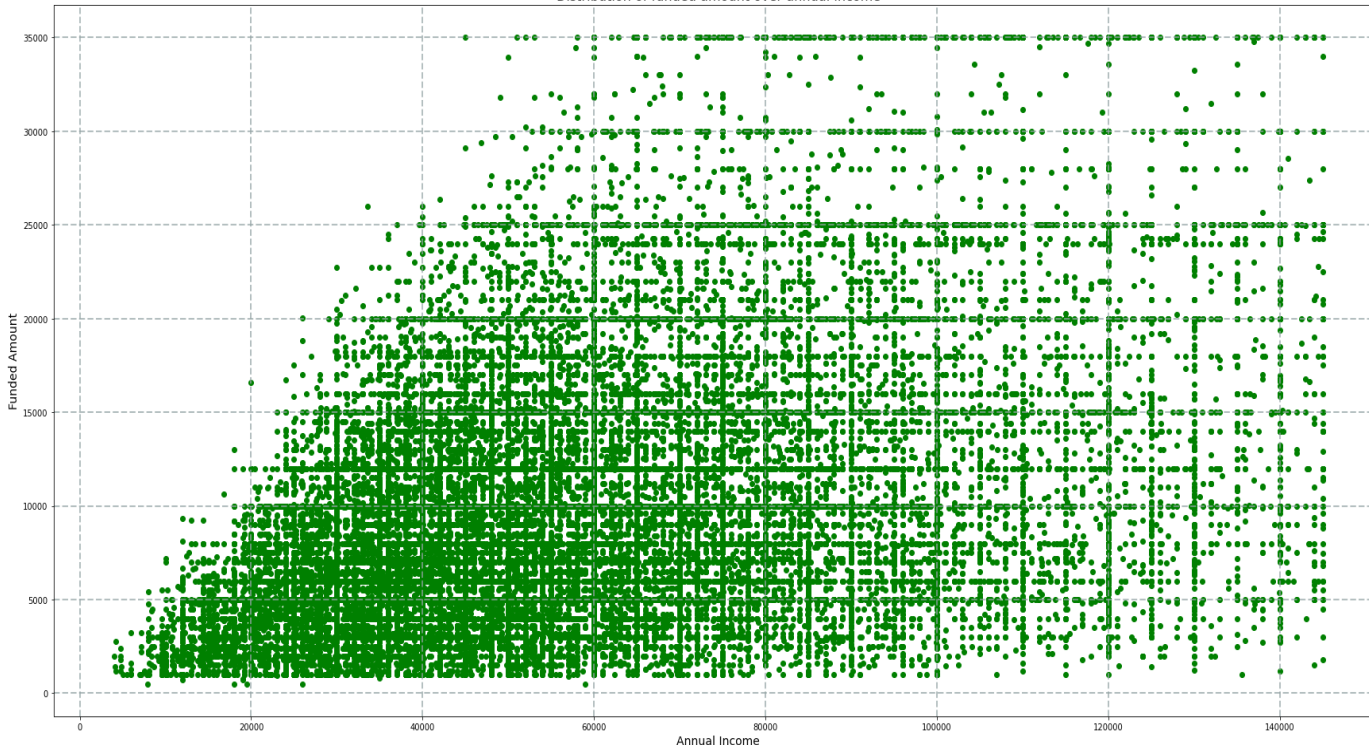
Annual Income Analysis

- Higher annual income groups definitely received larger amounts.
- High annual income group applicants are less likely to be defaulter but have caused substantial financial loss
- In case of being a defaulter, a low annual income applicant tends to cause a loss of around 4000 USD, whereas a very high annual income applicant tends to cause a loss of around 10,000 USD

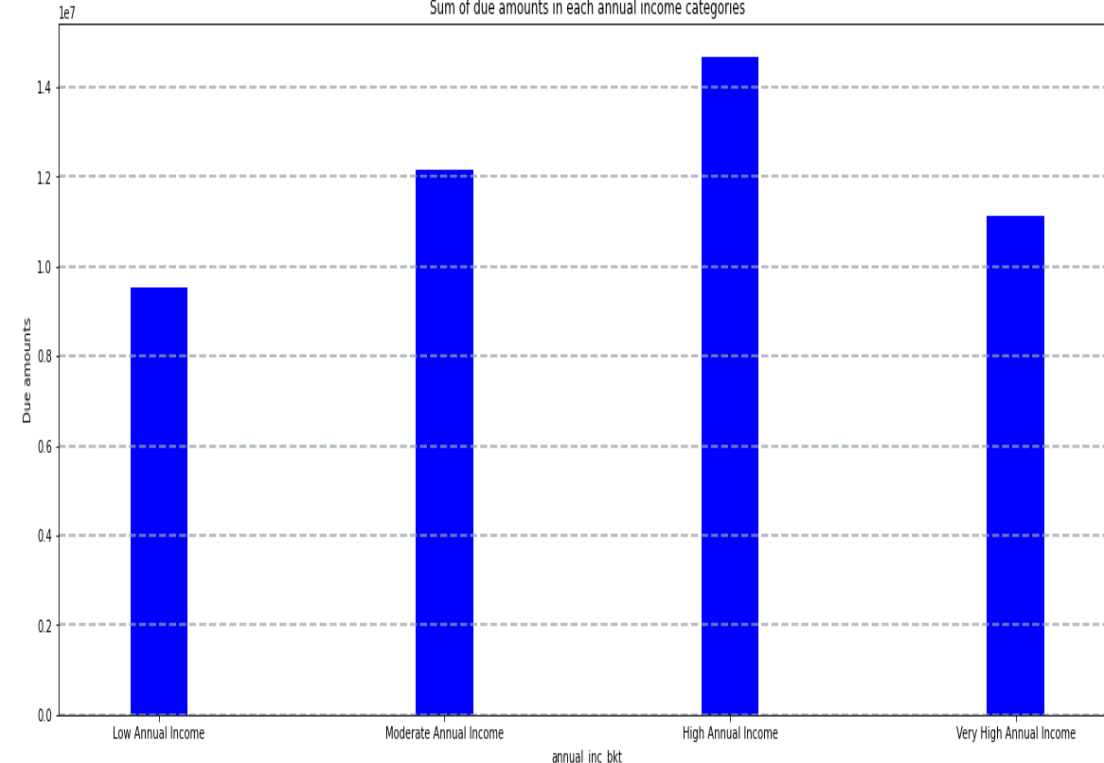
Percentage of count each annual income group contributes in being charged off



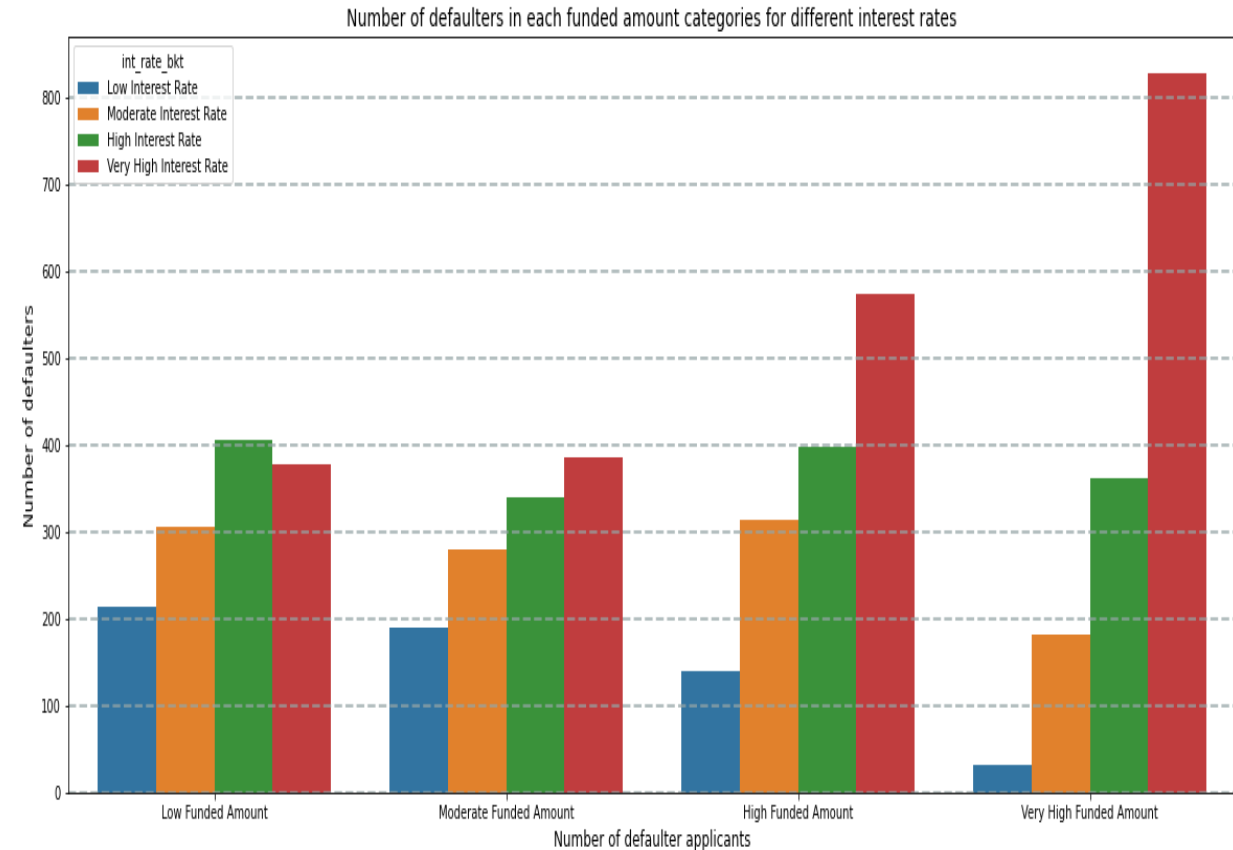
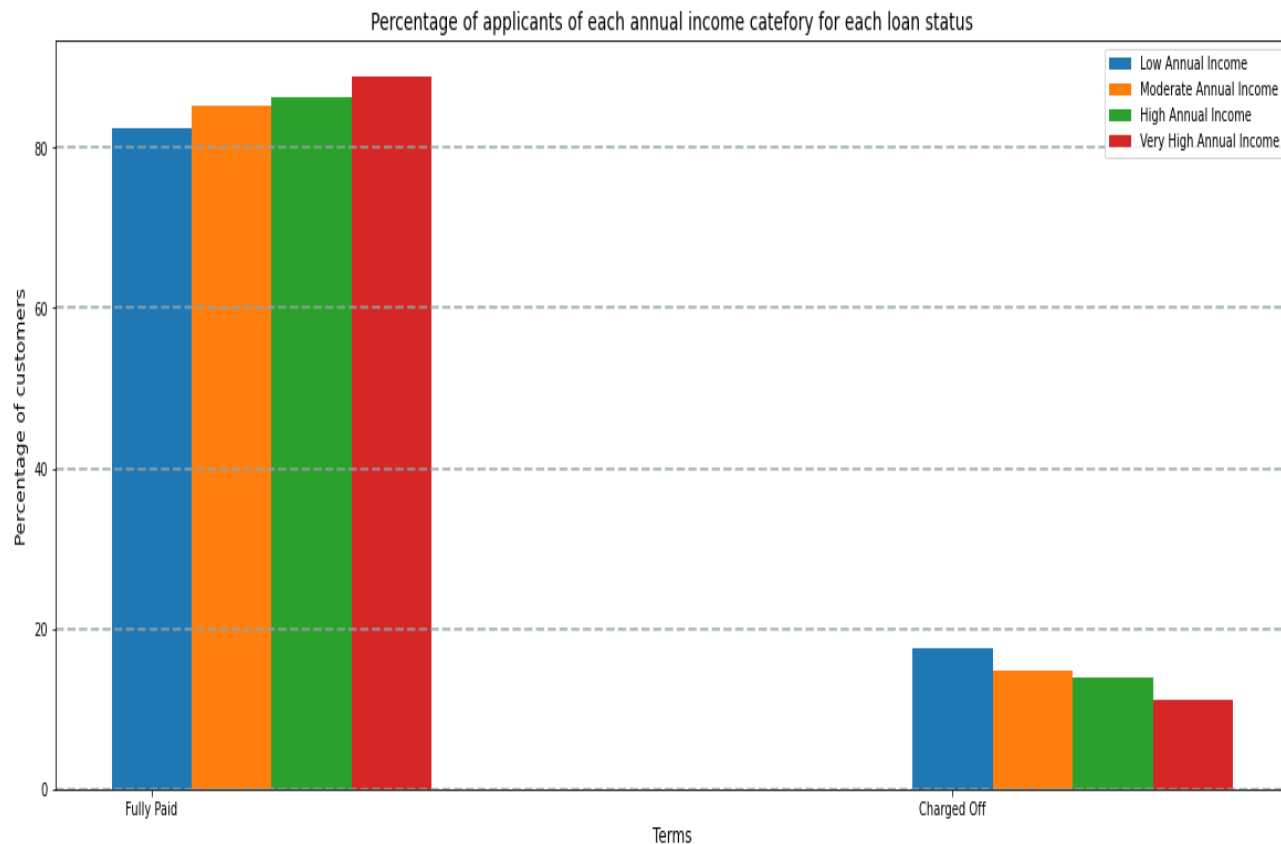
Distribution of funded amount over annual income



Sum of due amounts in each annual income categories

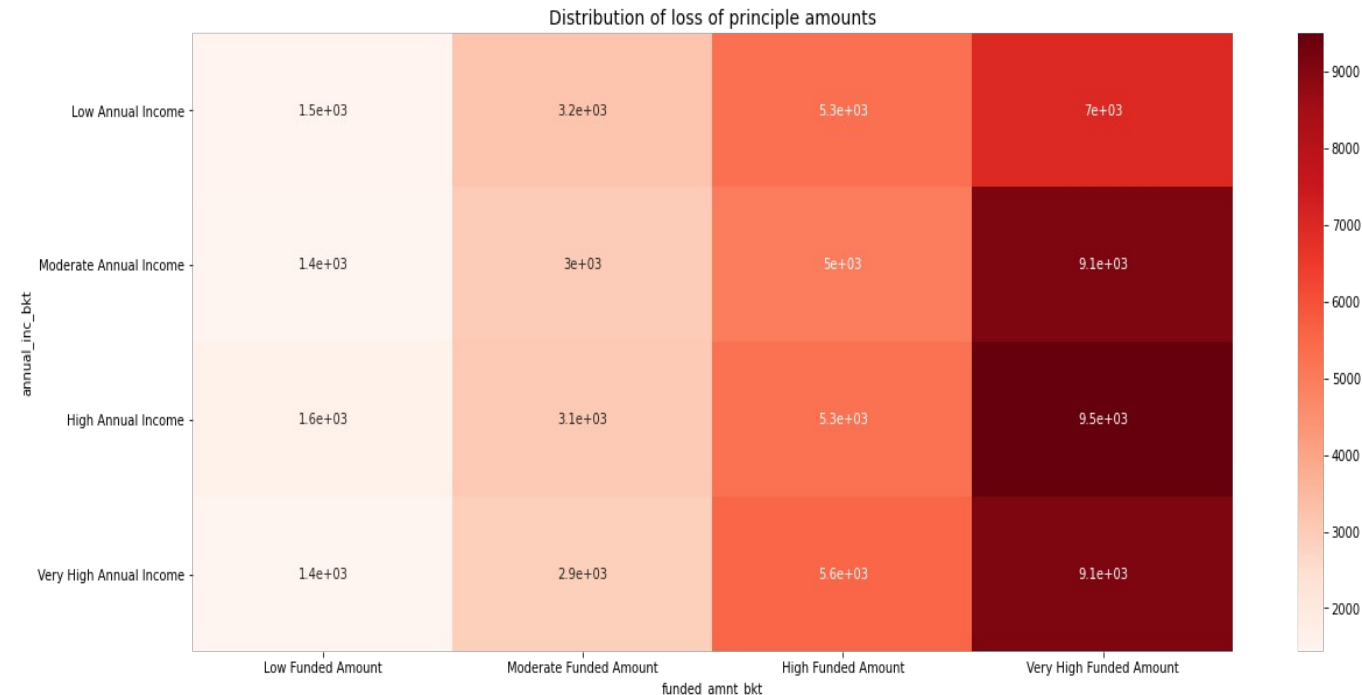
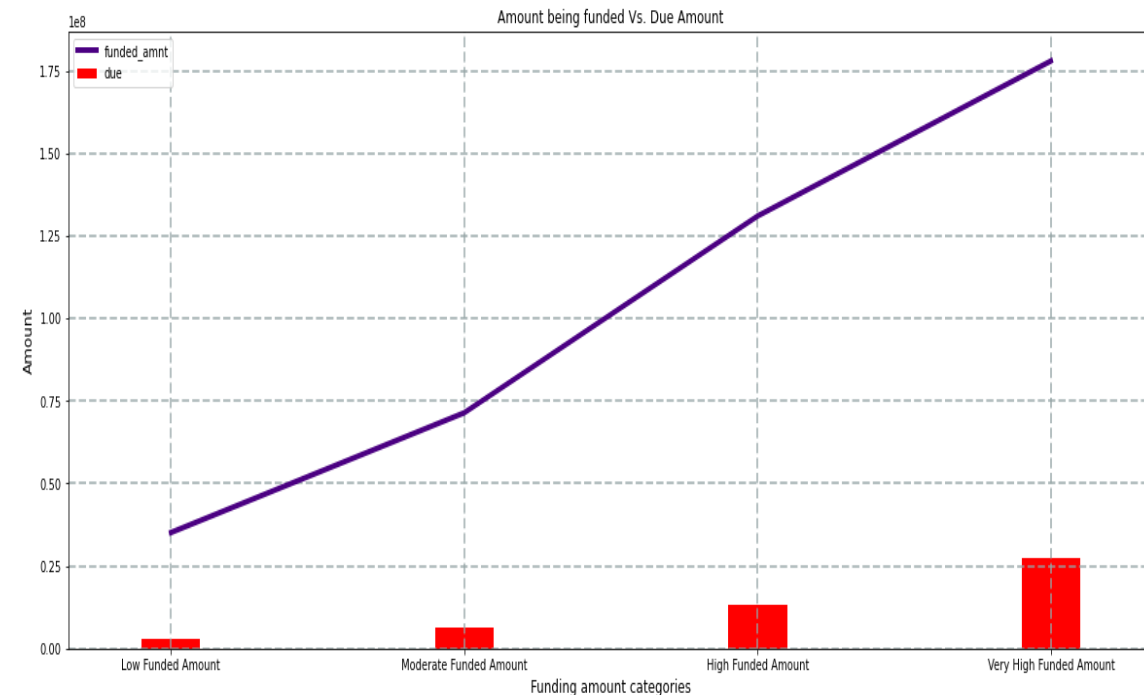


- 82.32 % of Low annual income customers tend to fully pay their loan whereas 17.68 % of customers tend to be charged off.
- 85.18 % of Moderate annual income customers tend to fully pay their loan whereas 14.82 % of customers tend to be charged off.
- 86.02 % of High annual income customers tend to fully pay their loan whereas 13.98 % of customers tend to be charged off.
- 88.68 % of Very high annual income customers tend to fully pay their loan whereas 11.32 % of customers tend to be charged off.
- There is more chance of being defaulter if interest rate is very high high in high and very high funded amounts.

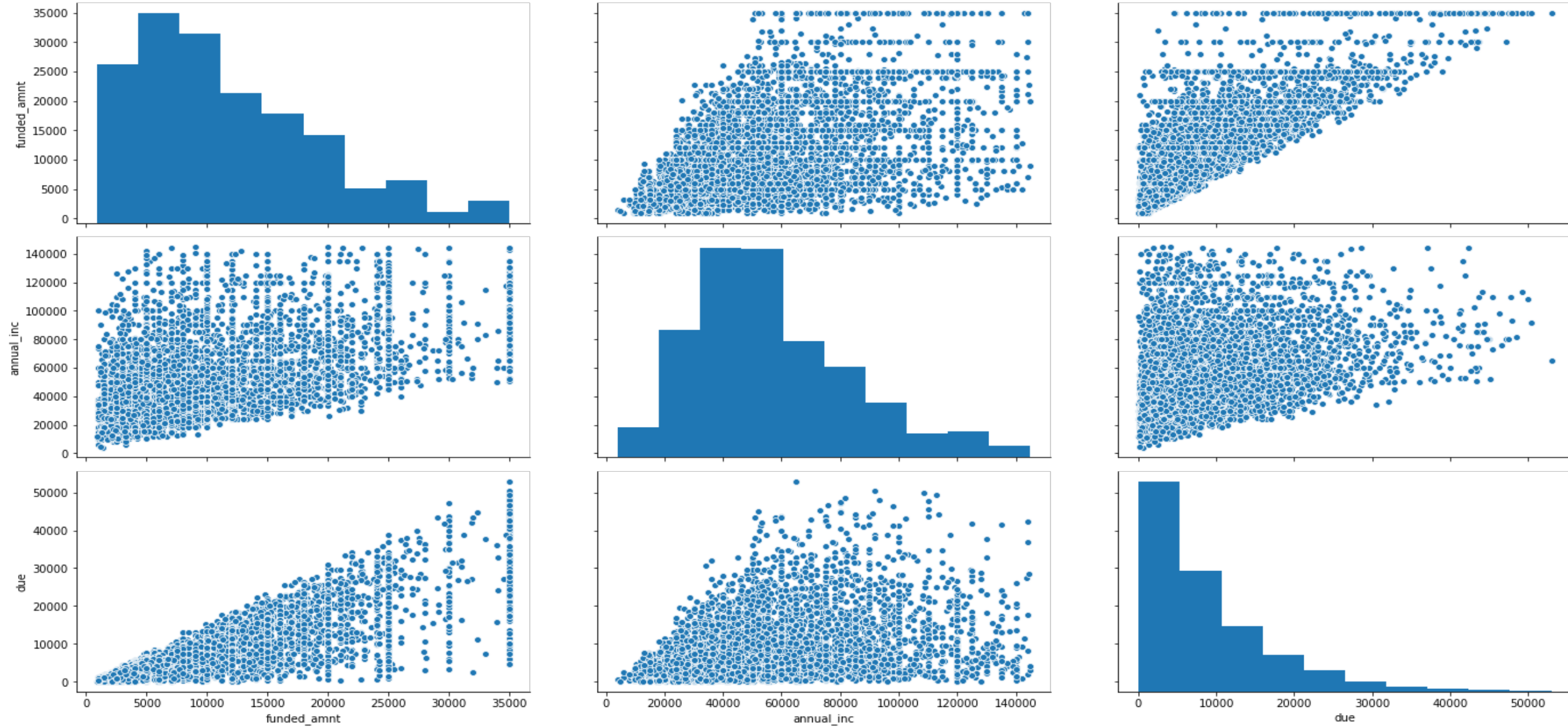


Due Analysis

- Annual Income over 50,000 USD tend to be defaulter of high amounts.
- Most of the dues are committed by candidates with annual income within 60,000 USD funding amount within 10,000 USD.
- Eligibility of funded amount is directly proportional to annual income.
- Lower Annual Income group of applicants often tend to be defaulters while probability of higher annual income applicants of being defaulter is less but if in case they end up being defaulter, they may cause huge loss.
- Very high funded amounts have caused most of the monetary losses.
- Most principle amount losses have occurred in very high funding amount category.

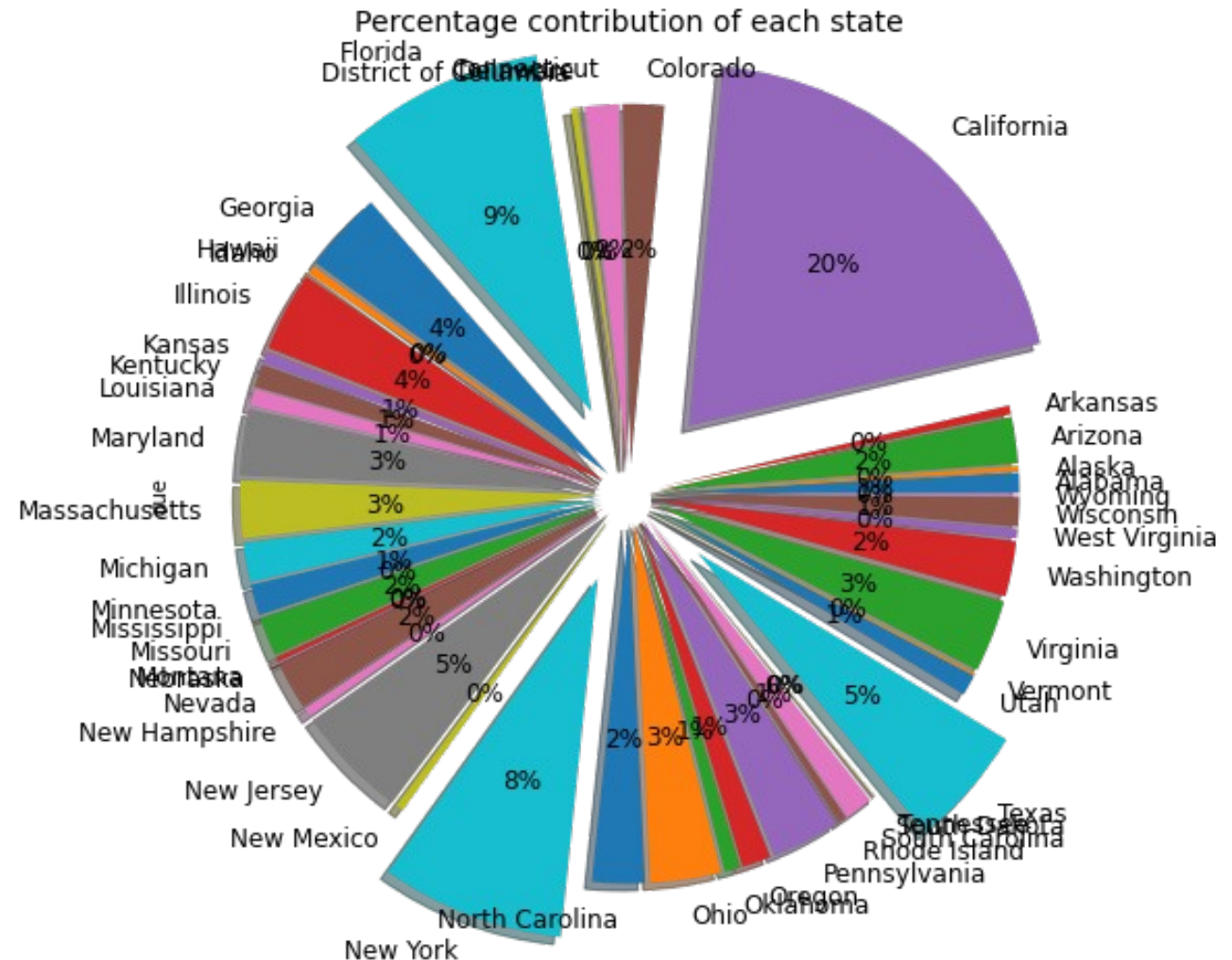


Correlation of funded amount, annual income and dues :-



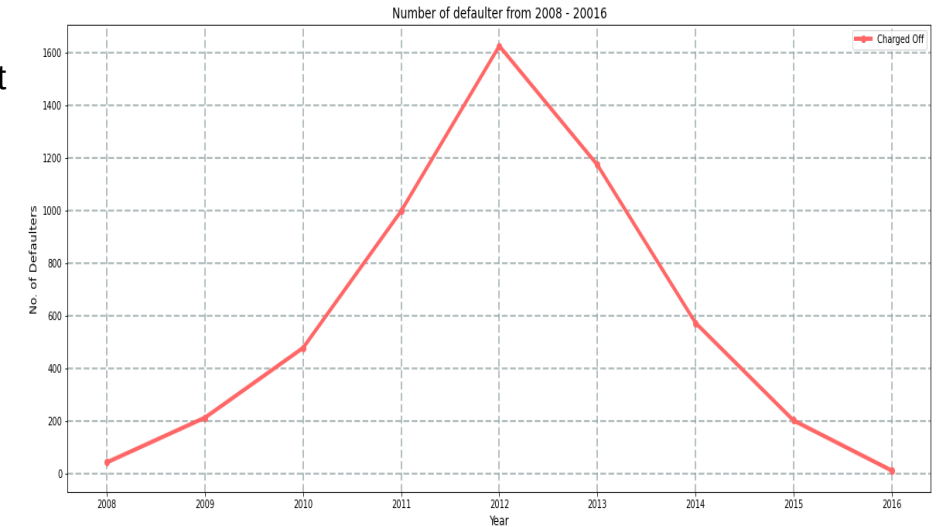
State Analysis

- The largest state to be charged off is California followed by Florida, New York and Texas based on both count and amount
- These top 4 countries alone contribute for 42% of total loss.
- Most customers from California have a High Annual Income receives a High Funded Amount and together contribute to a loss of around 20 %
- Most customers from Florida have a Low Annual Income receives a Low Funded Amount and together contribute to a loss of around 9 %
- Most customers from New York have a Moderate Annual Income receives a High Funded Amount and together contribute to a loss of around 8 %
- Most customers from Texas have a Moderate Annual Income receives a High Funded Amount and together contribute to a loss of around 5 %

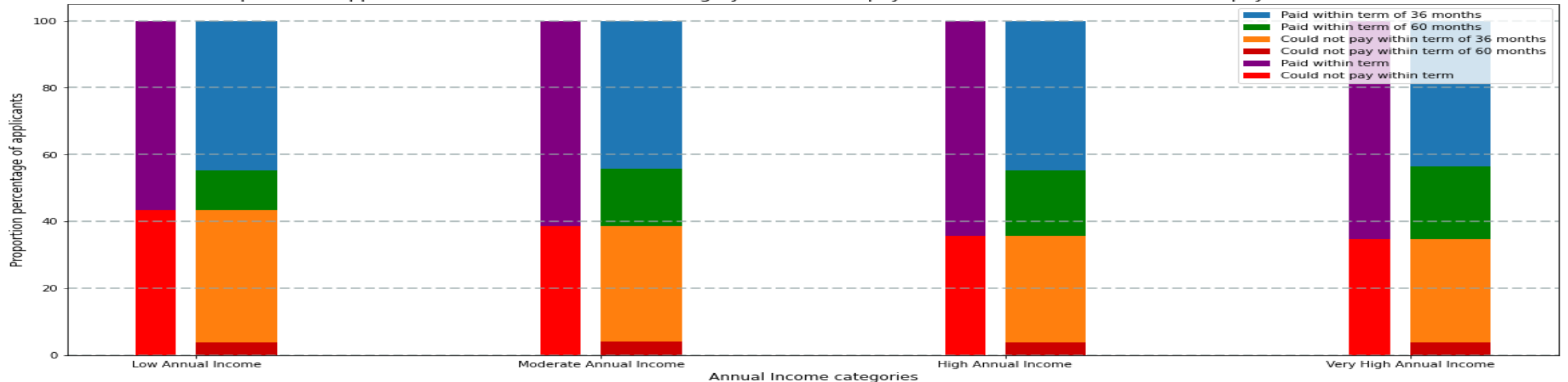


Date and Time Analysis

- No. of applicants who stopped their installment and became defaulter increased drastically from 2008 to 2012 and then decreased till 2016. There is evidence which satisfies this.
- 61.75 % of applicants could repay the loan within time whereas 38.25 % of applicants could not repay the loan within time.
- 56.34 % of applicants who opted for 36 months term could pay the loan within time whereas 81.84 % of applicants who opted for 60 months term could pay the loan within time.
- Percentage of applicants opted for 36 months who failed to repay within time decreases with increase in annual income.
- Percentage of applicants opted for 60 months who could repay the within time increases with increase in annual income.
- Percentage of applicants opted for 36 months who could repay the loan and the percentage of applicants opted for 60 months who failed to repay within time does not have any significance relateness with the annual income.



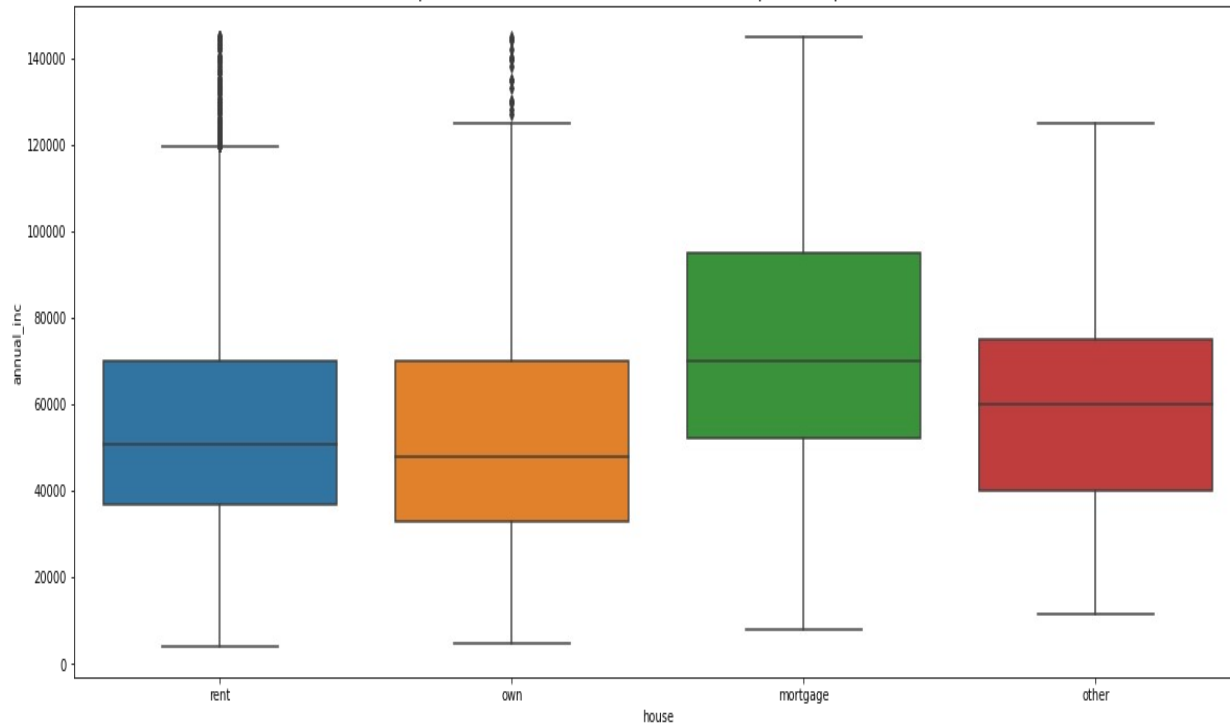
Proportion of applicants in each annual income category who could repay within the term Vs. who could not pay within the term



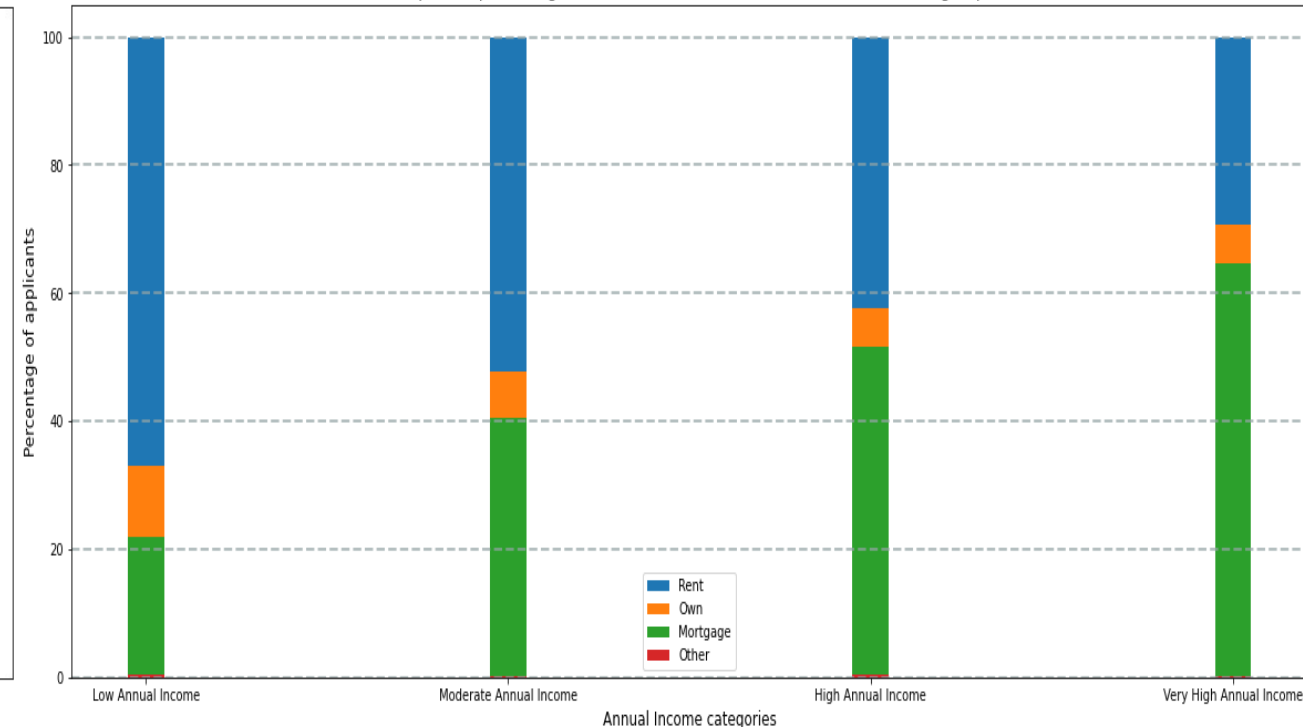
Home Ownership Analysis

- Trend of house ownership is almost similar all over.
- We can say, that high salaried applicants usually have their house mortgaged or vice-versa. From the past experience in Annual Income analysis, it is obvious that Higher annual income groups definitely received larger amounts of loan. Therefore, there is a slim chance that due to burden of the mortgage, they end up being defaulter.
- With increase in annual income, rented house owners decrease whereas mortgage house owners increase.

Comparison of annual income with home ownership in the top 4 countries

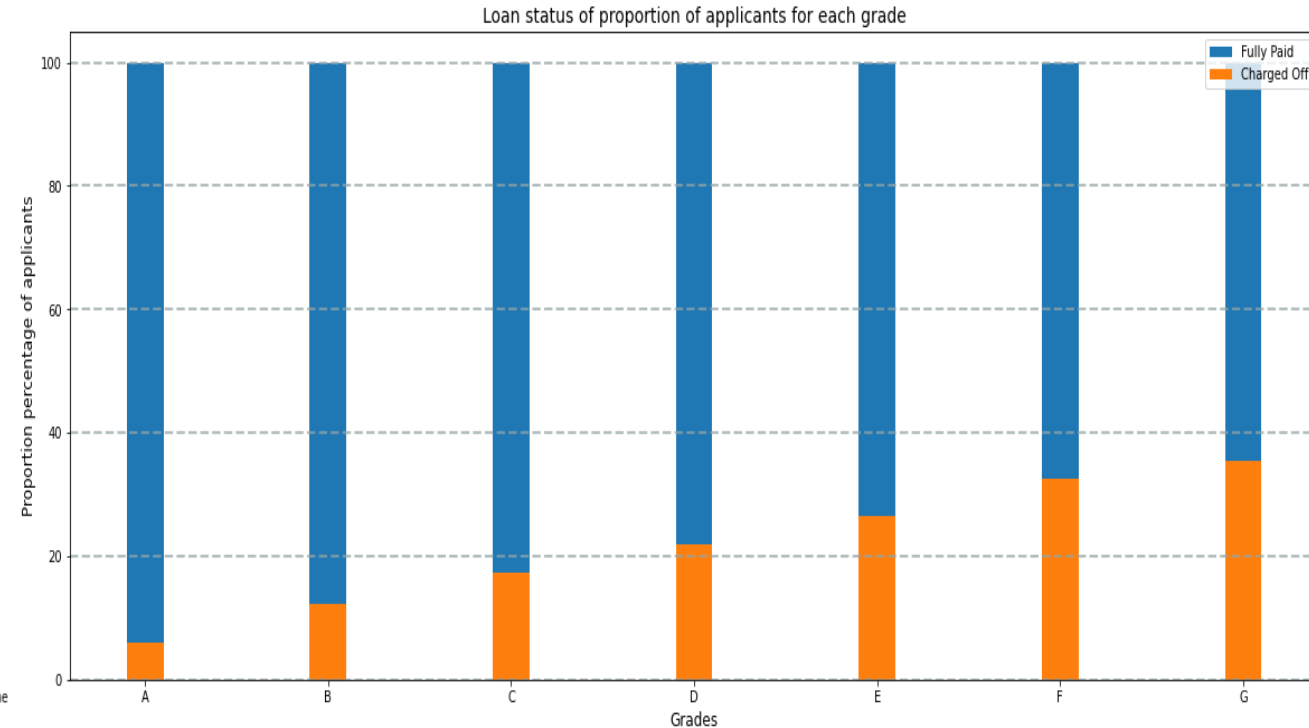
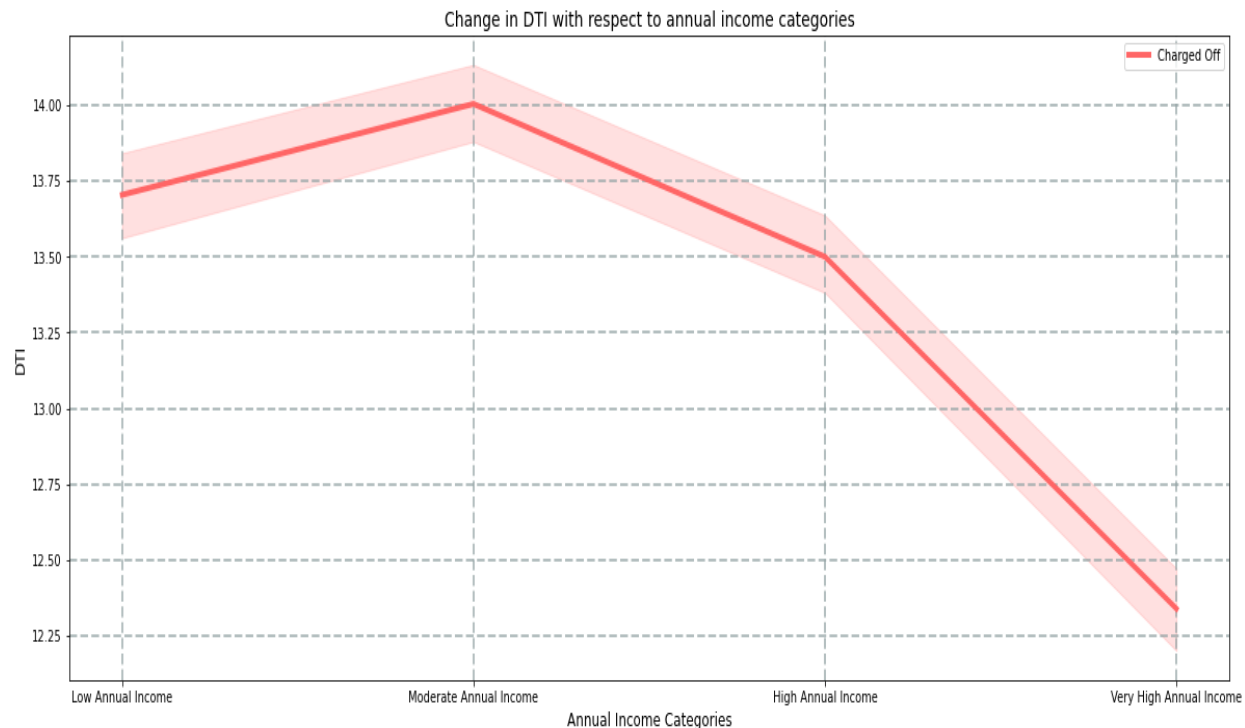
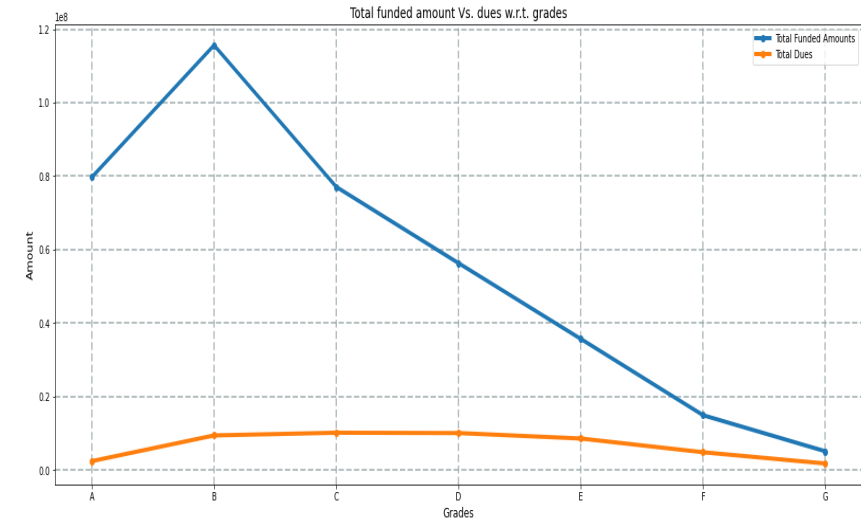


Proportion percentage of house of defaulters' for each annual income group



Grade Analysis

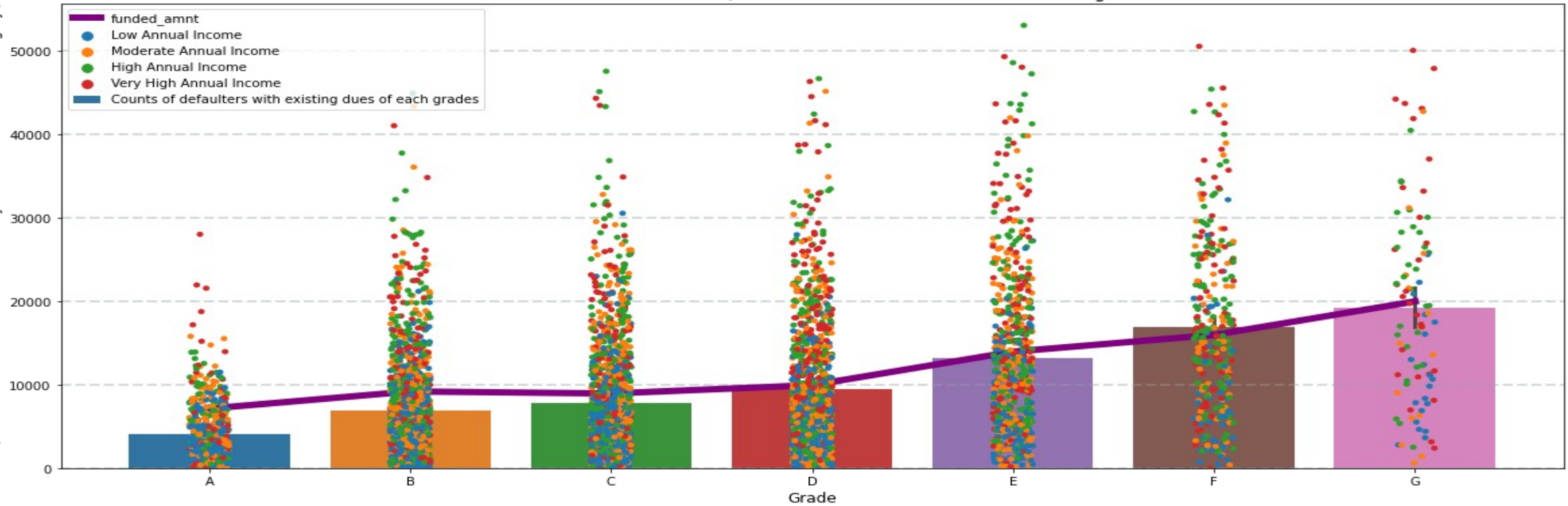
- Debt to Income ratio (DTI) decreases with increase in annual income.
- Percentage of defaulters increases with deterioration of grades.
- With deterioration of grades, total funded amount decreased but the dues remained almost consistent.
- Grades 'E', 'F' and 'G' usually receive significantly larger funded amounts (greater than median of non defaulter applicants) who have caused fairly large number and amount of dues



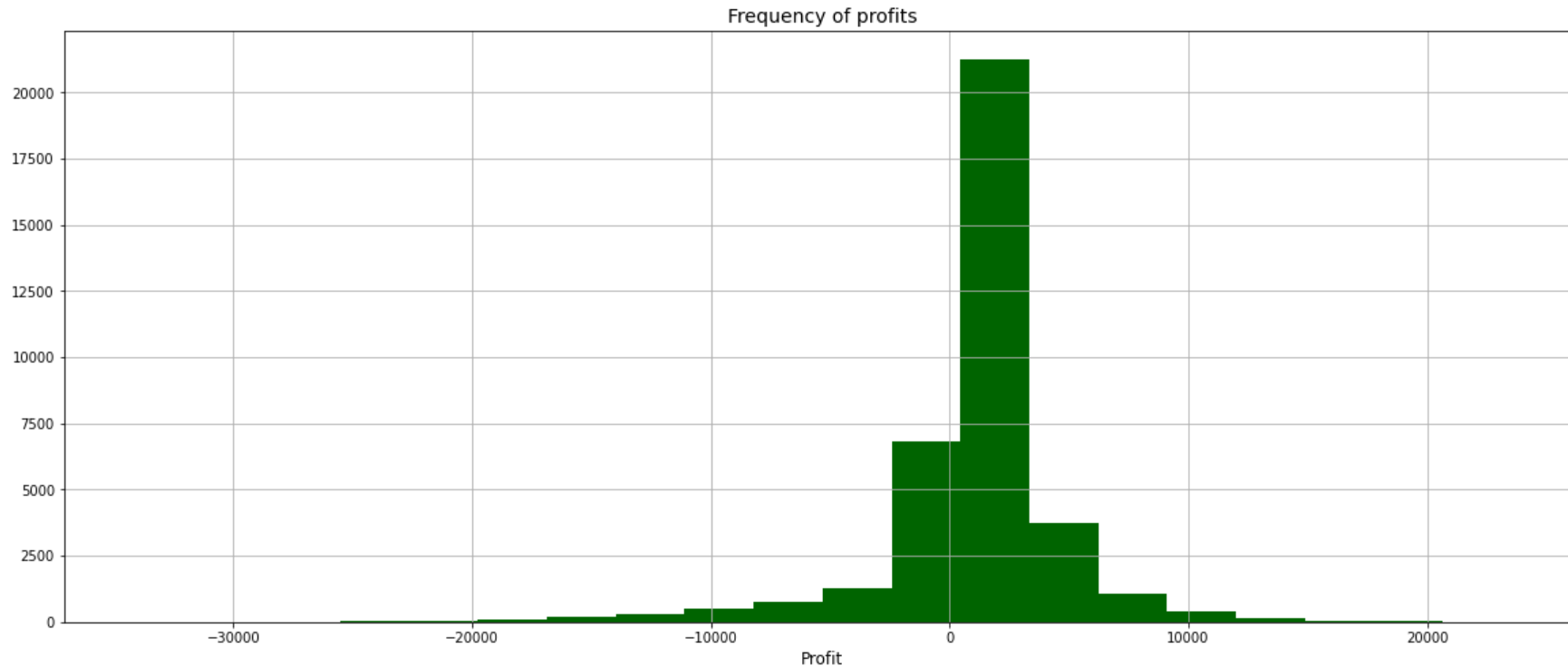
With deterioration of grades, funded amount, overall due amounts value and due counts increase.

Amount of funded amount / Amount and distribution of dues by each annual income category / Counts of defaulters

Vaiations of funded amount, due amounts and count of dues with grades



Conclusion



After all, since the company is not going through huge profits, it should consider reviewing its policies based on all the observations and trends.