

Gramener Case Study

By: Rohan Daniel

The Problem Statement:

Company

Lending Club is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures.

Borrowers can easily access lower interest rate loans through a fast online interface.

Context


Lending Club wants to understand the **driving factors** behind loan default, i.e. the **driver variables** which are strong indicators of default.

The company can utilise this knowledge for its portfolio and risk assessment.

Problem Statement

As a data scientist working for Lending Club analyze the dataset containing information about past loan applicants using EDA to understand how consumer attributes and loan attributes influence the tendency of default

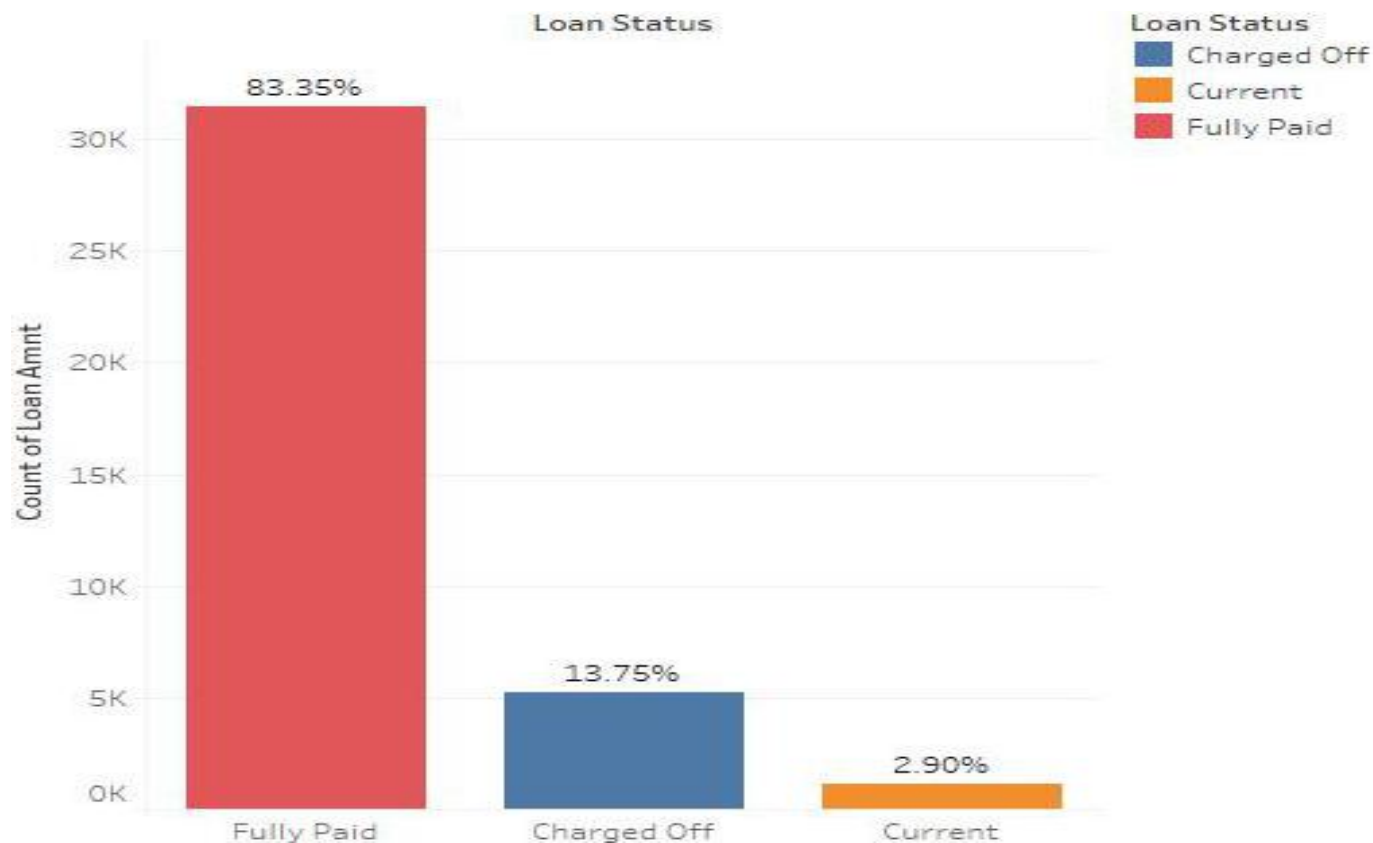
Data Cleaning Steps

- Delete columns: Delete unnecessary columns.
 - Remove outliers: Remove high and low values that would disproportionately affect the results of your analysis.
 - Missing values: Treat missing values with appropriate approach.
 - Duplicate data: Remove identical rows, remove rows where some columns are identical.
 - Filter rows: Filter by segment, filter by date period to get only the rows relevant to the analysis.
- 

Analysis

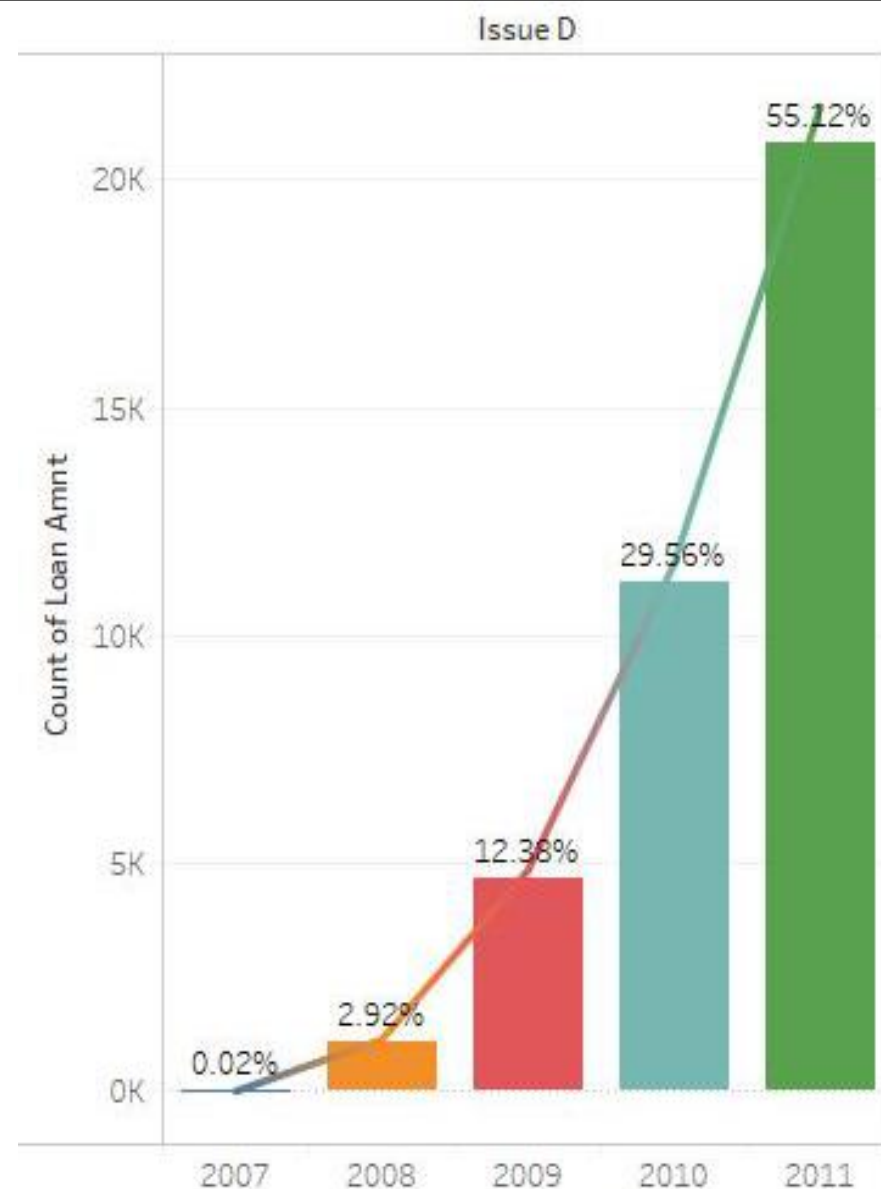
- The essence of the whole project is to analyze and understand how consumer attributes and loan attributes are influencing the tendency of defaulting.
- We performed data cleaning and preparation on the Loan dataset:
 - Created two new columns:
 - First called Balance Payment in which we know how much amount is pending to particular loans.
 - Second made year and month columns of all the datetime columns.
- During analysis we have created:
 - Performed the Univariate Analysis, made histograms and Bar charts to check out the distribution of all the driver variables
 - Performed the Segmented Analysis, made box plots to detect the Outliers
 - Performed the Bivariate analysis to understand how different variables interact with each other

Overall Loan Status

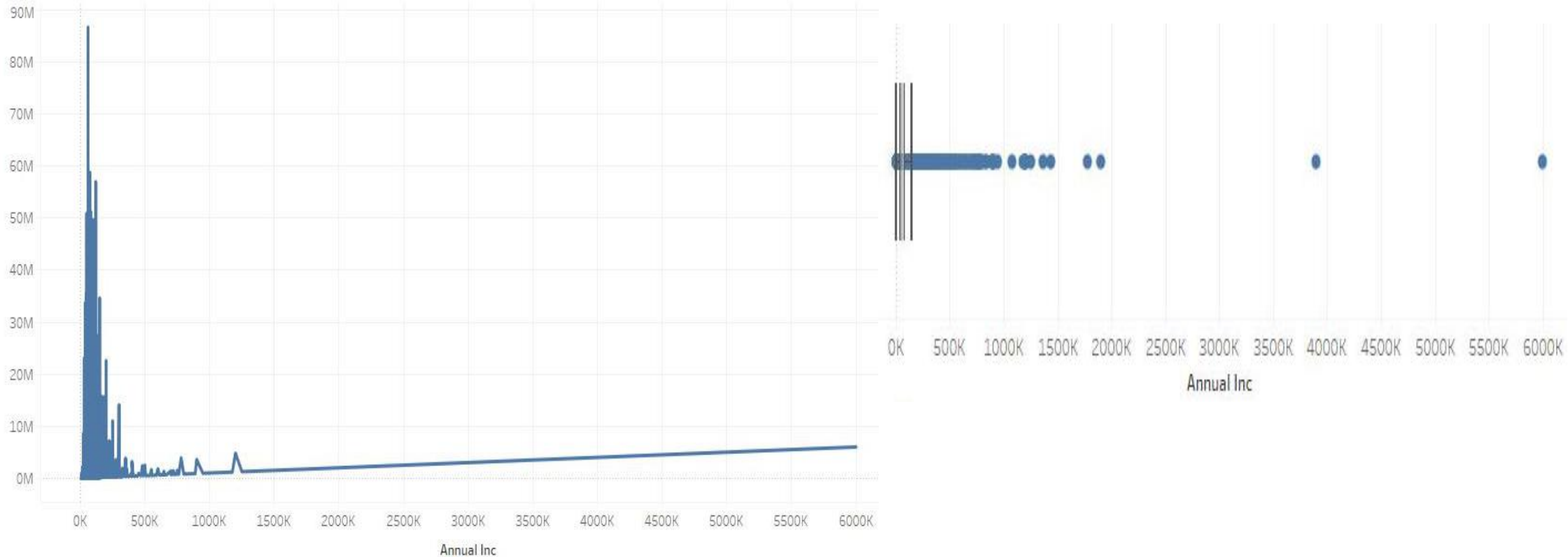


Conclusions:

1. If we see above plots 83.35% of the loans have been paid.
2. Around 13.75% of the loans have defaulted in this dataset.
3. Over the years there has been significant increase in number of loans.



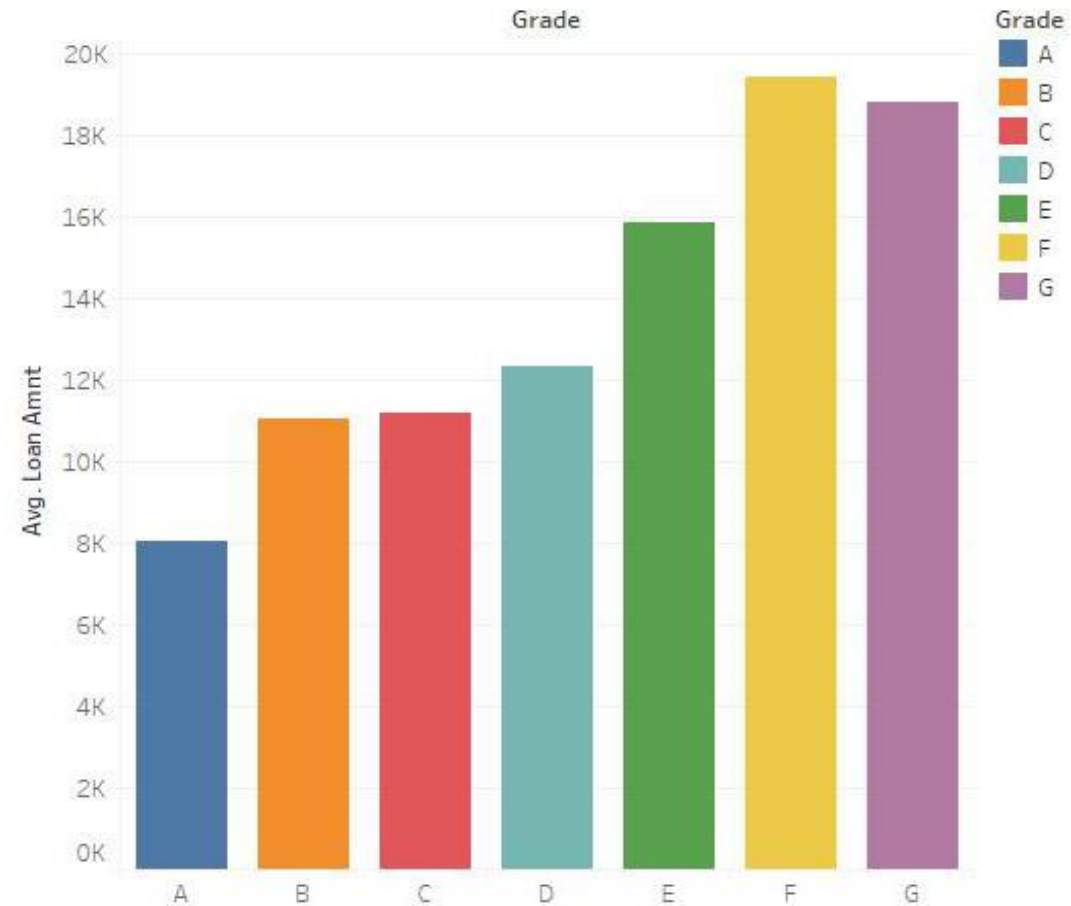
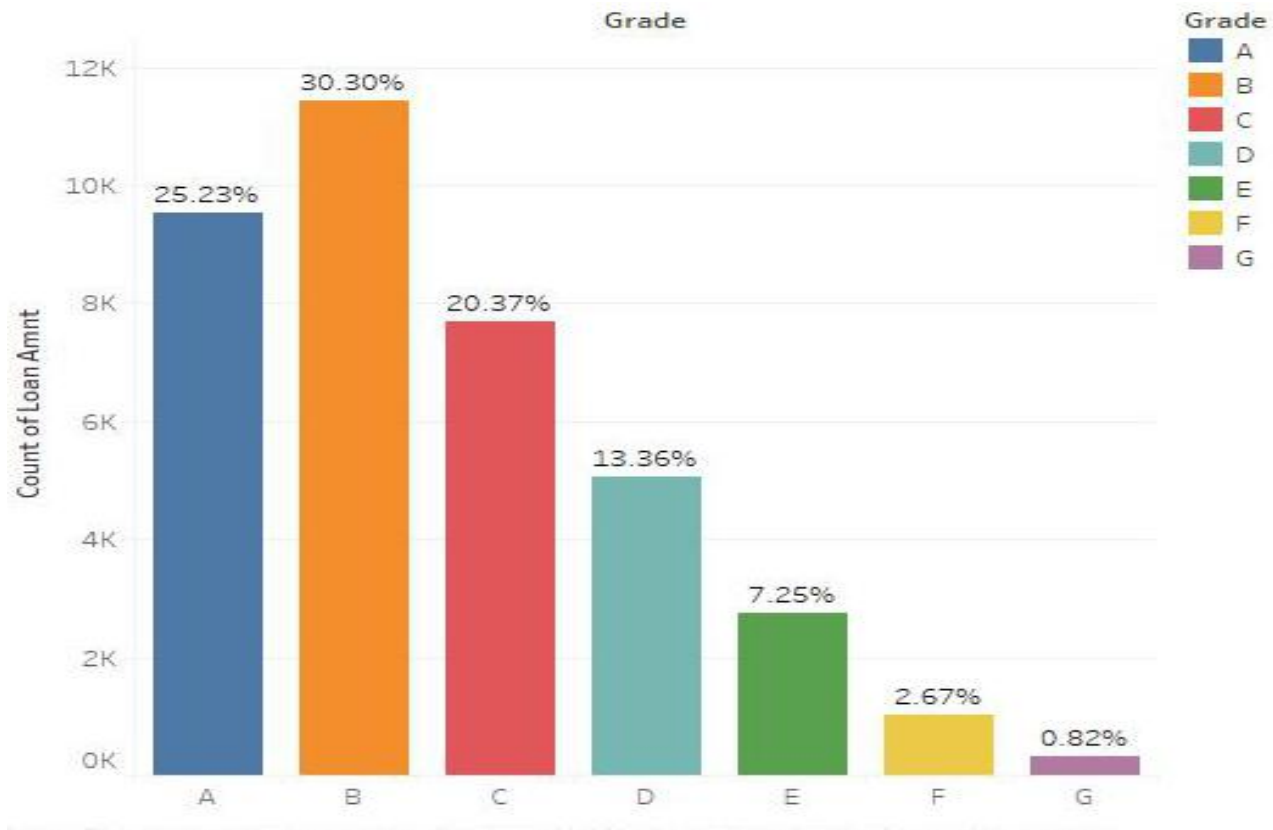
Income Analysis of the Applicants



Conclusions:

1. From the above plots we say that most of loans have income between 0 to 5 lakhs .
2. There are only 2 loans applicants having income more than 30 lakhs.

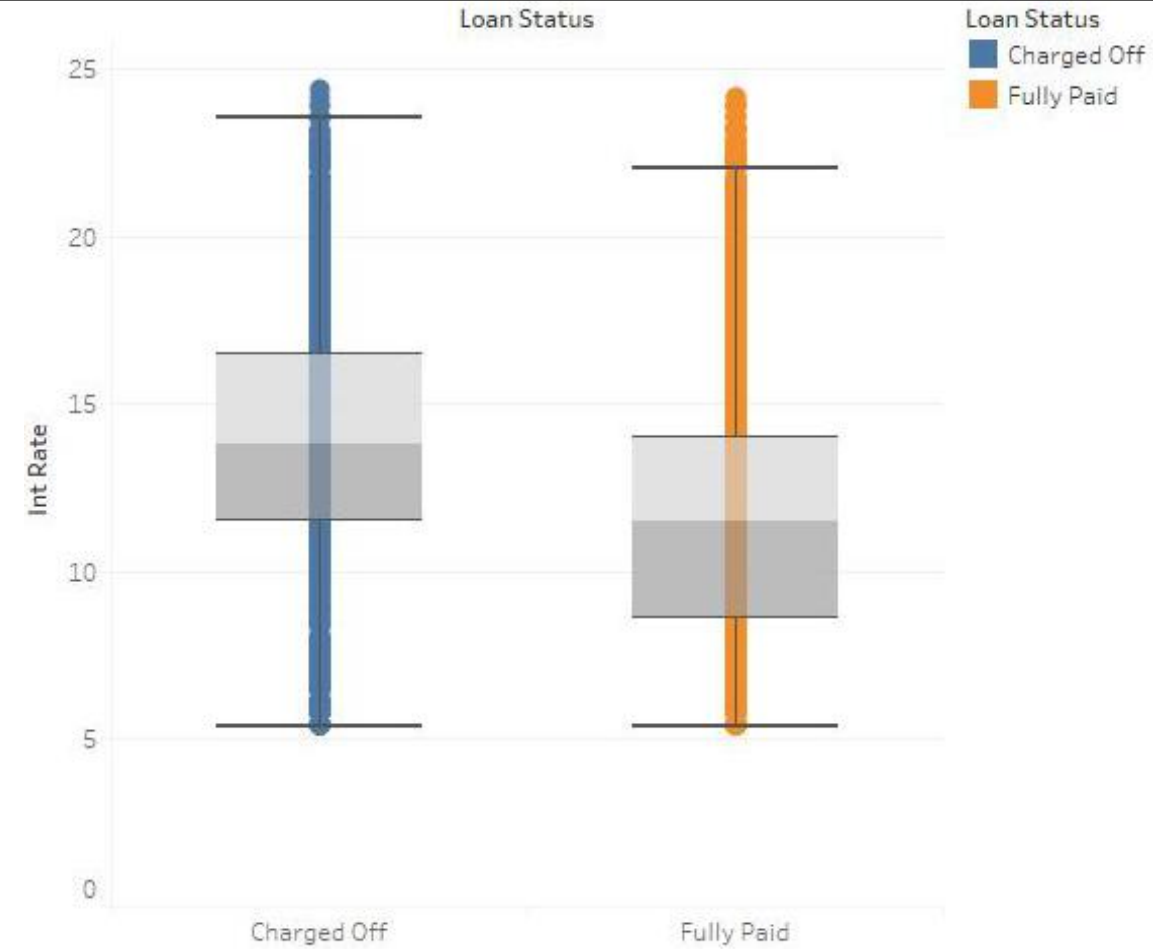
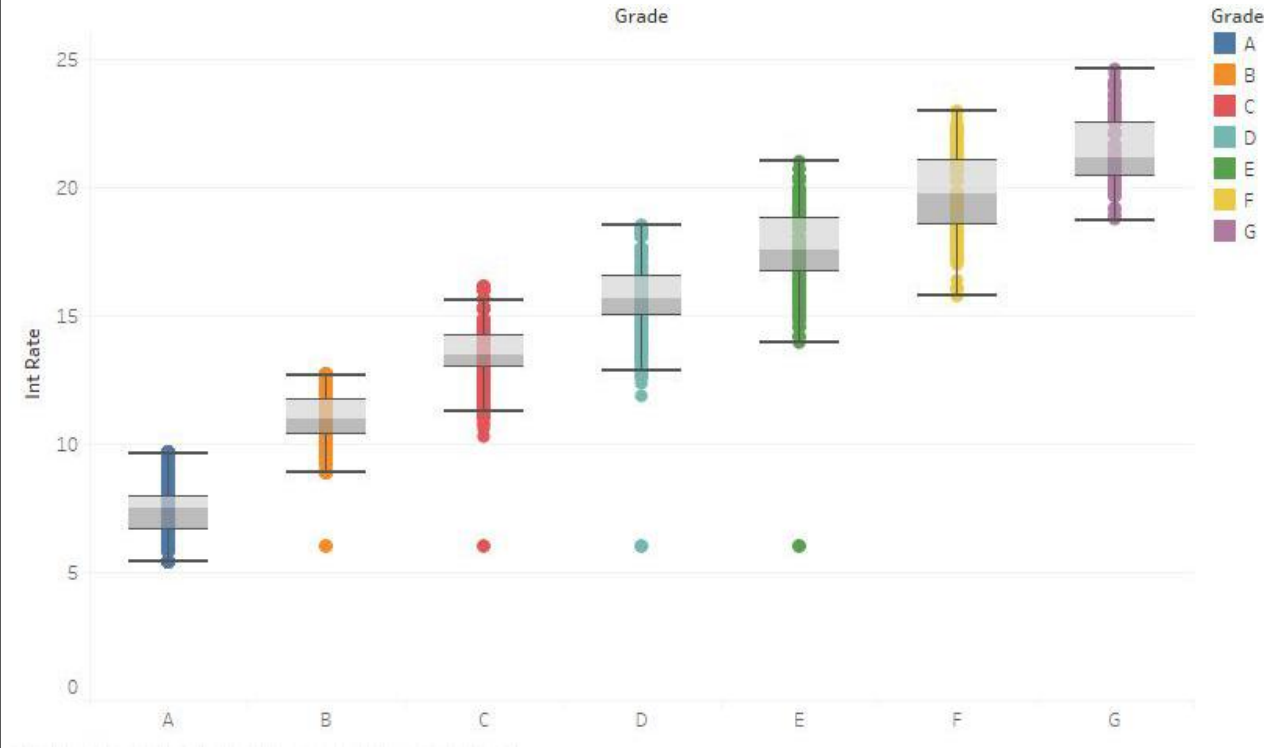
Grade Analysis



Conclusions:

1. If we see above plot most have of the loans have high grades like A and B. Around 46% of the loan are highly graded in this dataset.
2. If we see the plots most of the default loans are been occupied by the low graded applicants.

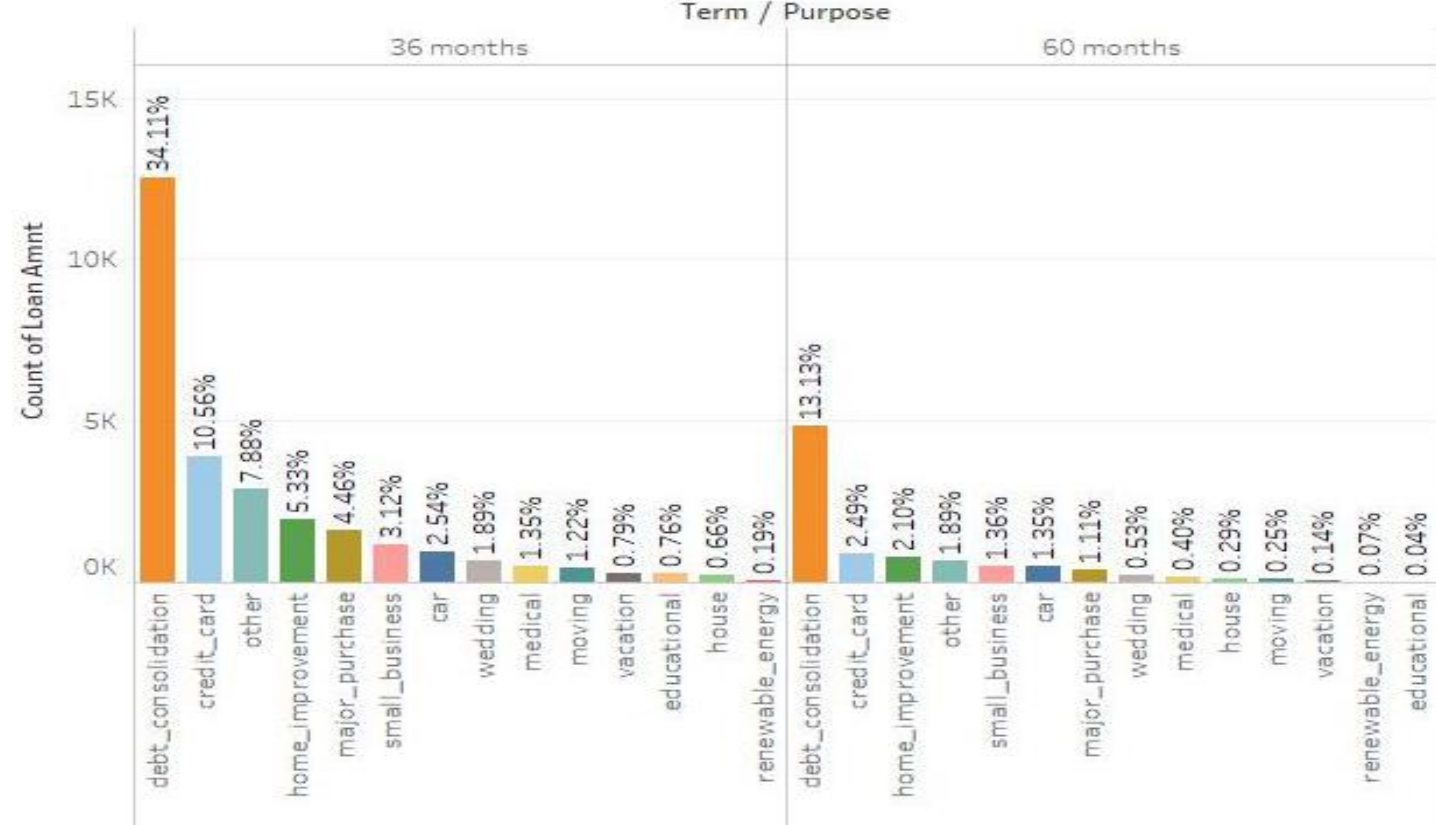
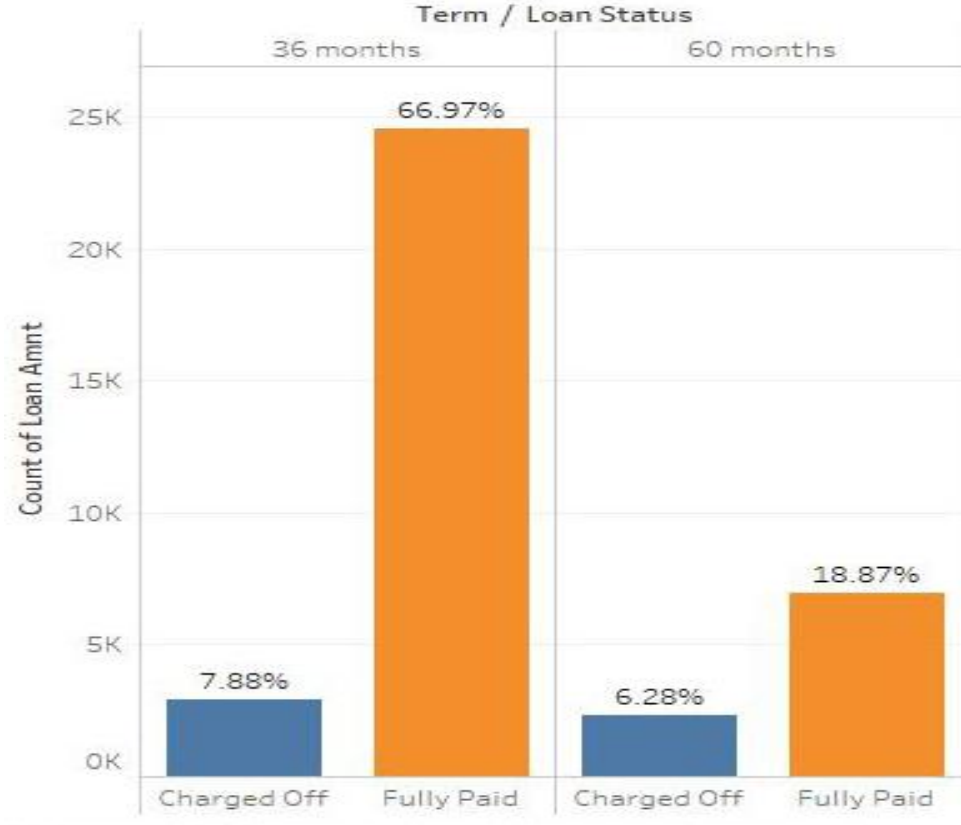
Interest Rate Analysis



Conclusions:

1. From the plot low quality grades have high interest rate as compare to high quality grades.
2. In the earlier slide also we he seen higher default rate were from low graded loans because of high interest rate in the loan, higher the interest rate more will be the default loans.

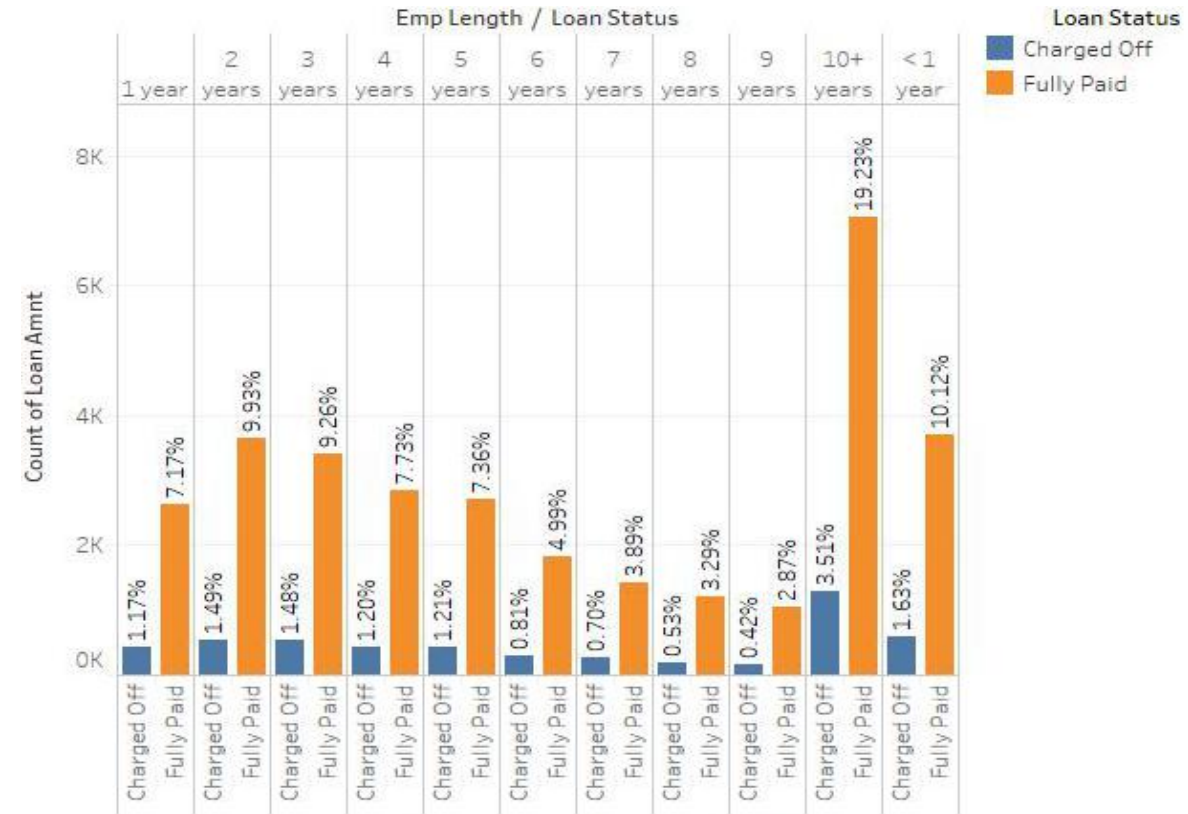
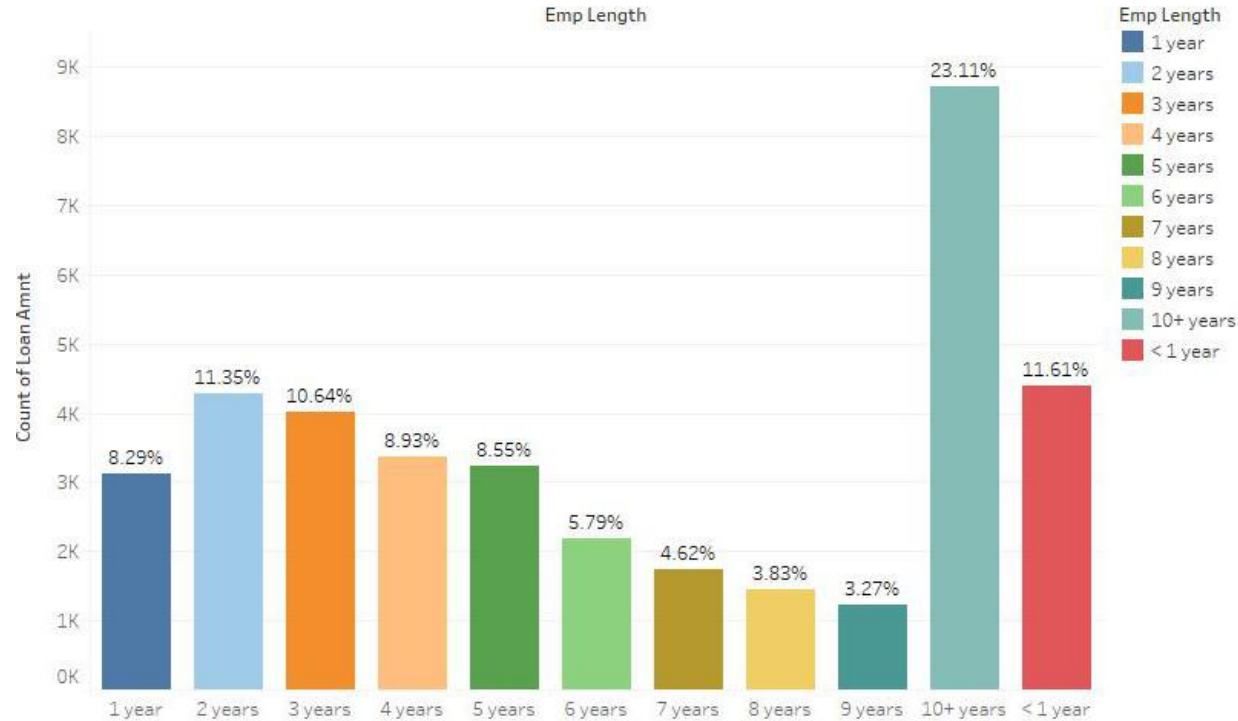
Analysis on Term and Purpose



Conclusions:

1. Maximum number of the loans are granted were for 36 months.
2. Loans granted for 36 months are tend to default more as compare to other values.
3. Debt Consolidation is the highest in the number in both the term as purpose.

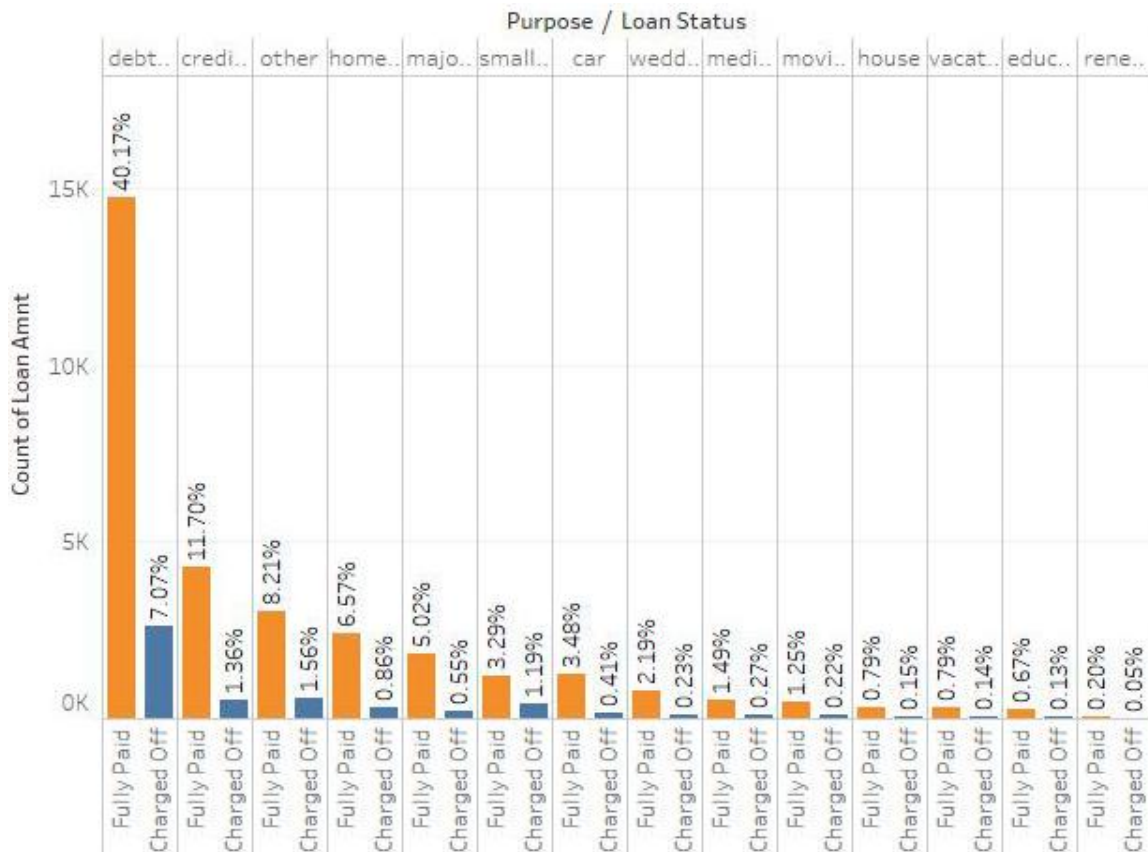
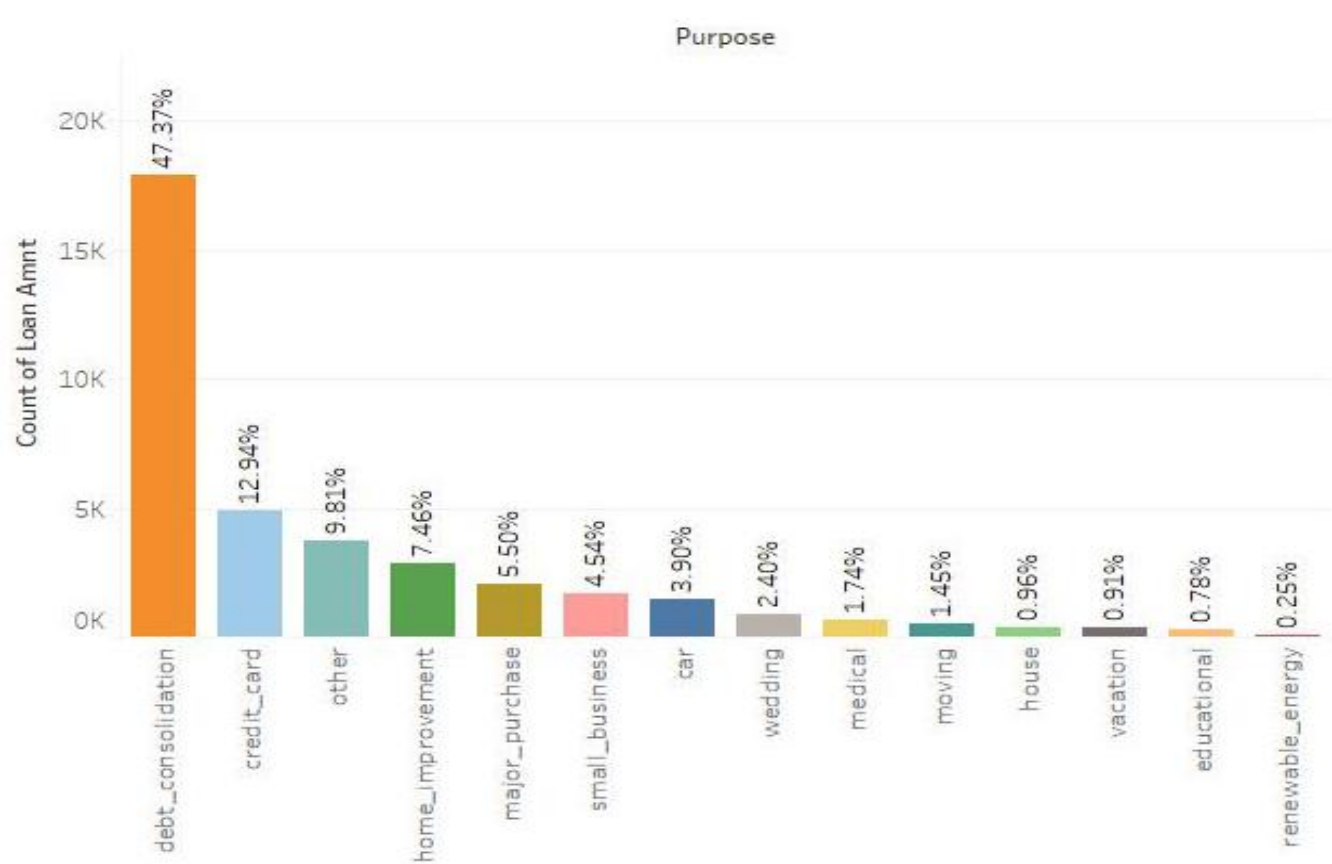
Applicants Work Experience Analysis



Conclusions:

1. From the plot we can see that majority of the loans who are applied have more than 10 years work experience.
2. But if we see in the next plot loans having more than 10 years work experience are defaulting more as compare to other work experience loans, so people with more than 10 years work experience should be given less number of loans.

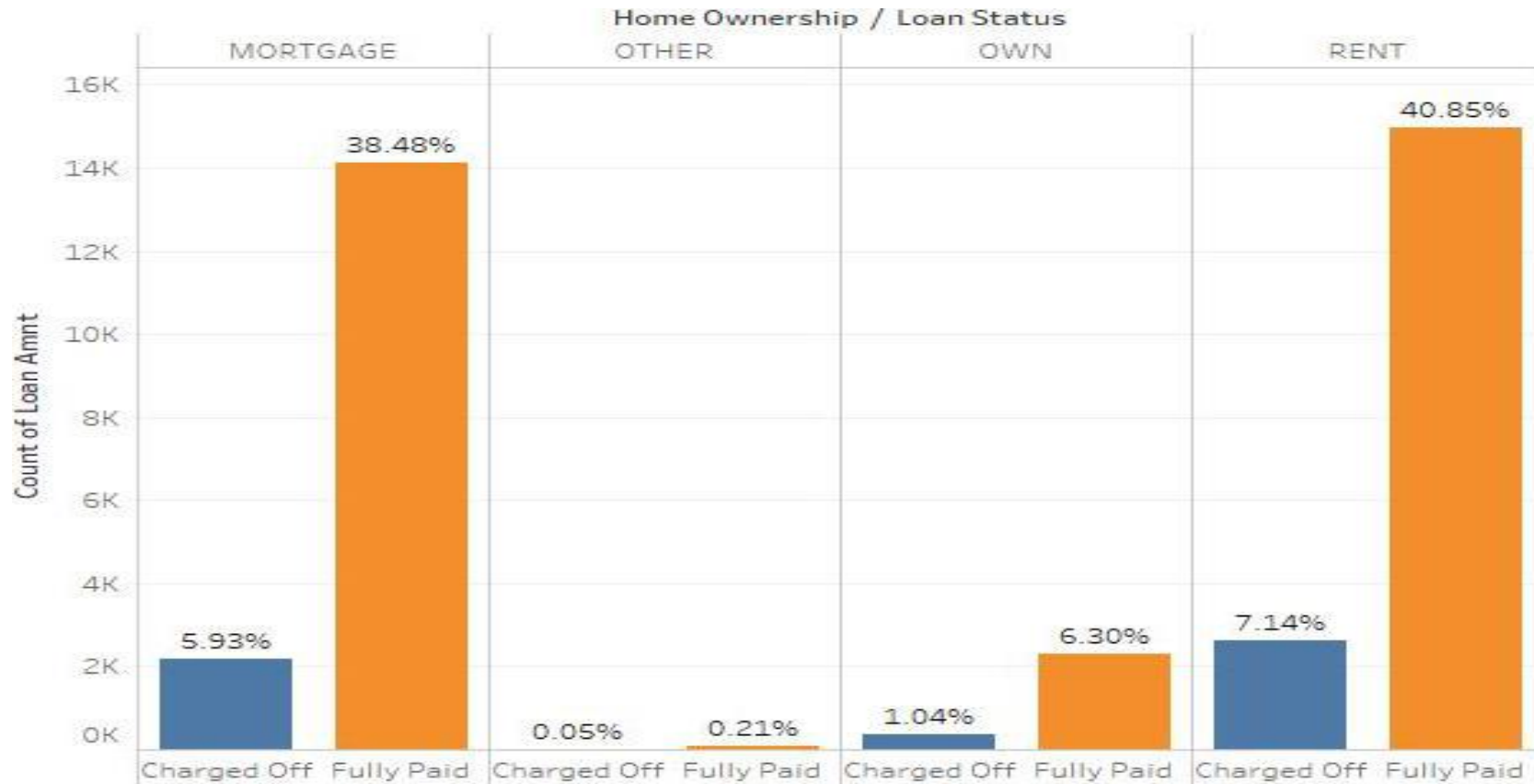
Analysis on Purpose for applying Loan



Conclusions:

- 1. Maximum number of the loans are granted were for Debt Consolidation.
- 2. Maximum number of the defaulters are also from Debt Consolidation.

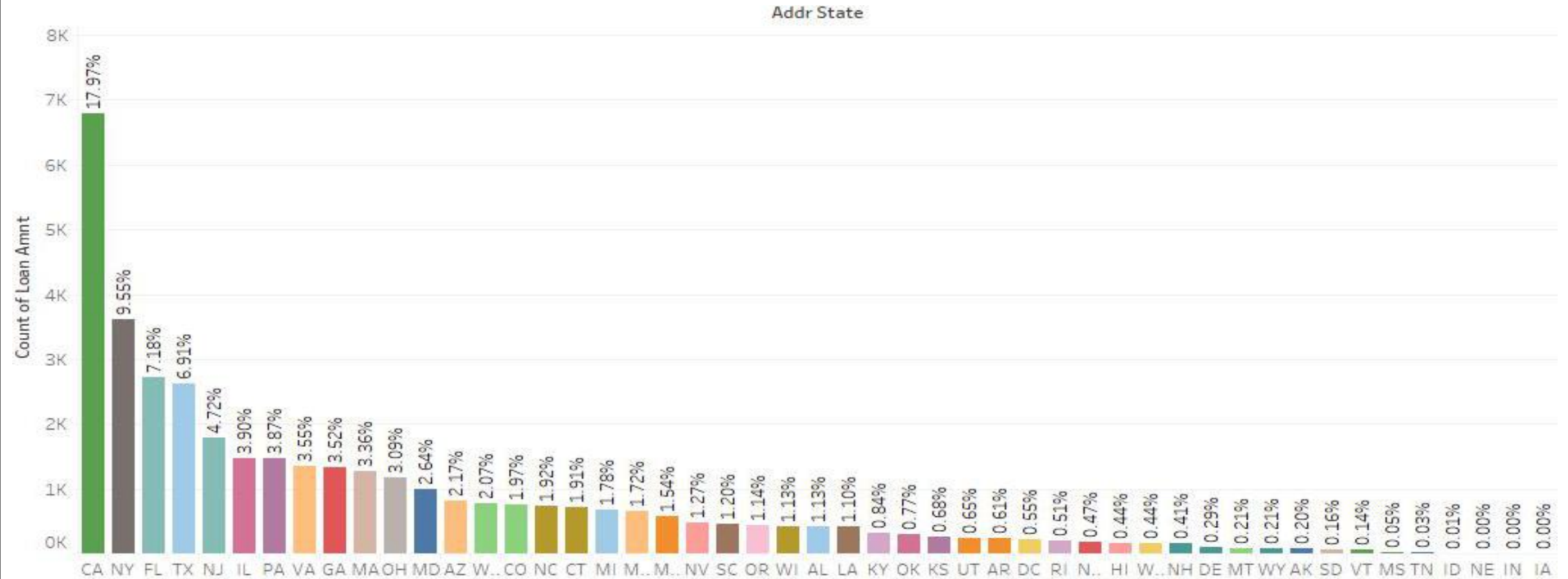
Analysis on Home Ownership with Loan Status



Conclusions:

1. Loans having Mortgage and Rent as their Home Ownership have almost similar tendency to default as compare to other values present in the column.

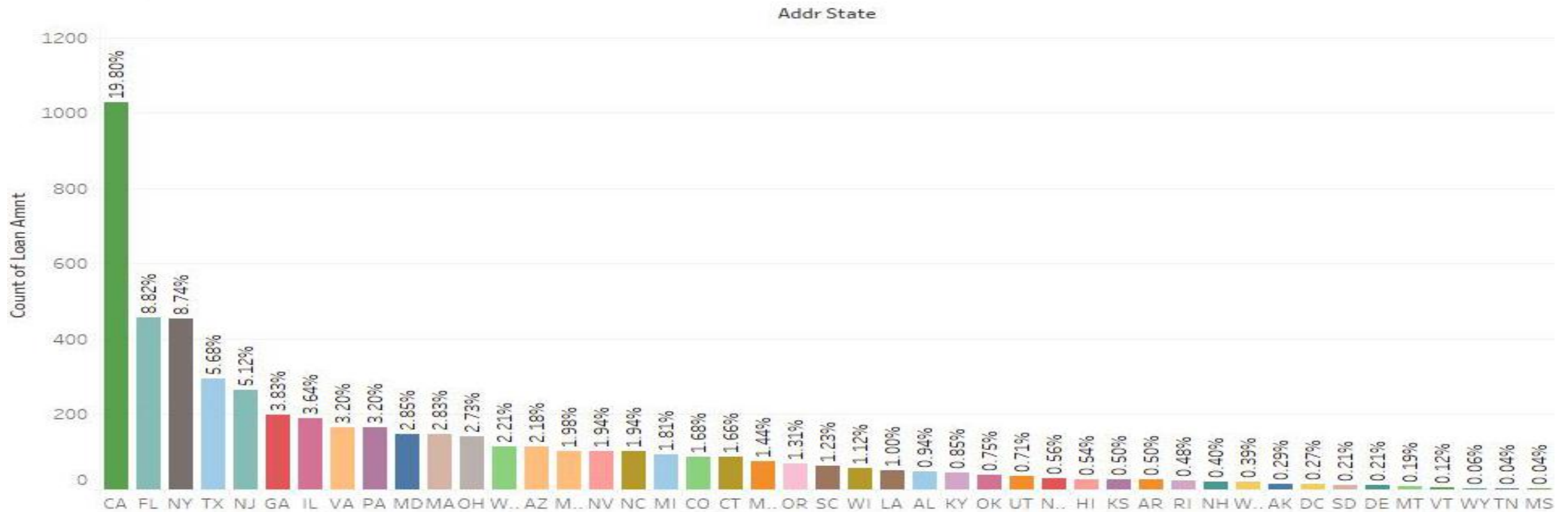
Loan Applicants According to State Analysis



Conclusions:

1. Most of the loan applicants are from the state Carlifornia .

Loan Defaulter Applicants According to State Analysis



Conclusions:

1. Most of the defaulters are also from the state California .

Final Conclusions / Recommendations:

1. Low Grade loans are highly tends to default. Less loans should be given to the loans having low grades, it might be because of higher interest rate.
 2. Loans having higher interest rate are more to default. History of the applicants should be checked by company before giving loans at higher interest rate.
 3. Applicants having work experiene more than 10 years should be given less loan amount as there are more likely tend to default in terms of loan and on applicants having home ownership as Rent an Mortgage are having high tendency to default.
 4. Purpose having Debt Conslidation should be given less loan amount as they have high tendency to default.
 5. Careful Examination should be done for applicants belonging to Carlifornia state, as applicants from that state are more likely to default in term of loans.
- 