

PGDDS – Sep 2018

Gramener Case Study

Group Members:

1. Ankit Saini
2. Ankita Girdhar
3. Arihant Sethia
4. Aditi Nag

- **PROBLEM STATEMENT**

- To analyze the driving factors (driving variables) behind loan default i.e. variables which are strong indicators of default.
- Based on these factors company is able to identify risky loan applicants. By reducing these loans they can reduce the credit loss eventually reducing the financial loss.

- **ASSUMPTIONS**

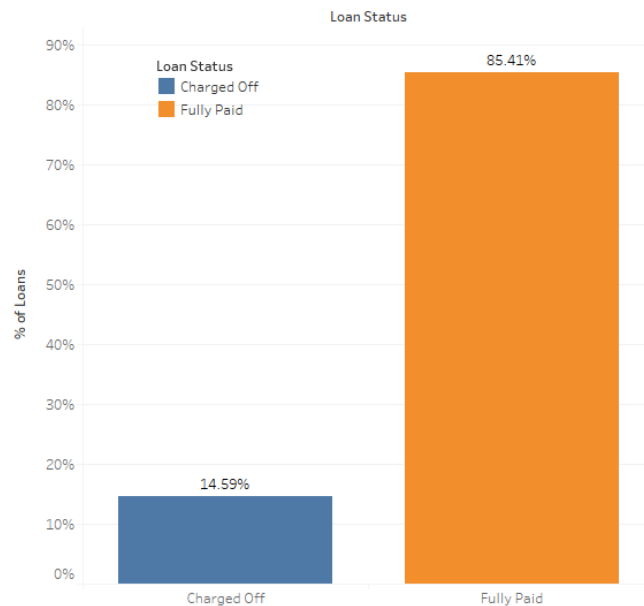
- There are many columns in the dataset which have all values as null or just have a single value populated across it. We have not considered these in our analysis and dropped such columns from the dataframe.

Problem Solving Methodology

1. Data Exploration.
2. Dropping the columns with all nulls or single values.
3. Analysing further and dropping columns which are not relevant for the analysis.
4. Deriving other metrics and variables from given data.
5. Plotting data to visually perform the univariate and multi-variate analysis.
6. Based on above analysis identifying the driving variables behind loan default by the applicants.

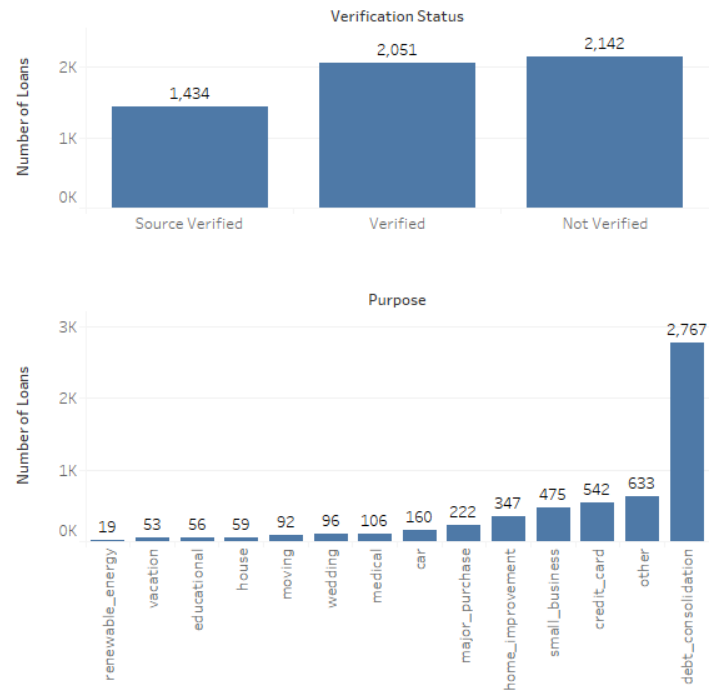
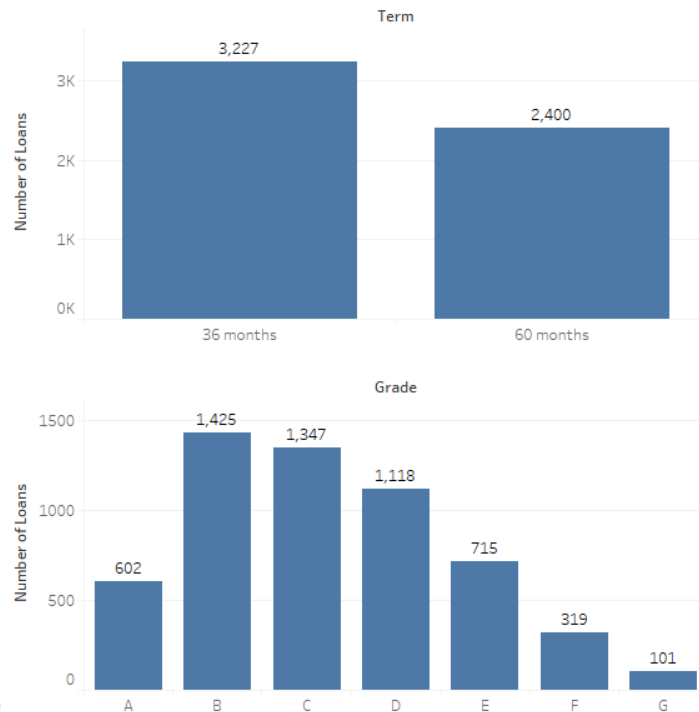
5627 no. of approved loans have been defaulted.

14.59% of approved loans have been defaulted.



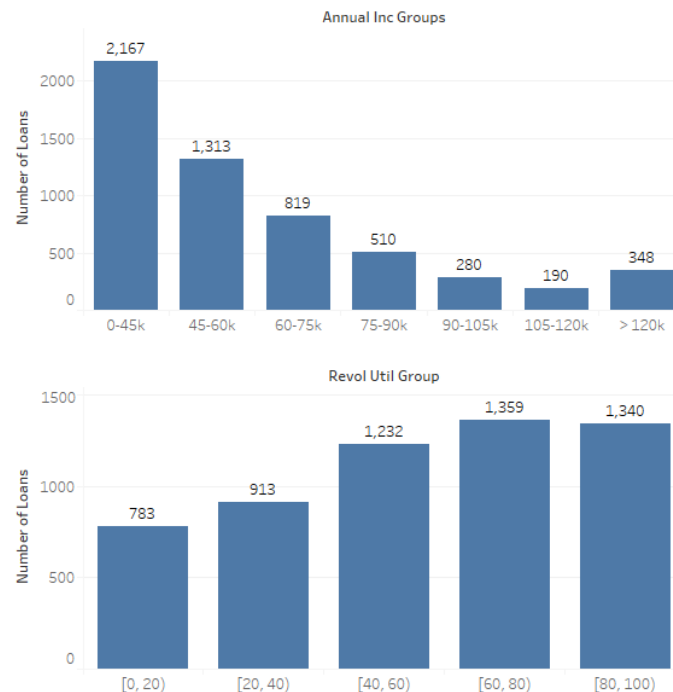
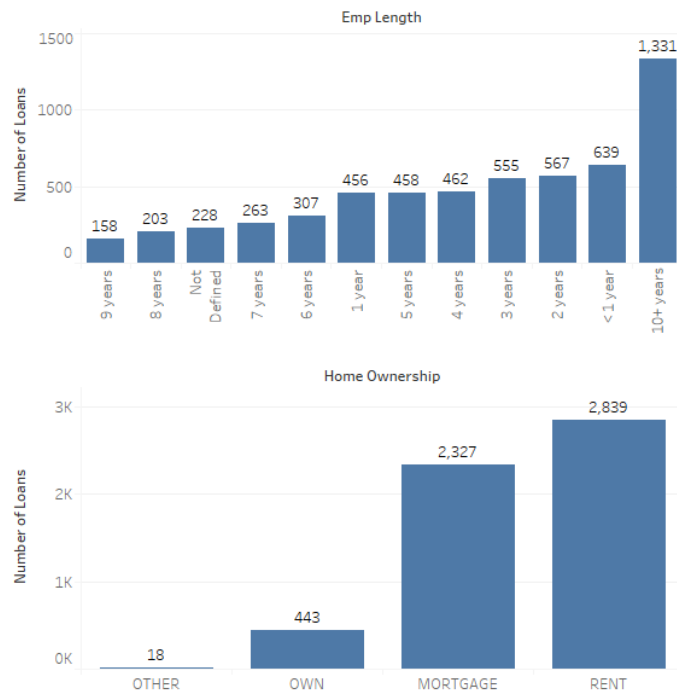
Number of defaulters are plotted for

Term, Verification Status, Grade and Purpose of Loan



Number of loan defaulters plotted for:-

Home Ownership , Employment Length , Annual Income Group, Revolving Utilization Rate



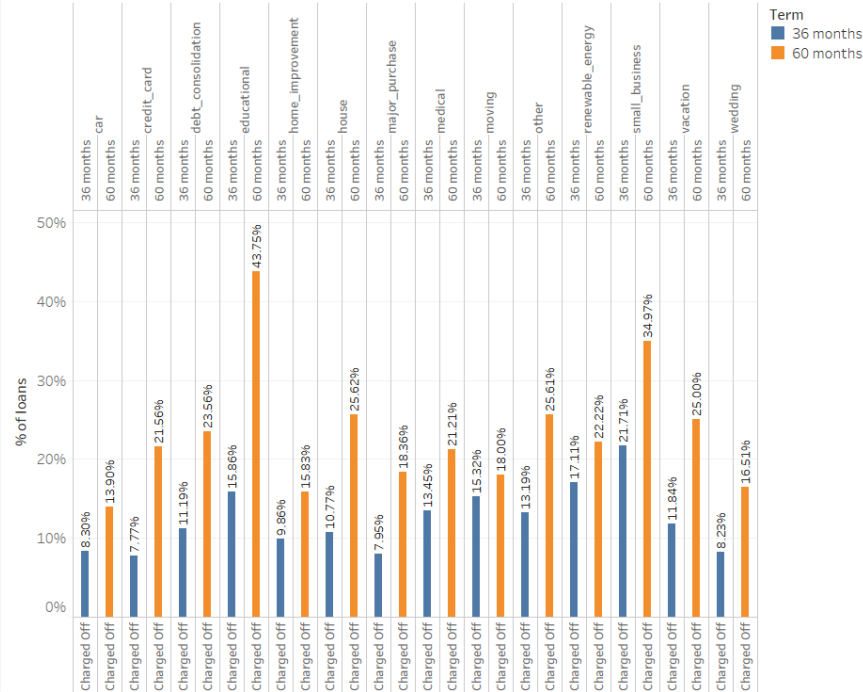
For every purpose, 60 months term loans have maximum percentage of defaulters.

For every emp length, 60 months term loans have maximum percentage of defaulters.

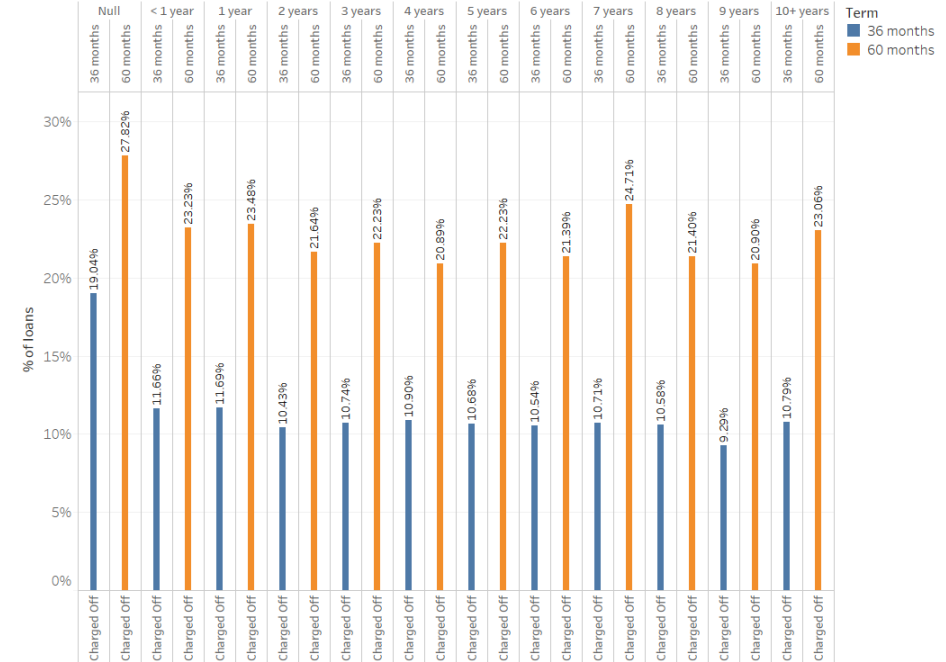
Educational loans have maximum percentage of defaulters in 60 months term.

Emp length of null is having maximum percentage of defaulters in 60 month term.

Percentage of loans defaulted by purpose in each term



Percentage of loans defaulted by emp length in each term



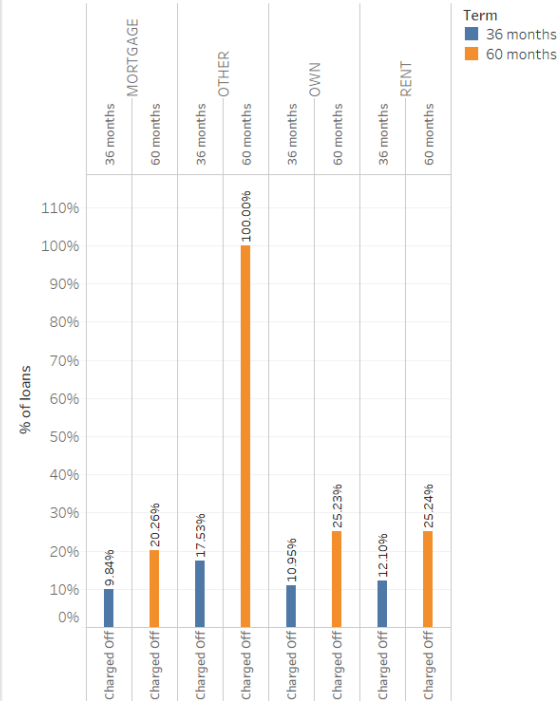
For every house ownership, 60 months term loans have maximum percentage of defaulters.

For every grade except G, 60 months term loans have maximum percentage of defaulters.

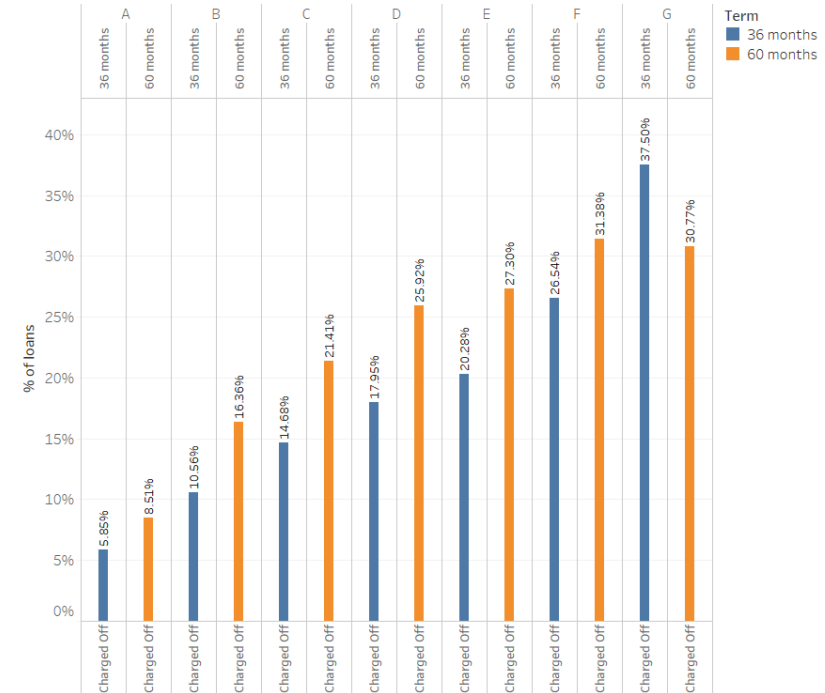
Other House Ownership have maximum percentage of defaulters in 60 months term.

F Grade is having maximum percentage of defaulters in 60 month term.

Percentage of loans defaulted by house ownership in each term



Percentage of loans defaulted by grade in each term



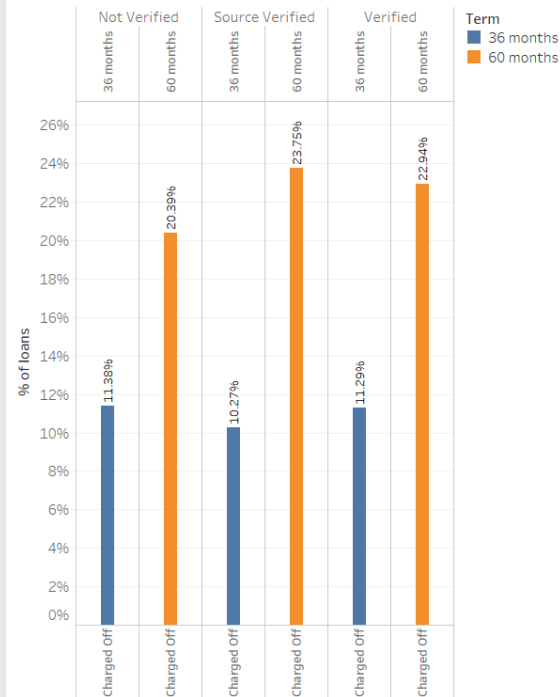
For every verification status, 60 months term loans have maximum percentage of defaulters.

For almost every debt to income ratio, 60 months term loans have maximum percentage of defaulters.

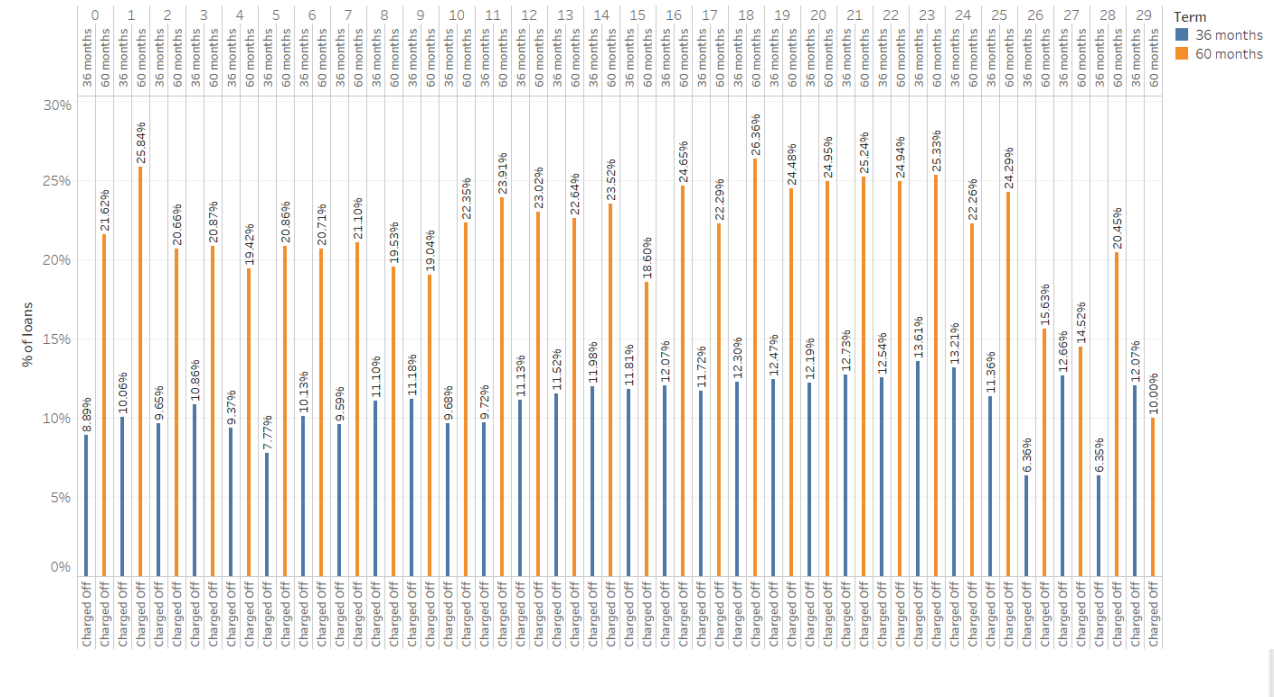
Source Verified loans have maximum percentage of defaulters in 60 months term.

Debt to Income ratio of 18-19 is having maximum percentage of defaulters in 60 month term.

Percentage of loans defaulted by verification status in each term



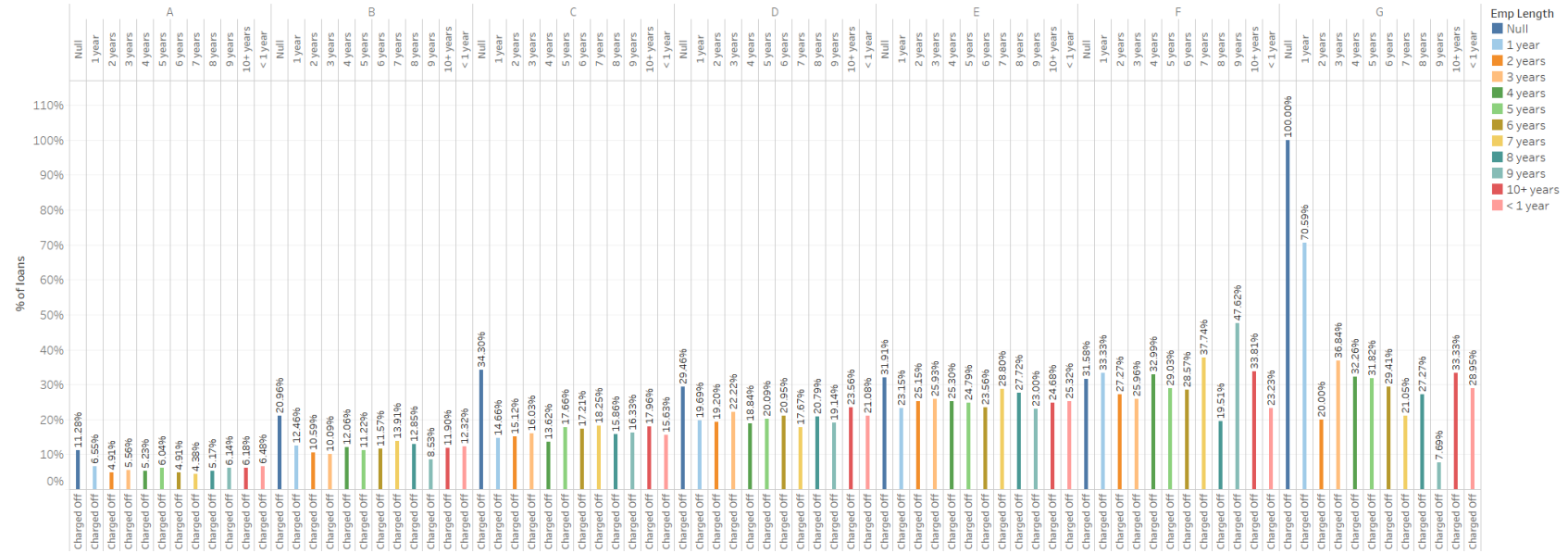
Percentage of loans defaulted by debttoincome ratio in each term



For almost every grade, emp length of null has maximum percentage of defaulters.

Grade G is having maximum percentage of defaulters for emp length of null.

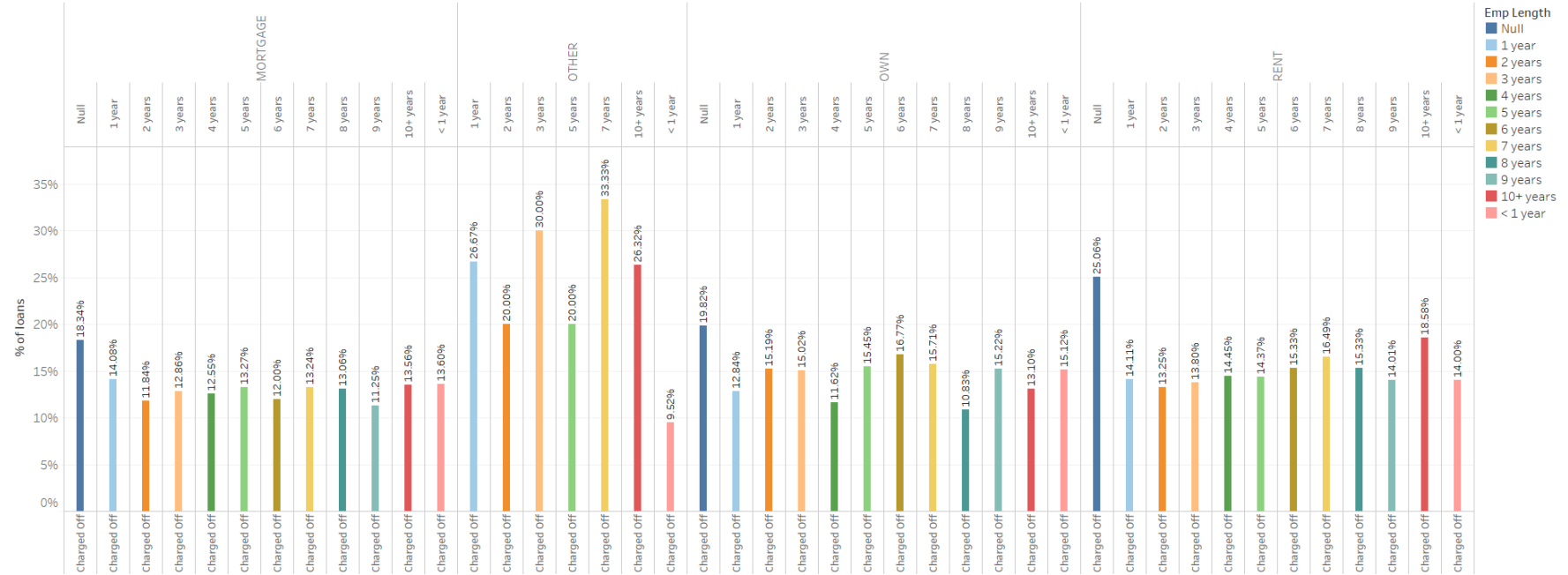
Percentage of loans defaulted by emp length in each grade



For almost every house ownership, emp length of null has maximum percentage of defaulters.

Rent House Ownership is having maximum percentage of defaulters for emp length of null.

Percentage of loans defaulted by emp length for each house ownership



CONCLUSION

Based on the analysis performed we can conclude that following variables are influencing the loan status to a large extent. Lending Company need to carefully analyse following traits of a loan applicant in order to understand if the applicant is likely to pay the loan and based on that decide the interest rates for the applicants.

Variables:

1. Annual Income: Applicants with lower income tends to default more.
2. Grades: Grades and rate of interest show a similar trend. Grades increases alphabetically and corresponding rate of interest also increases. Applicants having highest rate of interest (Grade G) tend to default most. Most of the applicants lie in grade B to D.
3. Revol_Util: Rate of revolving utilization is directly proportional to loan default. As this rate increase there is high chance that applicant will default.

CONCLUSION

4. Some States are safe having no defaulters (IA,IN, ME). Whereas in State like NV and CA there are high number of defaulters.
5. Term: Applicants with 60 months term duration tend to default more.
6. Home Ownership- Applicants Other types of home tend to default more.
7. Purpose: Applicants taking loan for Small Business, Debt Consolidation and Credit Card tend to default more than any other purposes.
8. Verification Status: Verification process needs to be improved. Applicants with source verified and verified status tends to default more.