

Gramener Case Study

GROUP NAME:

1. ALOK SINGH
2. NISHAN PATEL
3. SRIDHAR KUMAR
4. HARSHAL ZELE

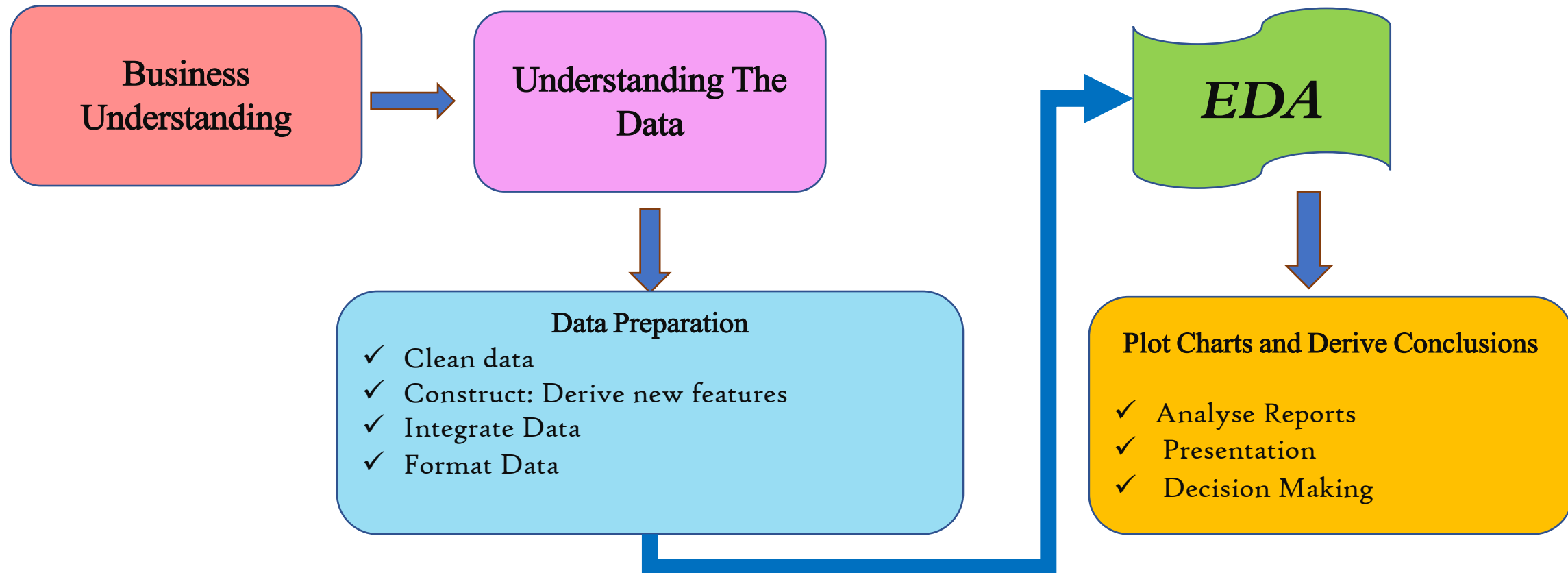
Business Objective

The consumer finance company specializes in different types of loans to the urban customer. The major problem they face is when the customer fails to repay the loan.

Two types of risks are associated with the bank's decision to provide loan to the applicants:

- ☐ 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 to the Lending Company.
-
- 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.

Problem solving methodology



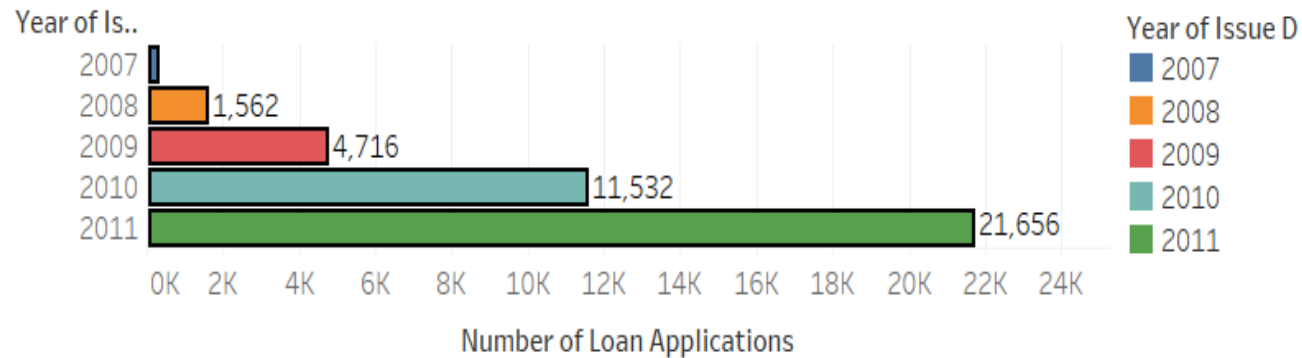
Data Cleaning:

- ✓ Remove all columns with 90% values are Null.
- ✓ Remove all columns which has only 'NA' and/or zero's.
- ✓ Remove Following Columns which do not provide any meaning to EDA analysis.

*'collections_12_mths_ex_med','tax_liens','delinq_amnt','acc_now_delinq','policy_code',
'chargeoff_within_12_mths','initial_list_status','pymnt_plan','url','member_id',
'zip_code'*

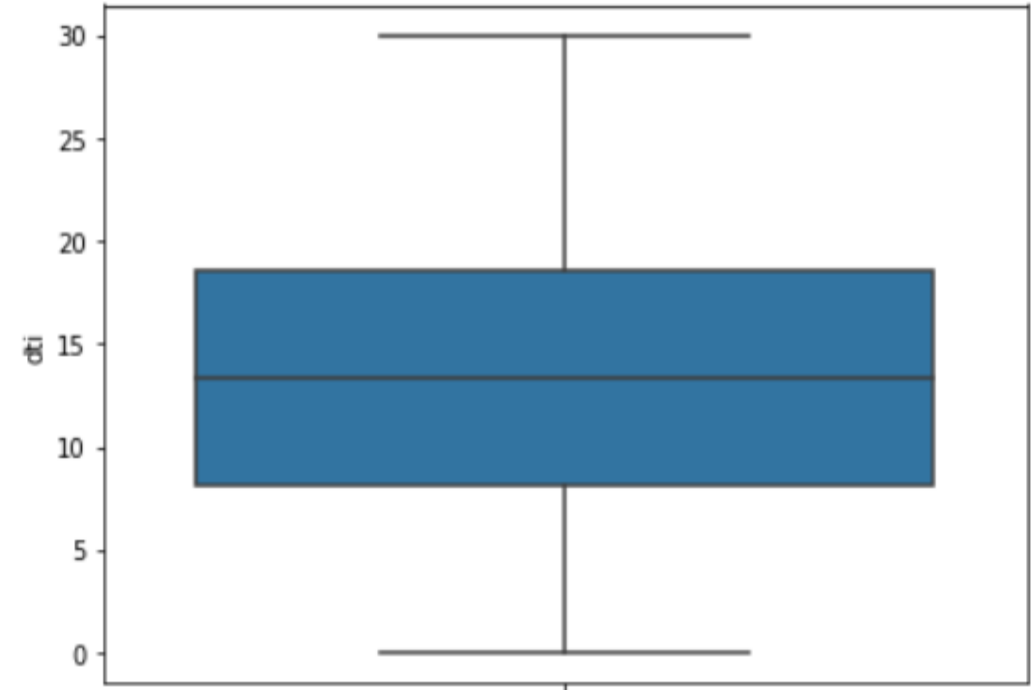
- ✓ Application type has same value this shall be removed. i.e if any column has only one value then drop it.
- ✓ 'mths_since_last_delinq' : Since 65% of data is NaN, which is high, so dropping this column.
- ✓ Convert dates to python datetime object.

Loan Company Issuing Loan per year



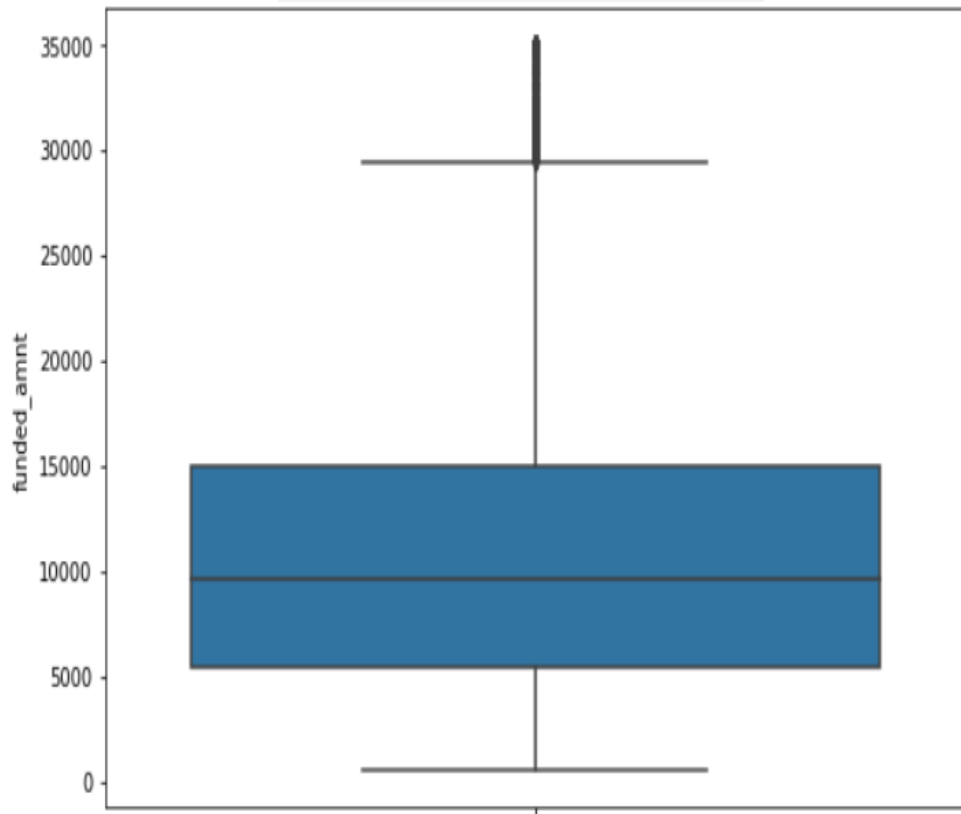
There is a considerable increase in the number of loan issued from the year 2007 to 2011.

Debt to Income

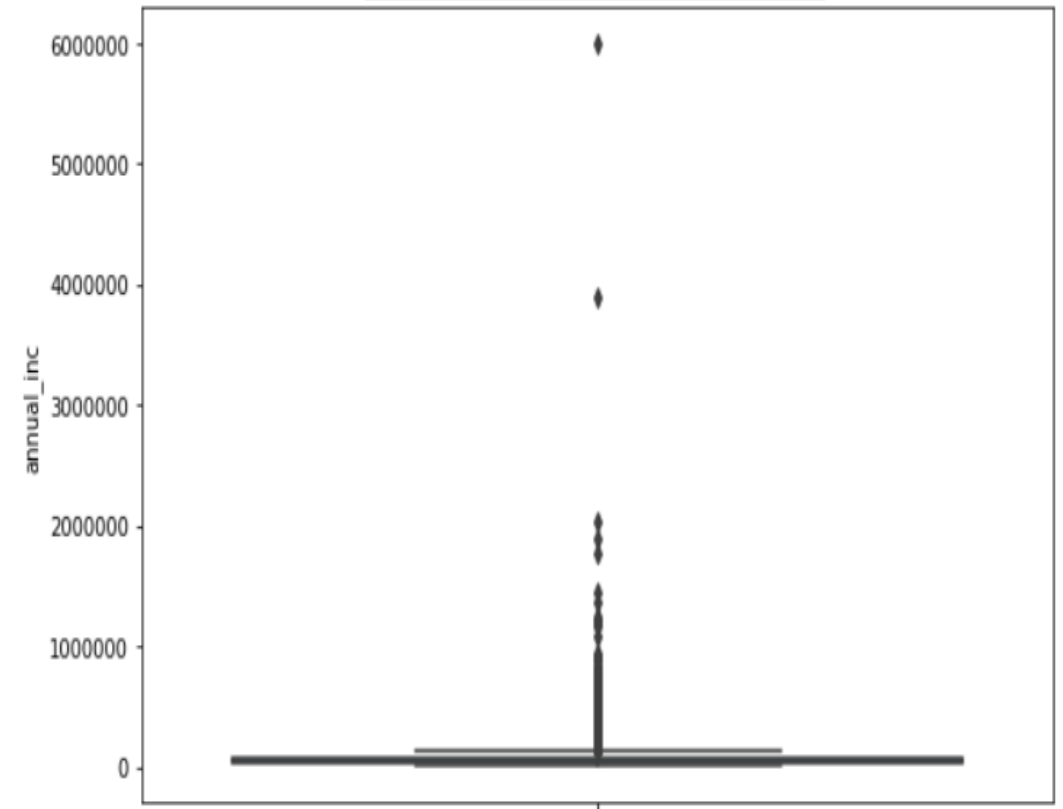


We Can see that max Debt to income ratio in provided dataset is 30% which is good. Anything more than 36 Debt to income ratio is considered as risky to provide loan. This suggests debt to income loan company is getting good borrowers.

Funded Amount



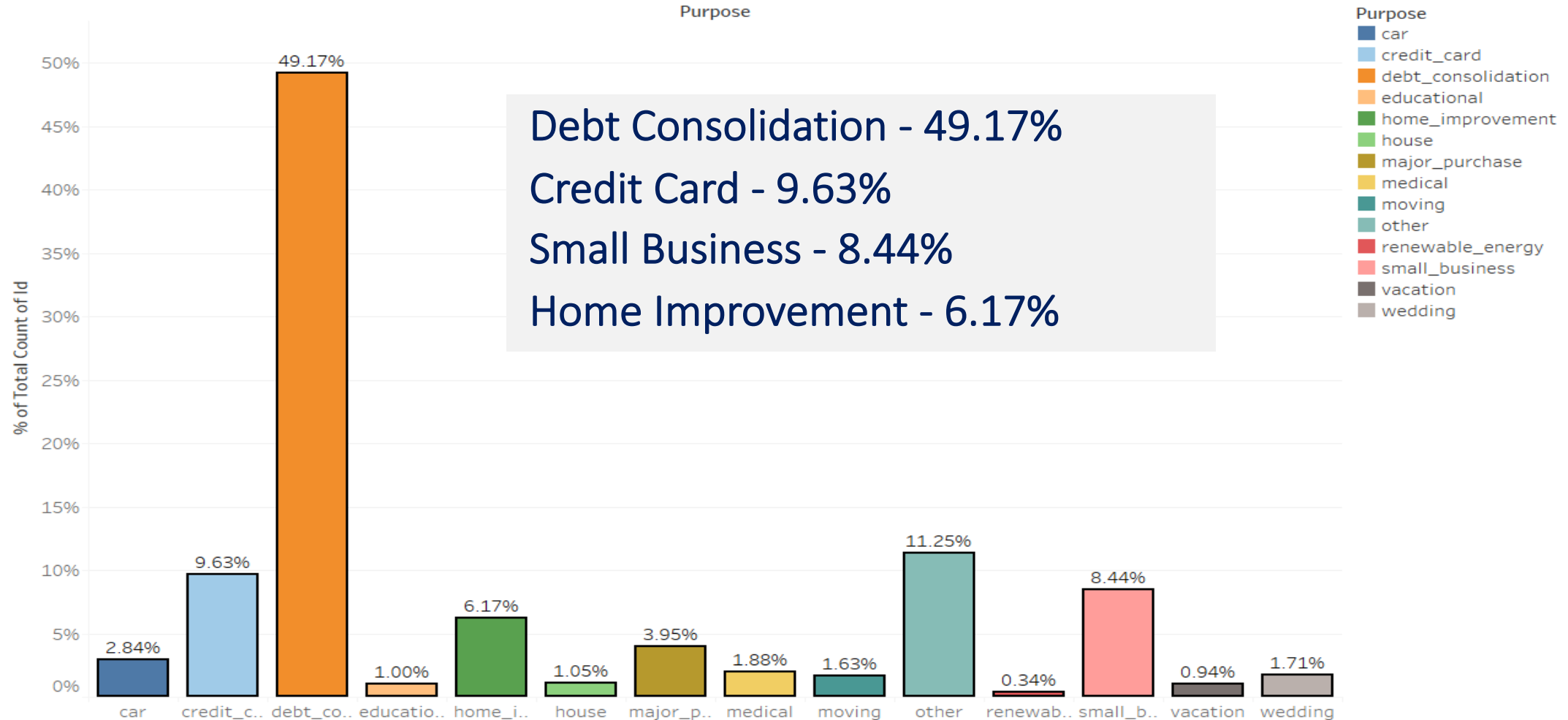
Annual income



Annual income and funded amount has lot of outliers.

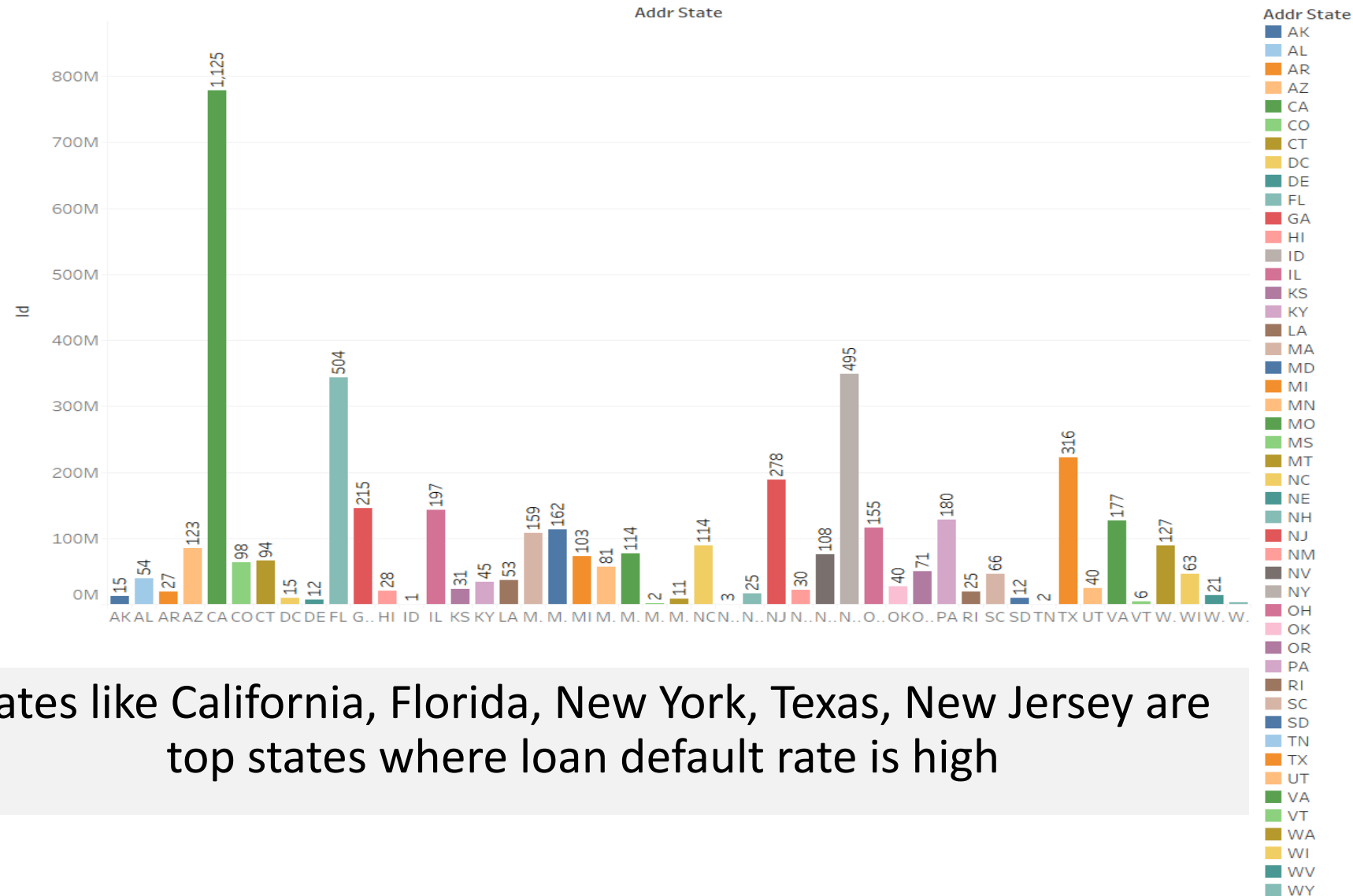
Univariate Analysis

Loan applications reasons which is charged of - Provided by Borrowers



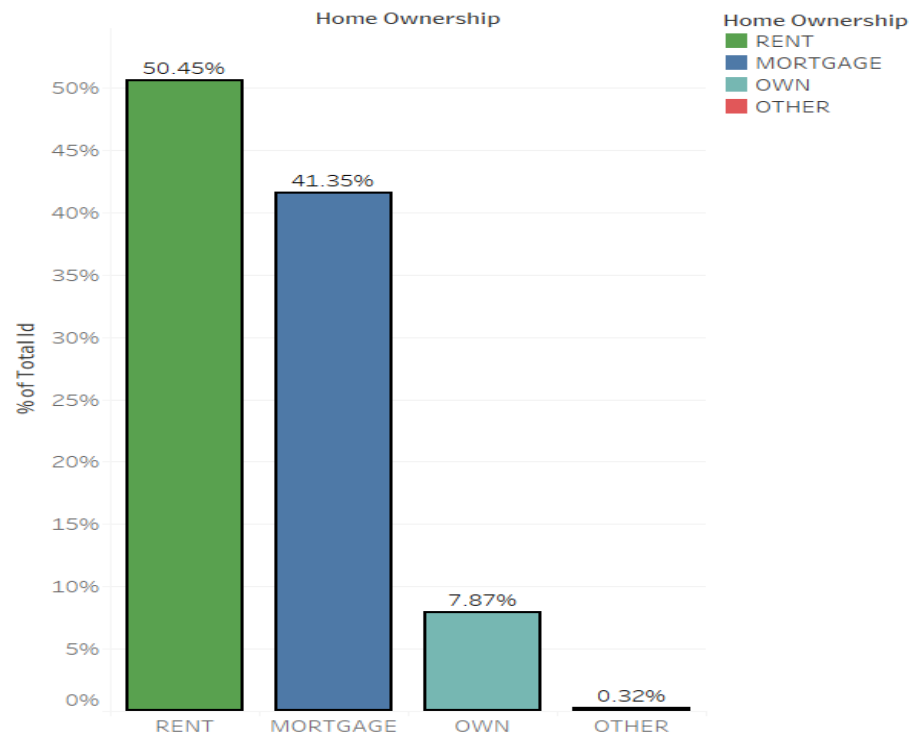
Univariate Analysis

State wise break up of charged of loans



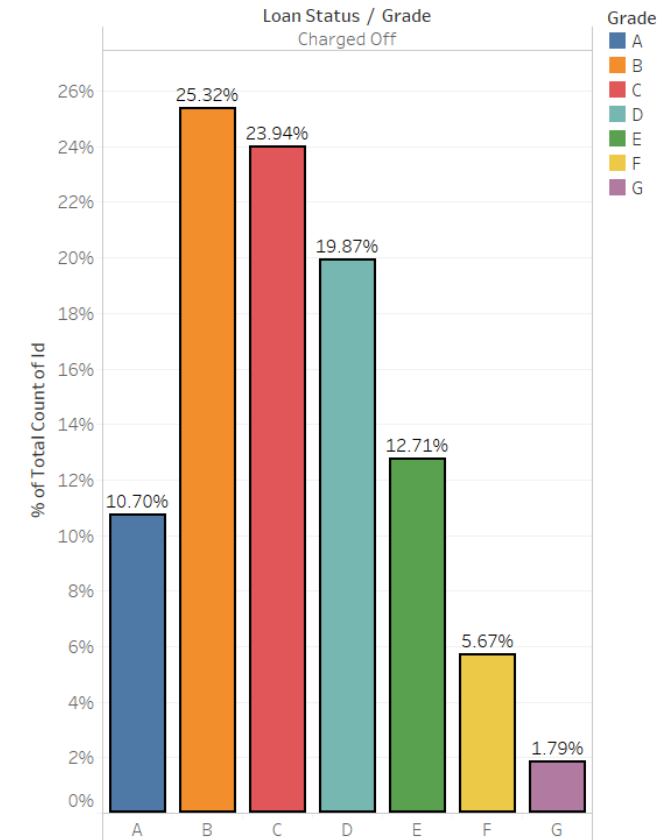
States like California, Florida, New York, Texas, New Jersey are top states where loan default rate is high

Defaulted Borrower - Home Ownership Status



Nearly 92% of the applicants who defaulted the loan are either living in mortgaged or rented houses.

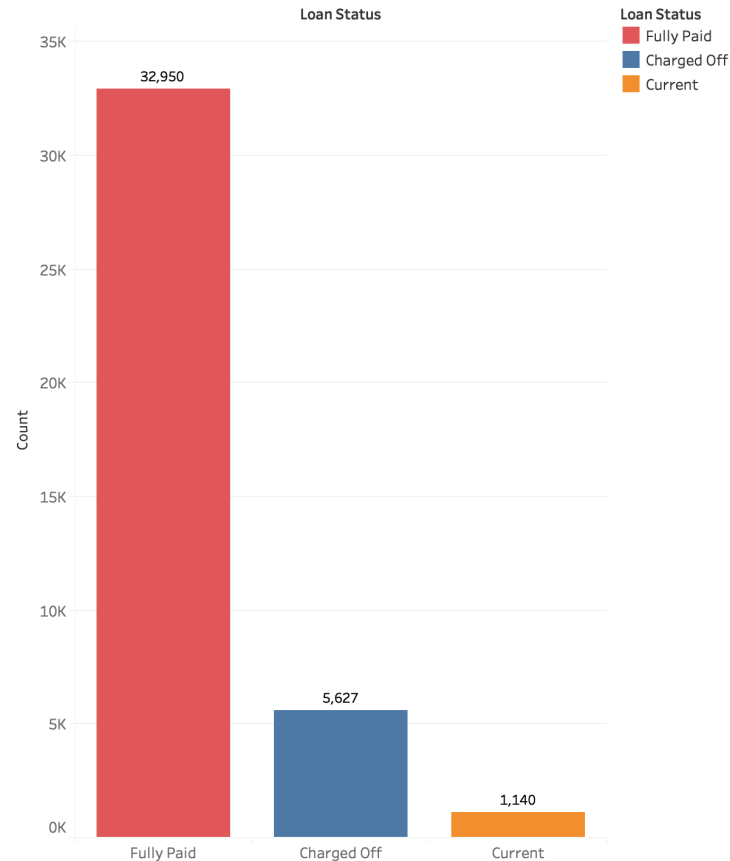
Loan Company grades vs Charged off



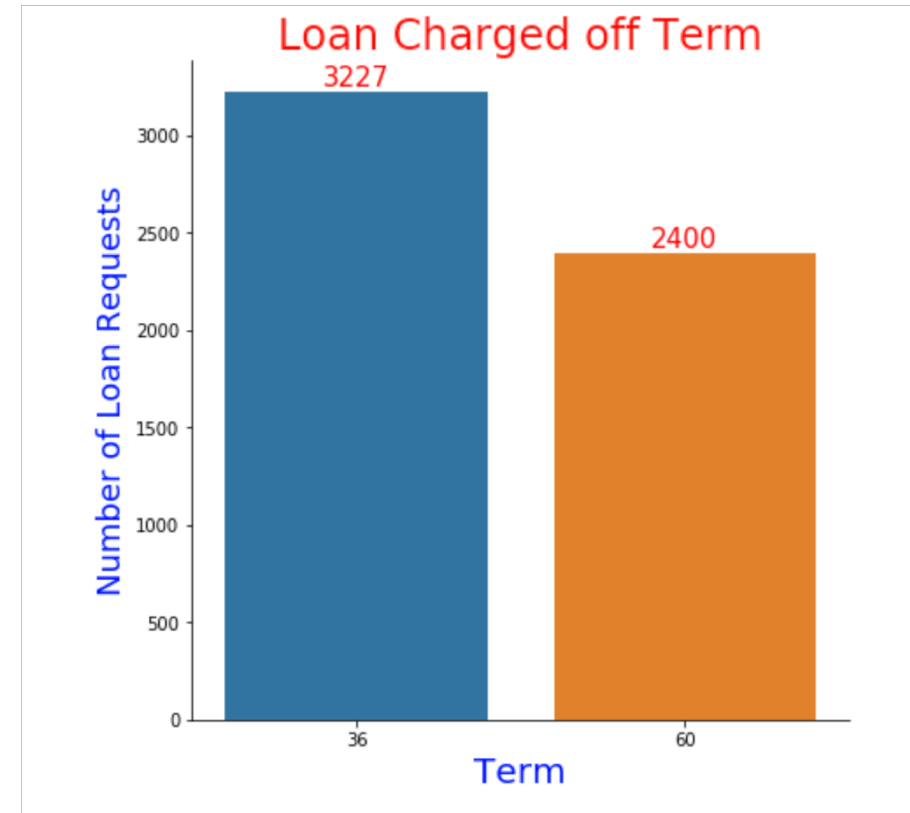
Grade B is having the highest number of loans, followed by Grade A, and the number decreases continuously from grade C all the way to grade F.

Univariate Analysis

Plot of Loan Status



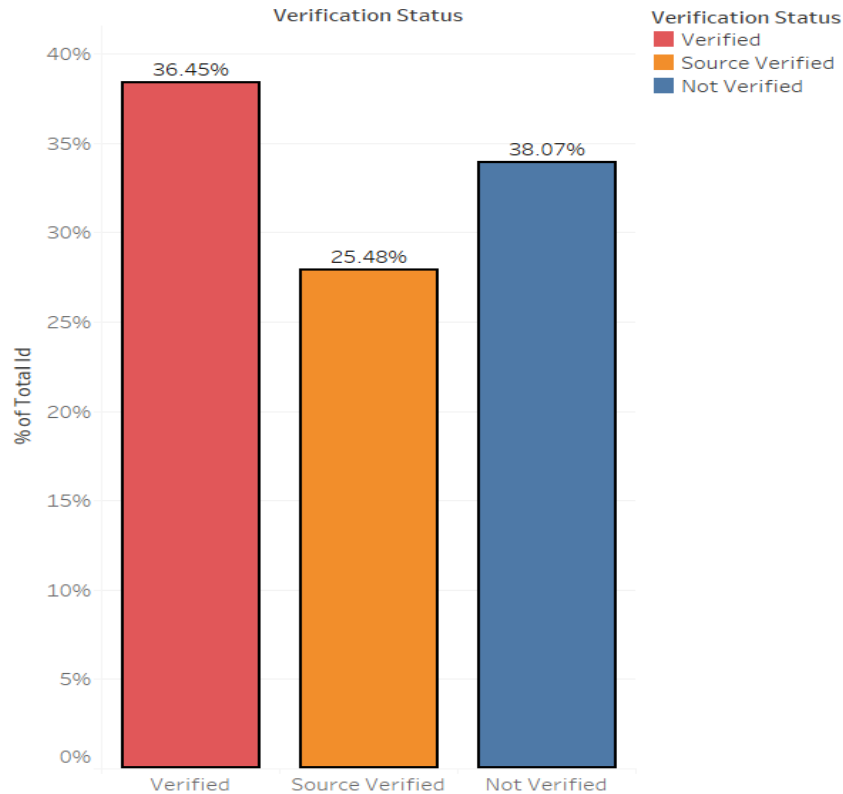
Around 14.8% of the applicants have Charged off (defaulted).



Most of the loan applications are for 36 months tenure.

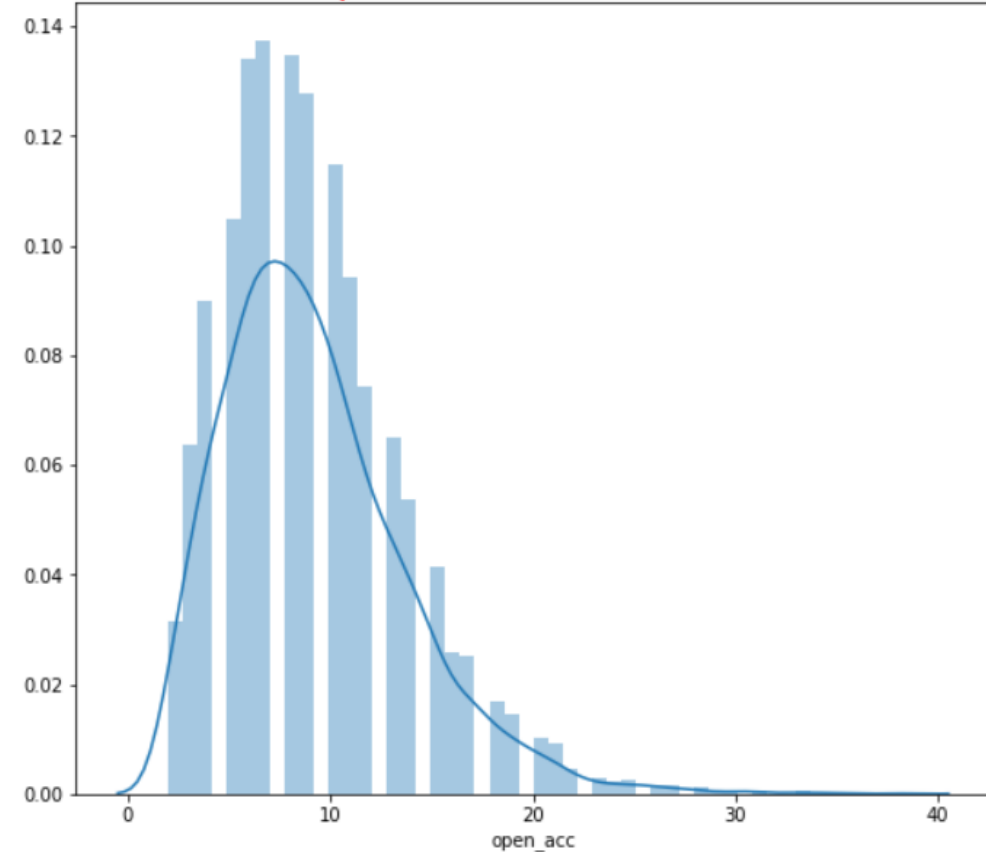
Univariate Analysis

Loan Application Verification Status which are defaulted



Most of the loan applications are not verified and the loans are given without the verification process getting completed.

Number of Open Accounts which defaulted



Borrowers who defaulted account has multiple open accounts

Univariate Analysis-Observations

❖ Grade

- ❖ 76% loan applications are graded in Category A, B, C.
- ❖ Grade B is having the highest number of loans, followed by Grade A, and the number decreases continuously from grade C all the way to grade F.
- ❖ 60% loan applications are graded in Category A, B, C defaulting the loan.
- ❖ Loan Company need to check their Grading system because 60% defaulters categorized as good borrowers.

❖ Home Ownership

- ❖ Majority of the borrowers are either living in Mortgaged or Rented houses.
- ❖ Majority of the loan applications are for 36 months tenure.

❖ Applications Verification

- ❖ Loan Company need to increase source verification which is only 25.48%
- ❖ Not Verified and Verified loan default contributes to 74.52% which is very high.
Loan Company can do additional source verification.

Univariate Analysis-Observations

❖ Other Observations -

- ❖ There is a considerable increase in the number of loan applications from the year 2007 to 2011.
- ❖ Highest number of loans are issued in the year 2011.
- ❖ Most of the applicants are applying for loan for debt consolidation purpose, followed by credit card, other, home improvement, major purchases. .
- ❖ Around 44% of the applicants are from the following states: 'CA', 'NY', "FL' , 'TX' and 'NJ'.
- ❖ The number of applications for loan are higher with the people having lesser years of experience and decreases as the years of experience increases. Note: The value in 10+ years of experience is more because the number of applications are spread over a larger group of people.

Loan Default reason Bi-variate Analysis -1 UpGrad

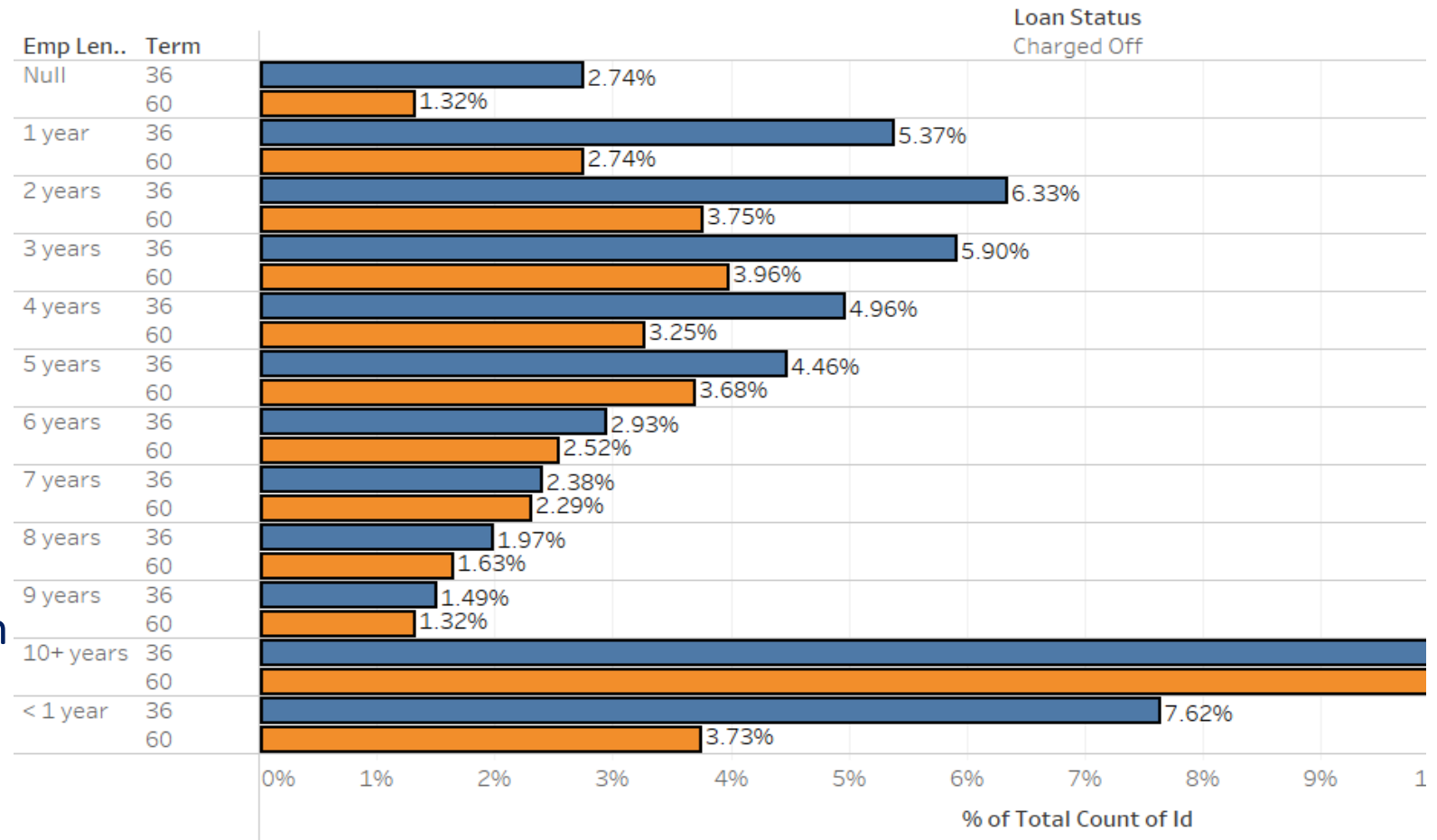
Loan Term and Years of employment

- ❖ Year of experience who defaulted the loan

Year of exp	Loan Term	
	36 Months	60 Months
< 1 Year	5.37%	2.74%
2 Year	6.33%	3.75%
3 Year	5.90%	3.96%
4 Year	4.96%	3.25%

- ❖ This clearly shows lesser the number of years of experience we need to provide higher loan term.

Loan Term vs Years of experience

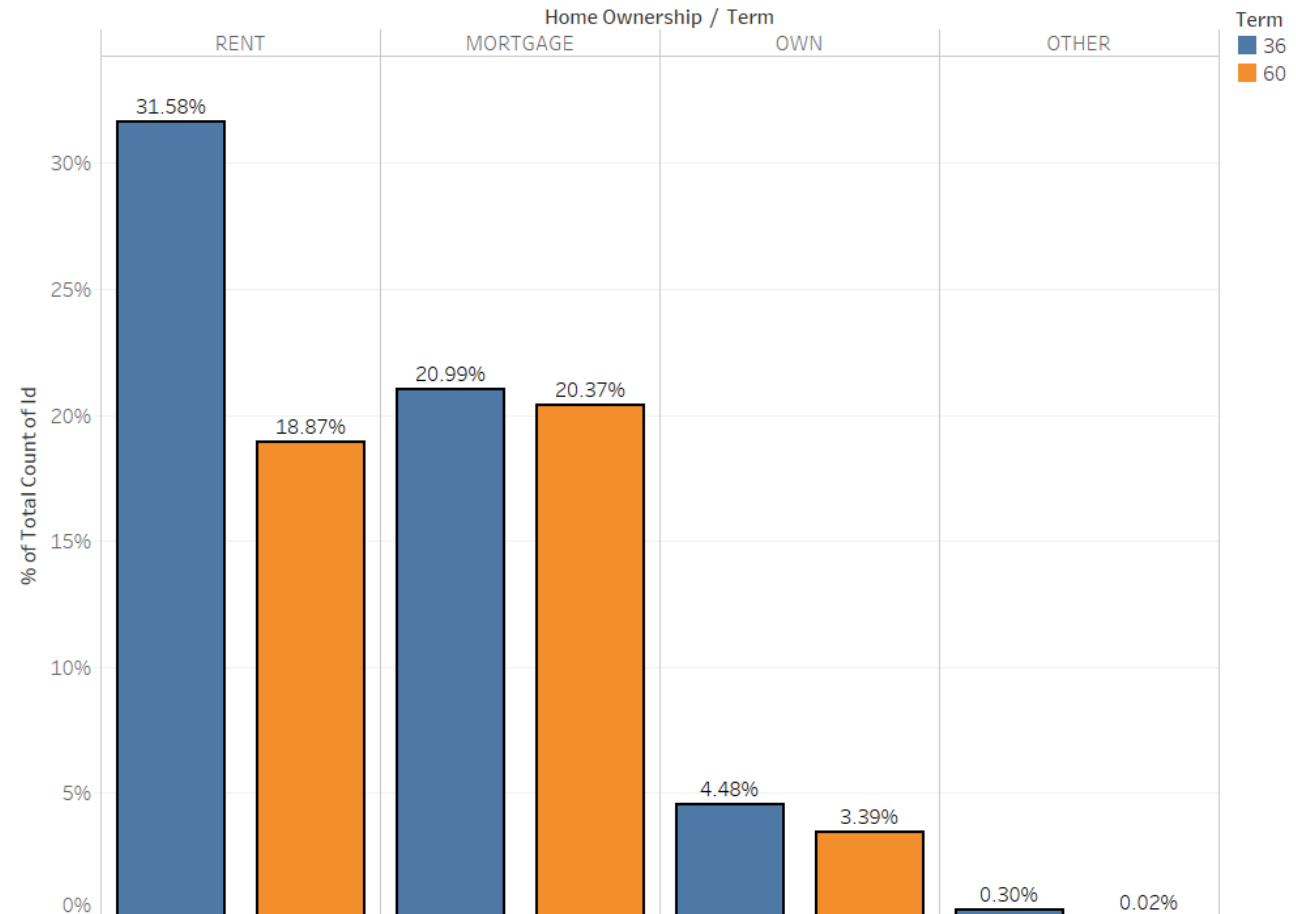


Loan Default reason Bi-variate Analysis -2 UpGrad

Analysis of Home Ownership

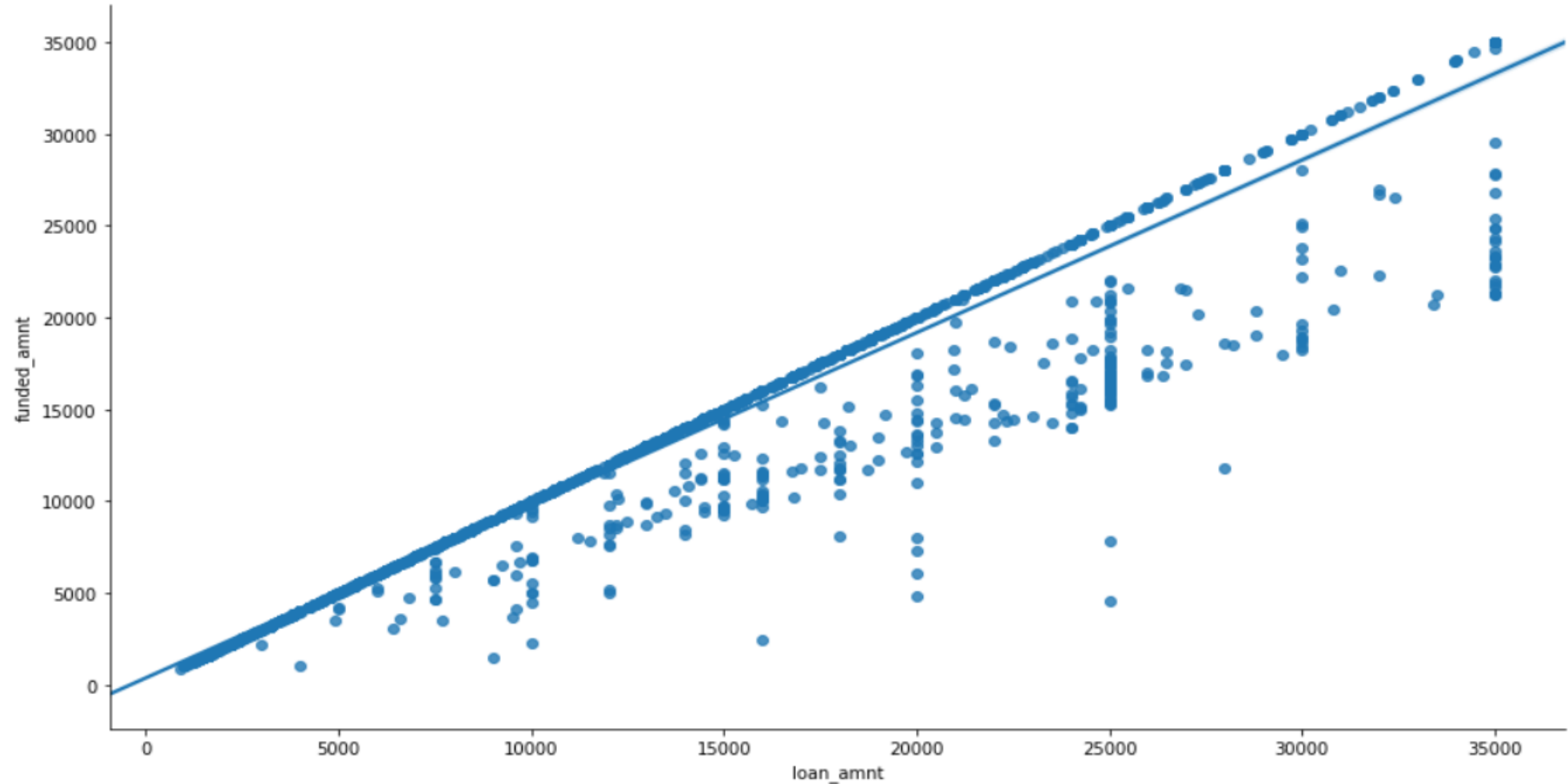
- ❖ If borrower is staying in rented home & term of loan is 36 Months, chances of default is 31.58 % against 18.87% for 60 Months term.
- ❖ This clearly shows that if borrower is staying in rented home, offer him higher loan term to avoid defaulting.

Home ownership which is charged of at time of loan application



Bi-variate Analysis -3

Analysis of Loan and Funded amount



There is high correlation between Funded amount & Loan amount.

The analysis shows that Following factors can have impact on risk of loans being defaulted :

☐ Loan Purpose

- Loans for Debt Consolidation and Credit Card make up to around 60% of the total loans. Loans taken for the purpose of Small Business are the riskiest having the highest rate of default.

☐ Grade

- There is a direct relation between the grade assigned and the defaulting rate. Around 94% of the B Grade loans were fully paid.
- This figure gradually reduces as the Grade degrades from B to G.

☐ Loan Tenure

- Loans granted for the higher tenure(60 months) are likely to be more risky than the loans granted for lower tenure.

☐ Employment Experience

- Applicants having lesser years of employment experience are the ones who take higher number of loans. This is mainly because of the lesser annual salary offered during the initial phase of the career. As the salary increases, the need for taking loan reduces.

☐ Annual Income

- Applicants having lesser salary are more likely to default on the loan. The applicants who have fully paid the loans are having higher amount of average annual salary.