CREDIT RISK ANALYSIS CASE STUDY

SUBMISSION

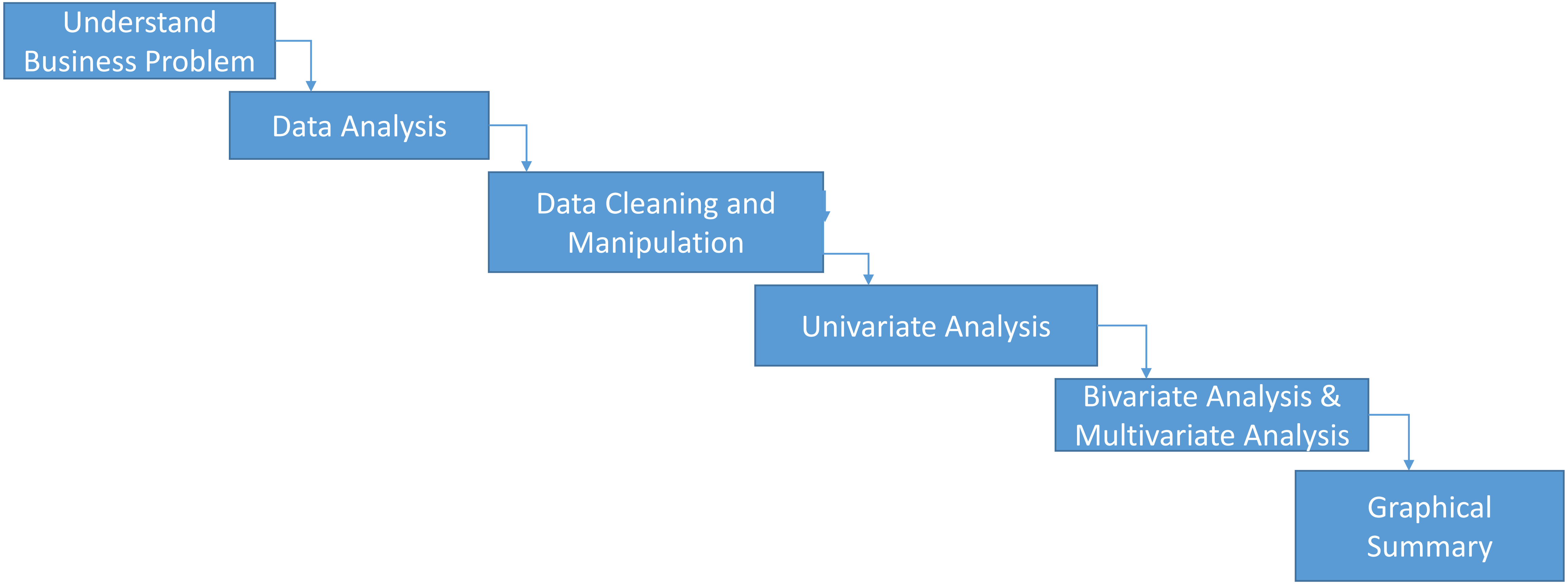
Group Name:

1. Manoj Kumar
2. Divyam Sinha
3. Shishir Pandey
4. Bhargav Kantipudi

Abstract

- The aim of the case study is to identify the credit risk for a consumer finance company
- Exploratory Data Analysis is done to understand how consumer attributes and loan attributes influence the tendency of default
- If one is able to identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss.
- The company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilise this knowledge for its portfolio and risk assessment.

Problem solving strategy



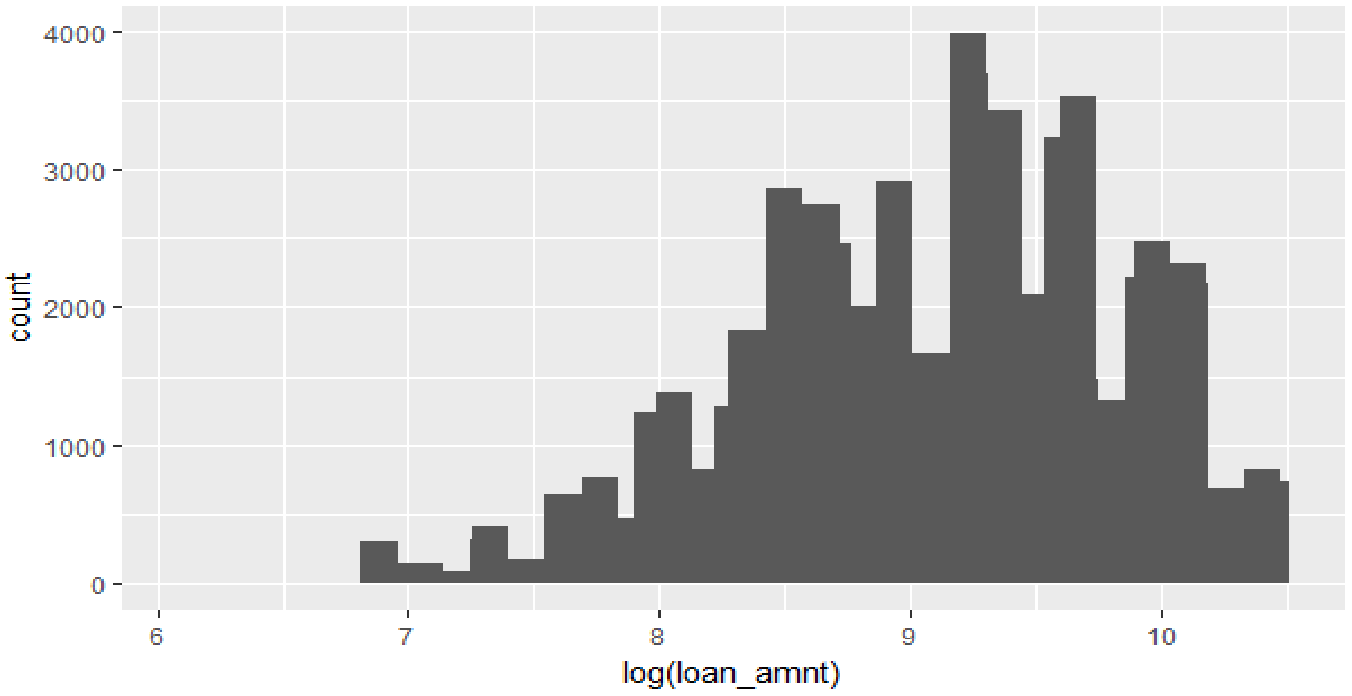
Data Analysis

- The loan dataset which is used for the EDA has 111 variables which include all the details of a loan application .
- There are lot of fields having NA values which were cleaned as a part of the data cleaning process
- There are few fields which only consisted of single values like ‘n’ , ‘f’ etc which will not help in EDA hence they were also dropped.
- There are some continuous variable like loan amount , funded amount etc and some categorical variables like loan status , home ownership etc which were critical during the EDA process
- The data formats for few fields like last payment date , issue date etc were also converted into standard date format for better analysis
- Some unwanted characters were found in string values like ‘%’ etc which were removed from the data.
- Date is mutated into different field to understand the time variance of the data.

UpGrad

Frequency Distribution of Loan mount

- Log transformations was applied to loan amount to reduce the skew and present better scaled plots as shown
- The loan amount skewed relatively small. Median is less than mean value.
- Majority of the loan amount is given on higher side

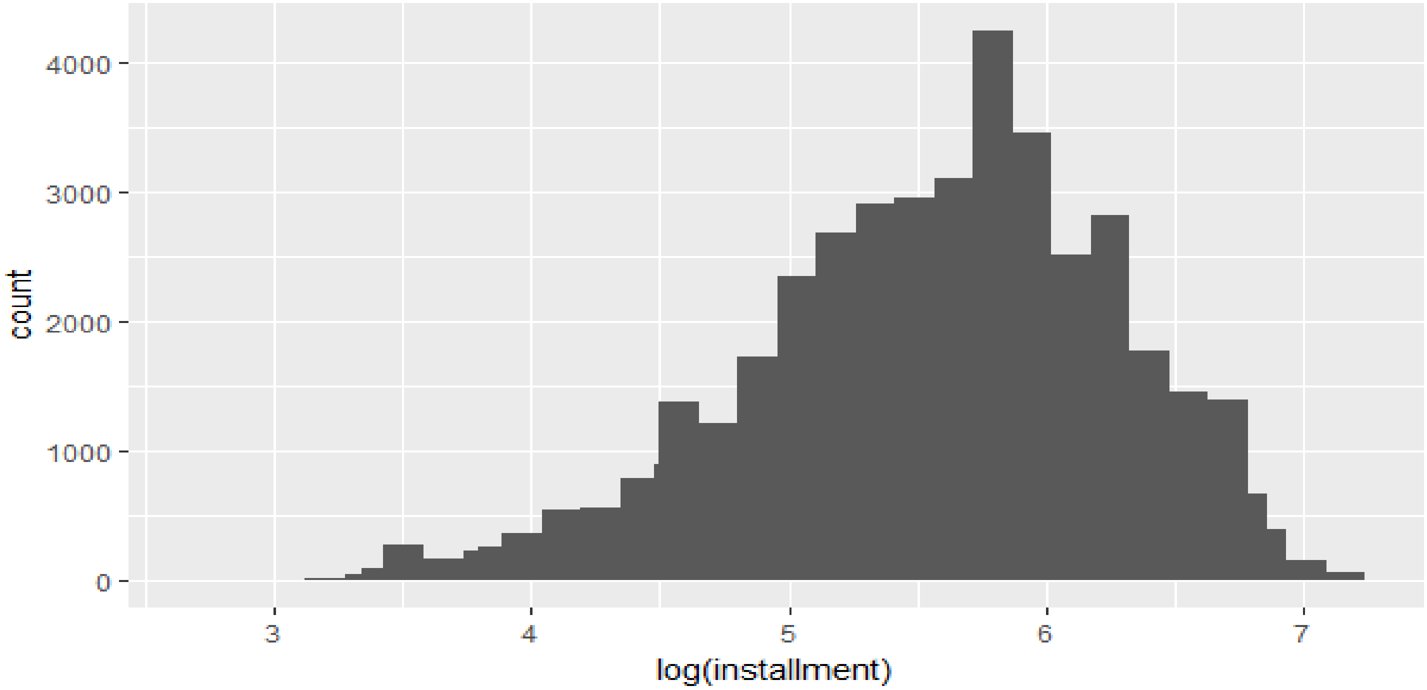


Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
500	5500	10000	11219	15000	35000

UpGrad

Frequency Distribution of Instalment

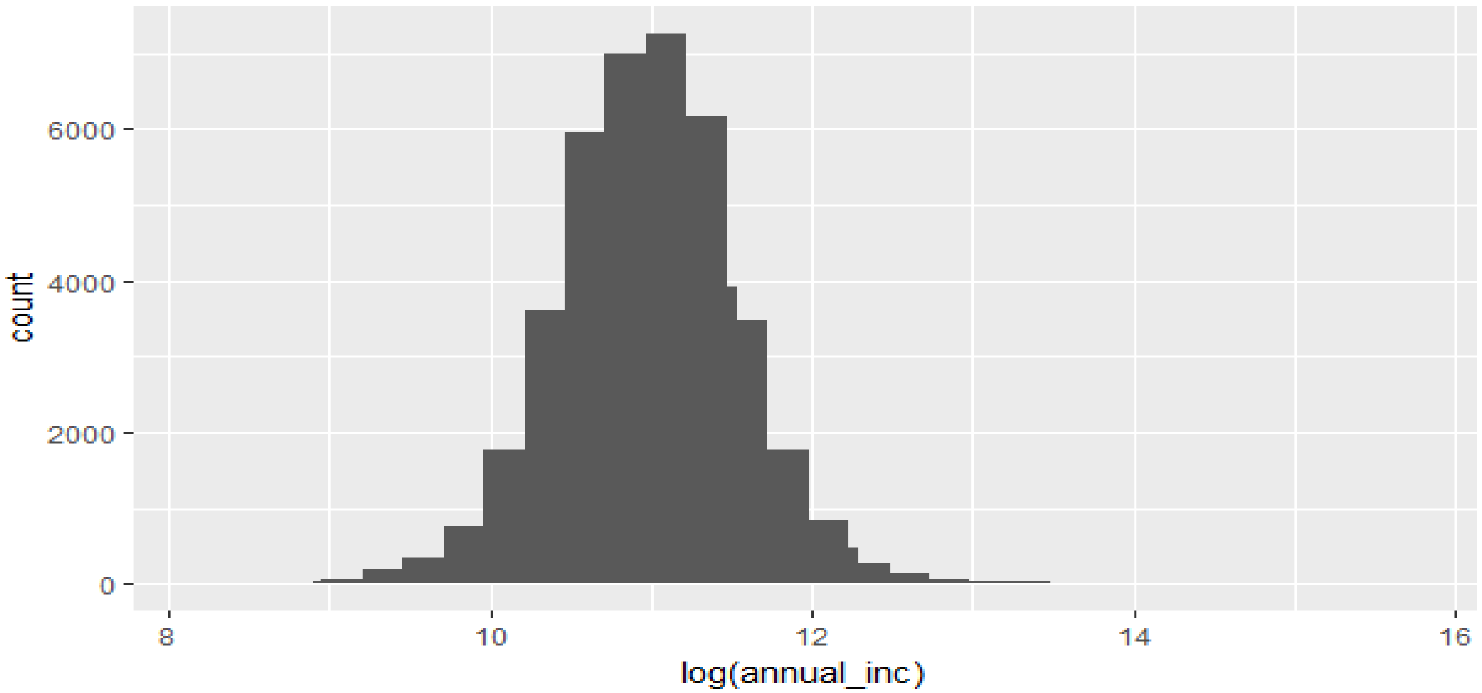
- This distribution is very close to normal distribution
- The instalment is skewed relatively small. Median is less than mean value.
- As expected majority of the instalment is given on higher side since the loan sectioned amount is on higher side.




```
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
15.69  167.02  280.22  324.56  430.78 1305.19
```

Frequency Distribution of Annual Income

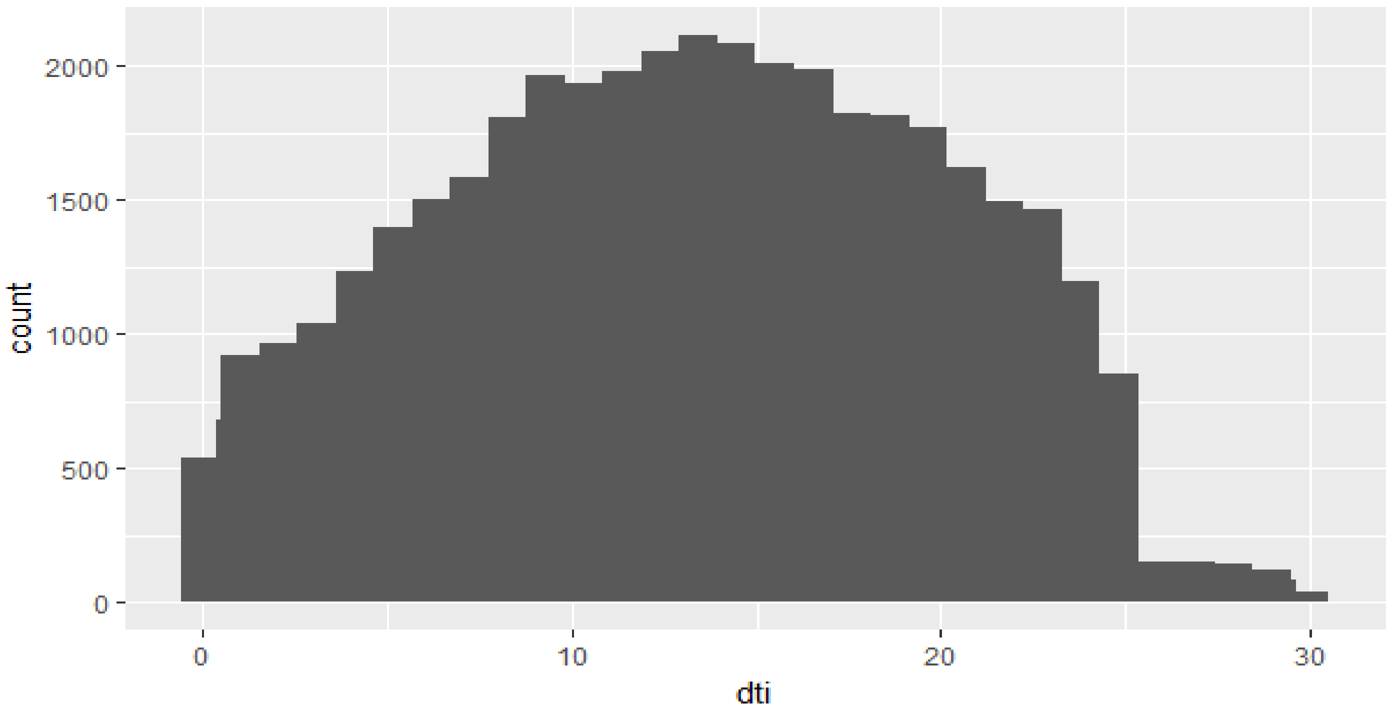
- This is very close to normal distribution



Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
4000	40404	59000	68969	82300	6000000

Frequency Distribution of Debt-To-Income Ratio

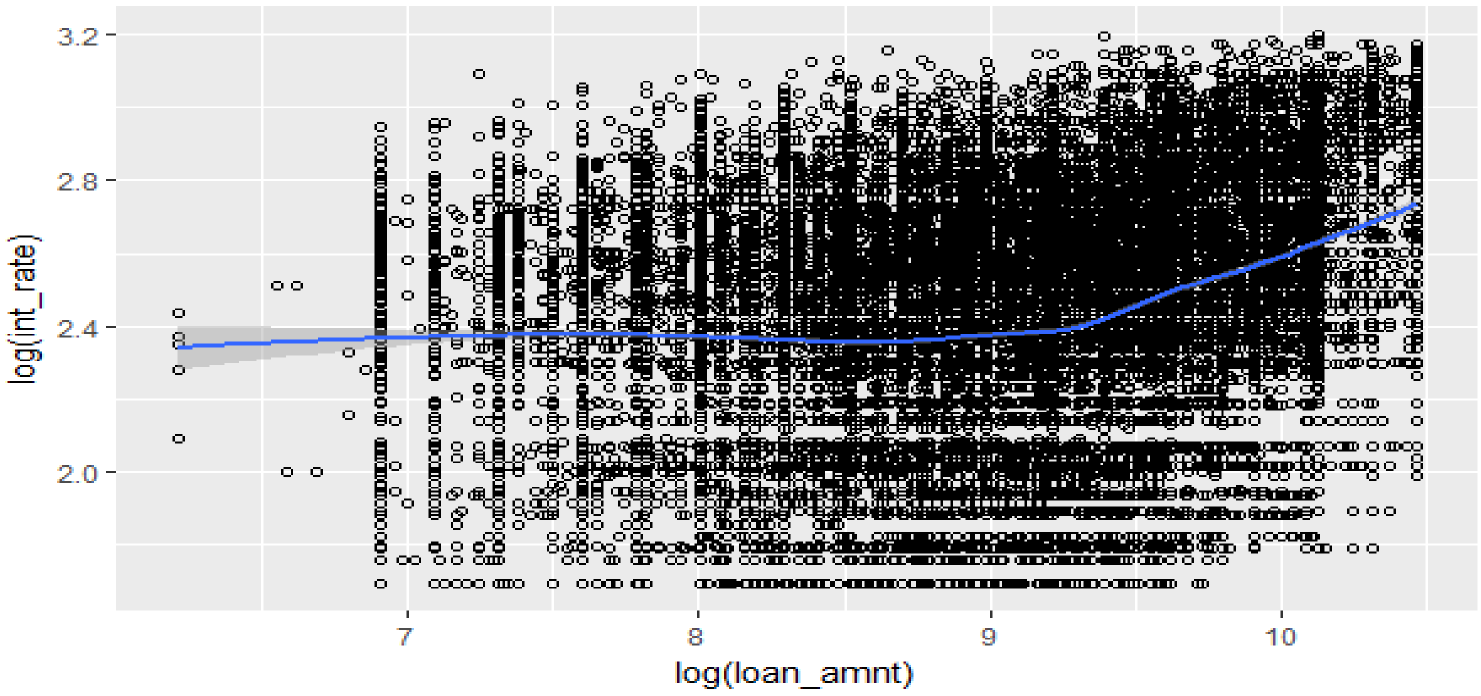
- This is very close to normal distribution
- Median is greater than mean and they very close each other.



Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.00	8.17	13.40	13.32	18.60	29.99

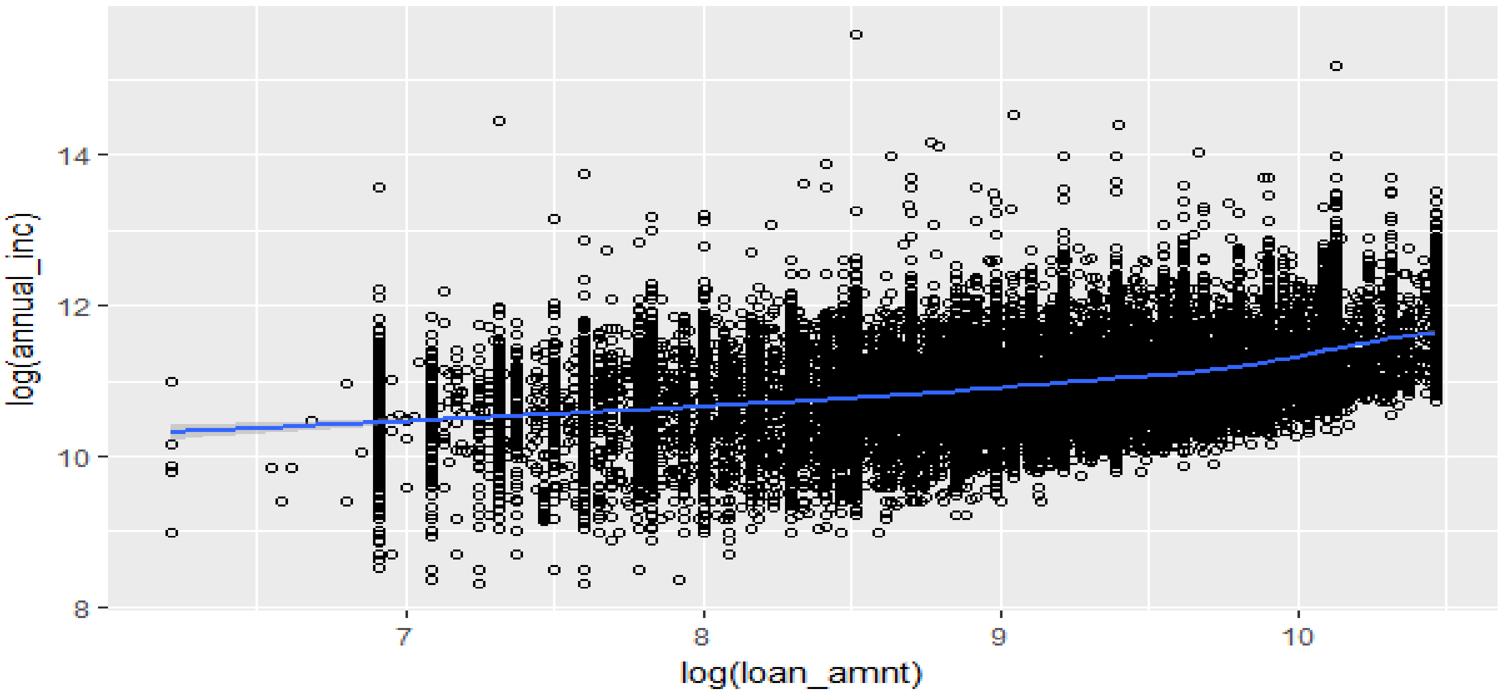
Scatter plot of Loan Amount vs Interest Rate

- We can conclude from this graph that interest rate tends to be on higher side for the higher loan amount
- It shows the directly proportional relationship between the two



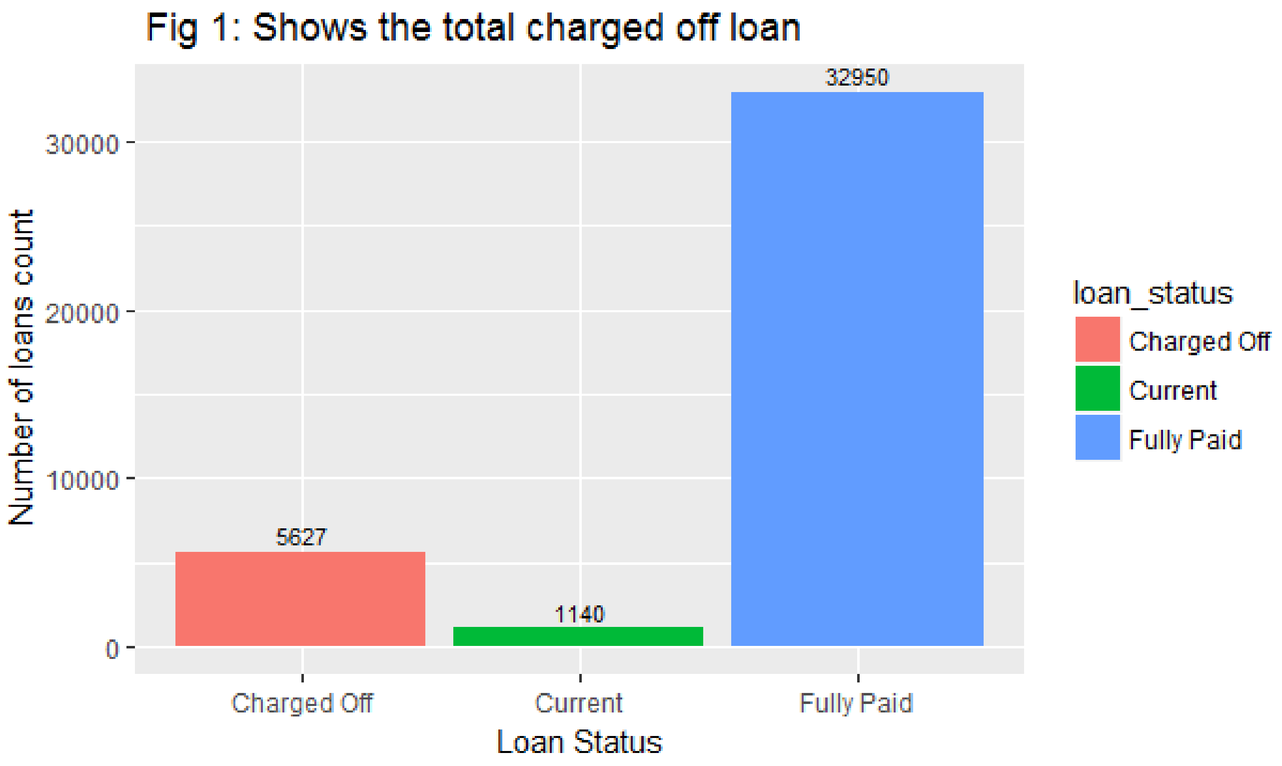
Scatter plot of Loan Amount vs Annual Income

- We can conclude from this graph that people who have higher income tend to take higher loan amount



Loan Status Analysis

- 85.8% are good loan and 14.16% are bad loans which have lead to great loss to bank
- Aim of our Analysis is to understand the root causes of this bad loan issuance

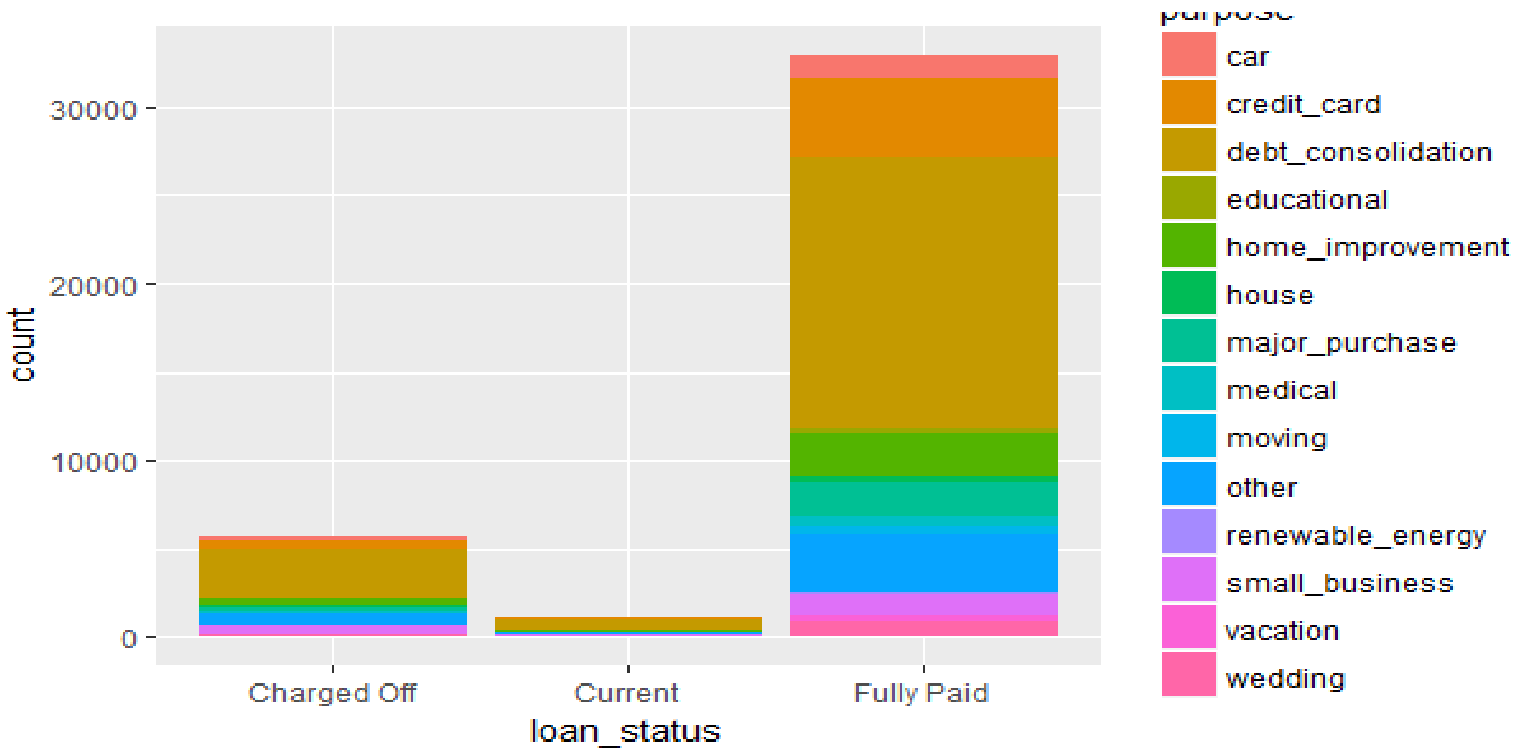


Loan Status vs Employee Year of Experience

- The plot suggest that applicants with lower professional experience tends to default higher.
- Applicant with employee length 0, 1, 2, 3, 4, 5, & 6 have number of defaults more than that of employee length 7, 8 & 9.
- There is an interesting insight that applicant with 10 and above years have higher loan approval rate but at the same time shows highest delinquency. This might be difficulty in getting new job quickly with higher experience when any applicant faces lost of job or a host of other reason outside the information the dataset provides.

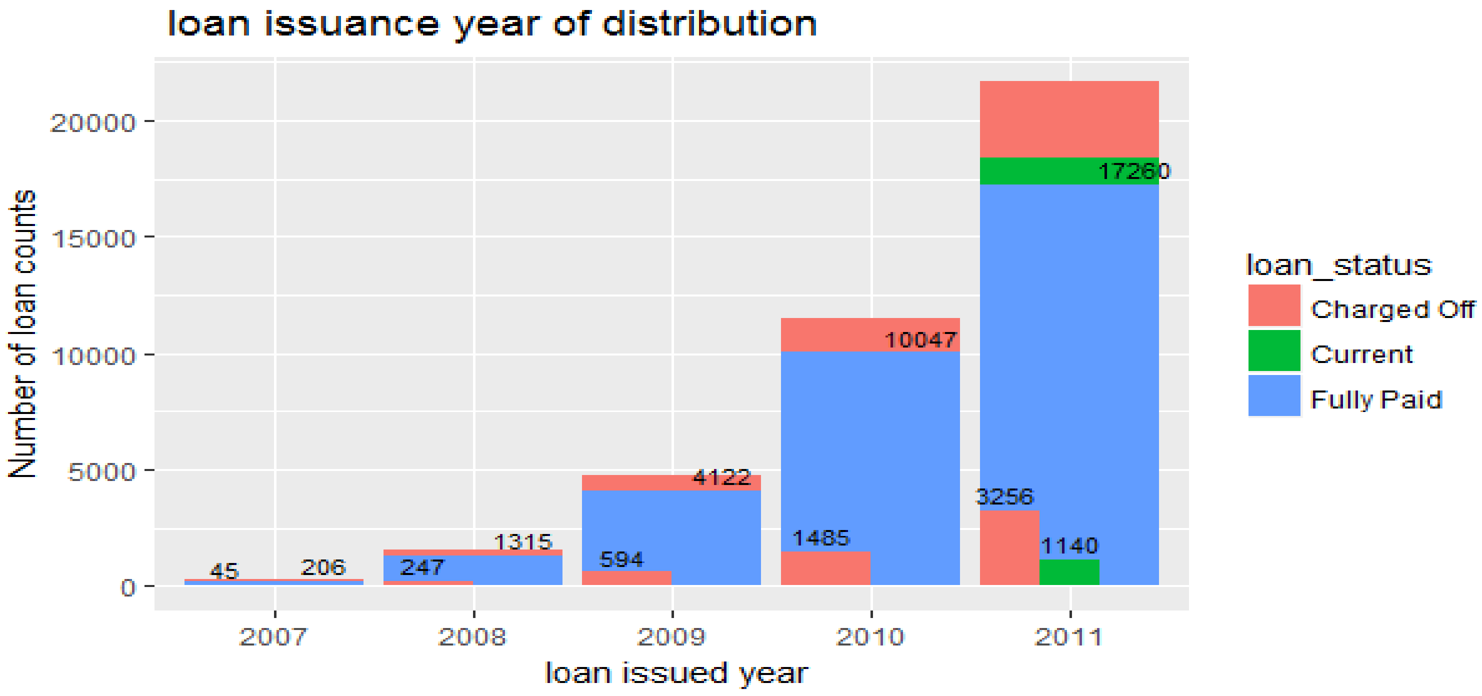
Loan Purpose vs Loan Status Analysis

- Highest Charged off People are those who have purpose of loan as Debt Consolidation though we can see the issuance rate of this loan is also very high
- People who have taken loan for wedding , renewable energy have not charged off much



Loan Issuance Year of Distribution

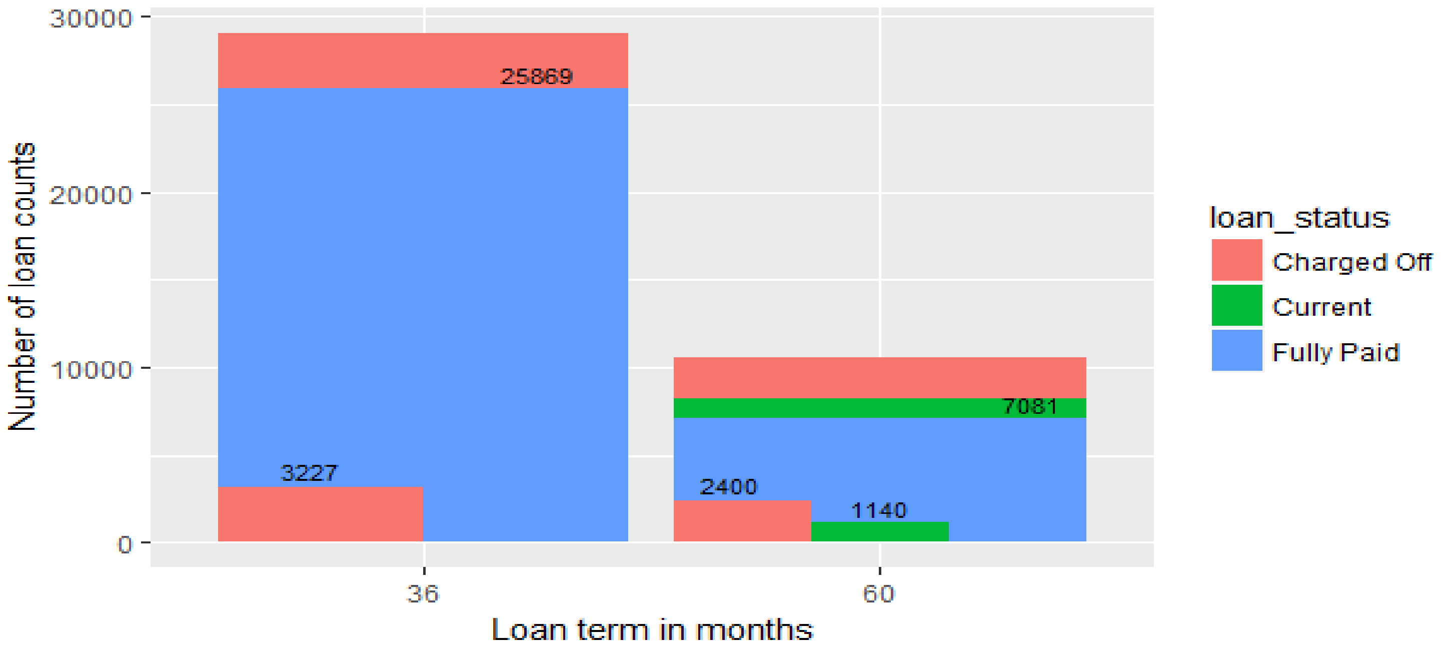
- With years the loan issuance has increased and also the bad loan rates
- There is a sharp increase in issuance of the loan in year 2011 and also the charged off people



Loan Term Analysis

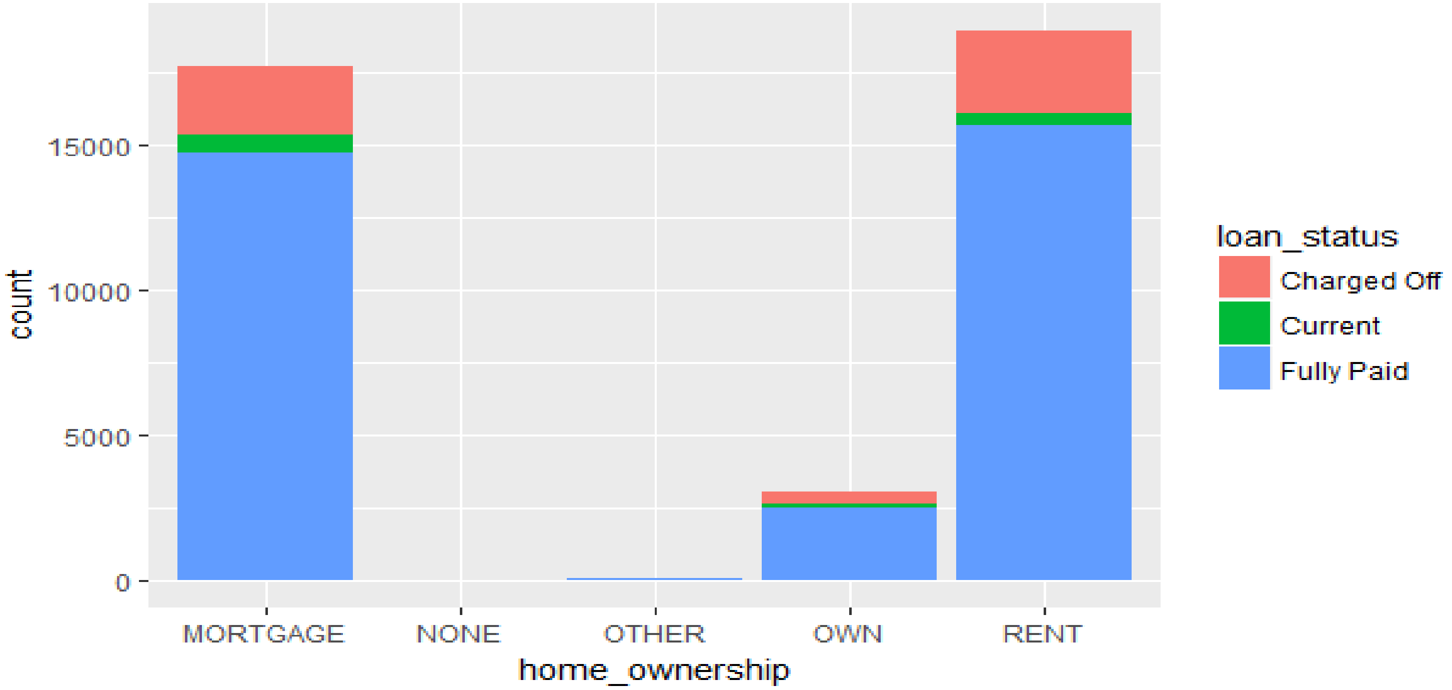
- Highest Charged Off people are for loan term 36 in months but loan with 60 terms shows higher percentage of charged off.

Fig 2: Loan term and its status



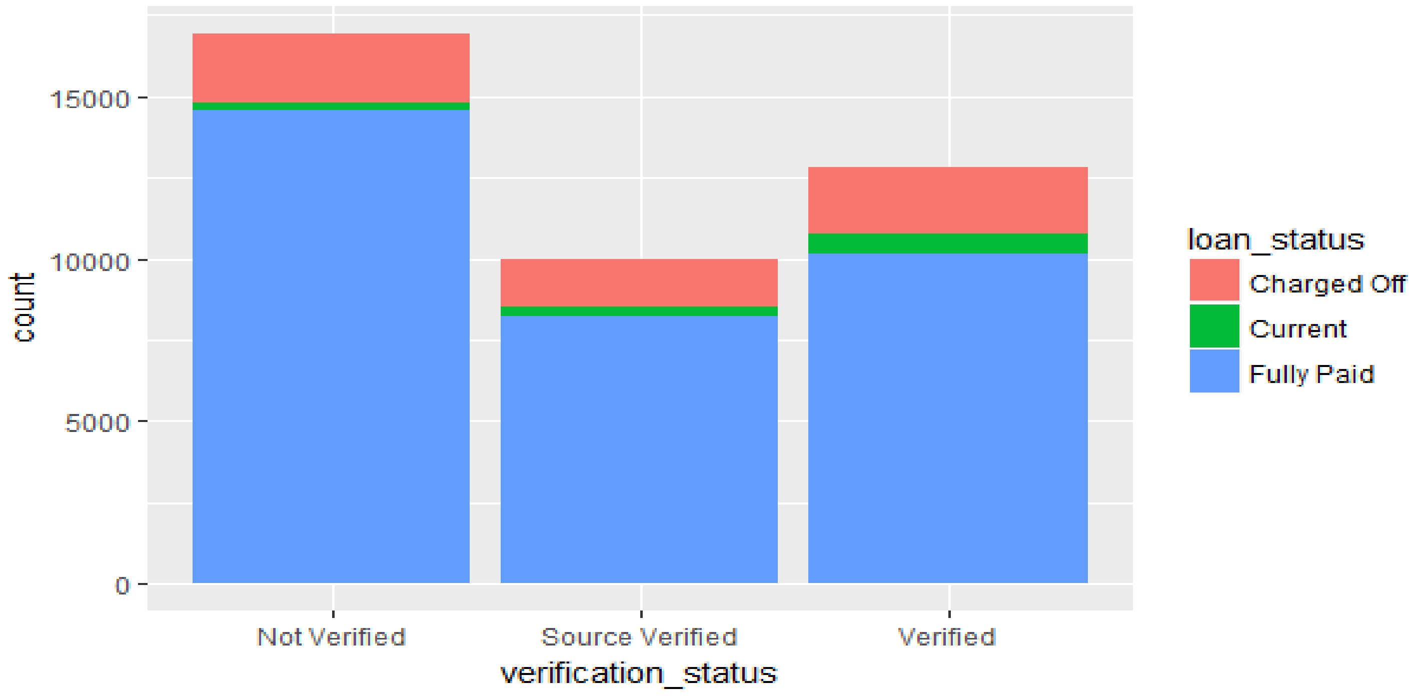
Home Ownership attribute Analysis

- People who have rented and mortgage property have Charged off more with respect to People who have there own house



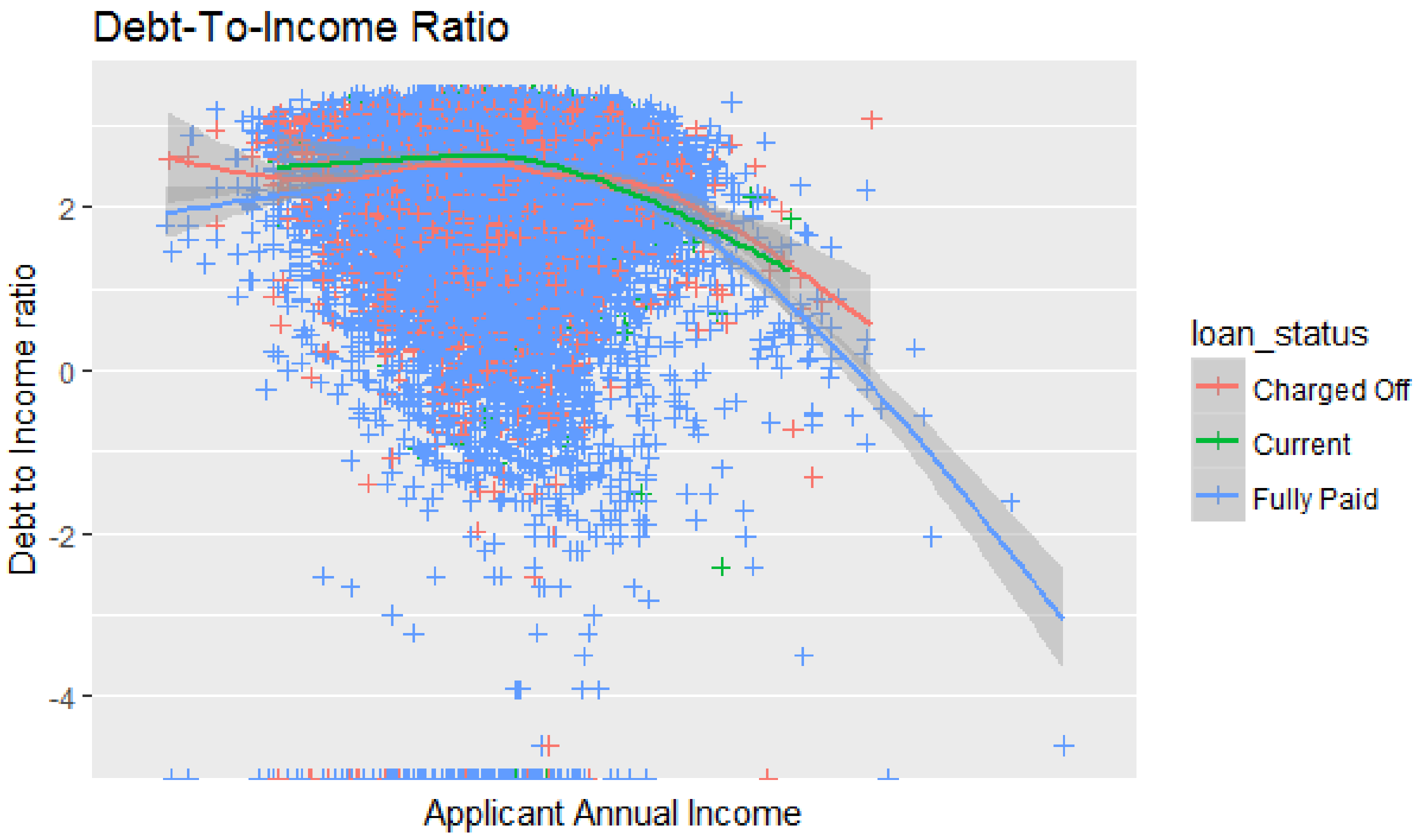
Verification Status Analysis

- Interesting part is when the source is verified the Charged off is less compared to when the Source is not verified

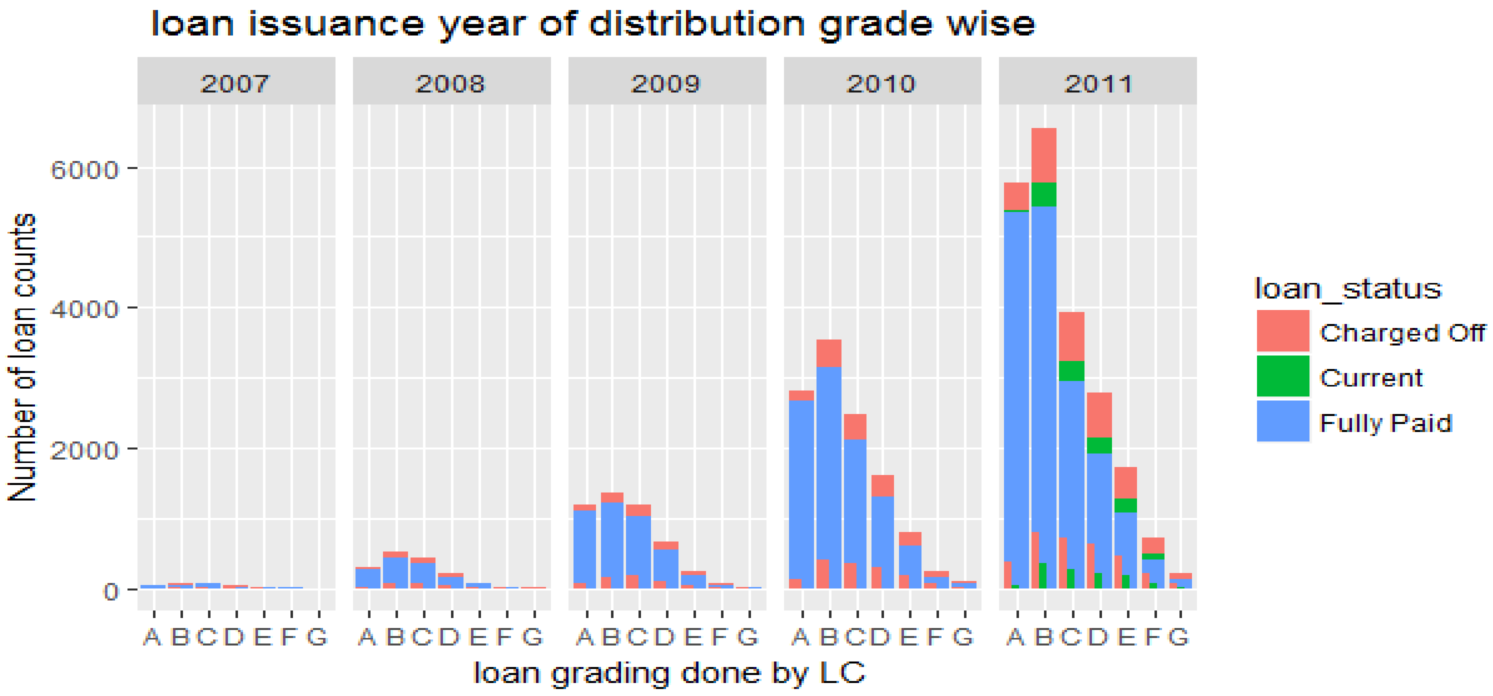


Debt-To-Income Ratio vs Applicant Annual Income

- The trend we can see that the delinquent applicants have lower annual income and higher dept to income ratio.
- The plot also suggest that with higher annual income of applicants leads to decrease in delinquent levels.



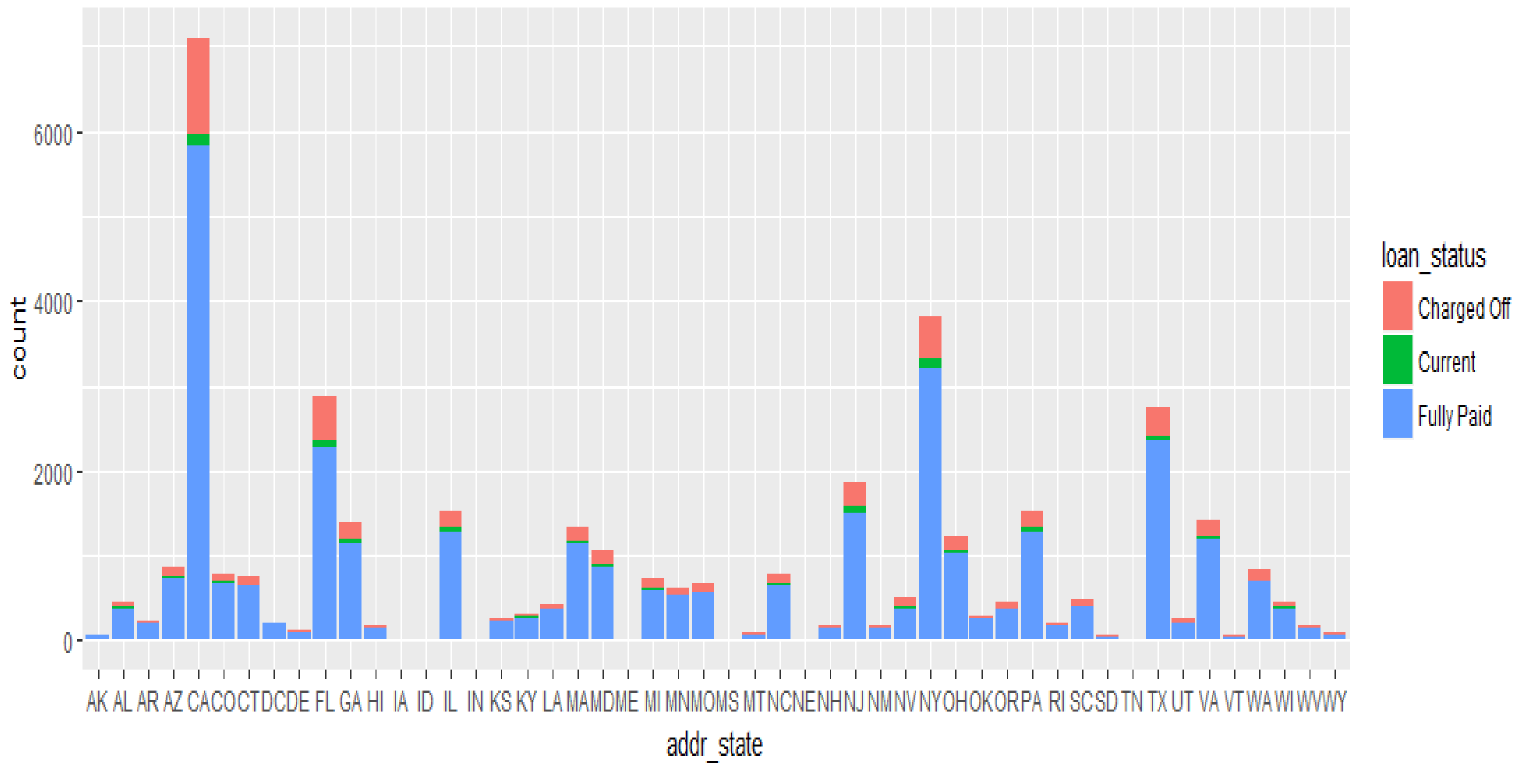
Grade Wise Analysis with Years



- As per the above trends which we have seen loan issuance has sharply increased throughout the Years
- Loan which were graded A were seen less likely to Charged off

Address State Analysis

- CA(California) shows the highest delinquents but it should be taken into consideration that the state is also more populous. NY, FL, TX follows the trend.
- The central American states & East coast region like WY, KS, ID, IA, NC, DE, FL e.tc. have lowest number of charged off loans applicants.
- There are no states which have more charged off loan applicant than the good loans.
- Few states like IA, ID, IN, ME, MS, TN, VT, MT, SD have negligible number of charged off loan applicants.



Conclusion

- The focus of our analysis is to understand loan and customer attribute which can affect Credit Risk
- The states of CA , NY , FL , TX are producing more defaulters but at the same time this states are more populous than others and shown promising locations for business. So while issuing loan to applicants in this area bank should do proper due diligence to avoid or reduce the loss due to loan defaulters.
- It is riskier to give a loan to people who have rented and mortgage property so bank should prefer to give to people who have own property.
- Applicants with lower income tends to show high debt to income ratio. Applicants with higher income ratio has shown reducing tendency in loan defaults. So banks should maintain proper due diligence while approving loans to low annual income applicants.
- Employees having less years of experience tends to default loans more than persons with higher years of experience. The only exception we saw with 10 and above years of applicants have high loan approval rates and higher delinquency rates as well than others. This would require more information for further analysis.
- People whose income source has not been verified are more likely to be Charged Off so it must be a norm to validate the source to avoid any Charge Off
- Loans which are graded A by LC are less likely to default so the bank should do proper due diligence before approving loan to lower grade applicants.
- Purpose of Loan also plays a key part in the credit risk , people who have taken loan for wedding are less likely to default than people who have taken for debt consolidation . A special wedding loan is a good bet for a bank. Bank should be cautious when issuing loan for ‘Debt Consolidation’. It is observed from data that cars and education loans are fully paid loans. It might be possible that the applicants who have car loans or students who have education loan might be paying off their loan using debt consolidation loans. So banks should do proper due diligence to avoid delinquency of loan repayment when taken with a reason of ‘Debt Consolidation’.