

Gramener EDA Case Study

Payel Das
Shalini
Sridar Ganesh
Ravishankar



Gramener Case Study



The company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access for loan request through a fast online interface. Lenders decide on terms, tenure, category of loan , interest percentage and etc.

Problem Statement:

Identify risky loan applicants, thereby cutting down the amount of credit loss.

Company's Business Objective:

1. To identify driving factors behind charged-off
2. The company can utilize this knowledge for its portfolio and risk assessment.



Metadata of Loan Data



Understanding loan data

- Data frame contains 39717 observations
- Data frame consists of 111 columns
- Has borrowers historical data from 2007 to 2011
- Contains quantitative, categorical, date and logical values
- The data has information about borrower,
Example: id, member_id, emp_length, home_ownership and etc.
- The data has information about the loan,
Example: funded_amnt, term, int_rate, loan_status, delinq_2years

Understand the data

- Identify the number of rows, NA, blank values
- Identify data for cleansing
- Check for duplicate values in the data
- Check for incorrect data formats

Cleansing the data

- Replacing with NA values for N/A,"","n/a
- Removing columns which has single value
- Converting categorical to numeric data type for columns such as interest_rate and etc.
- Removing Columns which contains only NA and 0 values
- Outlier detection and missing value imputation
- Selecting only the columns which are useful for analysis

Visually identify driving factors for charge-off

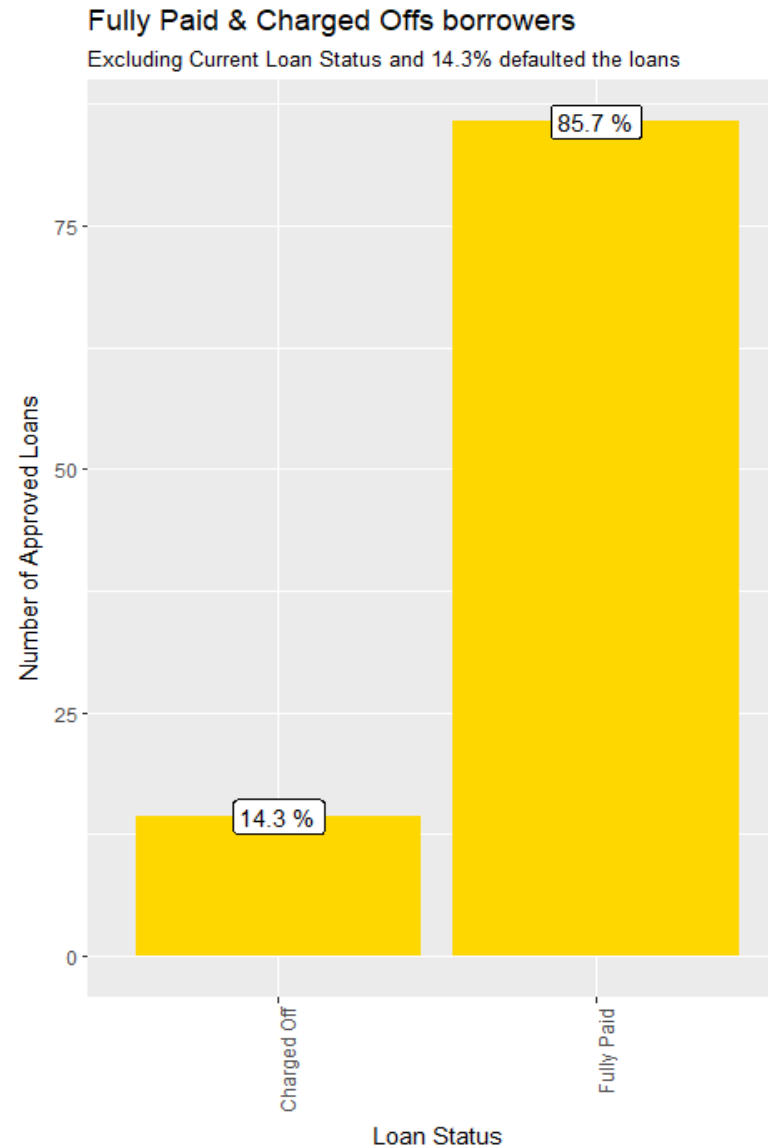
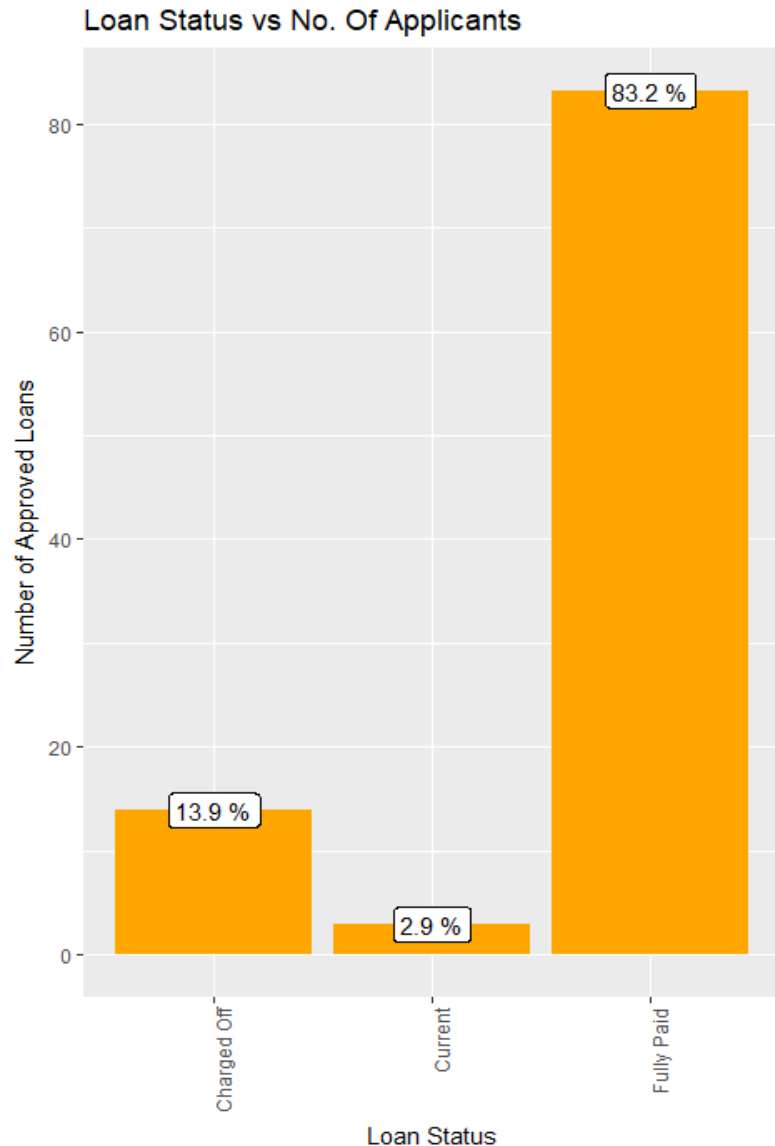
- Perform Univariate segmented analysis and Bivariate analysis for finding driving factors
- Plot visual graphs to understand proportionate for understand

Identify top 5 driving factors for charge-off

- Identify variables which are strong indicators of defaulters
- Use these indicators for providing recommendations

Recommendations

- Provide recommendations to the finance company to cut the credit loss

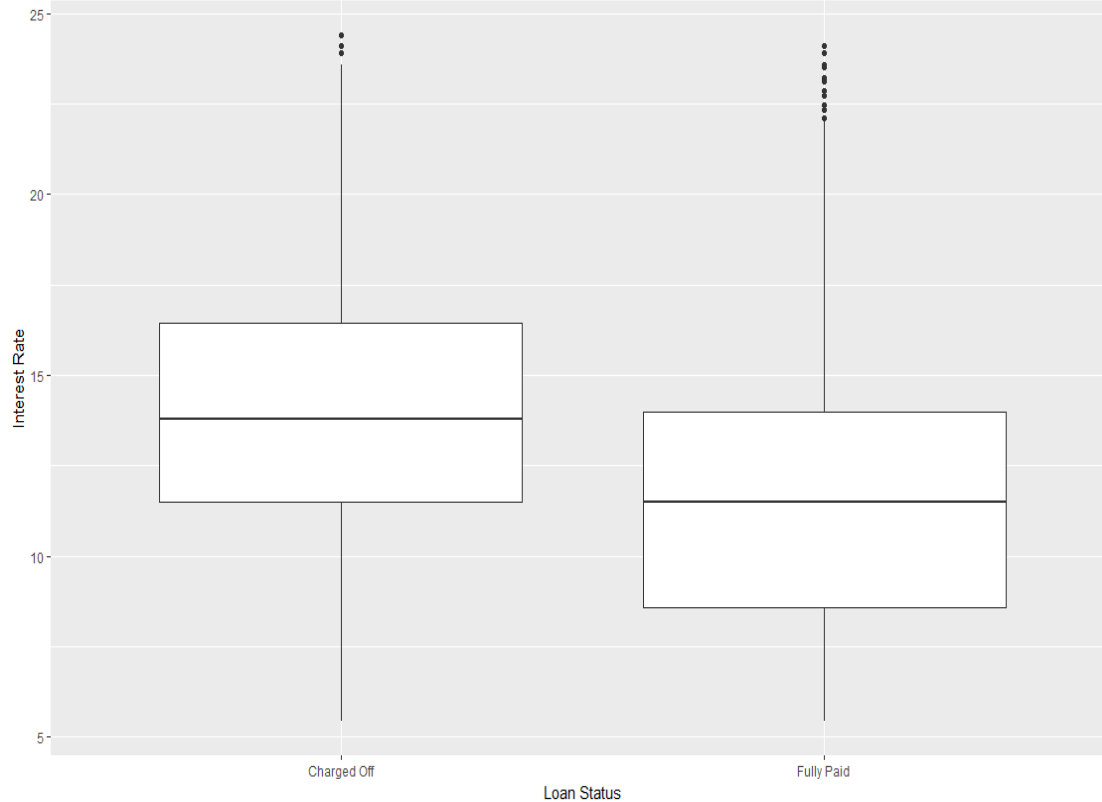


- There are 3 unique entries in loan status
 - Fully Paid
 - Charged Off
 - Current status of the loan
- Loan status with value "Current" (2.9% of total records) will be excluded in our analysis as we don't know whether those will be fully paid or become charged off in future
- After excluding Current loan status, charged-off status is at 14.3% of total population under consideration

1. Driving factor for charge-off: interest rate

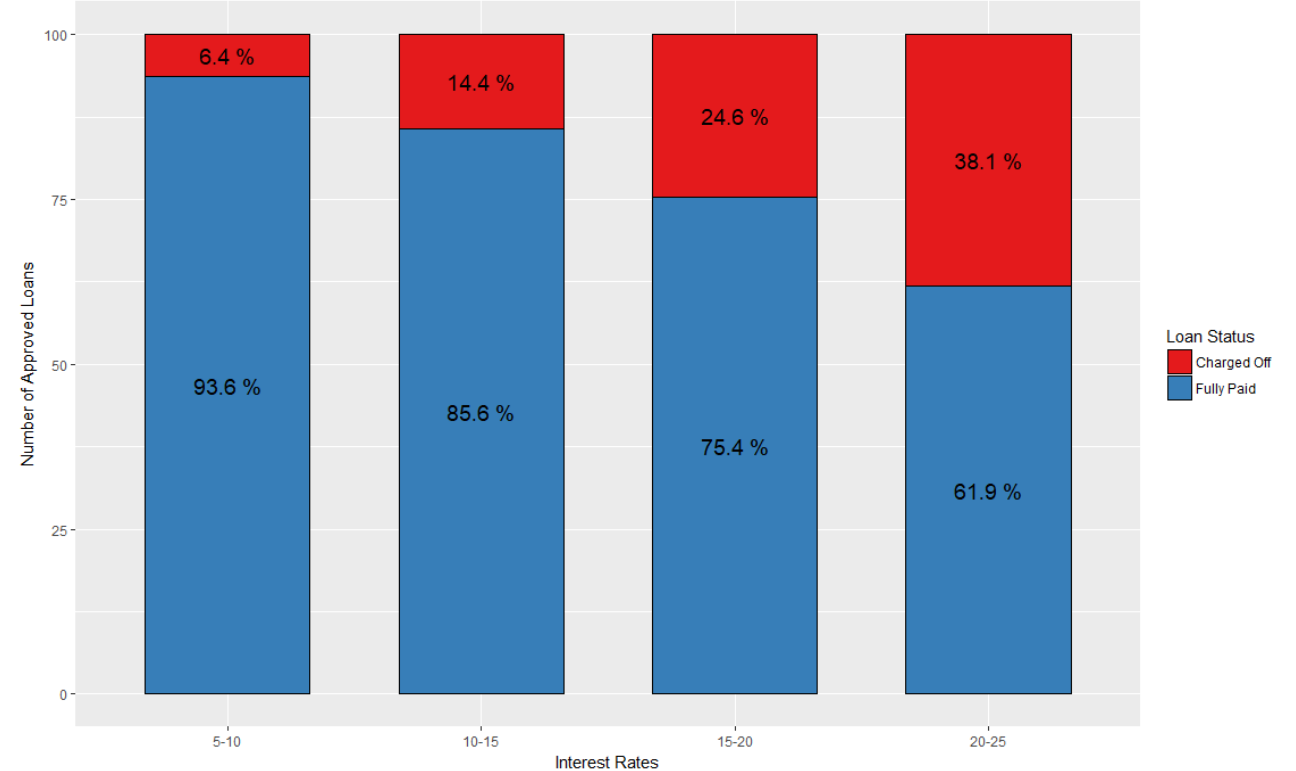
Loan Status vs Interest Rate

Borrowers who are paying high interest rate are more likely to Charge off



Charged Off & Fully Paid borrowers For Each Interest Rate Buckets

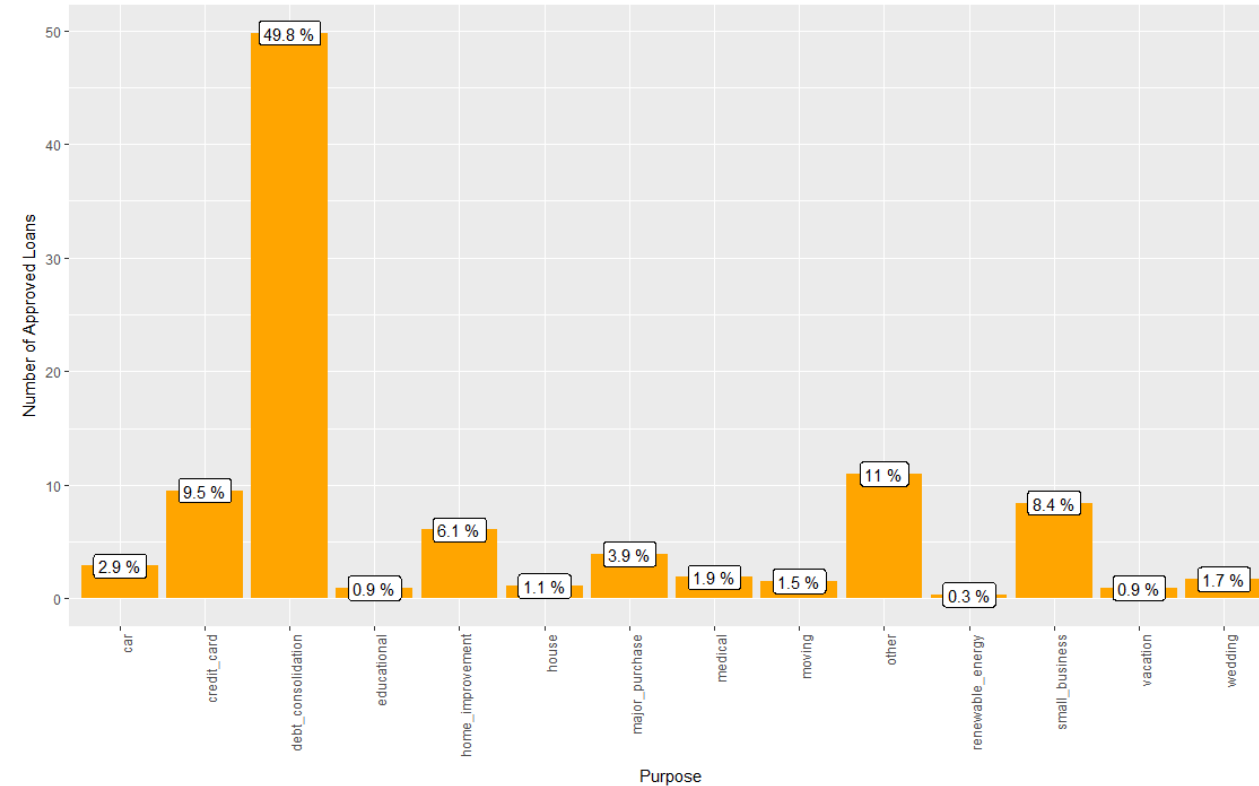
Increase in interest rates also increases defaulters



- Median interest rate is higher for the charged-off loan status (Ref: Box plot)
- The borrowers who are paying high interest rate are more likely to Charge-off(Ref: Image 2)

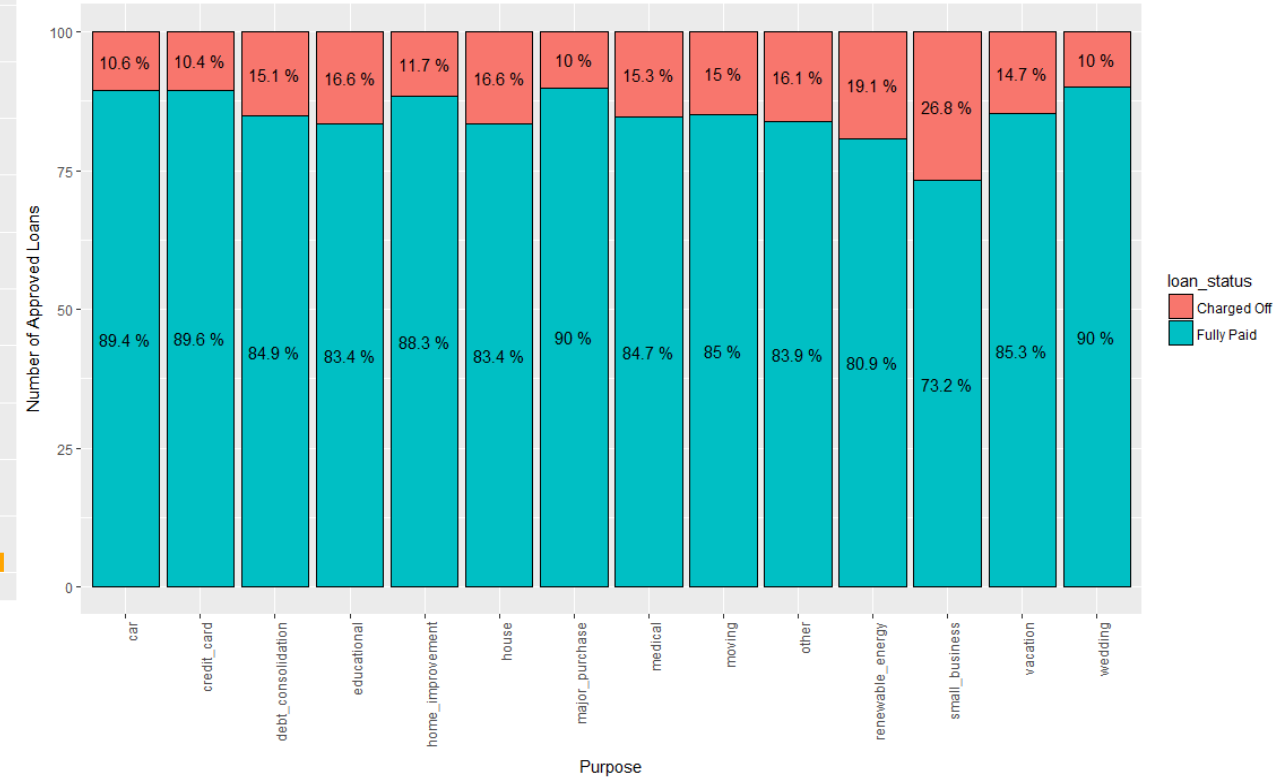
2. Driving factor for charge-off: loan purpose

Frequency of Charge Offs or various Loan Purposes



Frequency of Charge Offs or various Loan Purposes

Small business borrowers show 26.8% more likelihood to charge off than other borrowers

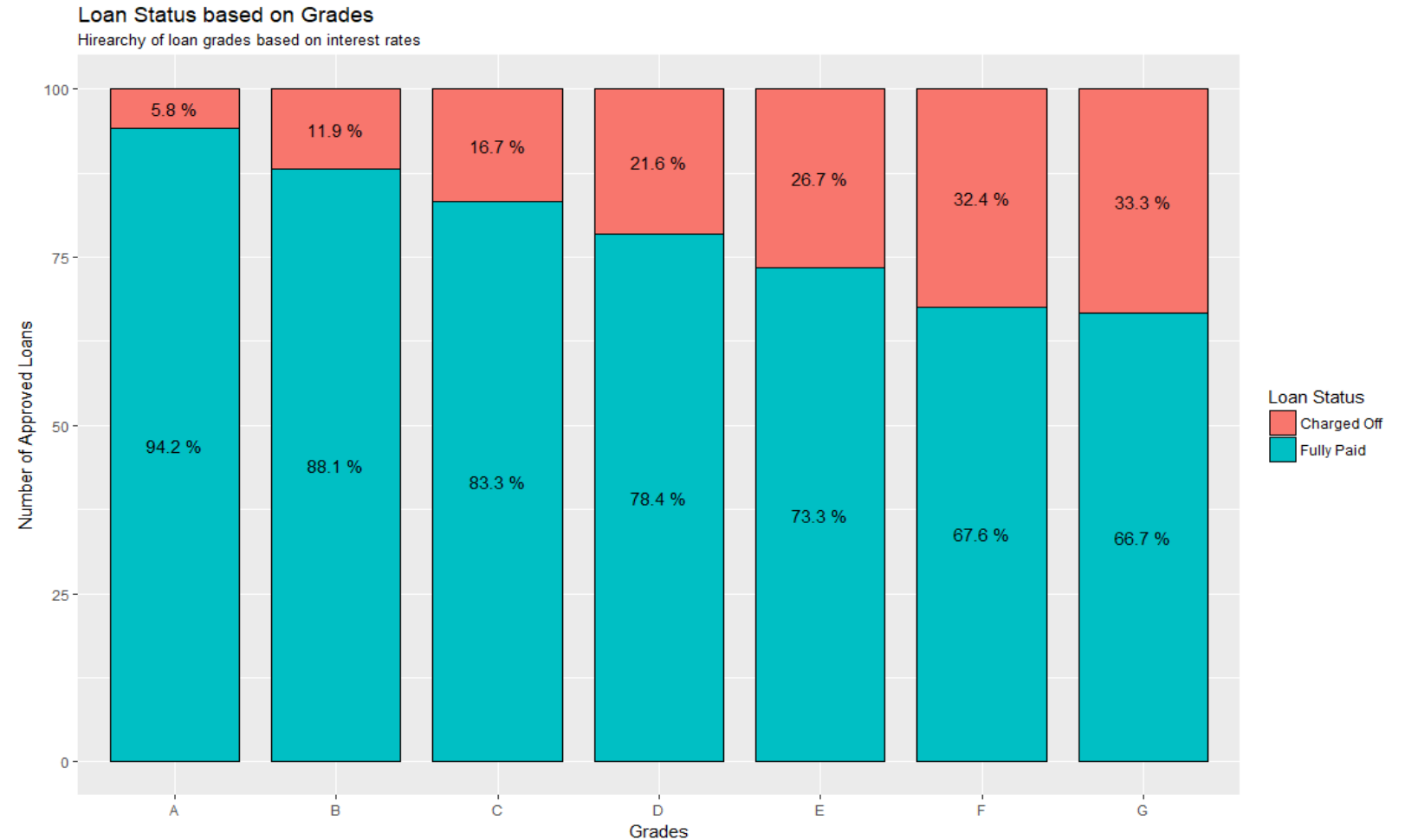


- We can consider top 5 loan request category that influences loan defaulters are debt consolidation, credit card, other, small business, other and home improvement (Ref: Image1)
- Results show Small business borrowers show 26.8% more likelihood to charge off than other borrowers (Ref: Image 2)

3. Driving factor for charge-off: Grades

Interest rate based on grades

A	B	C	D	E	F	G
7.49	10.99	13.49	15.62	17.51	19.41	20.90

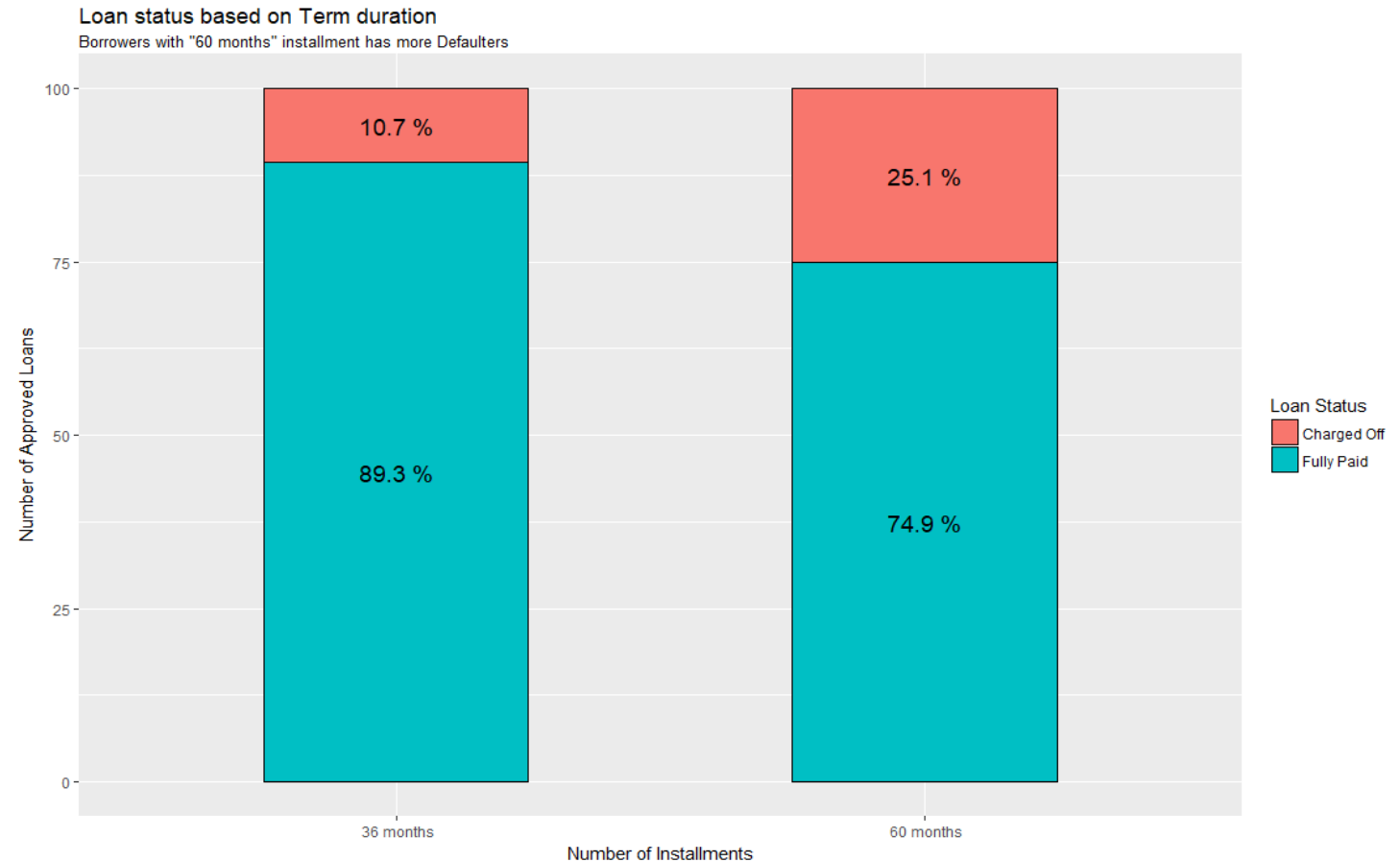


- Since there are extreme values, we are taking median instead of mean
- Results show as the grade of the loans go higher, interest rate increases and accordingly the likelihood of defaulter increases

4. Driving factor for charge-off: Term duration

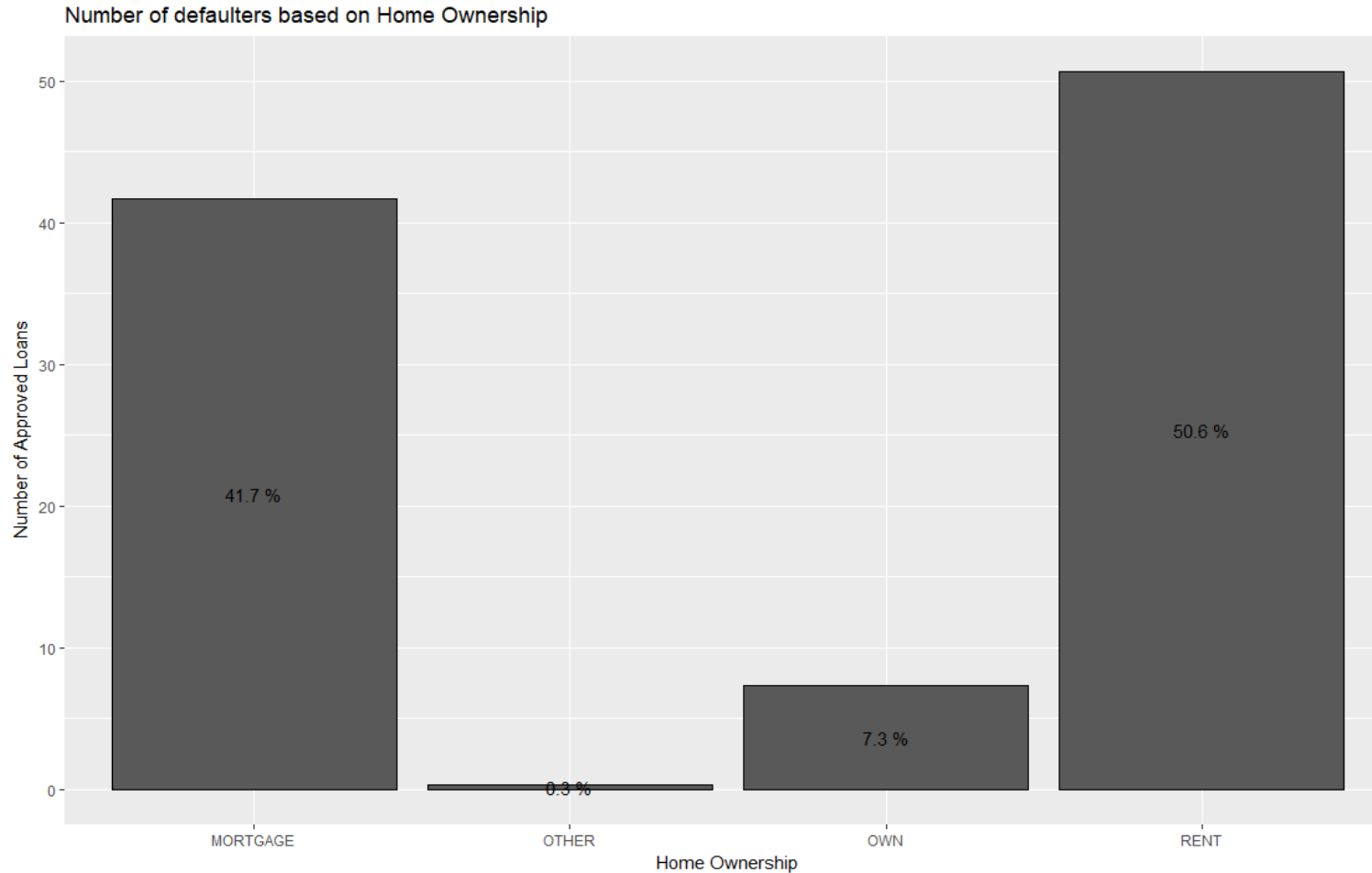
Interest Rate based on
term duration

36 months	60 months
10.99	14.91



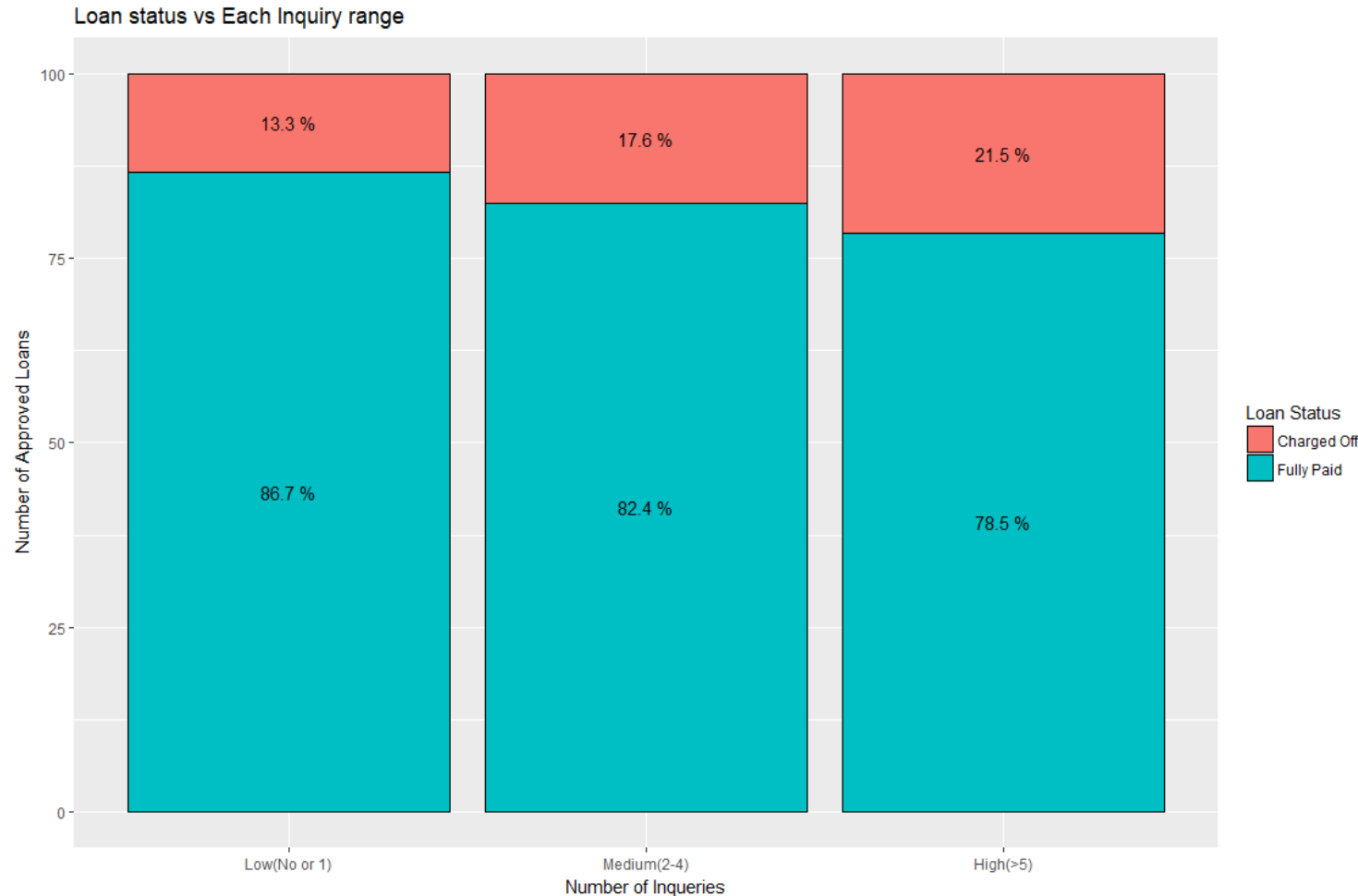
- Members who agreed to pay within 60 months are more likely to default (one reason being the higher interest rate for this category) than members who pay in 36 monthly installments,

5. Driving factor for charge-off: Home ownership



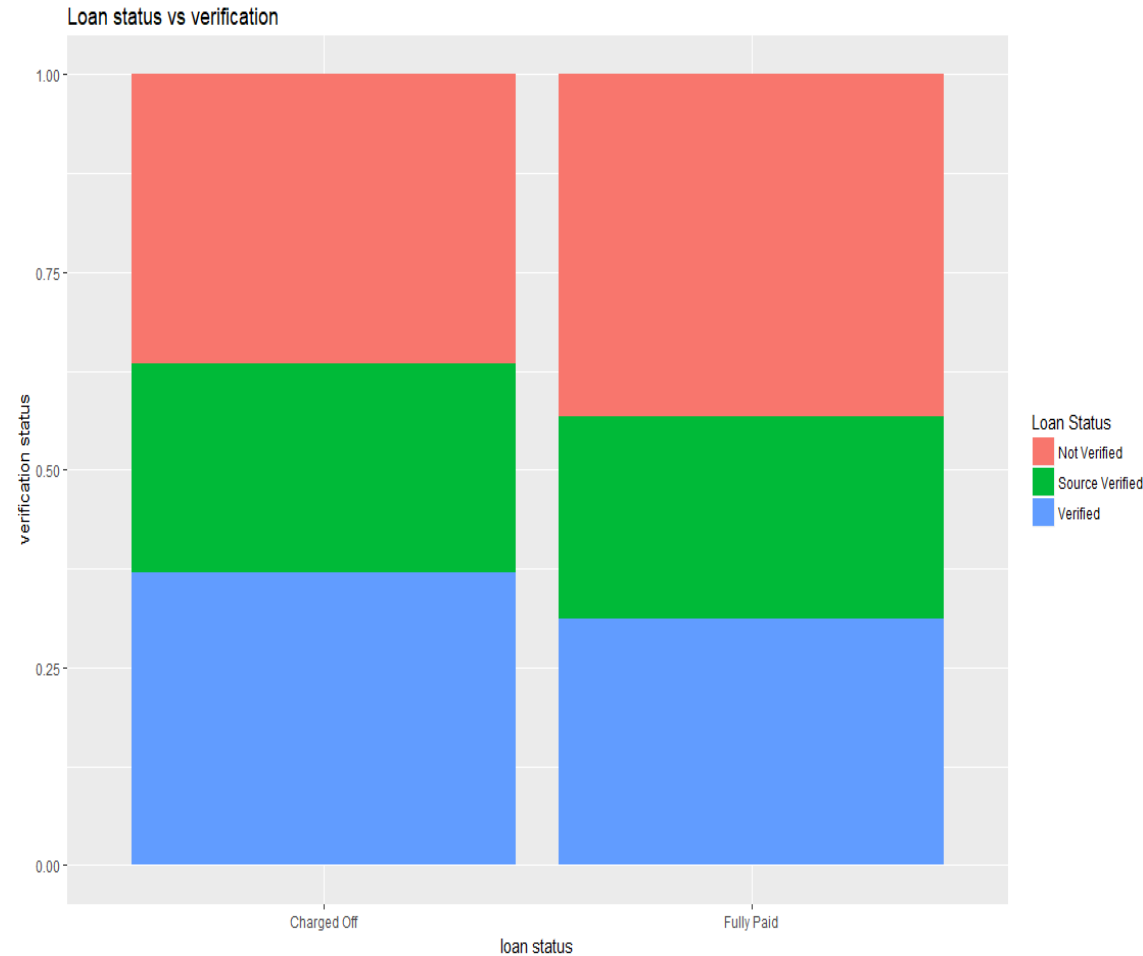
- People paying Rent for their home or who have mortgages are less likely to repay the loan within expected date and leading to defaulters

6. Driving factor for charge-off: Inquiries from past 6 months



- As the number of inquiries by a loan seeker goes higher, it indicates that he is desperately looking for a credit and borrower might not be capable of repaying it later

7. Driving factor for charge-off: Verification Status



- The default cases have more verified applications than fully paid cases
- Need to review the process of verification checks before granting the loan



Recommendations



- Disbursing loan with higher interest rates and higher grades are more likely to charge-off. Before approving company should verify borrowers portfolio and credit scores with special watch out on grade G
- Borrowers with excessive frequency for inquiries (>5 in 6 months) in a short timeframe are more likely to default, consideration to be made with regards to the customer's borrowing history
- Process of verifying annual income should be reviewed
- Borrowers for 'small business ' would most likely to default compare to other purposes, would recommend review of approval process for these borrowers
- Borrowers needs to be evaluated based on duration term as borrowers opting for longer duration are more likely to charge-off