

Background and Business Objective

Background:

Lending club is the largest peer to peer marketing place connecting borrowers with lenders. Borrowers apply through an online platform where they assigned an internal score. Lenders decide,

1. Whether to Lend
 2. The terms of loan such as interest rate, monthly instalment, tenure etc.
- some popular products are credit card loans, debt consolidation loans, house loans and car loans etc.

Business Objective :

To understand the driving factors behind loan default. In other words, variables that are more representative behind loan default. Lending company can utilize this knowledge for its portfolio and risk assessment and future loan approval/rejection.

Data Understanding

Type of variables

There are plethora of variables associated with each loan record. For ease of understanding, we have grouped them as below. We have considered only relevant variables for this analysis.

- Customer (applicant) demographic
- Loan related information and characteristics
- Customer behaviour (if the loan is granted)



Customer Demographic

- Employment length
- Employment title
- Annual Income
- Zip code
- Description



Loan related information

- Loan Amount
- Funded Amount
- Funded Amount Investment
- Interest rate
- Loan status
- Loan Grade
- Loan Term



Customer behaviour

- Delinquency year-2
- Earliest credit line
- Revolving balance
- Recoveries
- Application type
- Loan purpose

Data Understanding

- Basic analysis using significant columns

Univariate Analysis

- Impact analysis on one variable against loan default rate.

Bivariate/Multivariate Analysis

- Analysis on two or variables on Loan default rate

Loan purpose Distribution

- Selecting Top 5 products by Loan Purpose and dice other fields to reveal hidden insights.

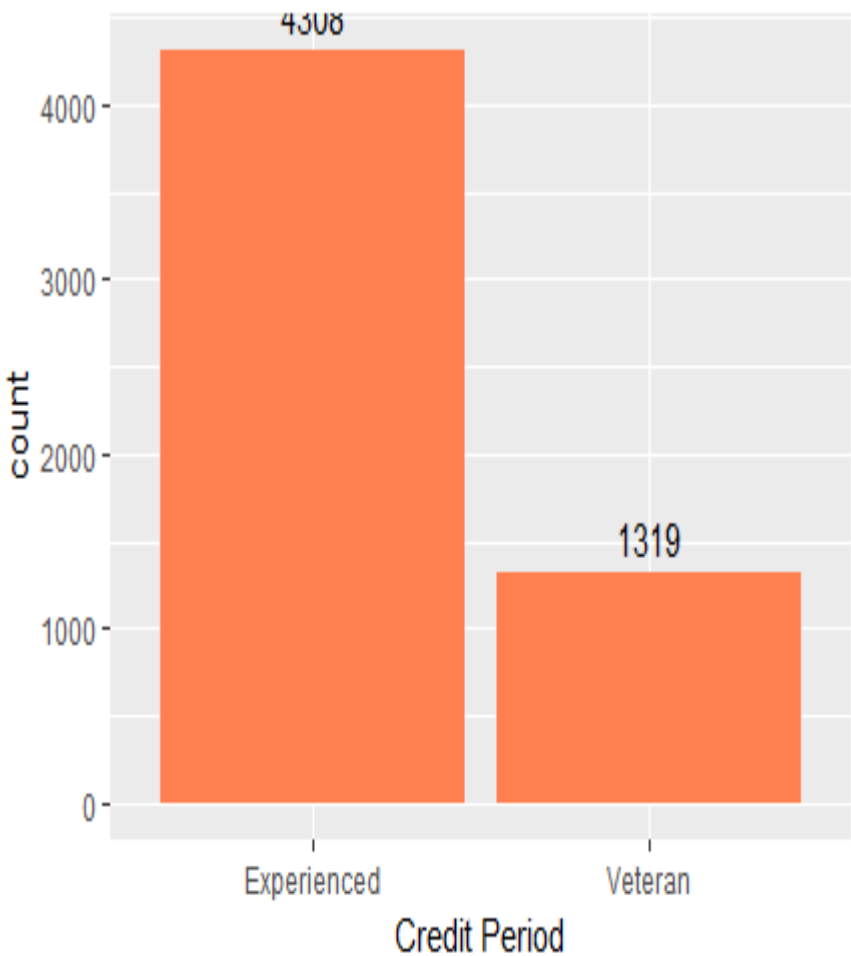
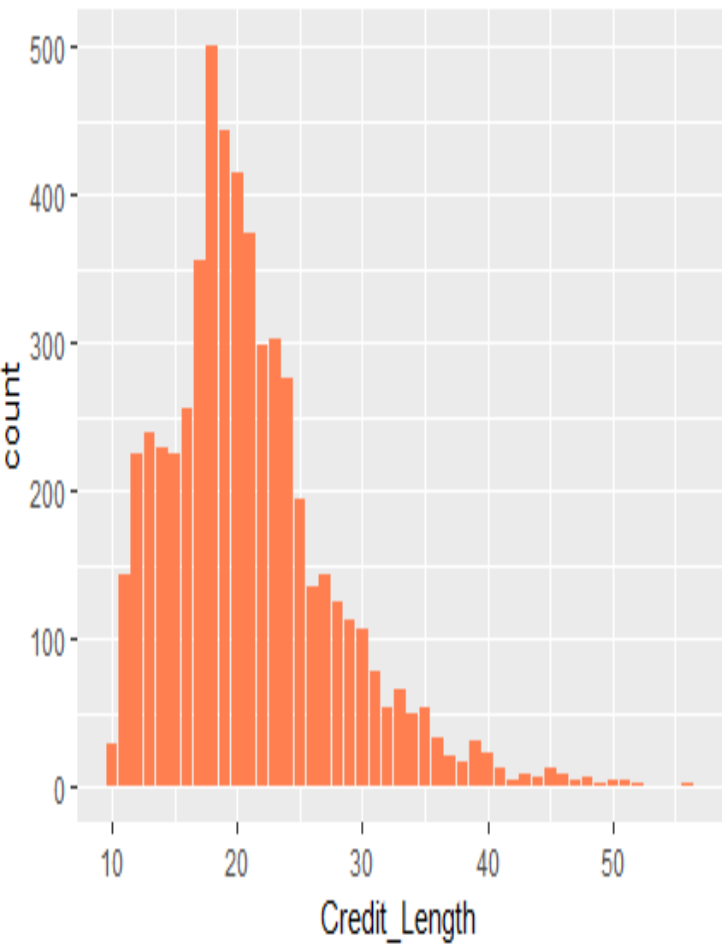
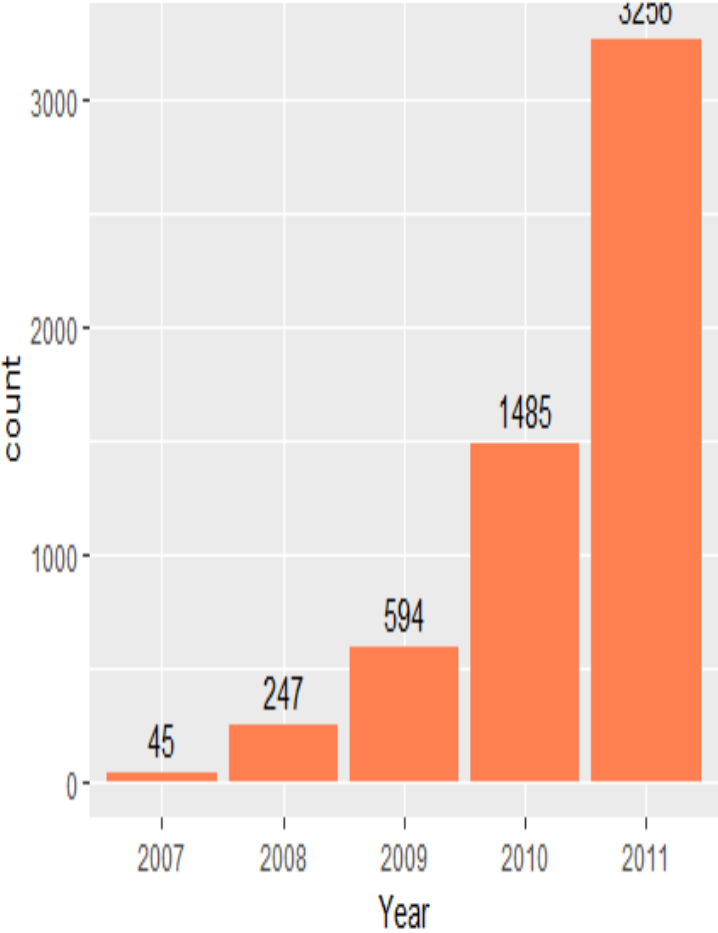
Various Analysis on Top5 Products

- Analysing Top 5 Loan Purposes against the fields that were included in Univariate and Bivariate analysis.

Important Variable – Top 5 Products

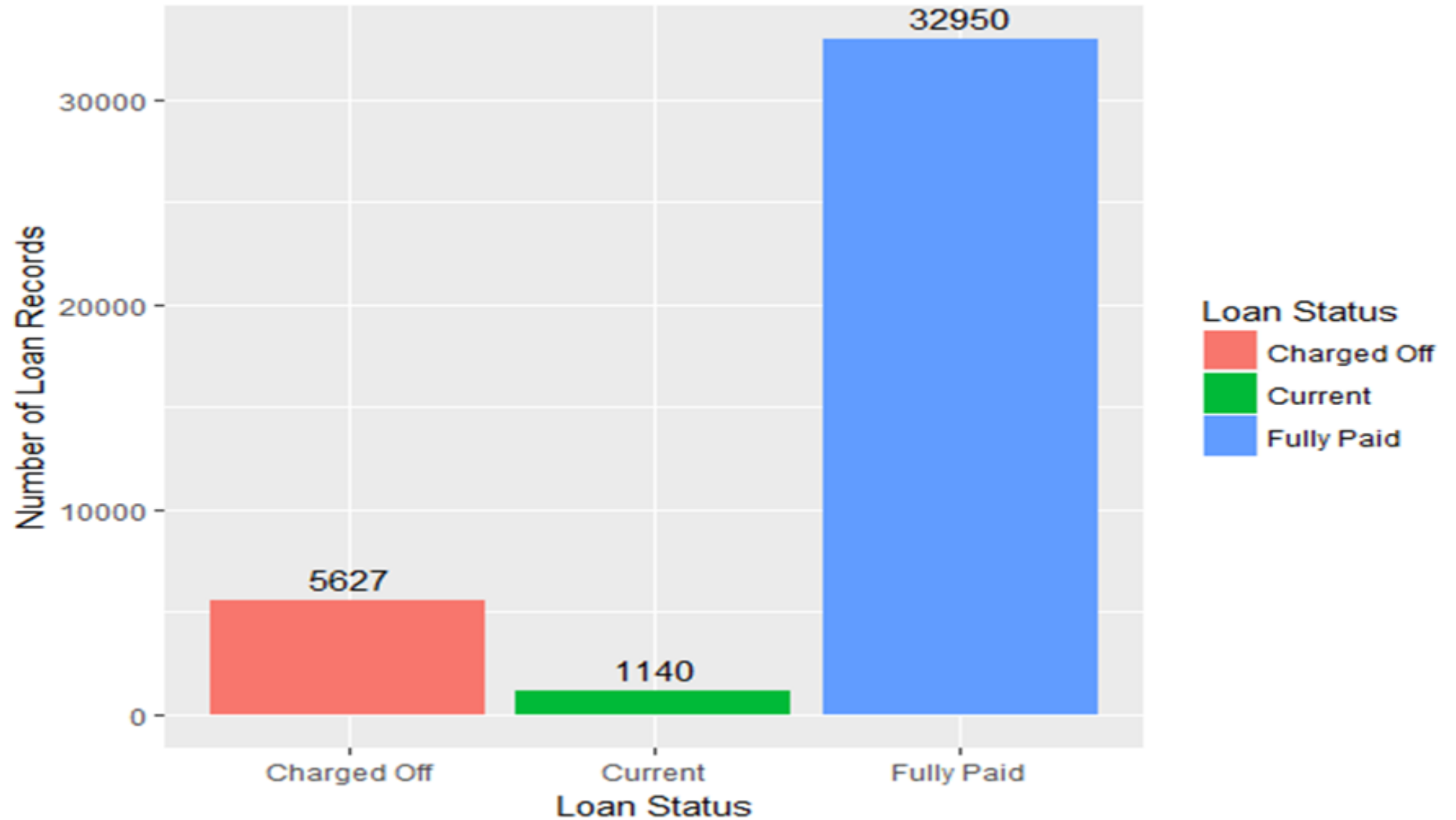
- Nail down more representative fields behind loan default eventually.

Data distribution by Years, Credit Length & Credit Length Group.



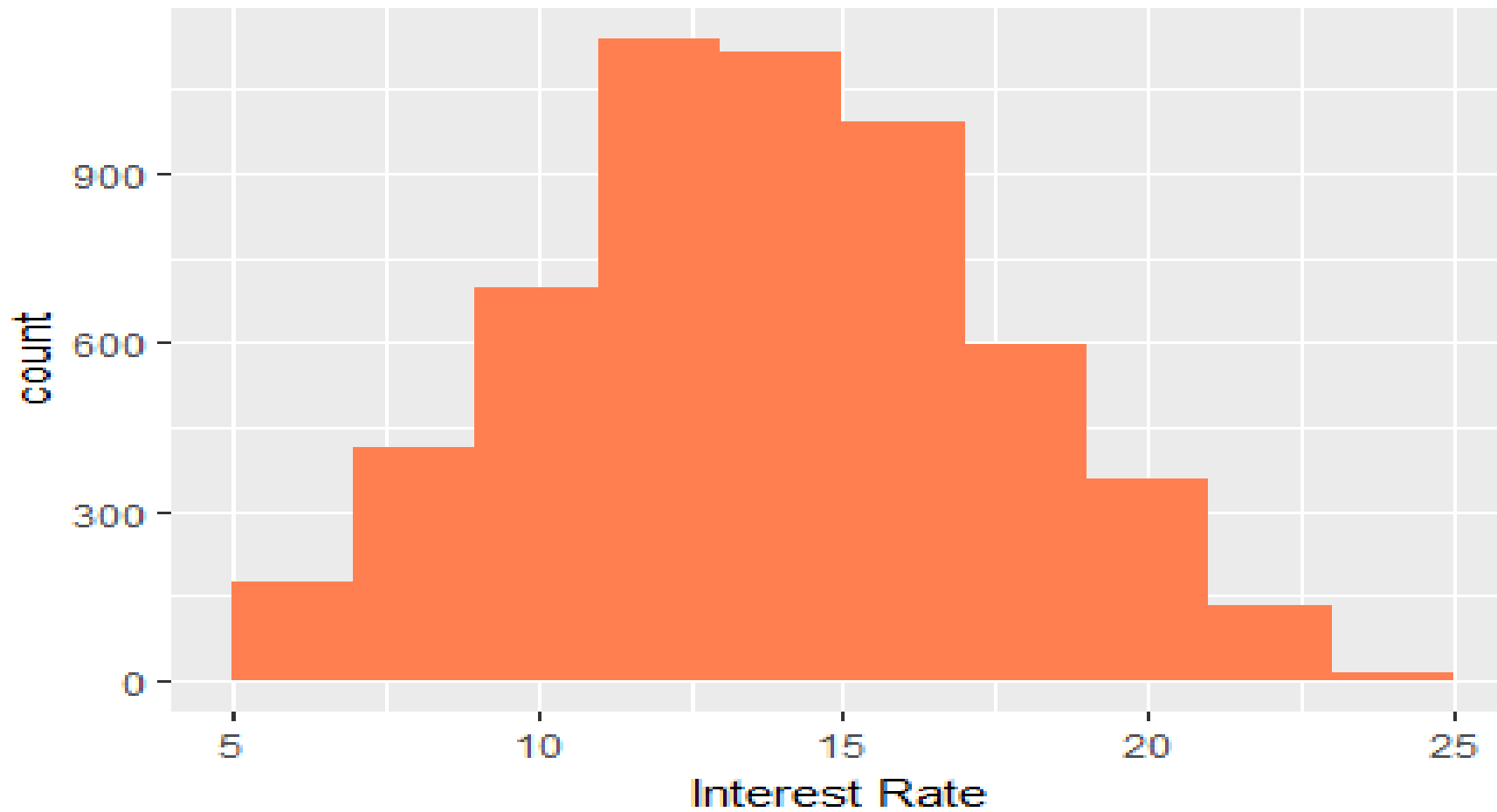
Distribution by Loan status

- 14% of total loans are defaulted.
- 83% of total loans are fully paid.
- Only “Charged Off” Loans are taken up for further analysis.



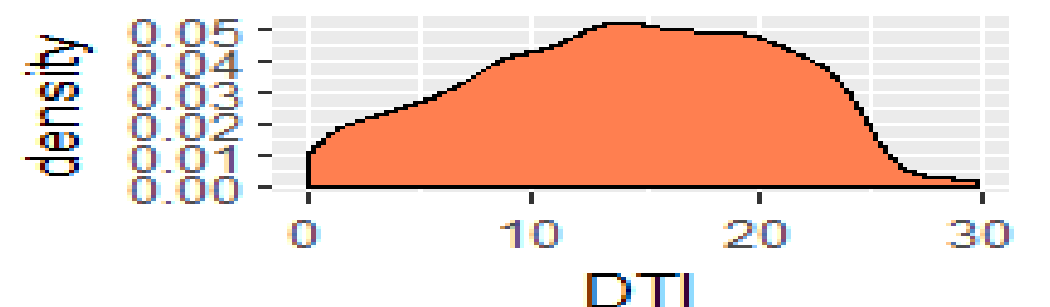
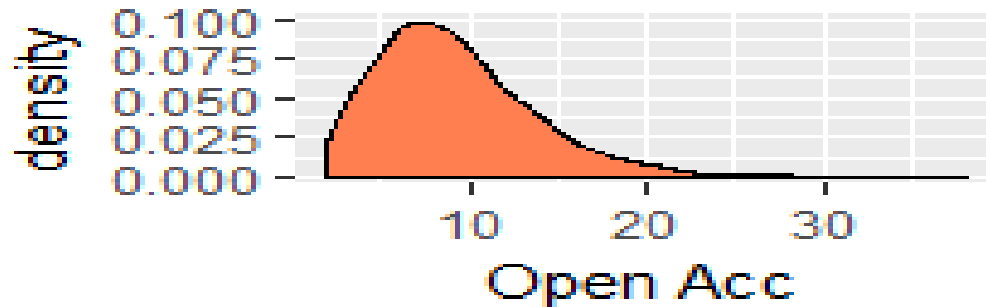
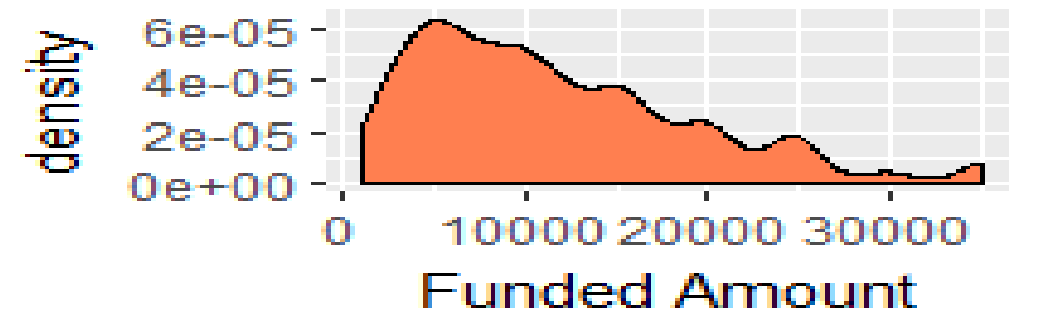
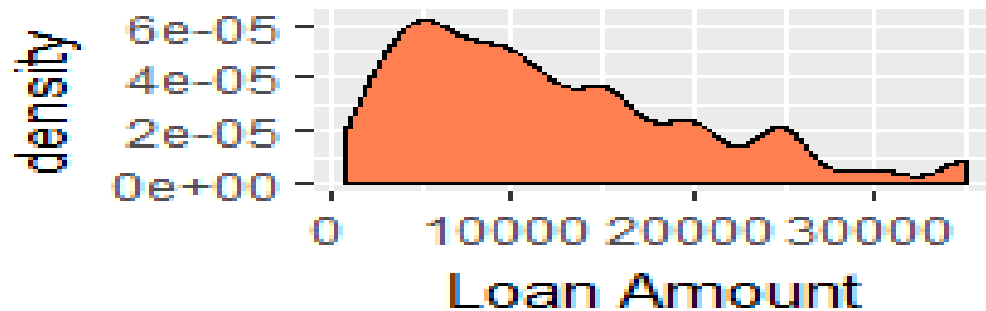
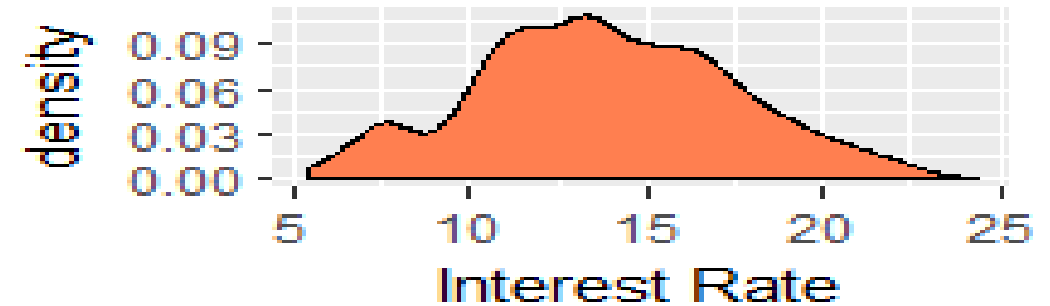
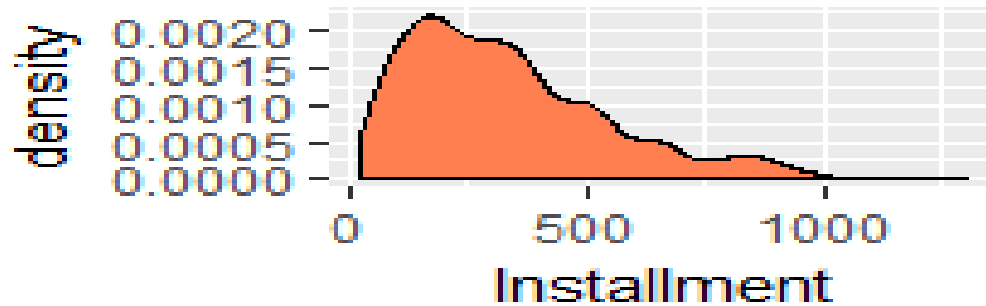
Interest Rate - Distribution

Inference : Interest rates between 10 and 15% are having highest number of loan default.

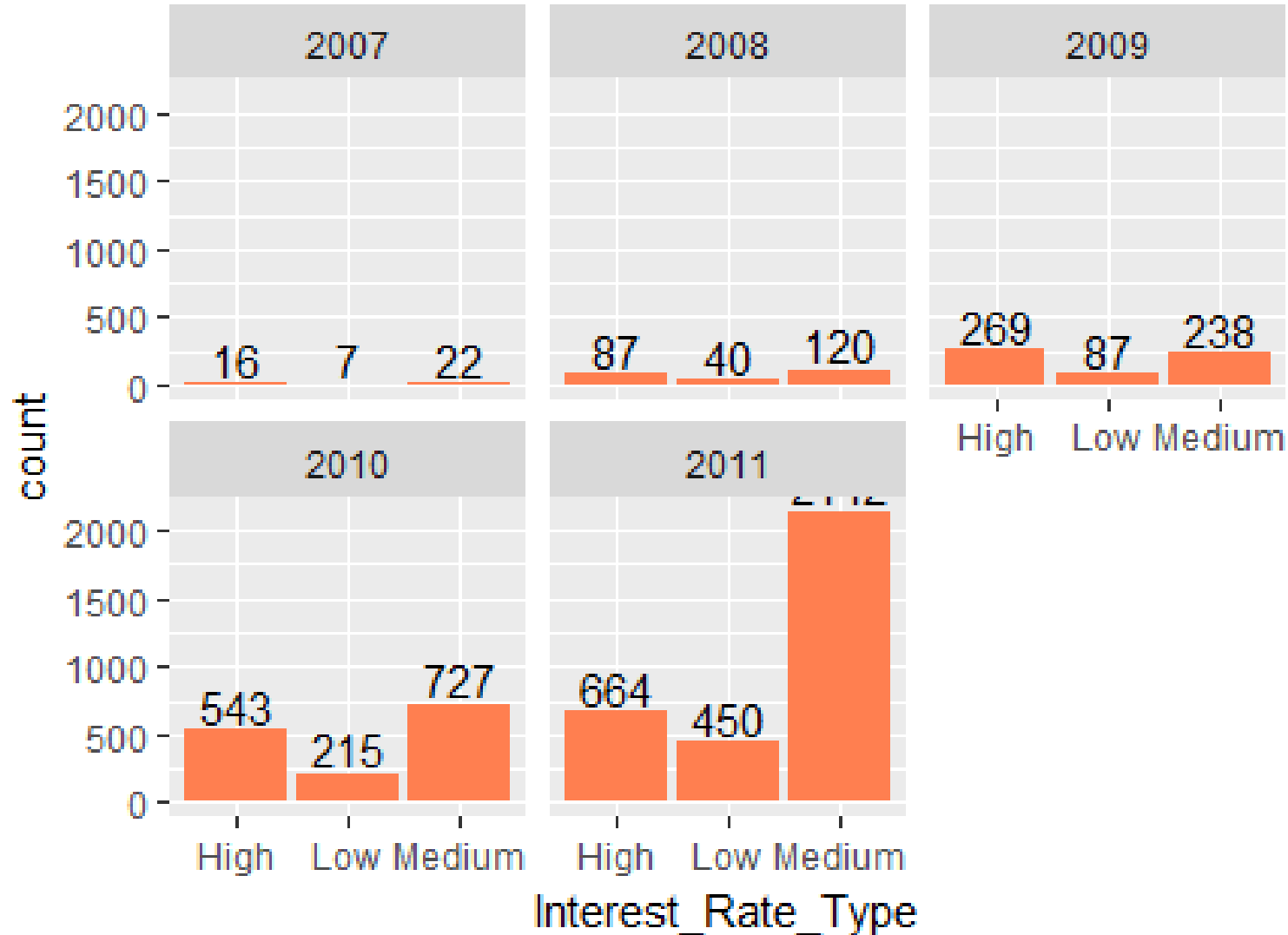


Density plots on quantitative variables

Except DTI & Interest rate, other continuous fields have data right skewed



Interest Rate Type over years

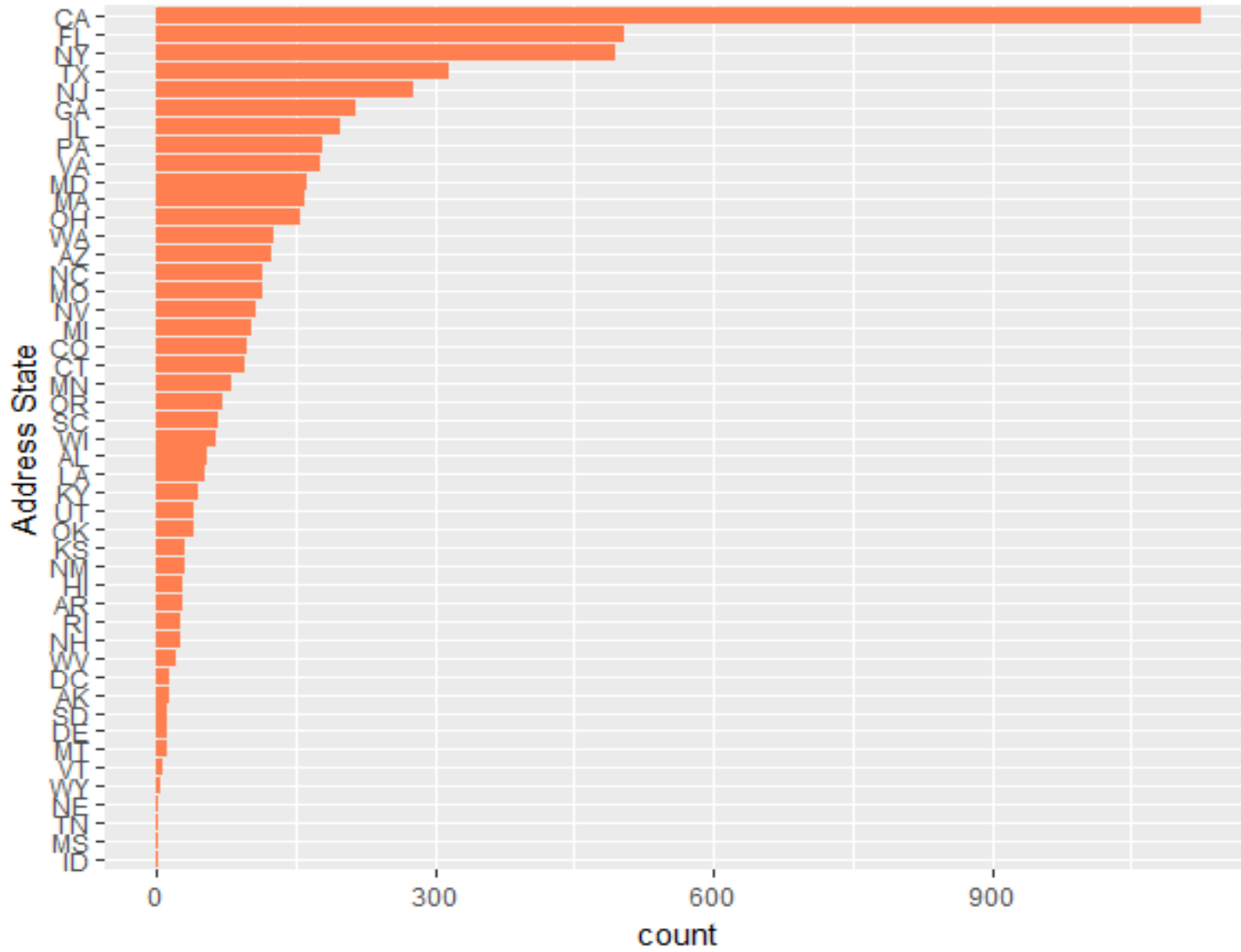


➤ Across the years, Medium Interest rated loans have mostly substantial amount of loan defaults.

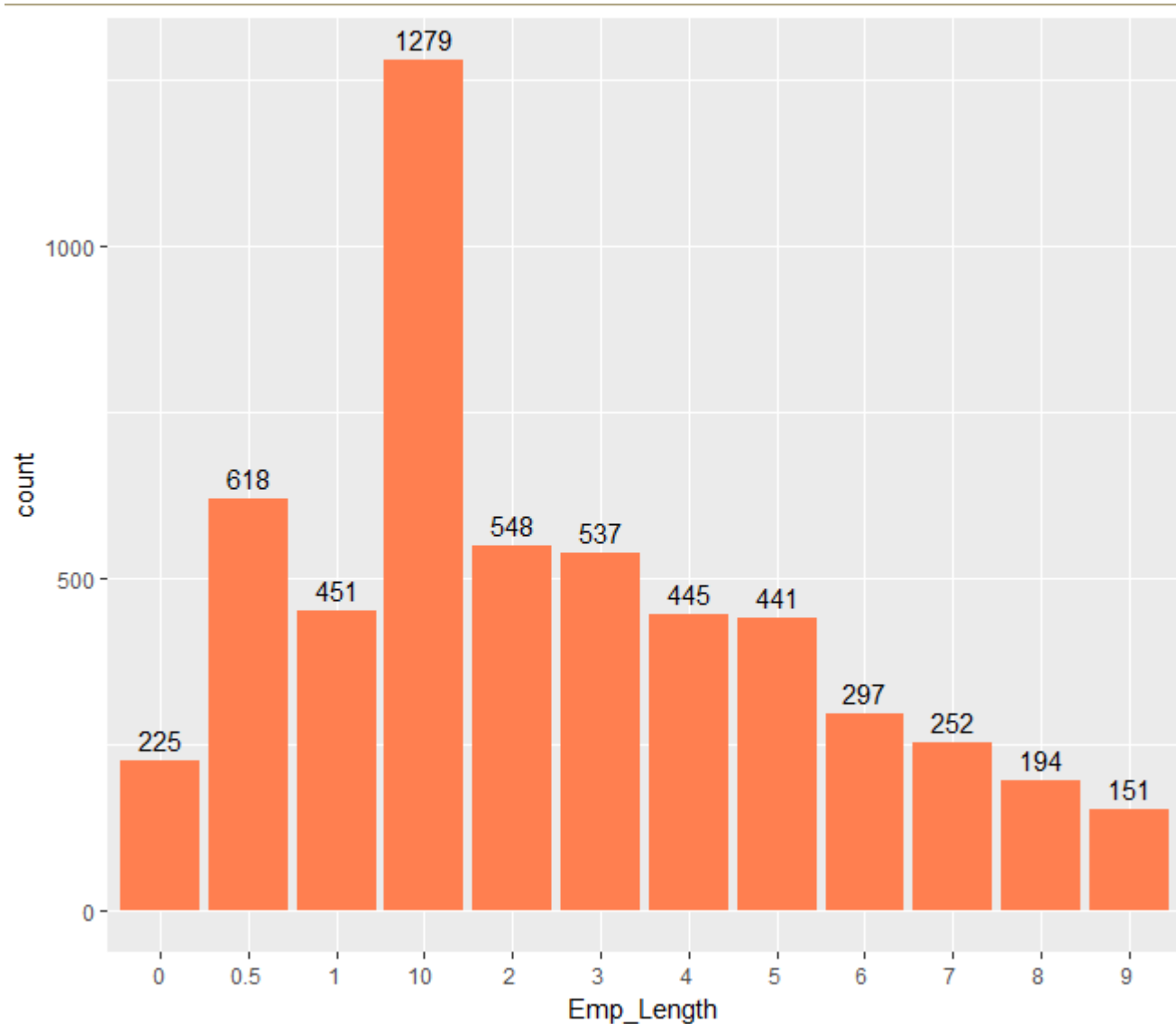
➤ For the ease of analysis, Interest rates are grouped Low, Medium and High based on below.

Interest Rate	Group
1-10%	Low
10-15%	Medium
15% and Above	High

Volume Analysis by Address state

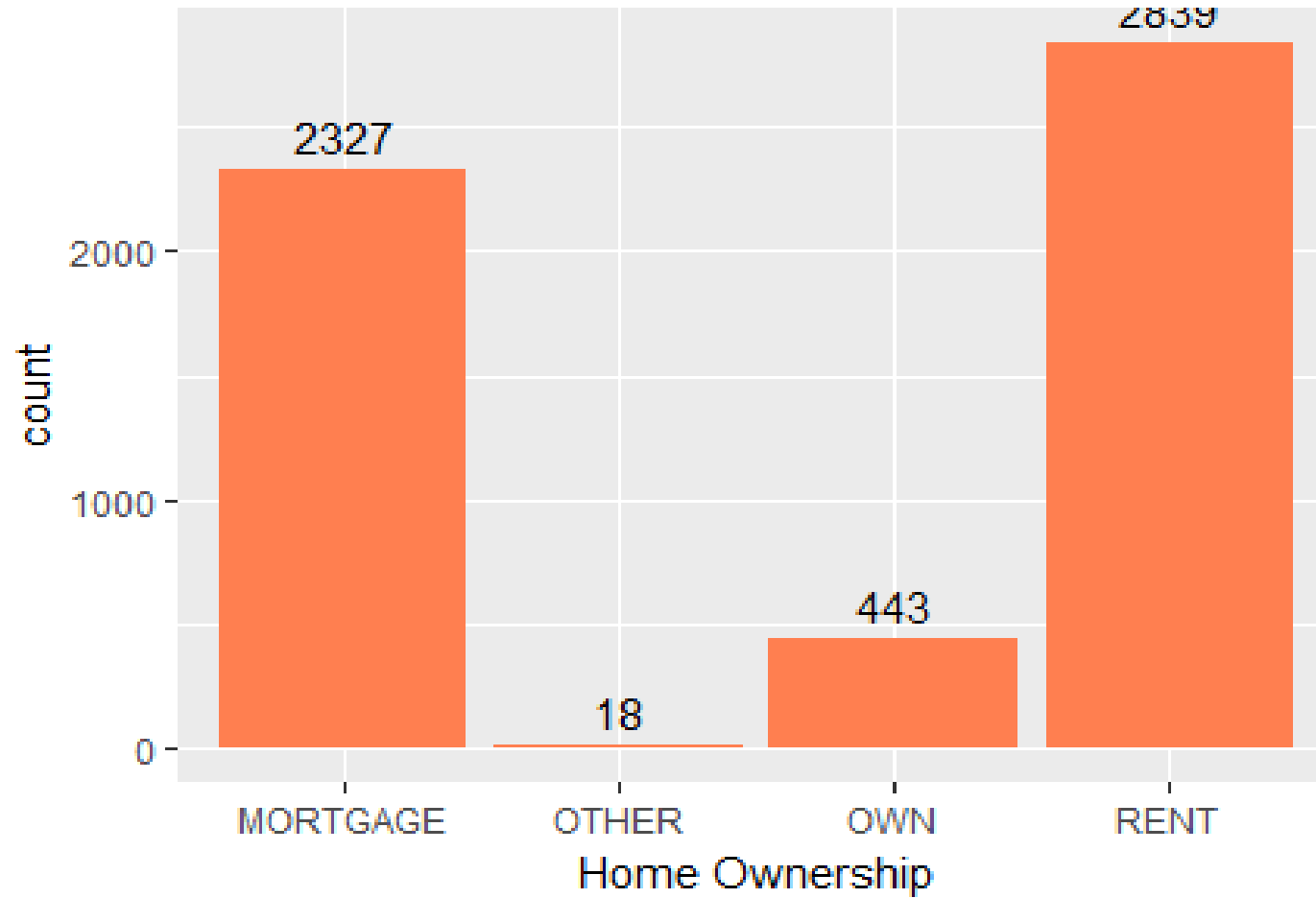


Interestingly, California based customers have significantly contributed more to loan default.



- 23% of total loans defaulted by borrowers with employment length more than 10 years
- 11% of total loans defaulted by those having less than a year of job experience
- 10% loan defaults caused by those with 2 years of job experience.

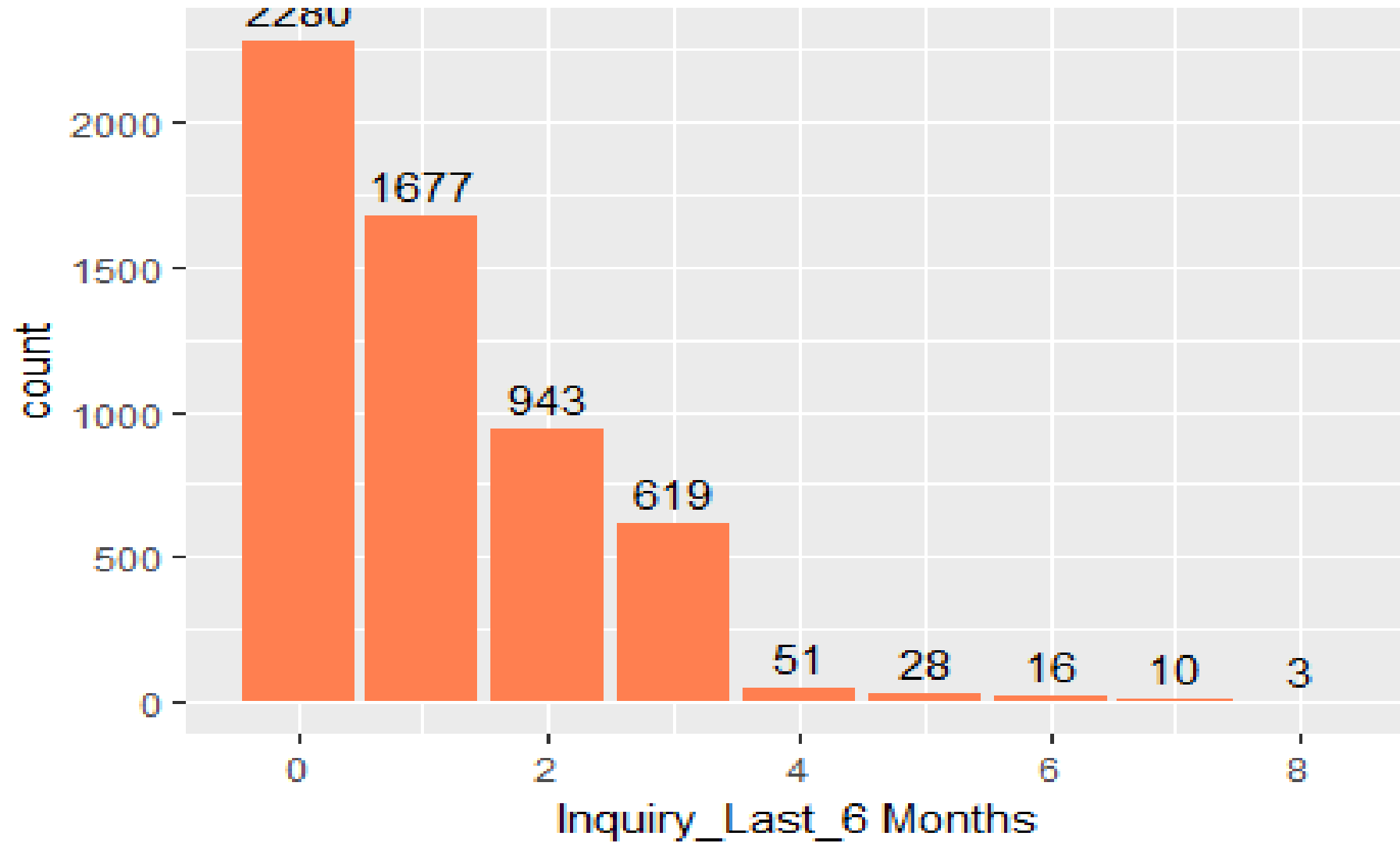
Home Ownership Vs Loan Default



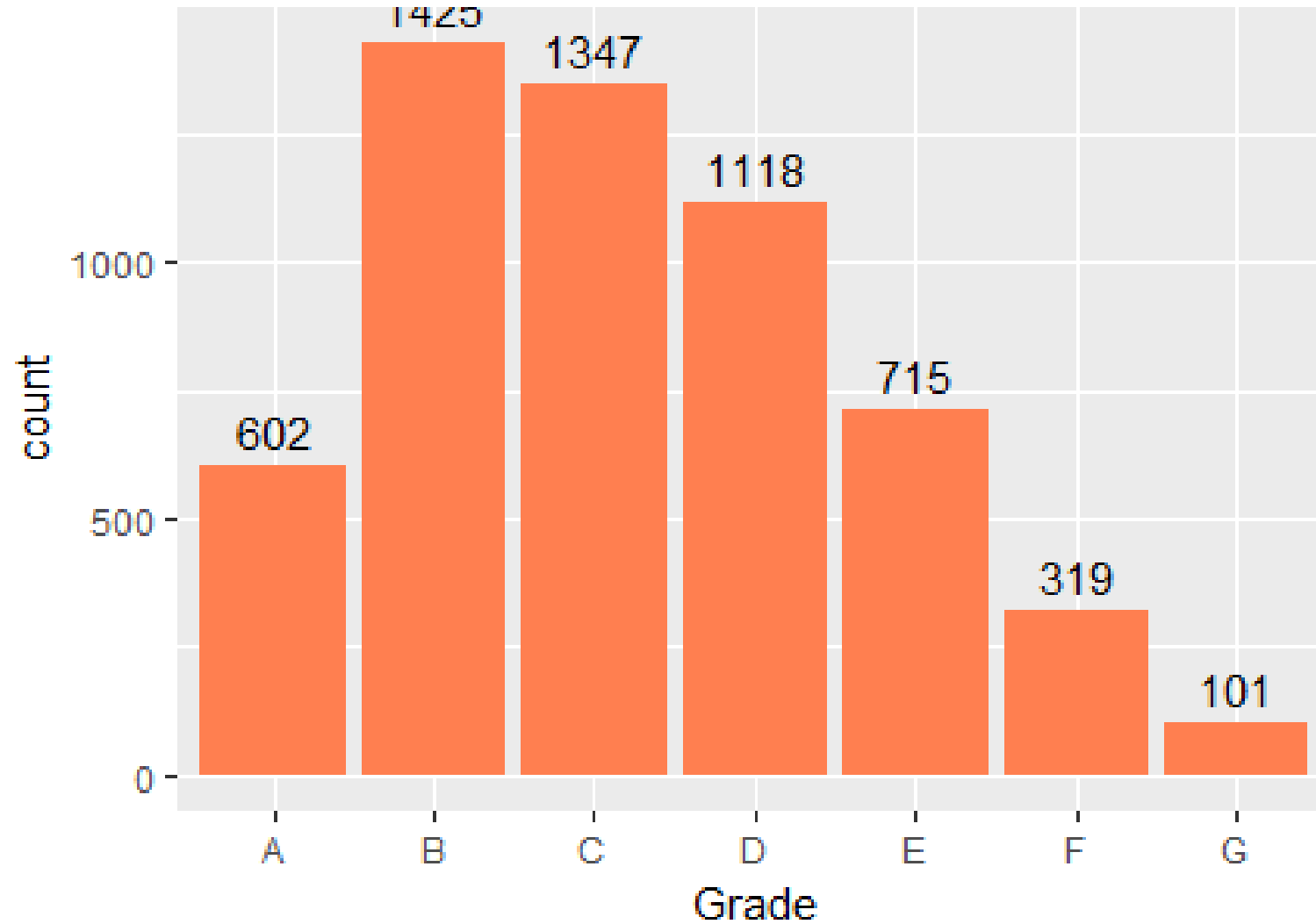
➤ 50% of default loan caused by those who rented the house.

➤ 41% of default loan inflicted by those who got it for mortgage.

Inquiry Last 6 months Vs Loan Default

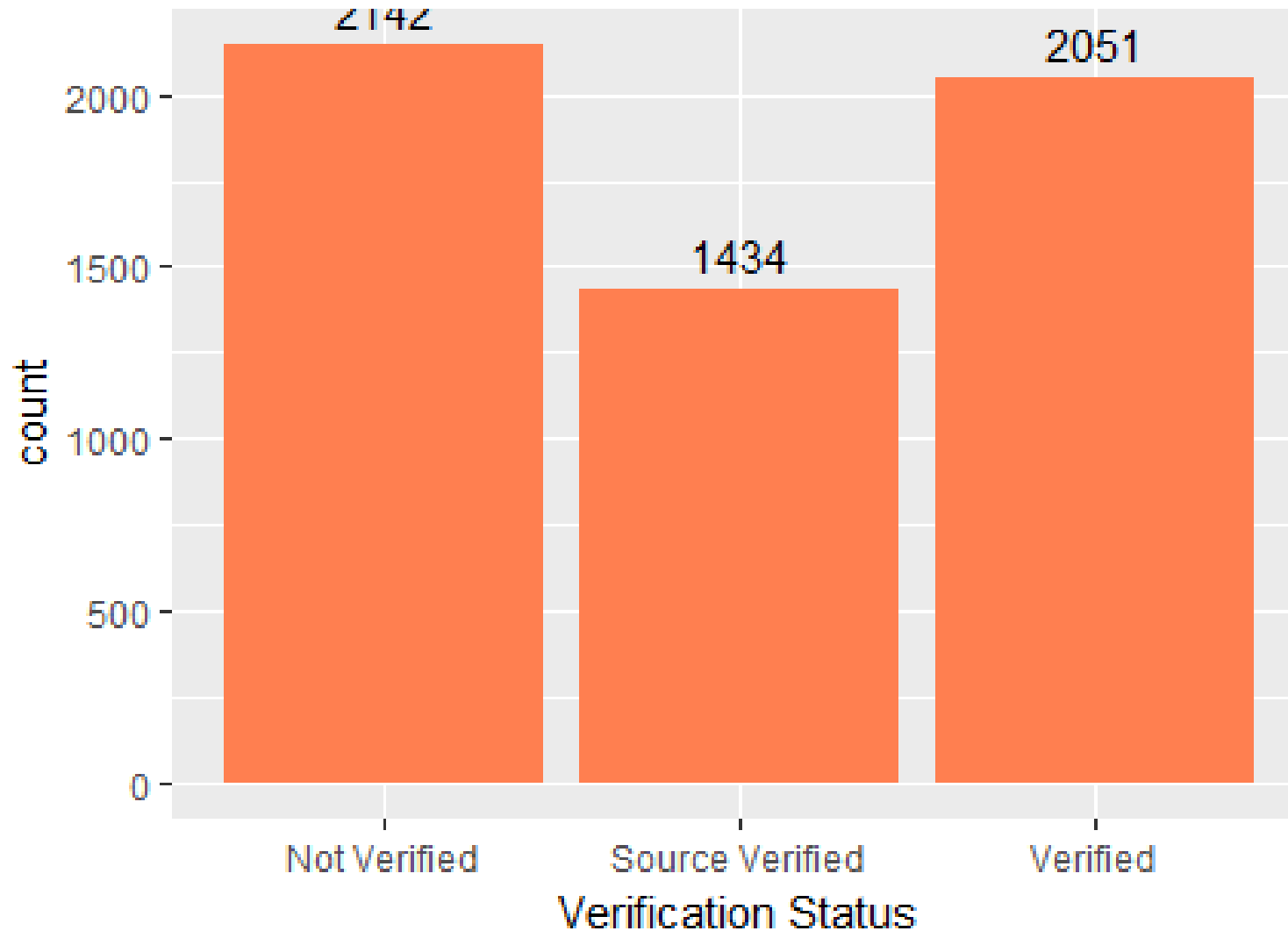


Grade Vs Loan Default



Customers categorized into Grade B,C,D and E are most likely to default the loan.

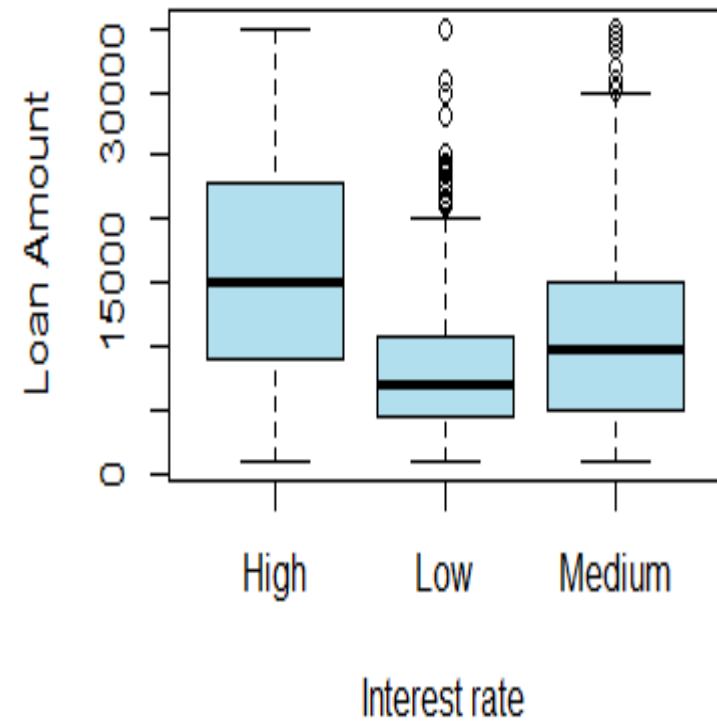
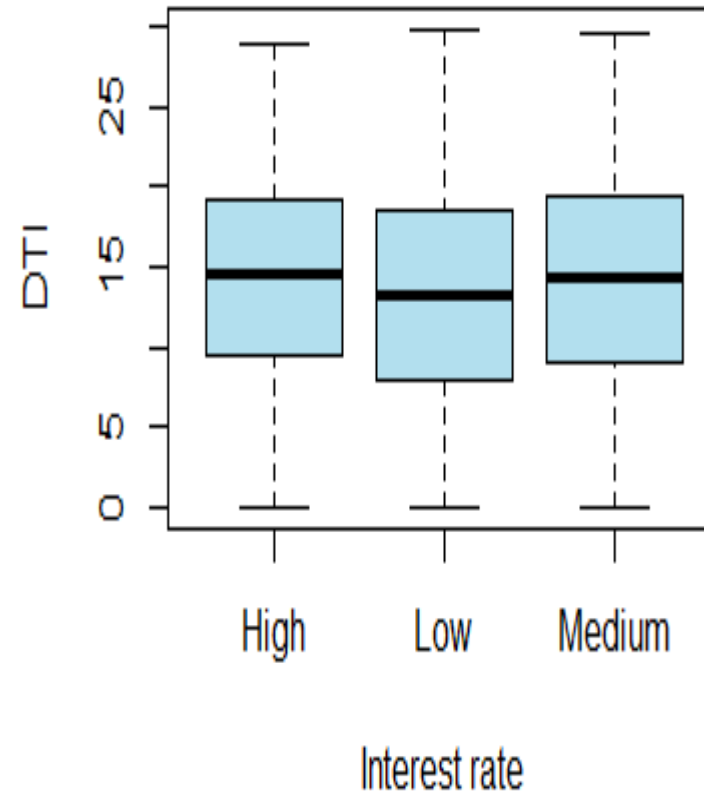
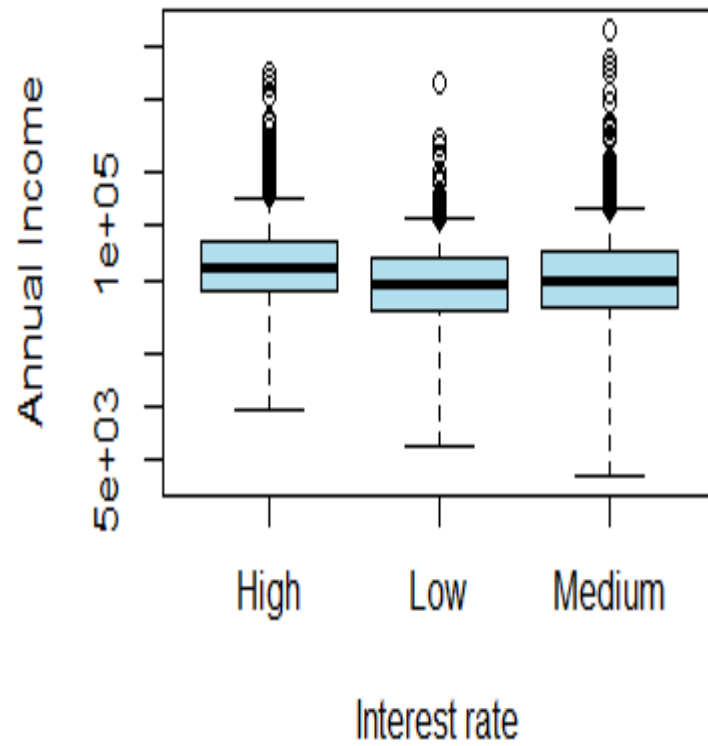
Verification Status



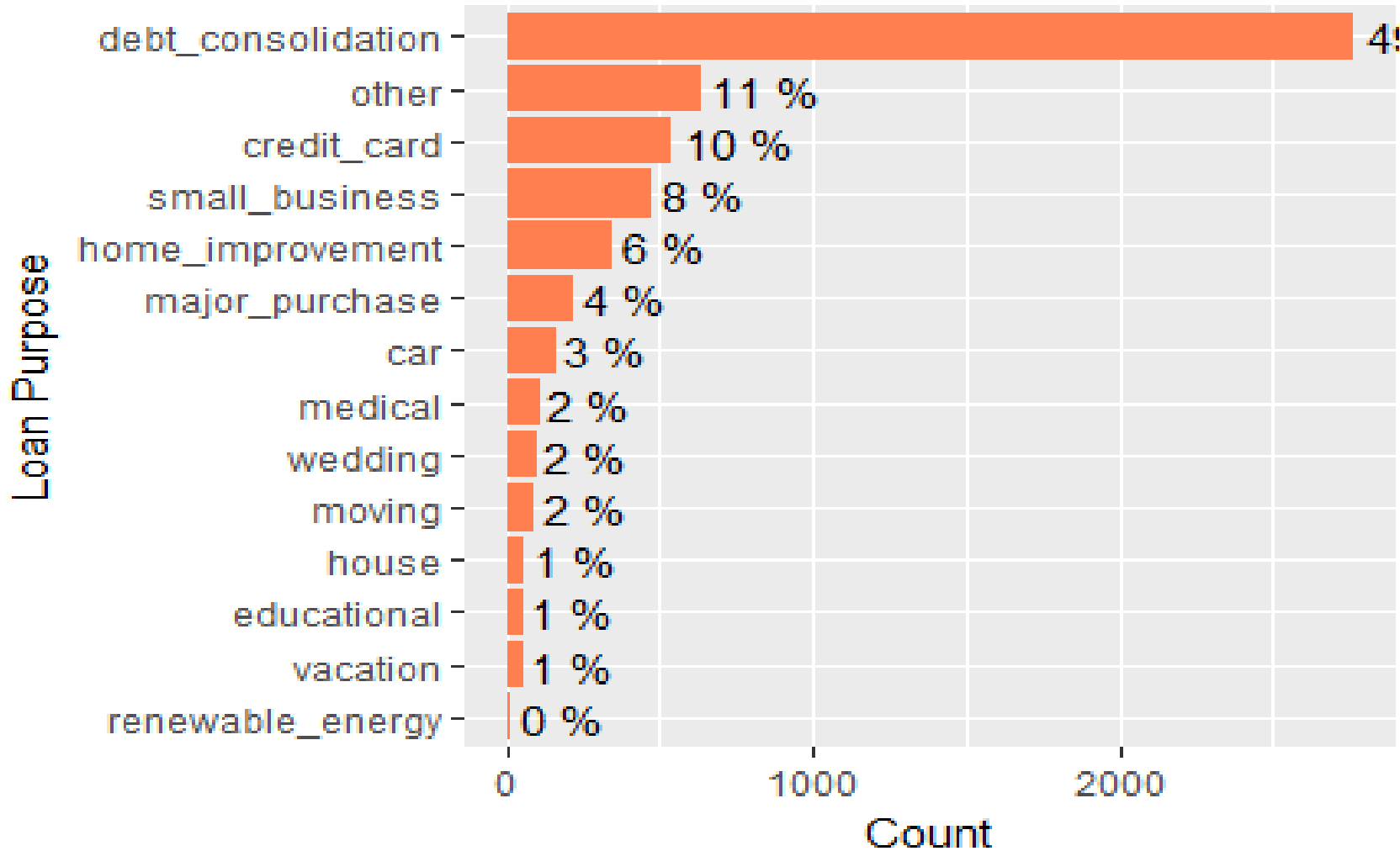
➤ 38% of total loans defaulted by the borrowers whose source of income was not verified.

➤ 36% of them was due to the borrowers whose source was verified.

Annual Income, DTI and Loan Amount Vs Interest Rate



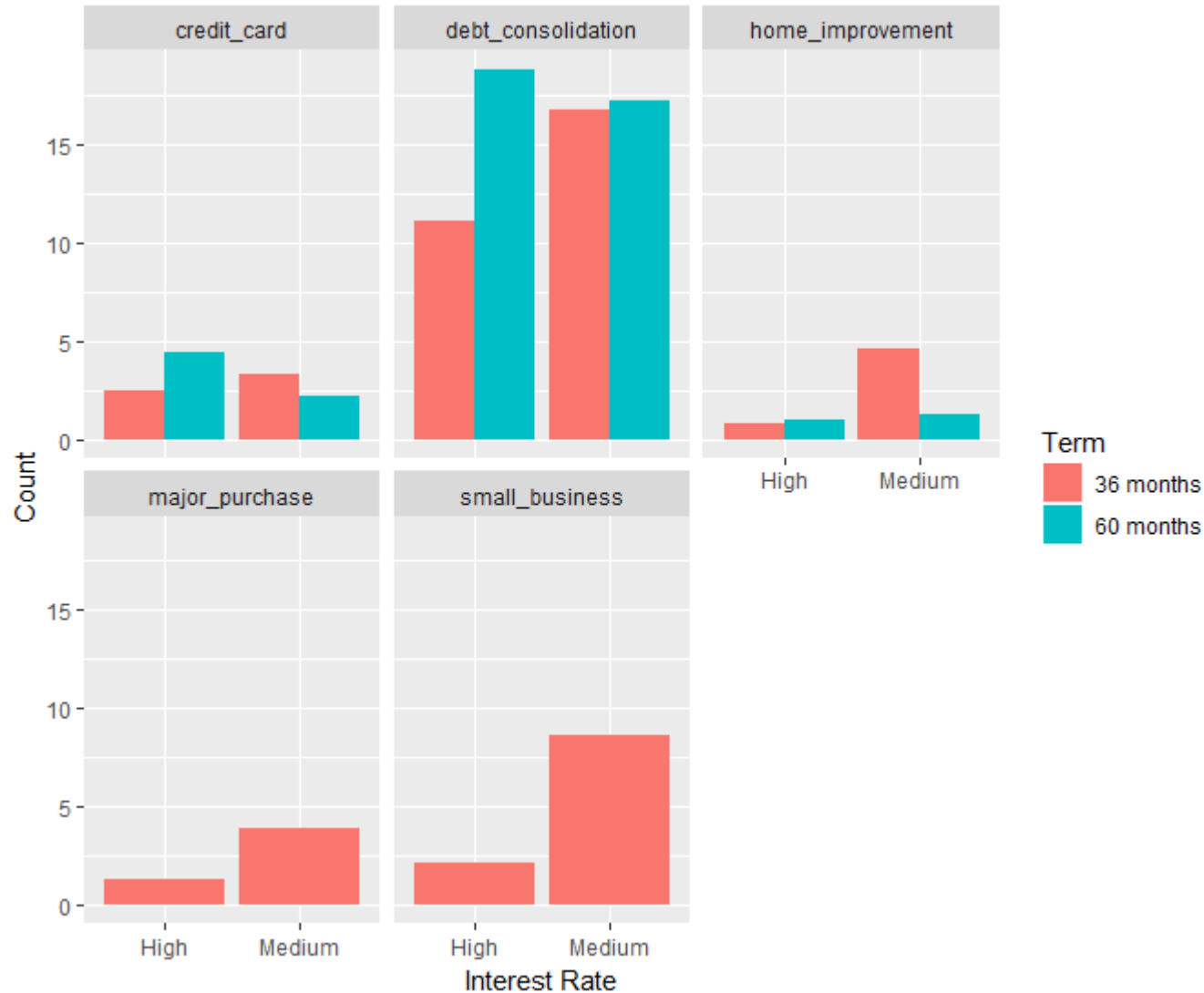
Loan purpose distribution – Top Loan Purposes



It is quite evident from this plot that loan taken for “Debt Consolidation” purpose has contributed significantly to loan default with 49%.

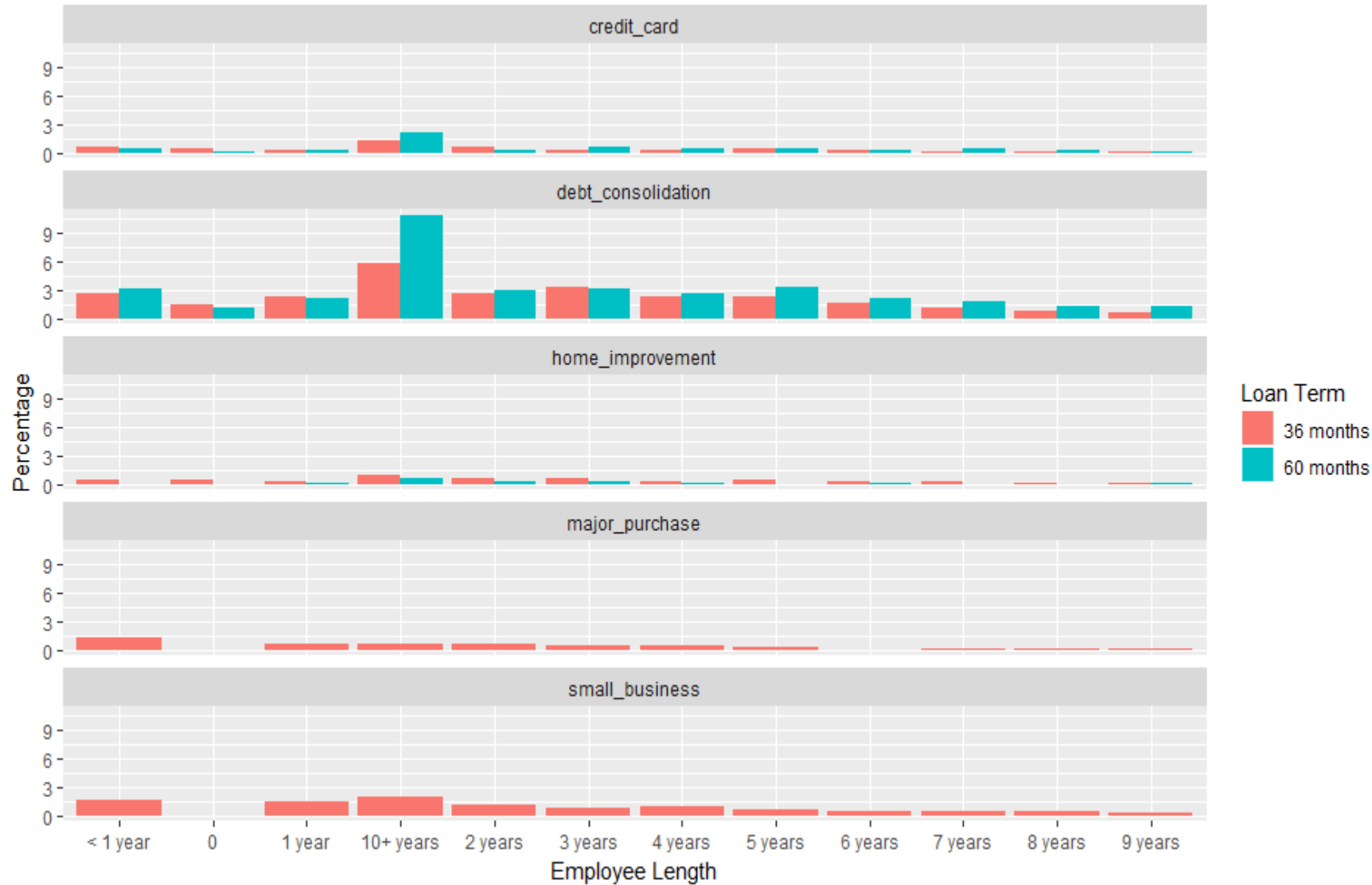
From this point, we will analyse top5 loan purpose records.

Interest Rate Vs Term across purpose



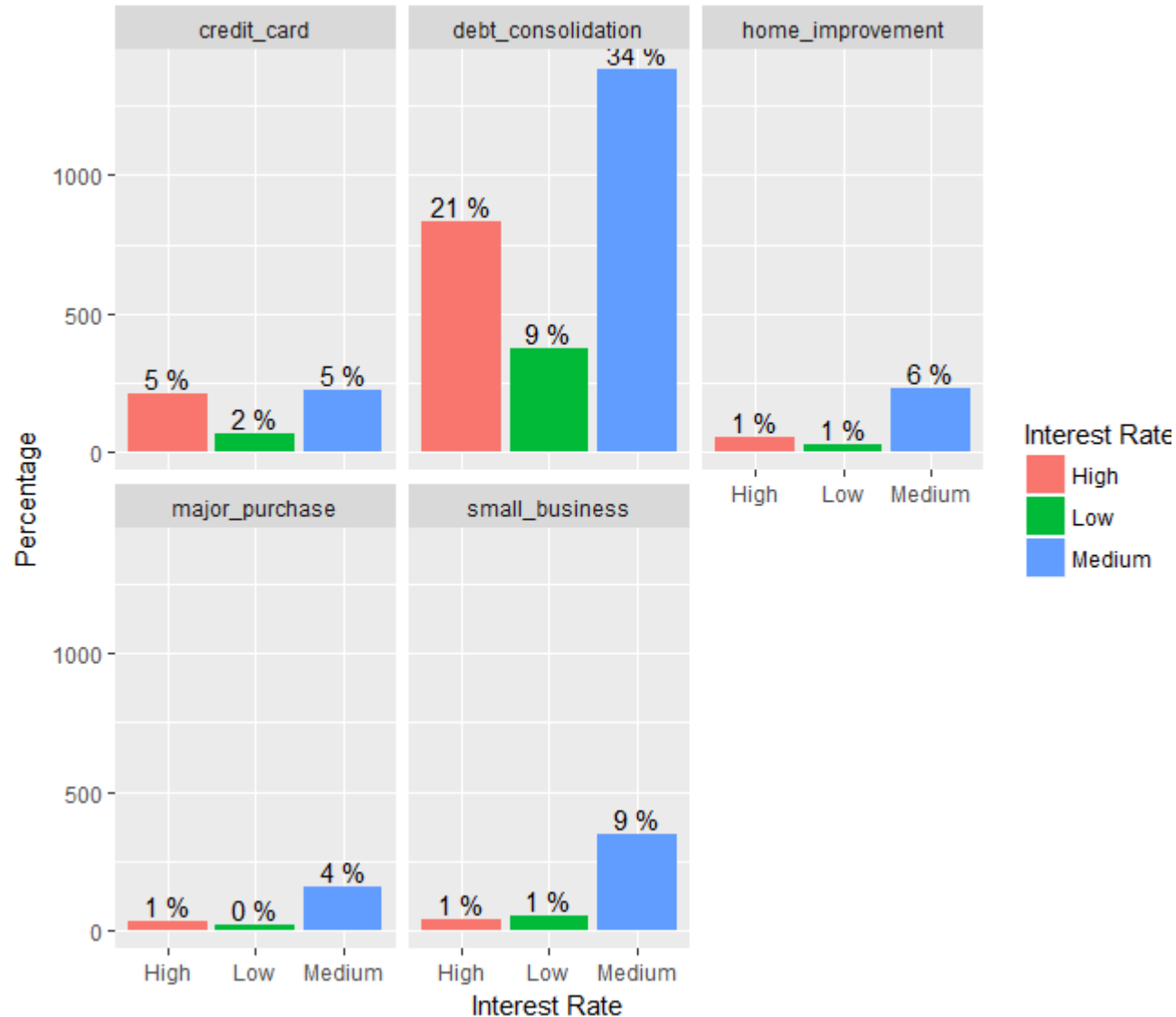
For Debt consolidation purpose, majority of borrowers with both high and medium interest rates have defaulted when loan term was 60 months.

Employee Length Vs Term of Loan across Purpose



More than 10% of customers with 60 months loan term have been behind loan default. All of them have employment length of more than 10 years.

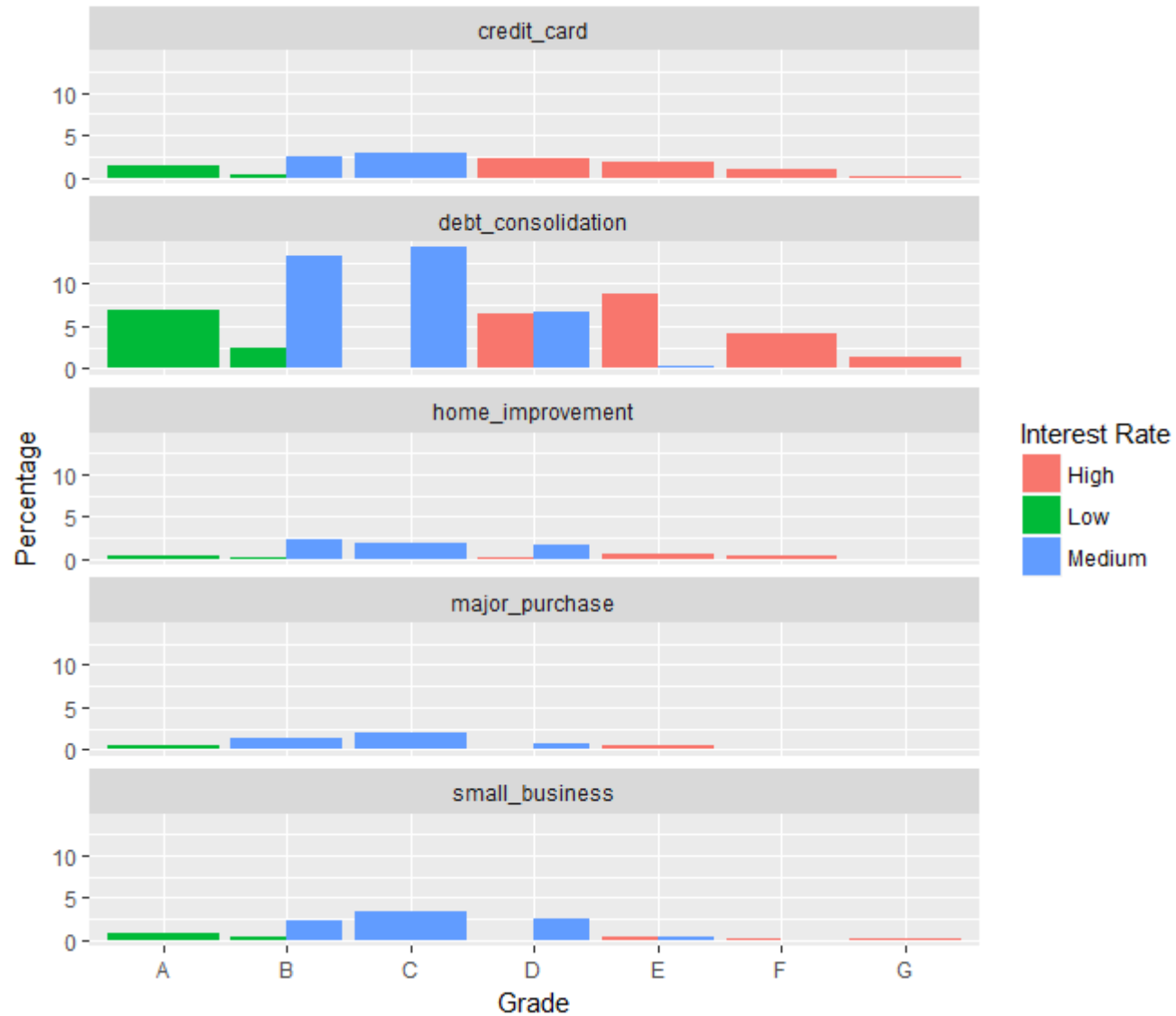
Purpose Vs Interest Rate



➤ Loans with medium interest rate have more contributed to loan default rate i.e. 58% across top 5 loan purposes.

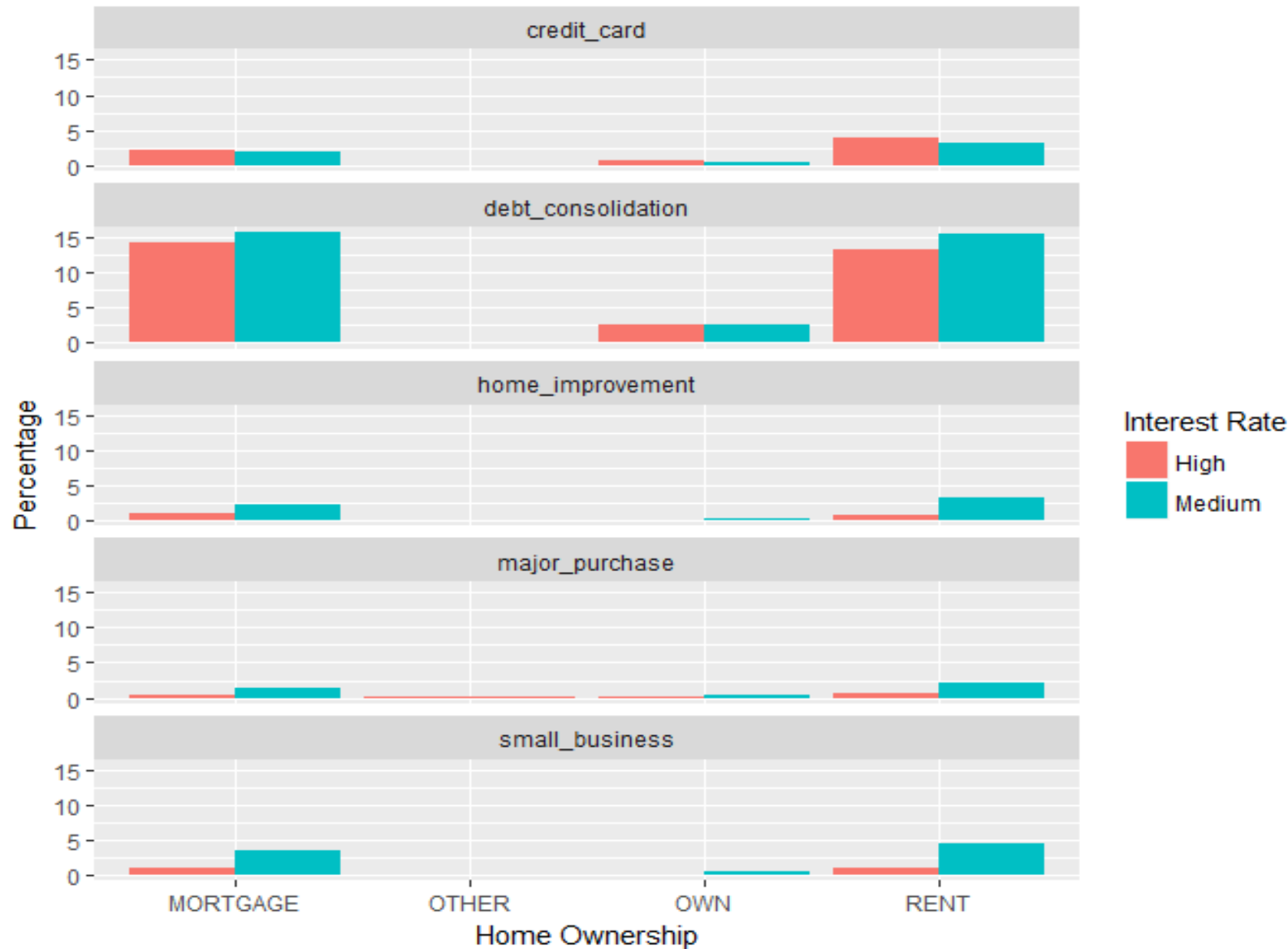
➤ 29% are from high interest rated loans.

Grade Vs Purpose



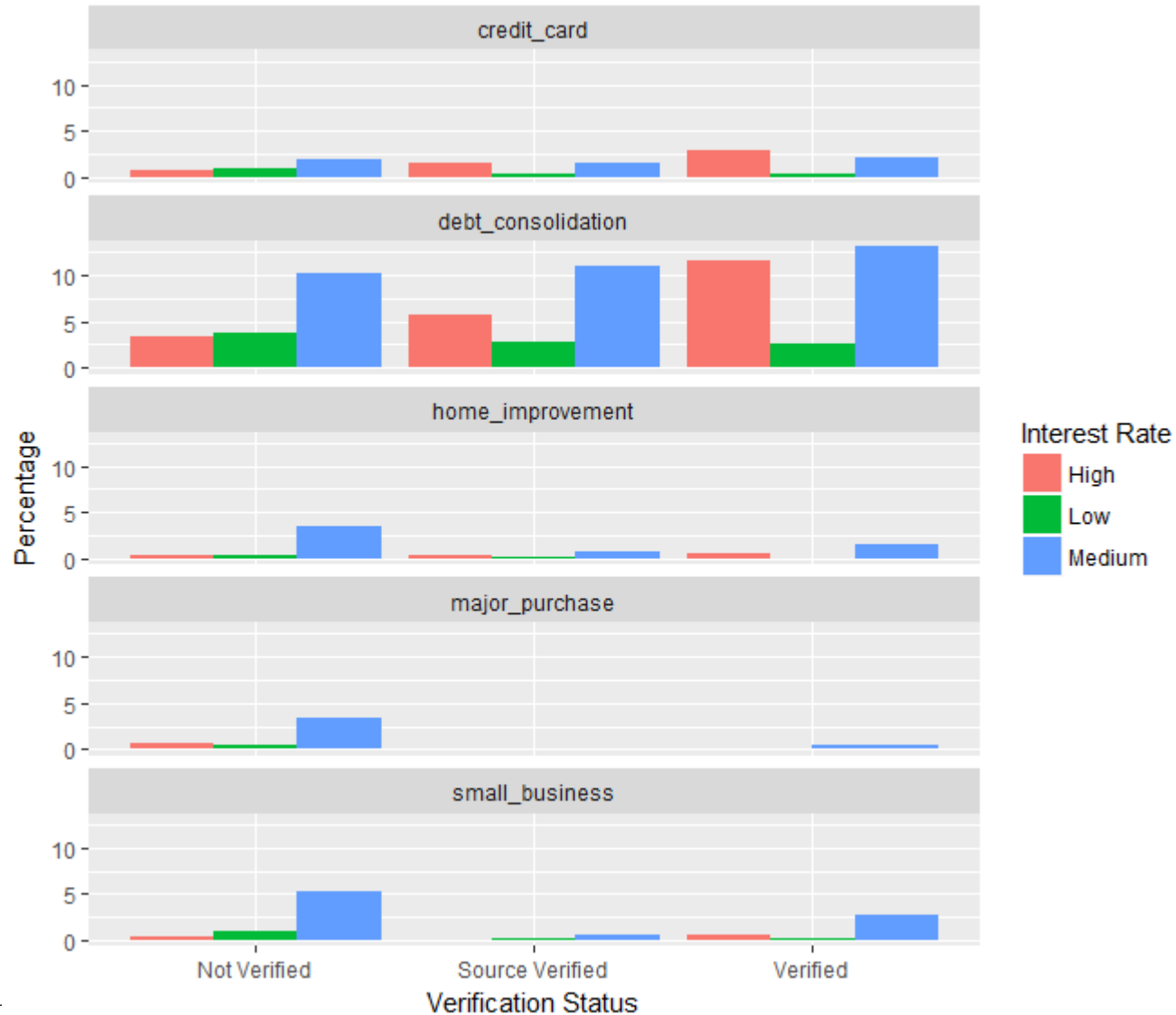
Customers with Grade B,C and D with medium interest rate have contributed significantly to loan default rates.

Home ownership Vs Purpose



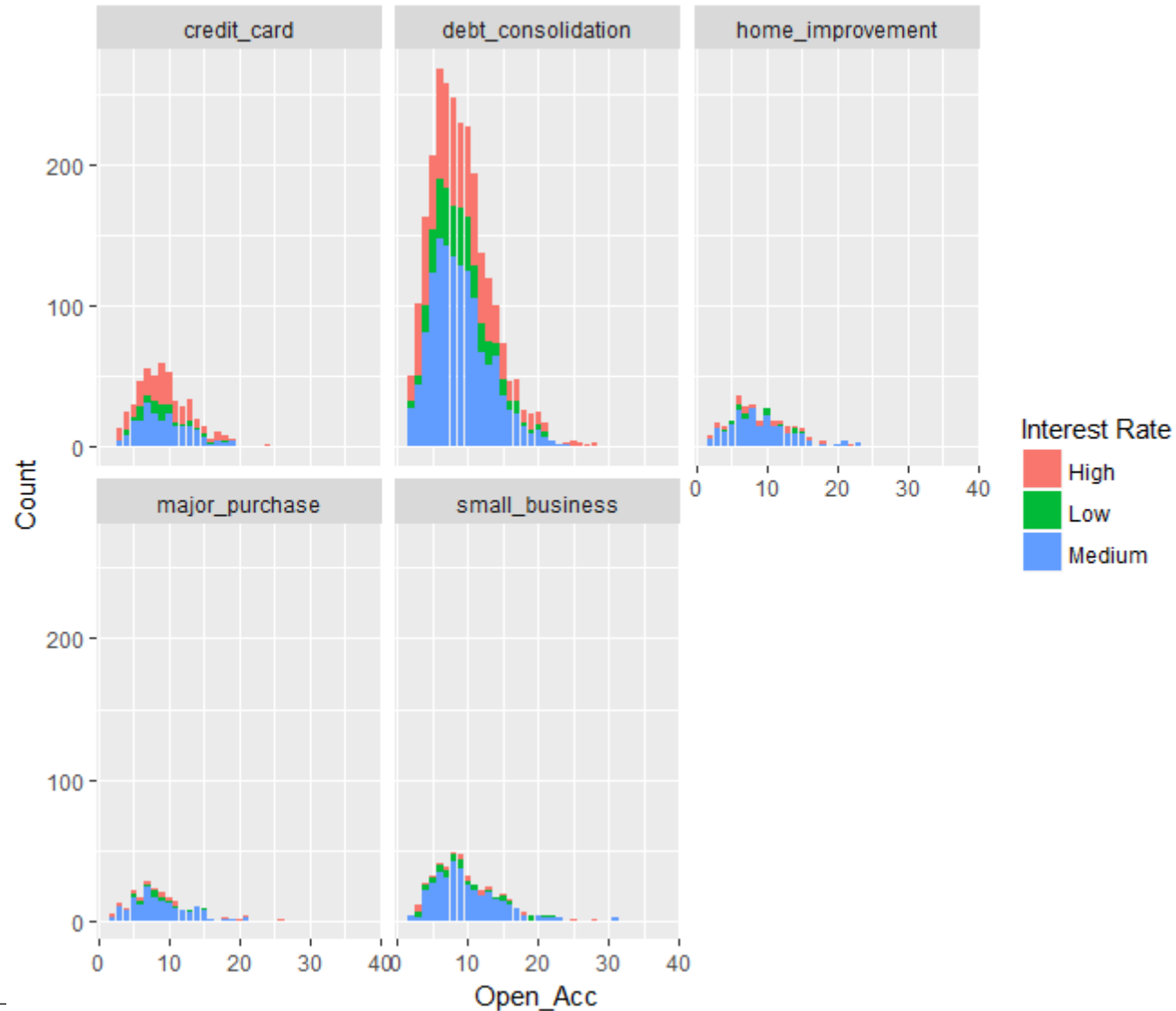
Both mortgage and rented ownership with both medium and high interest rates have contributed almost equal to loan default.

Verification Status Vs Purpose

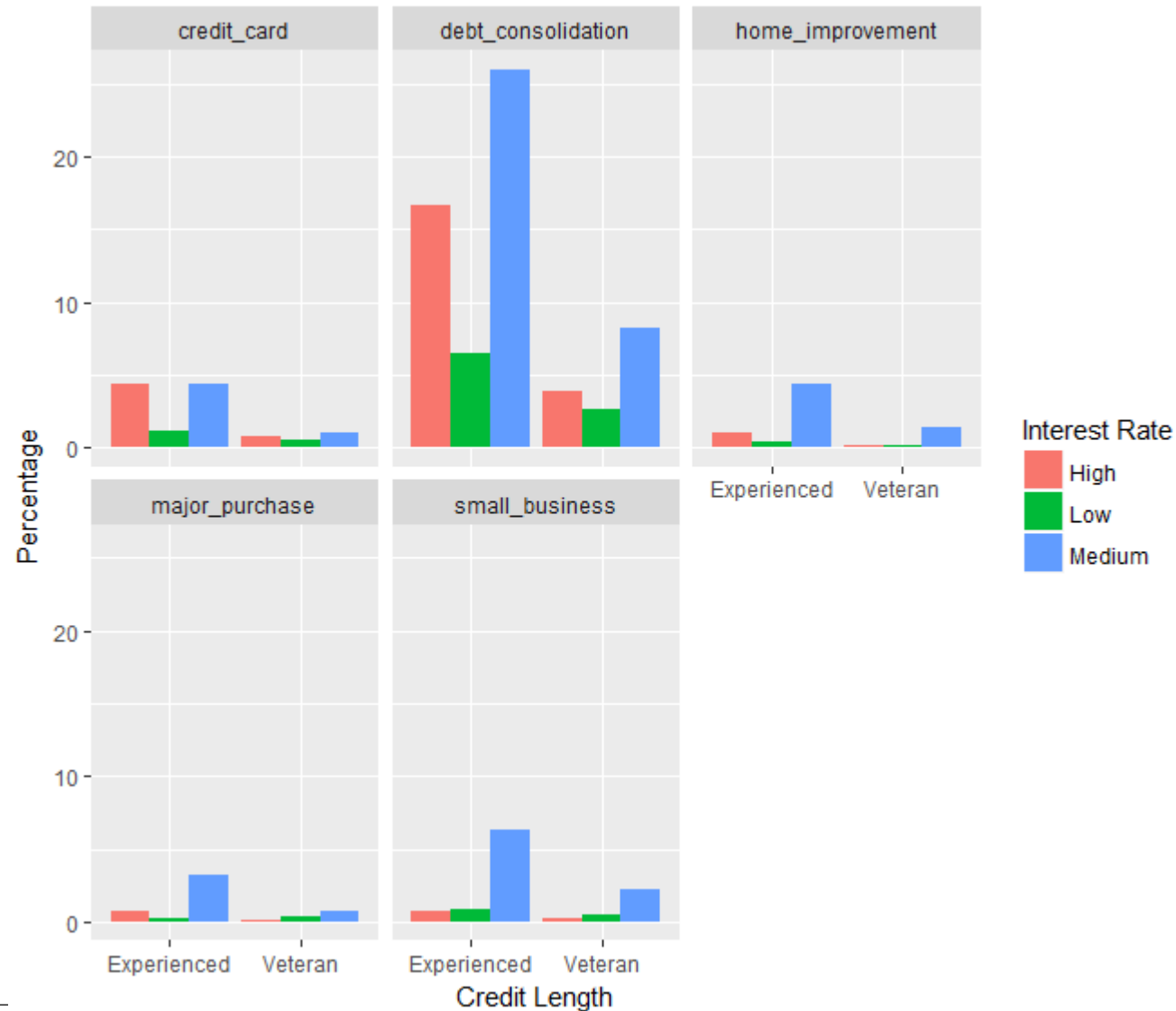


Even with verification status, both “Not Verified” and “Verified” status loans have high degree of default with medium interest rates.

Open Acc Vs Purpose



Credit card Length Vs Purpose

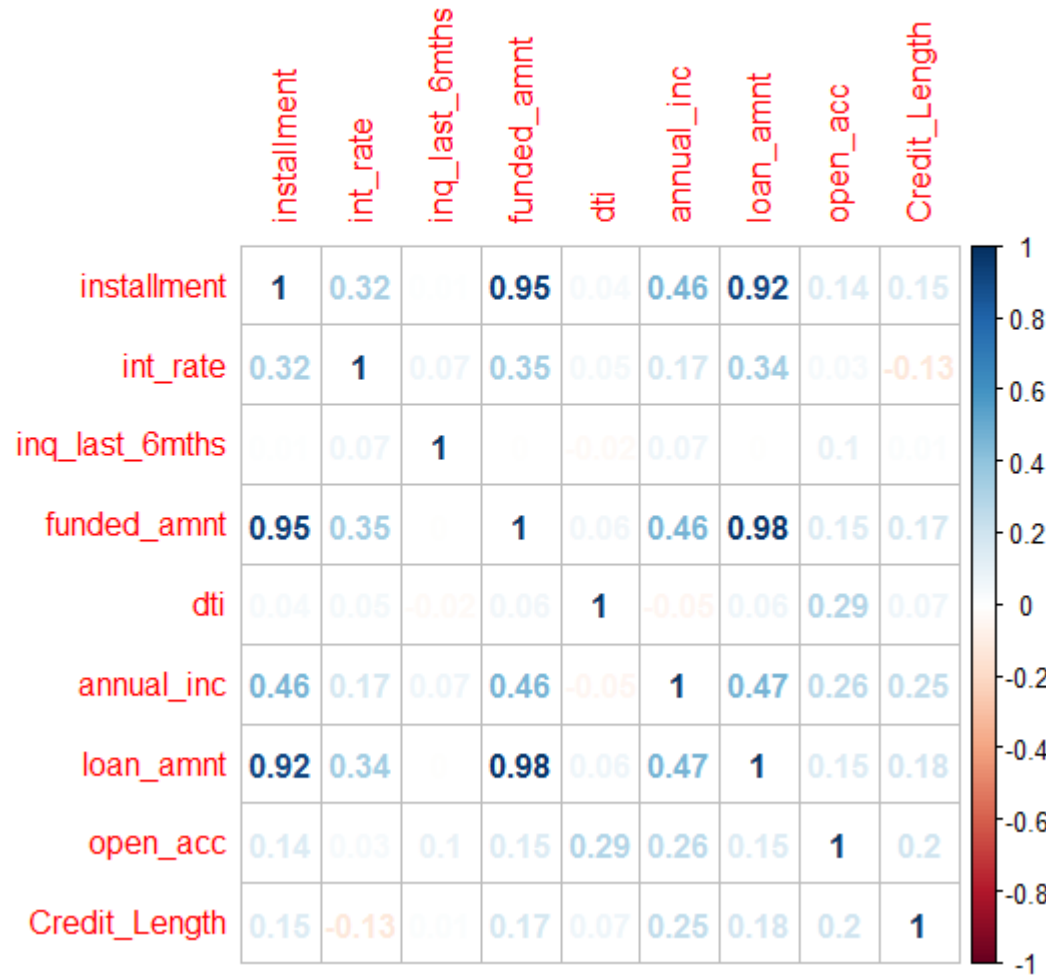


Under “Debt Consolidation”, those with Experienced credit usage and with medium interest rate have high number of loan defaults.

Credit length is grouped as below.

Credit Length	Group
10-25 Years	Experienced
> 25 years	Veterans

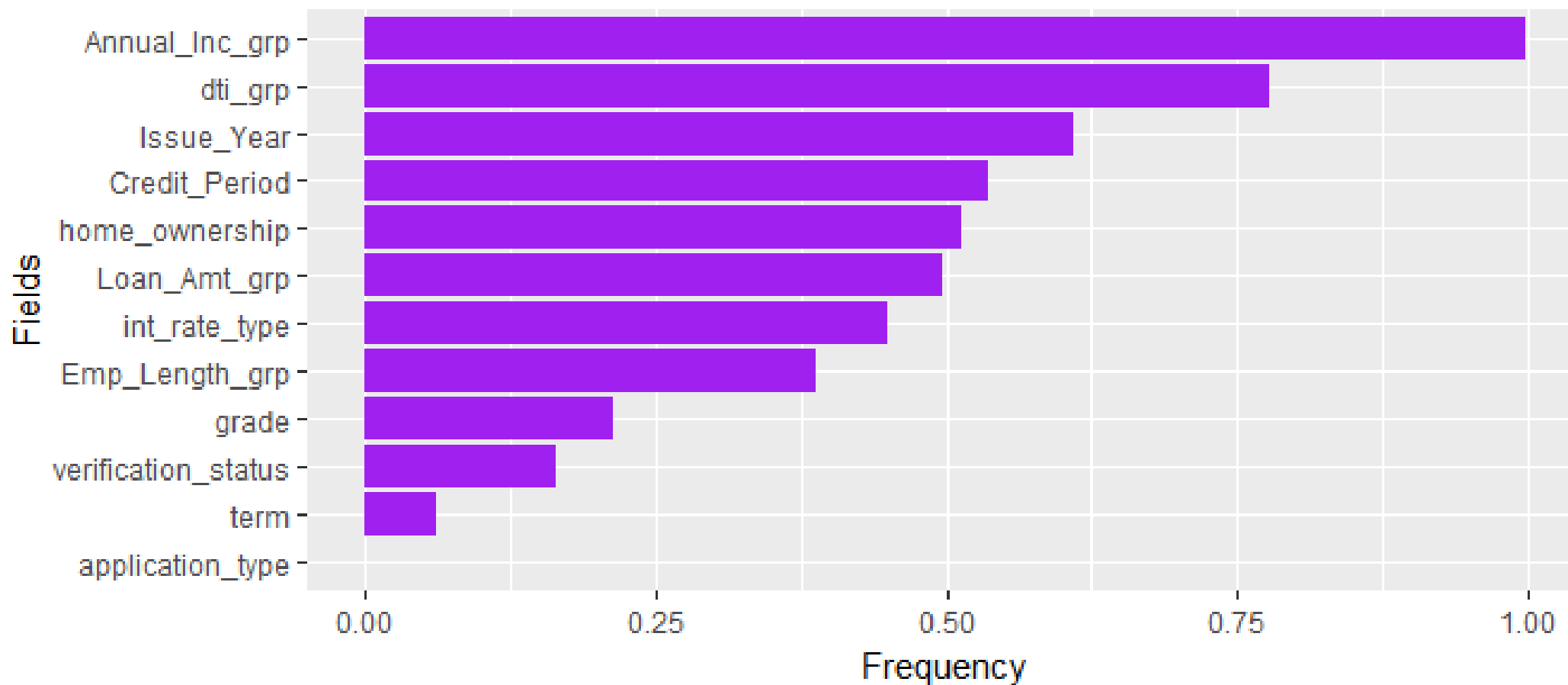
Correlation Matrix on quantitative variables



Loan Amount and Funded Amount are positively correlated with 0.98

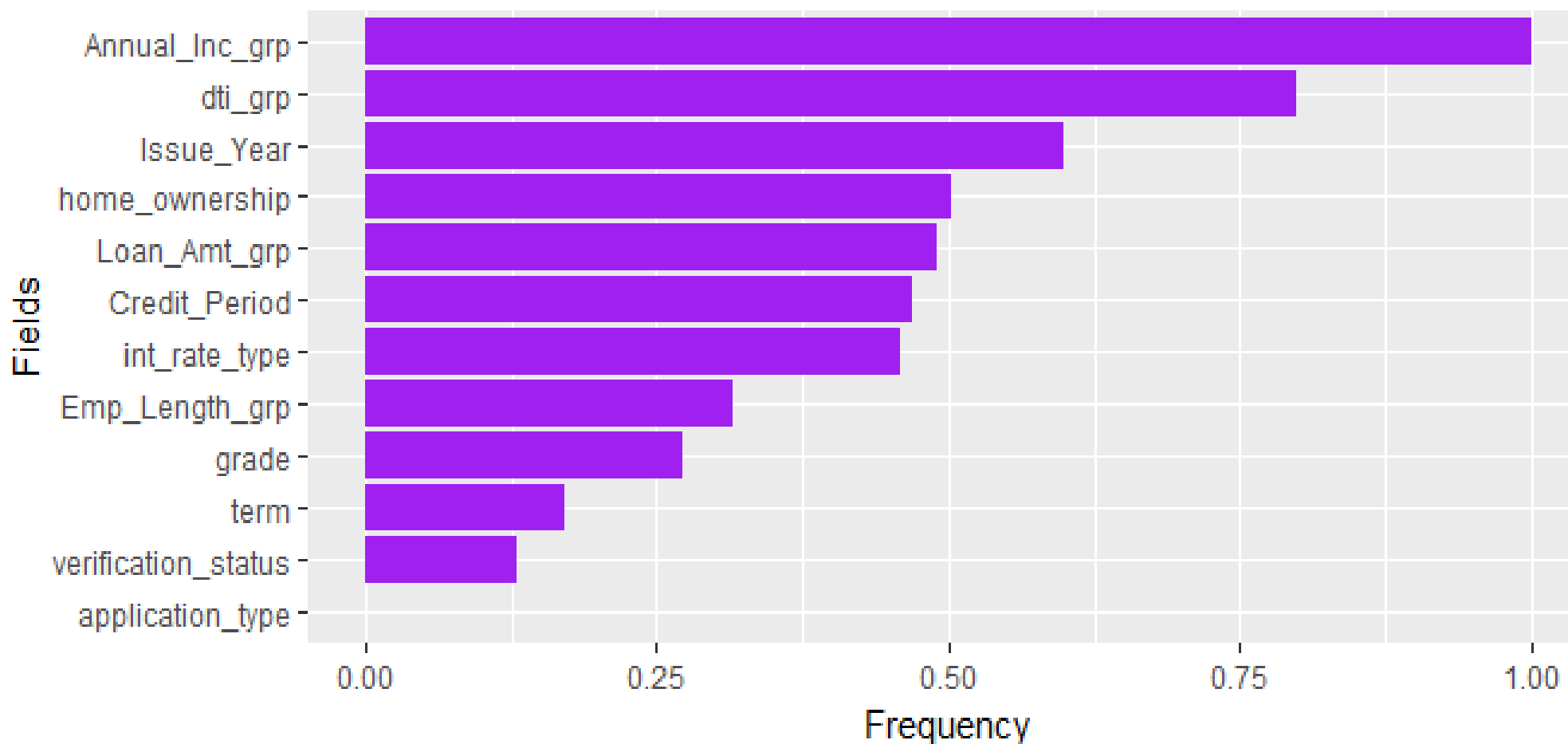
Important Variables : Debt Consolidation

Strong Drivers of Loan Default in Descending Order



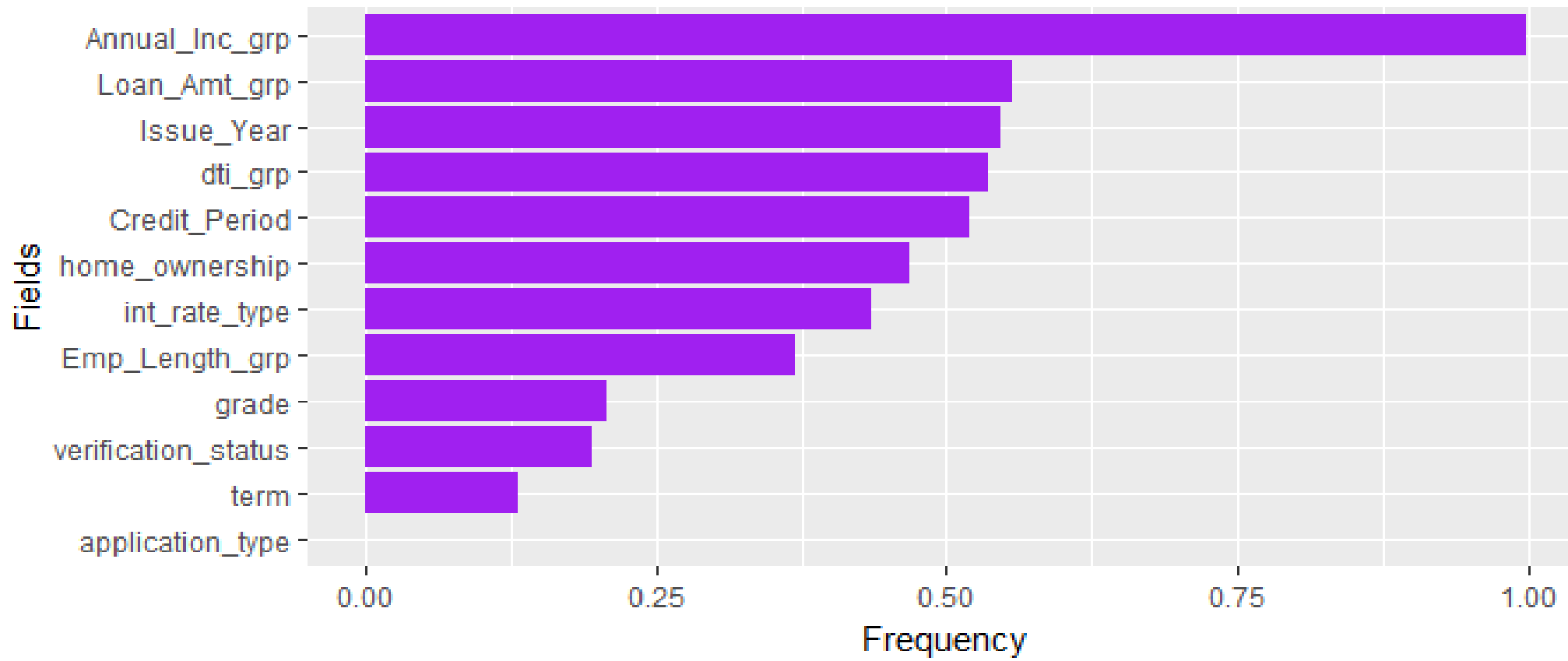
Important Variables : Credit Card

Strong Drivers of Loan Default in Descending Order



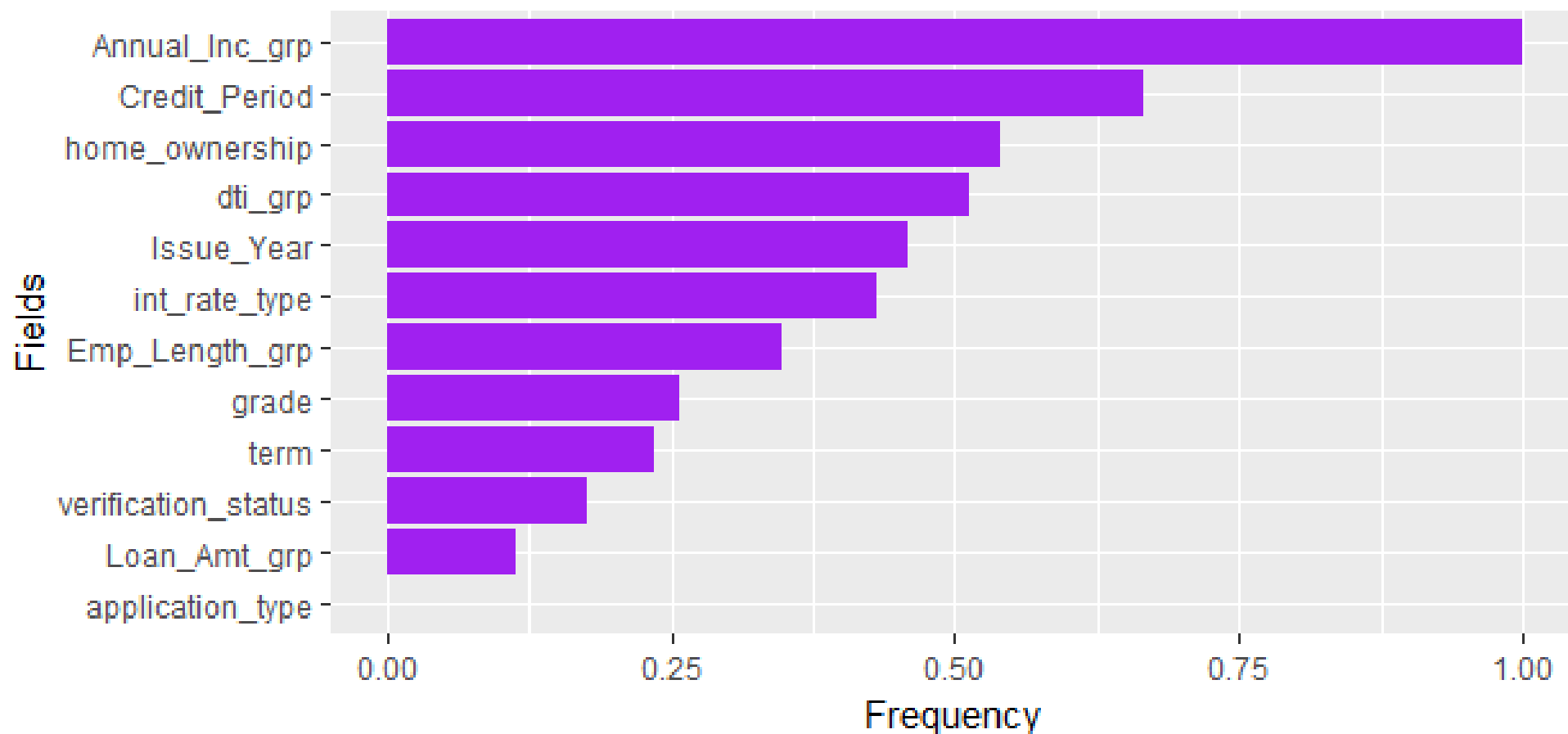
Important Variables : Small Business

Strong Drivers of Loan Default in Descending Order



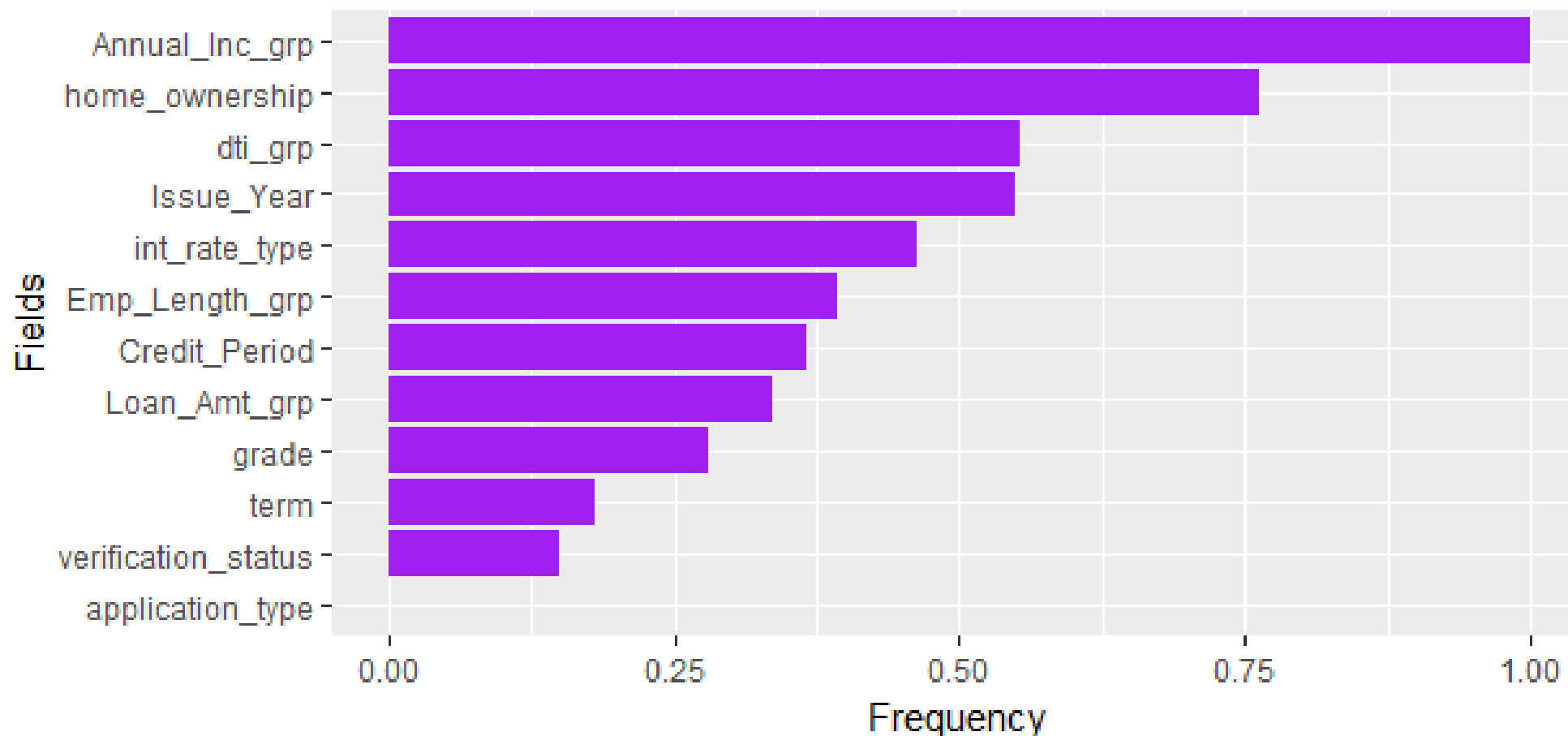
Important Variables : Major Purchase

Strong Drivers of Loan Default in Descending Order



Important Variables : Home Improvement

Strong Drivers of Loan Default in Descending Order



Conclusion

Based on the analysis on various variables, we can firmly disclose that the following loan parameters are strong indicators or driving factors behind loan default rate. In fact the order of these variables may vary depending on purpose of the loan. However, holistically they all contribute significantly to higher loan default. As it is clearly evident that many of the defaulted loans fall in medium interest rates, while computing the interest rates these strong indicators could be considered for giving higher interest rates.

- 1) Purpose of the Loan
- 2) Annual Income
- 3) DTI
- 4) Credit Length
- 5) Home Ownership
- 6) Employment Length
- 7) Loan Amount
- 8) Interest Rate
- 9) Grade