GRAMENER CASE STUDY

SUBMISSION

Group Name:

1. Kanupriya Kushwaha
2. Omkar Murthy Appala
3. Santhosh Dhanabal
4. Thogaruchiti Khasim Basha

Abstract – Identification of factors behind loan default

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

Objectives:

- Identify the risky loan applicants to reduce credit loss. Borrowers who default cause the largest amount of loss to the lenders. In this case, the customers labelled as 'charged-off' are the 'defaulters'.
- Understand the **driving factors (or driver variables)** behind loan default, i.e. the variables which are strong indicators of default.
- This knowledge will be used in the processing of future applications.

Problem Solving Methodology

Business Understanding

- Understanding the problem at hand (Identification of the risky loan applications to minimize the reduce loss)

Data Understanding

- Understanding the data at hand (Information about past loan applicants and whether they ‘defaulted’ or not. Data for all loans issued through 2007 to 2011)

Data Preparation

- Clean and prepare the data so it can be effectively utilized for EDA

Exploratory Data Analysis

- Perform Univariate and Multivariate analysis to understand the distribution and relationship of variables

Conclusion

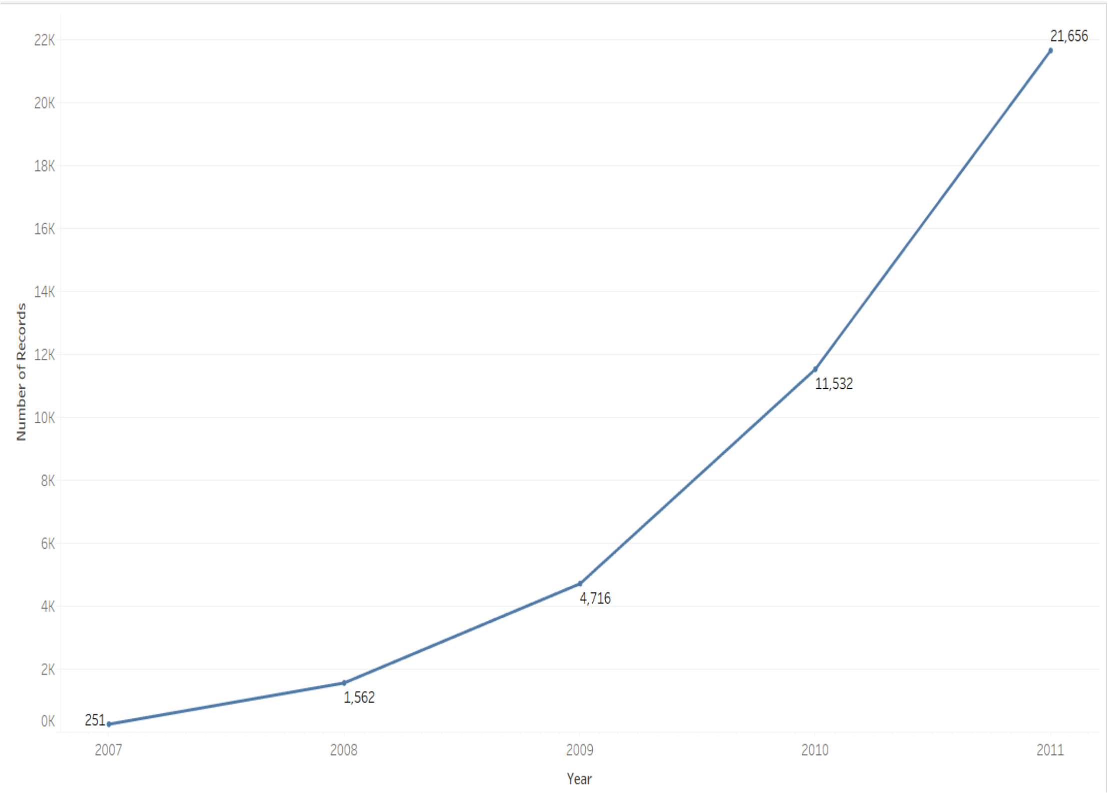
- Identify driving factors behind the loan default

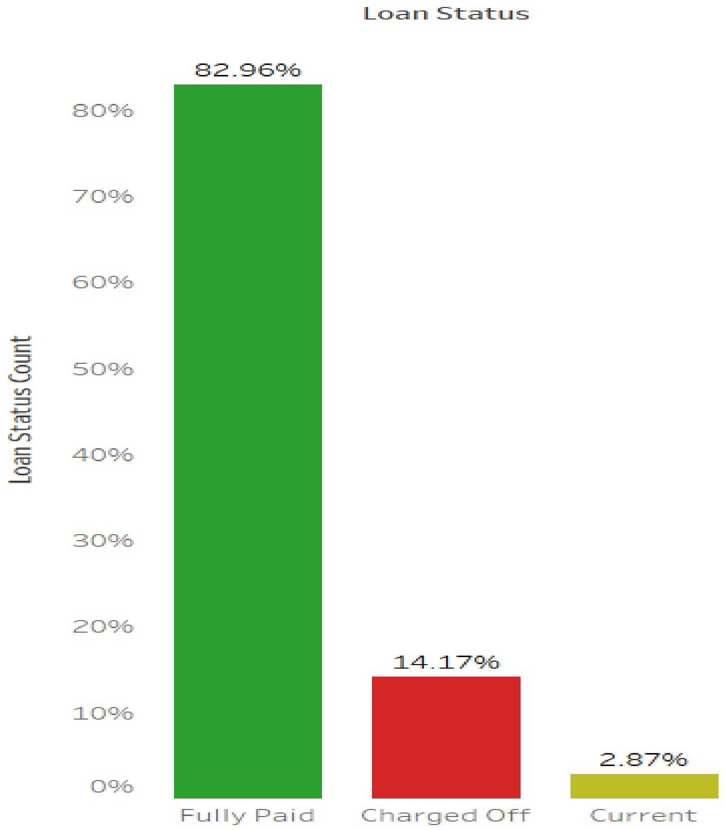
Data Cleaning

- Remove columns having all 0 or n/a values
- Remove descriptive (title, emp_title, desc etc.) or unique (id, member_id, URL etc.) columns
- Remove columns that relate to the current loan status (out_prncp, total_pymnt, last_pymnt_d etc.)
- Remove other columns that don’t contribute to the analysis

- Introduce new metrics as and when required
- Correct numeric, date and string formats
- Outlier detection

Issue at hand



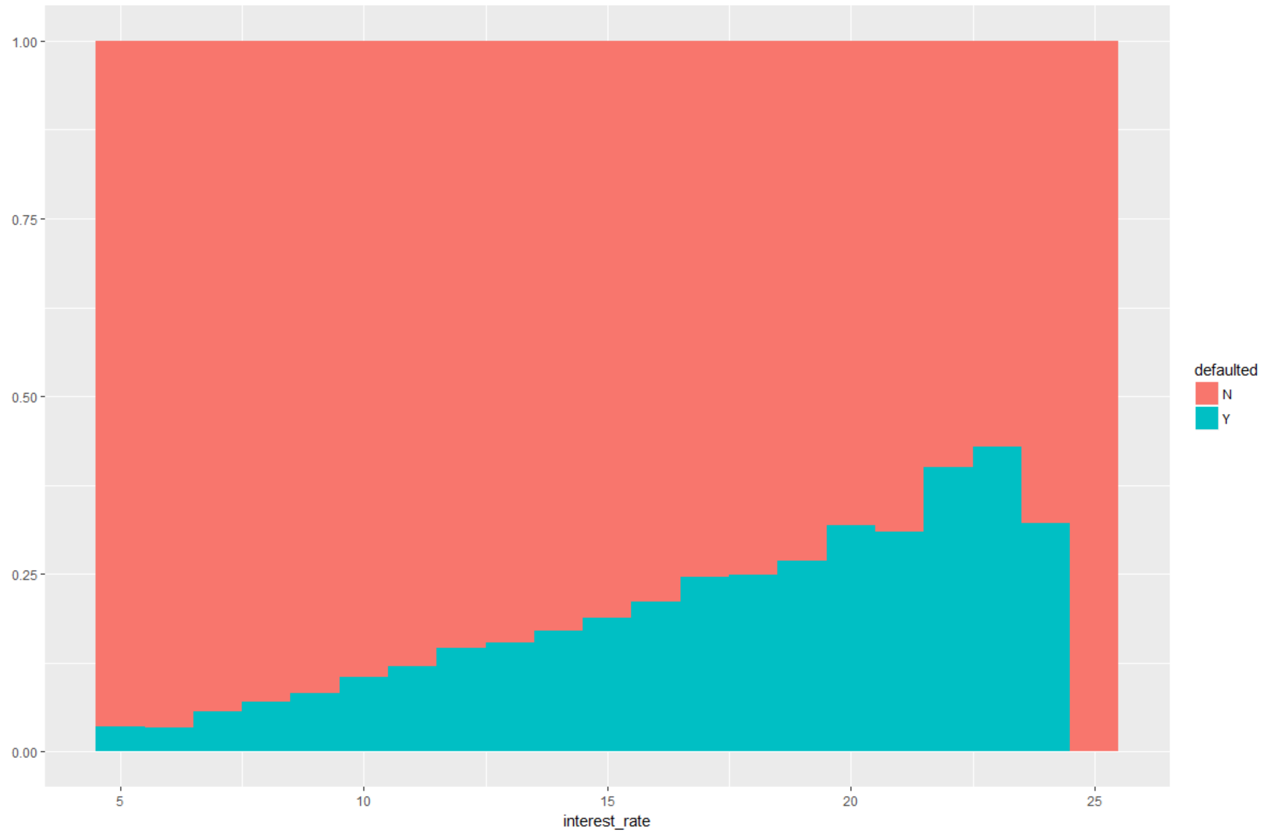


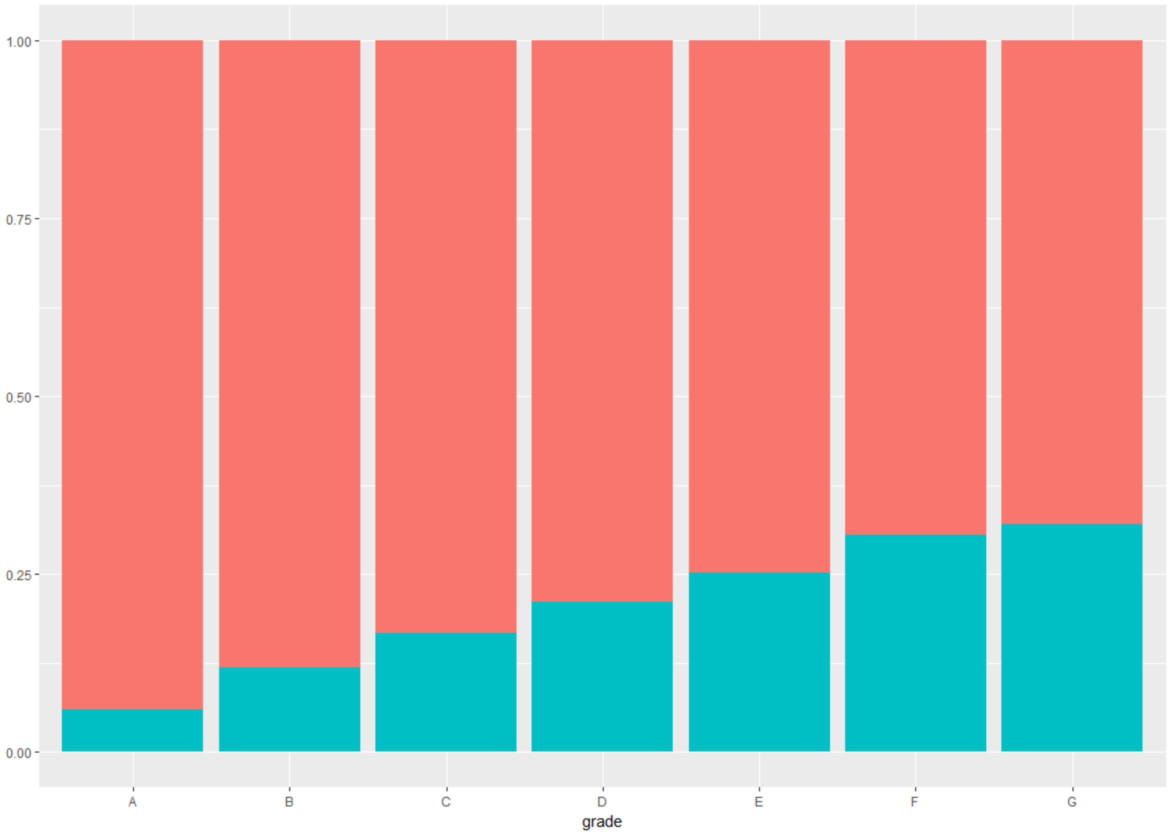
Numbers of loans taken are increasing almost exponentially

But around 14% of total loans are being Charged-off.

The aim is to find loan and consumer attributes that are strong indicators of loan default.

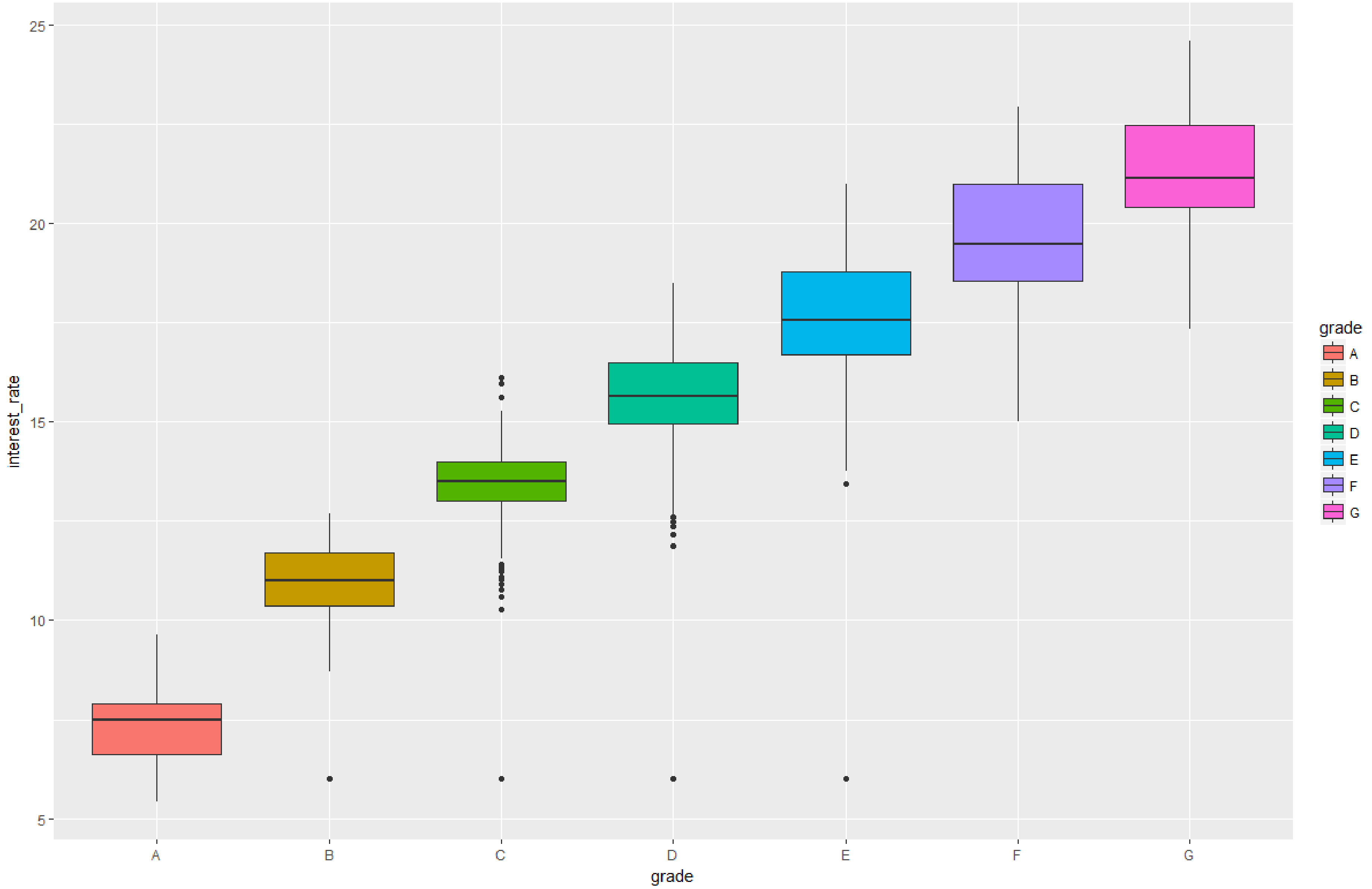
Effect of Interest Rate and Grades on Loan Default



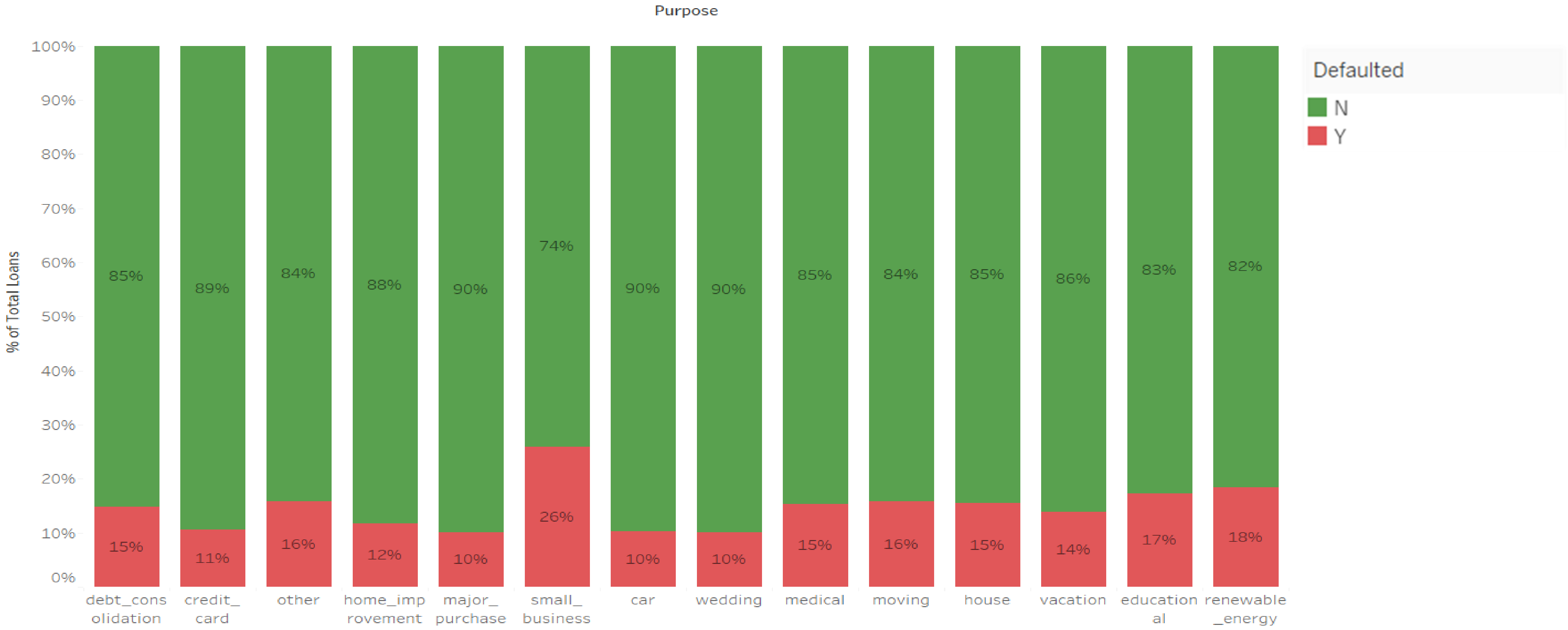


- Loan default increases with higher Interest Rate and worse Grades get.
- This is because Grades are driven by Credit Scores (like FICO, CIBIL) and Interest Rates are driven by Grades.

Relationship between Grades and Interest Rate

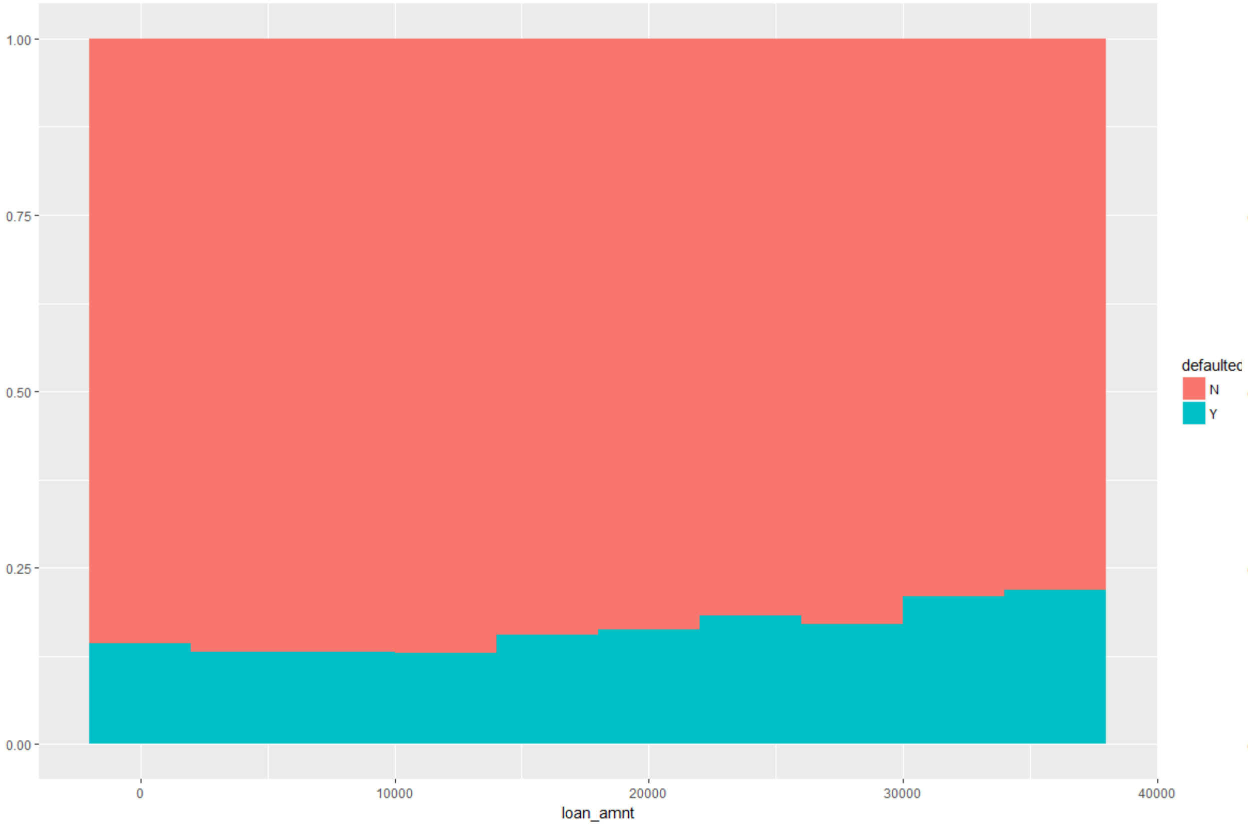


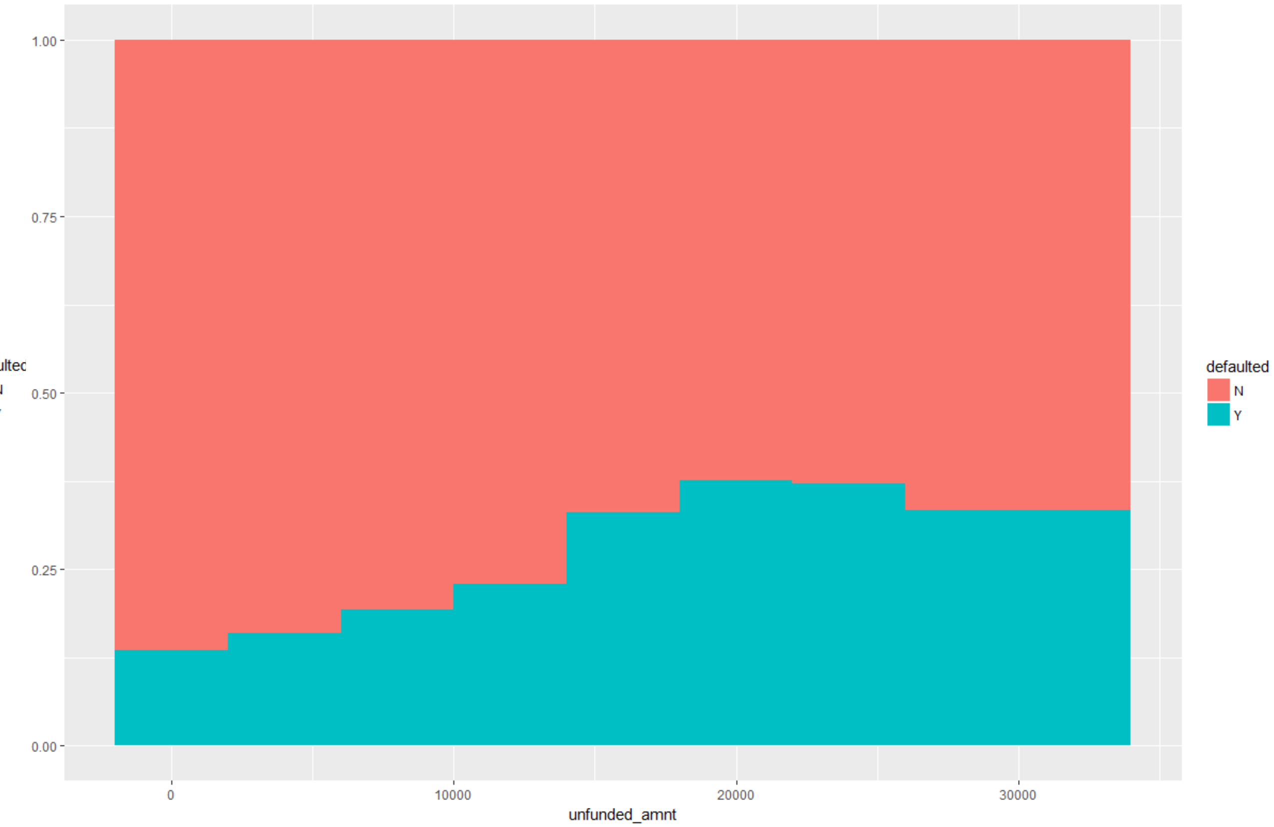
Effect of “Purpose” on Loan Default



Loan purposes like “Small Business", "Renewable energy”, ”Educational” have higher default percentage

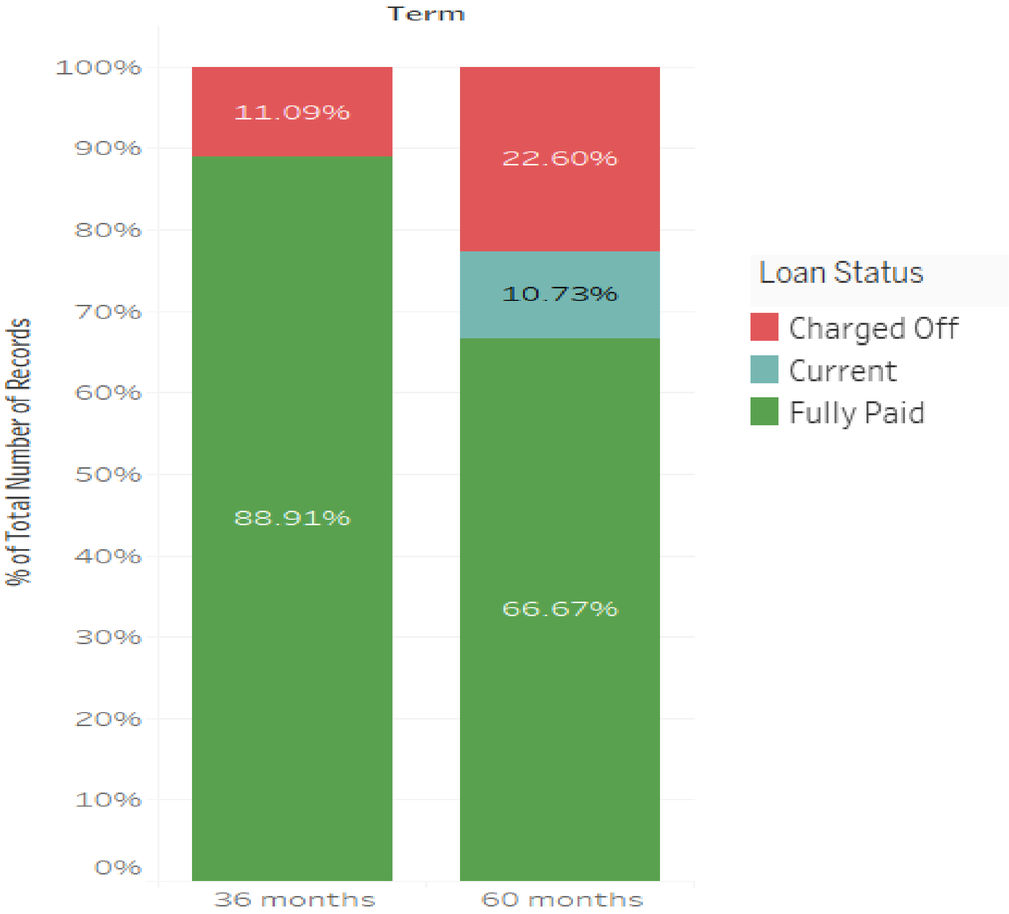
Effect of Loan Amount and “Unfunded Amount” on Loan Default

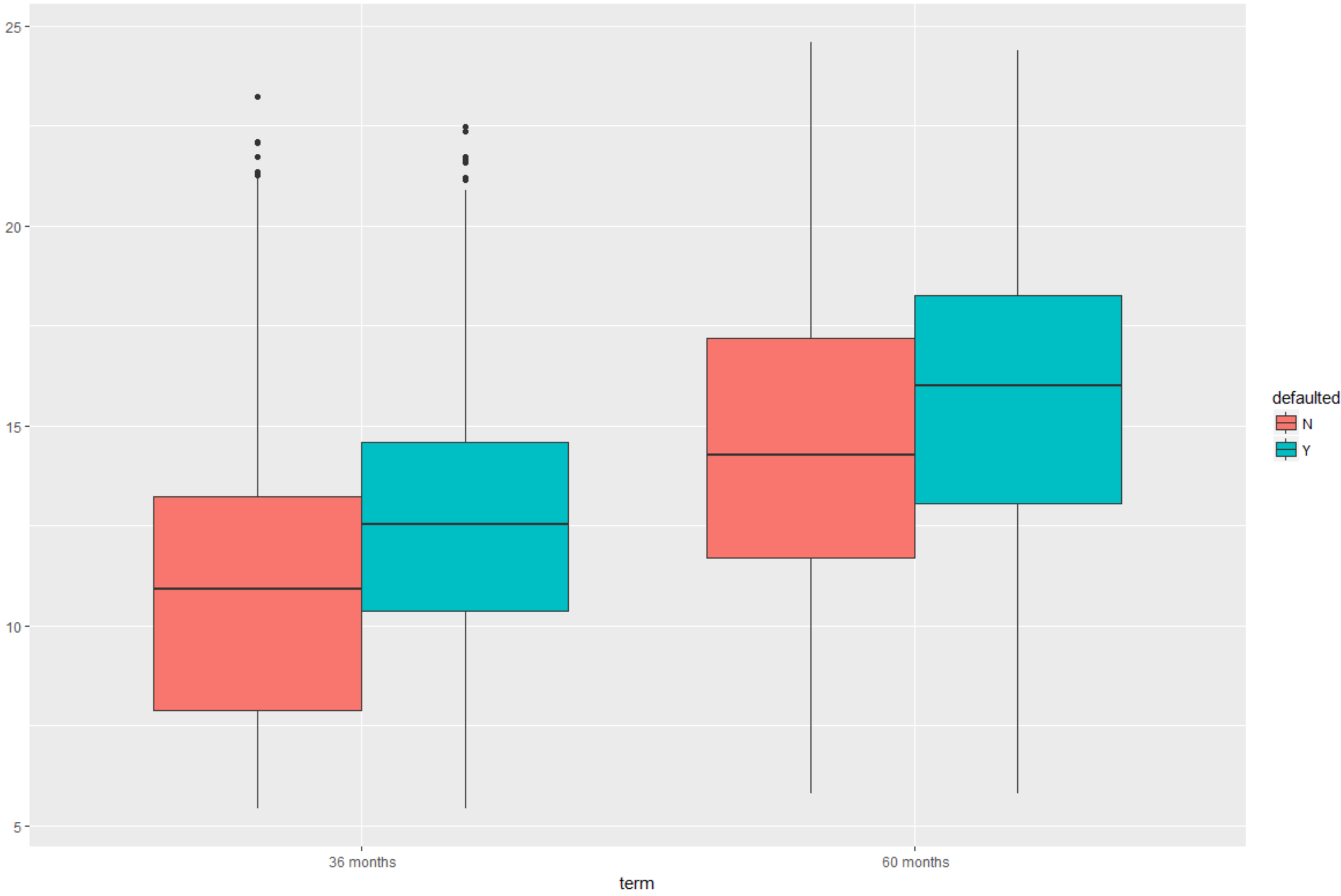




Higher loan amount as well as higher “unfunded amount” (loan_amnt – funded_amnt_inv) both impact loan default rate.

Effect of “Term” on Loan Default

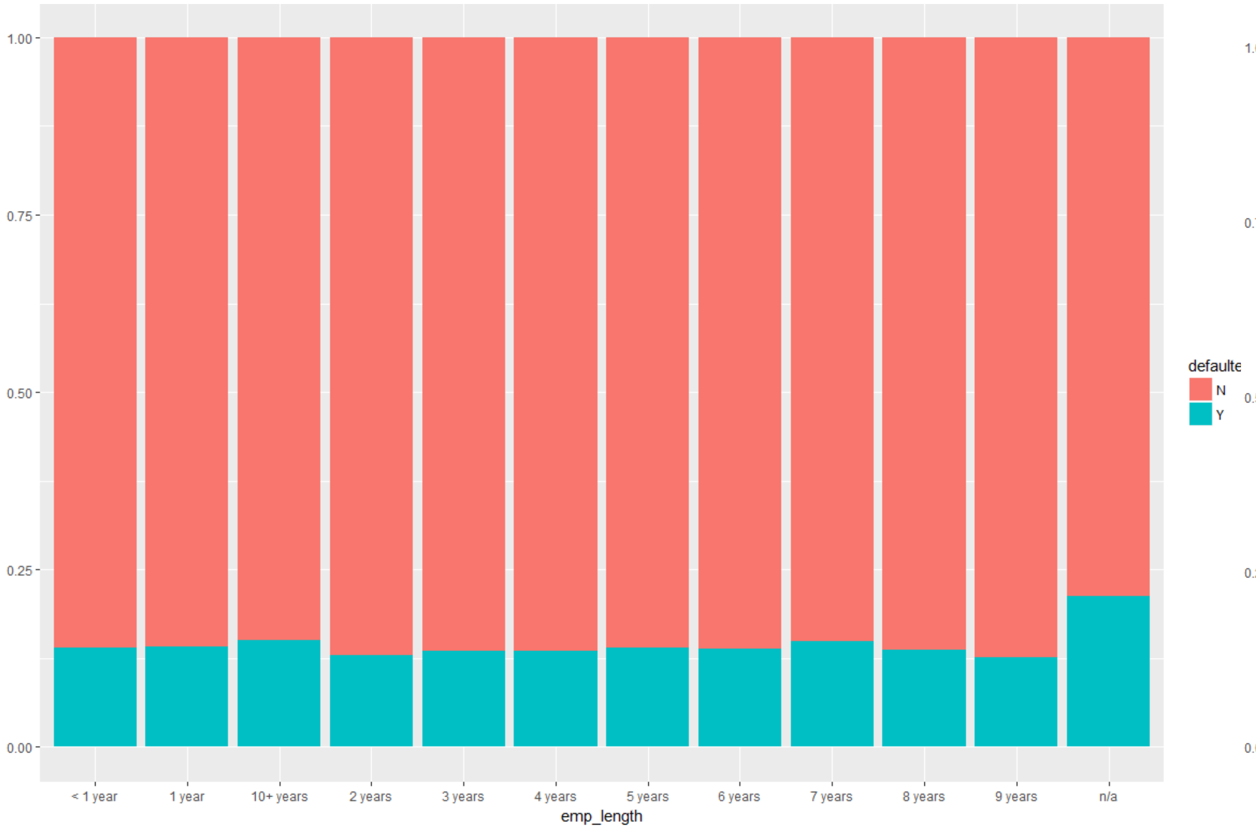


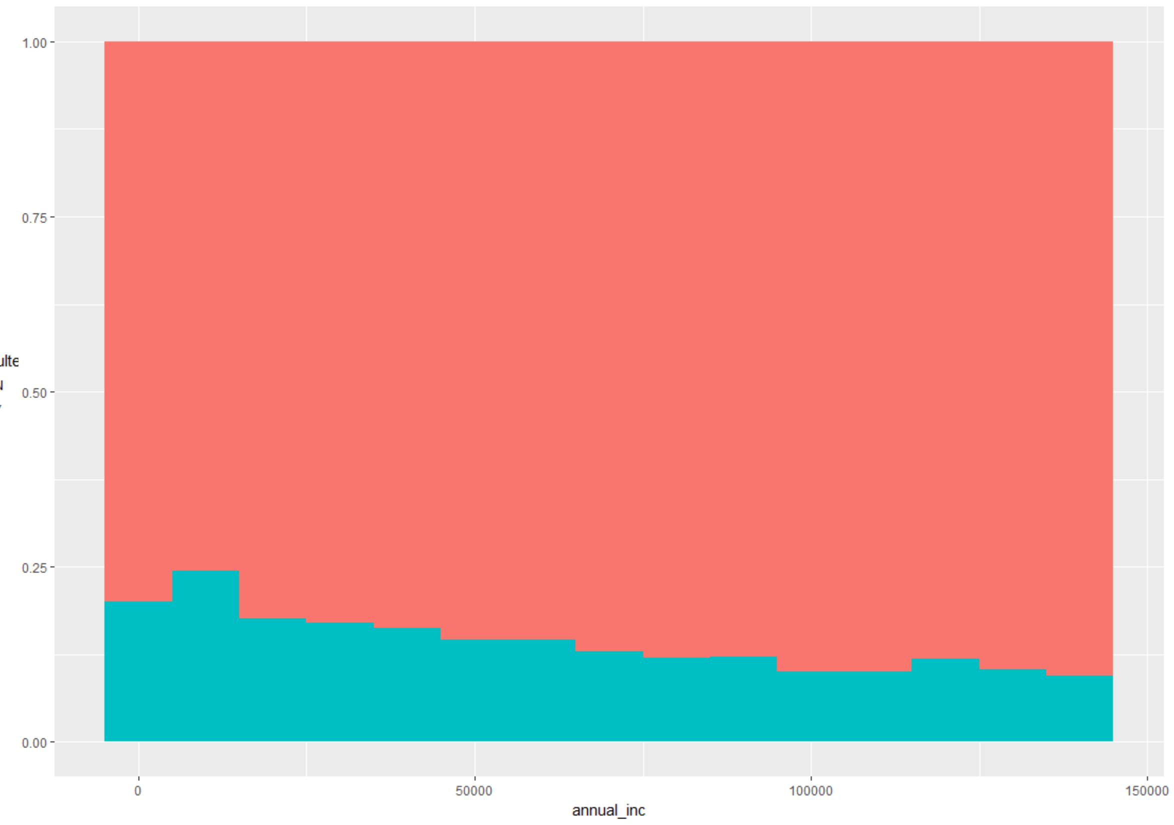


Loans with a term of 60 months are defaulted twice more than the loans with a term of 36 months.

Defaulters had a higher interest irrespective of the term.

Effect of Years of Employment and Annual Income on Loan Default

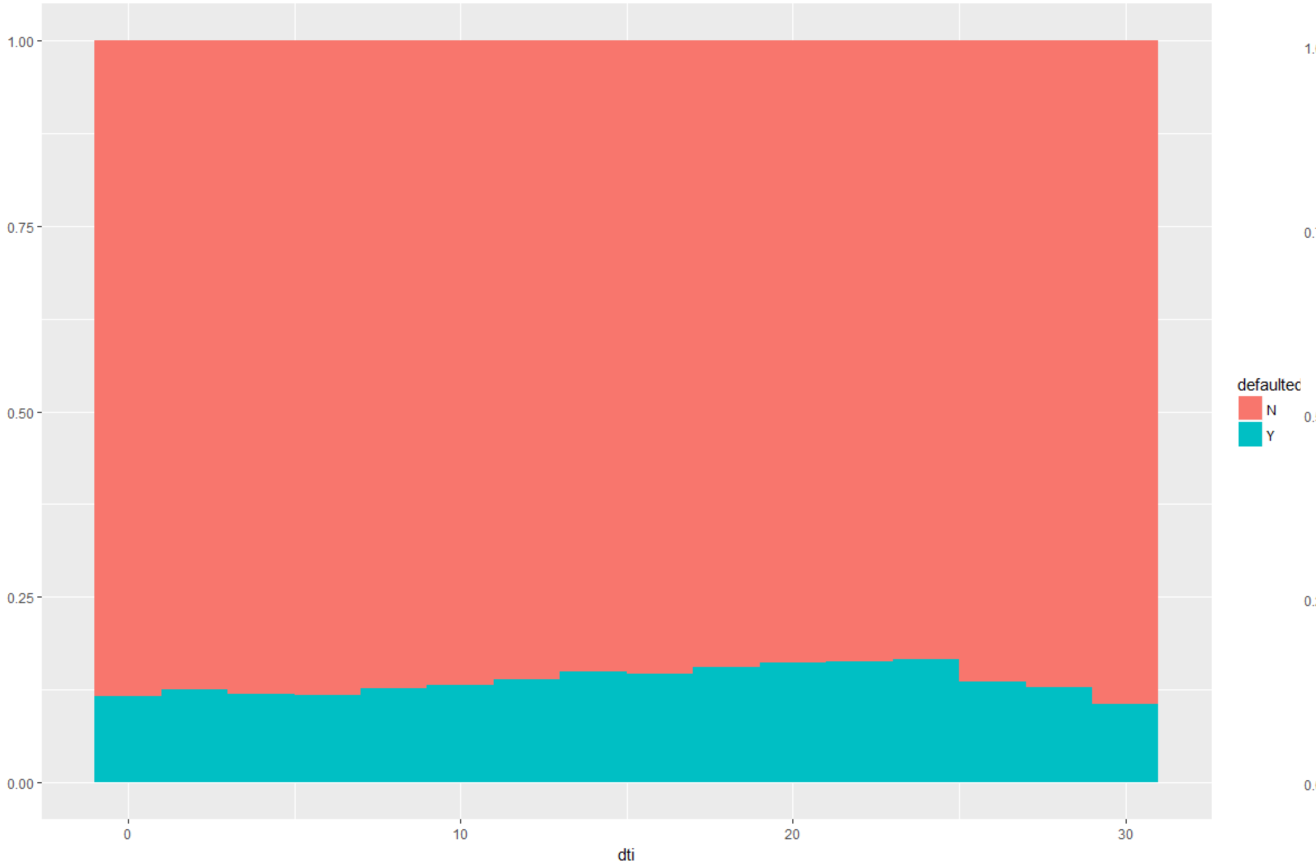




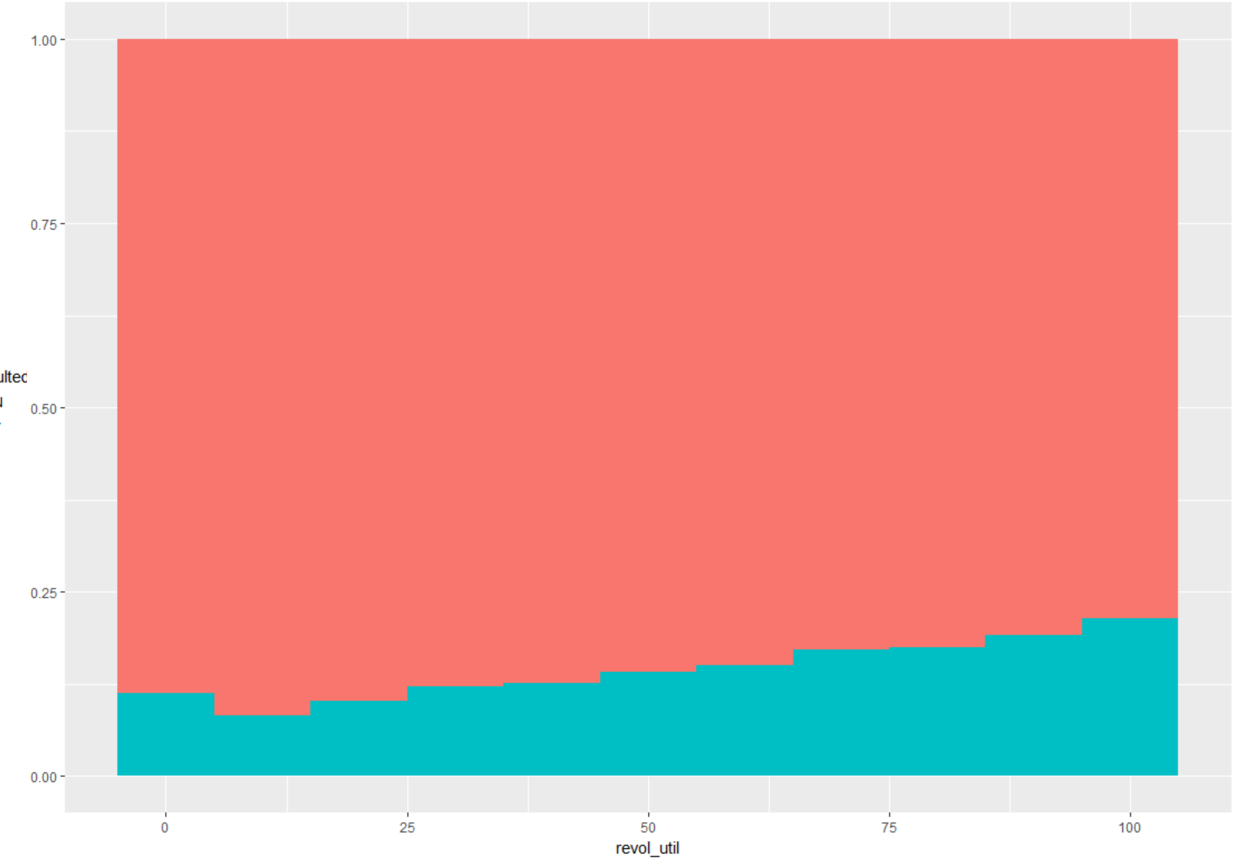
Years of Employment don't make a difference; unless the borrower has chosen to not disclose their employment years (these borrowers have higher default rates)

Lower income means higher chances of loan default.

Effect of Debt-to-income ratio and Revolving line utilization rate on Loan Default

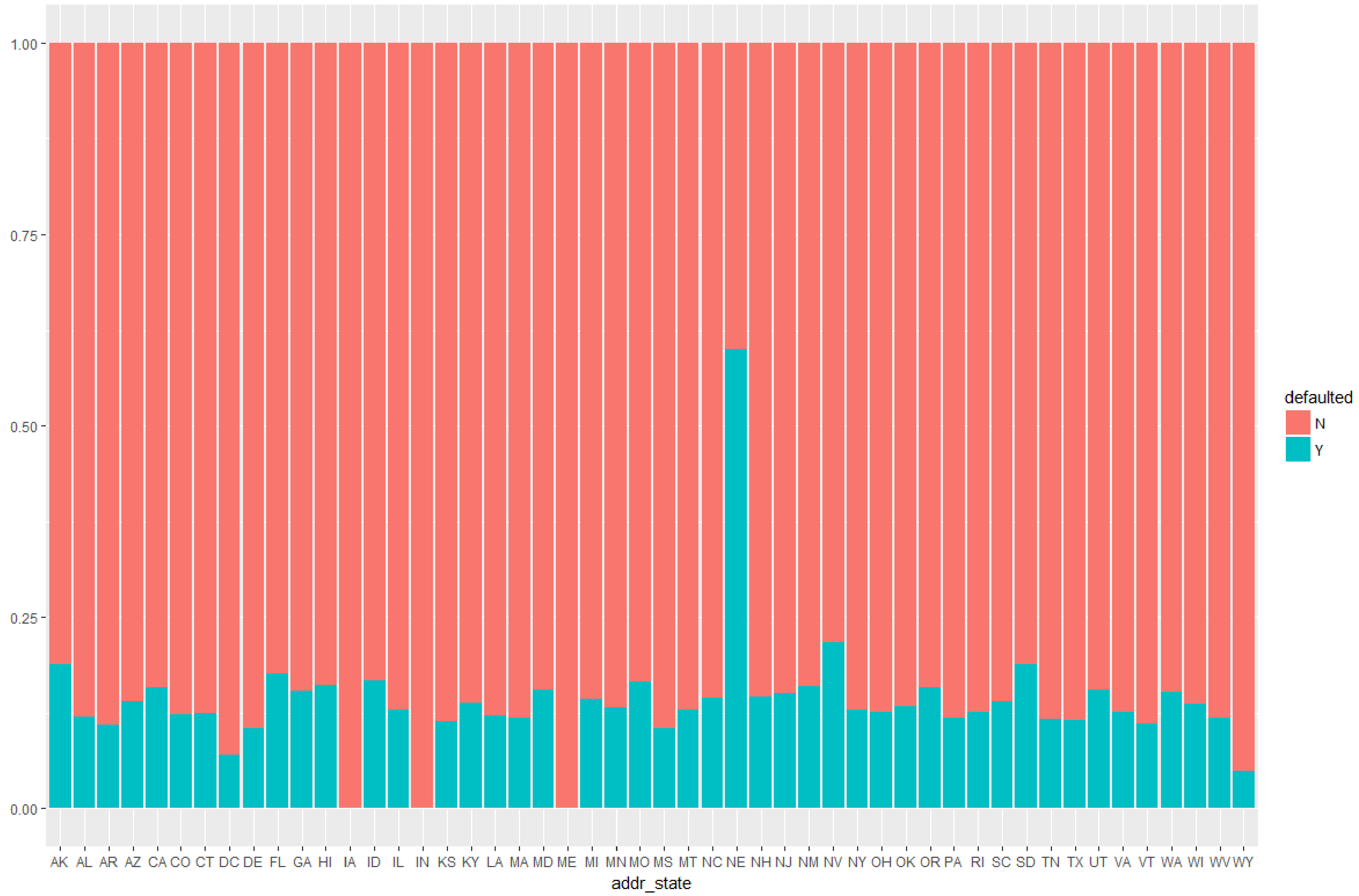


Default rate increases with higher DTI. There is a sudden drop in loans with DTI higher than 25; these can be considered as aberration.



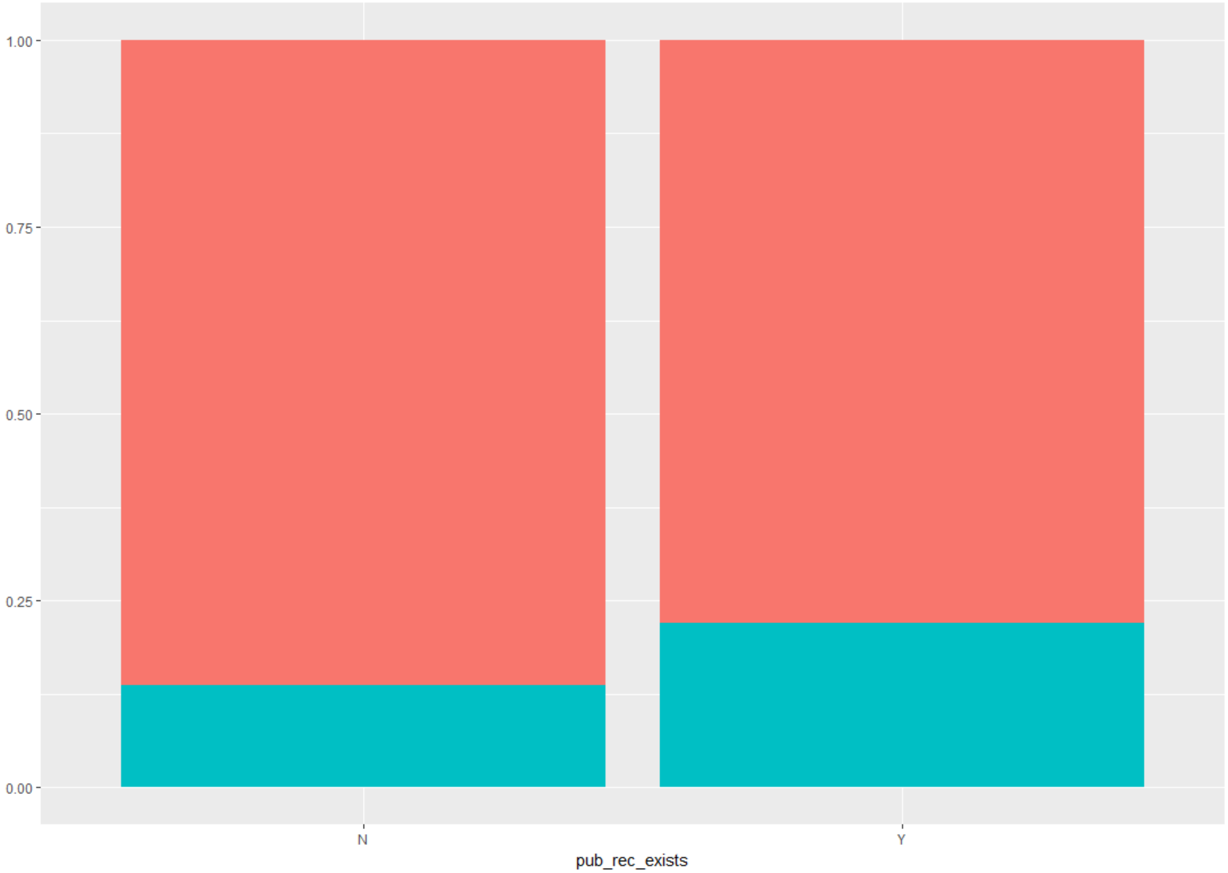
Default rate increases with higher revolving line utilization rate.

Effect of State on Loan Default

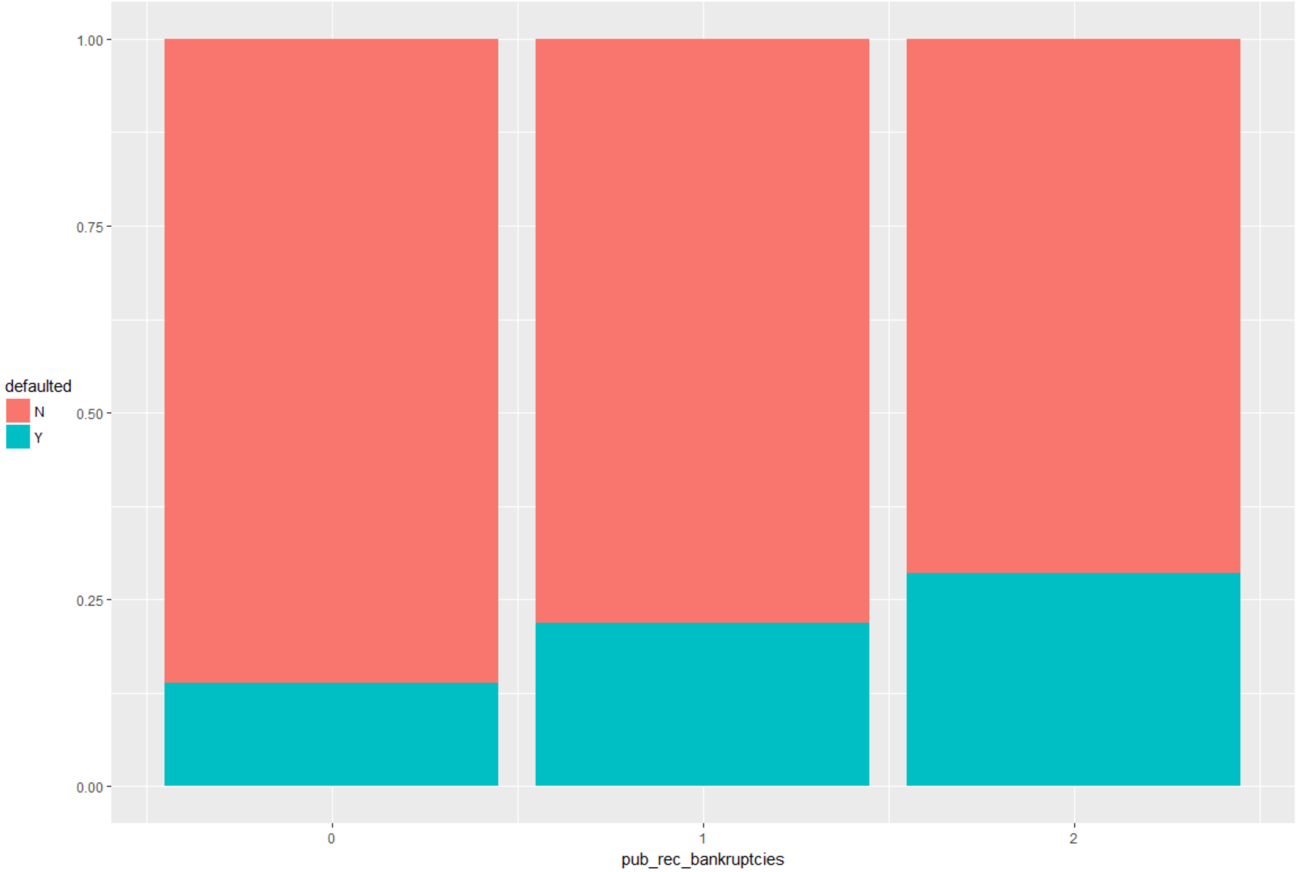


Depending on different states; the loan default rate varies.

Effect of Public Records and Bankruptcies on Loan Default

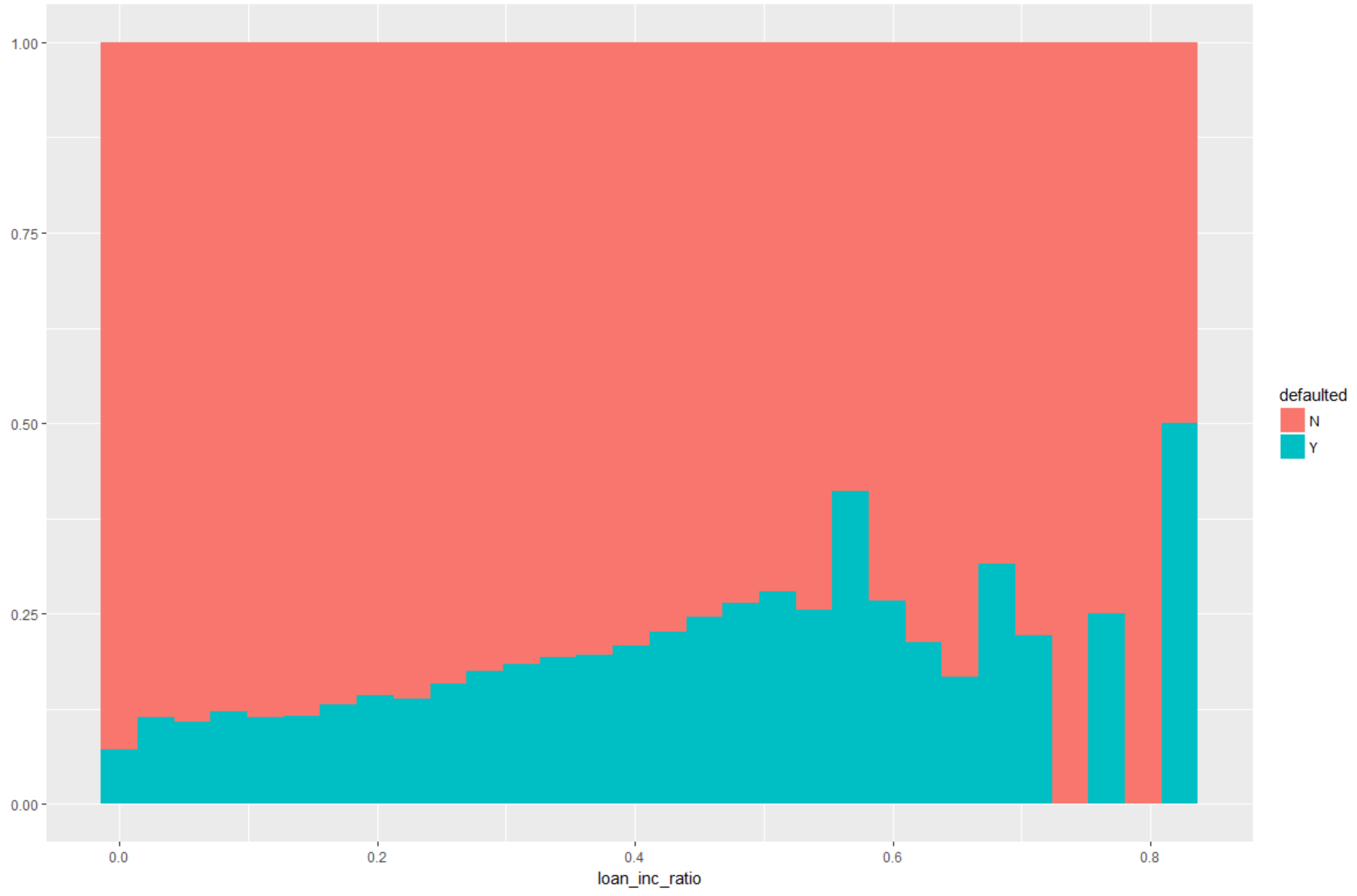


If someone has a public record; the chances of their defaulting increases.



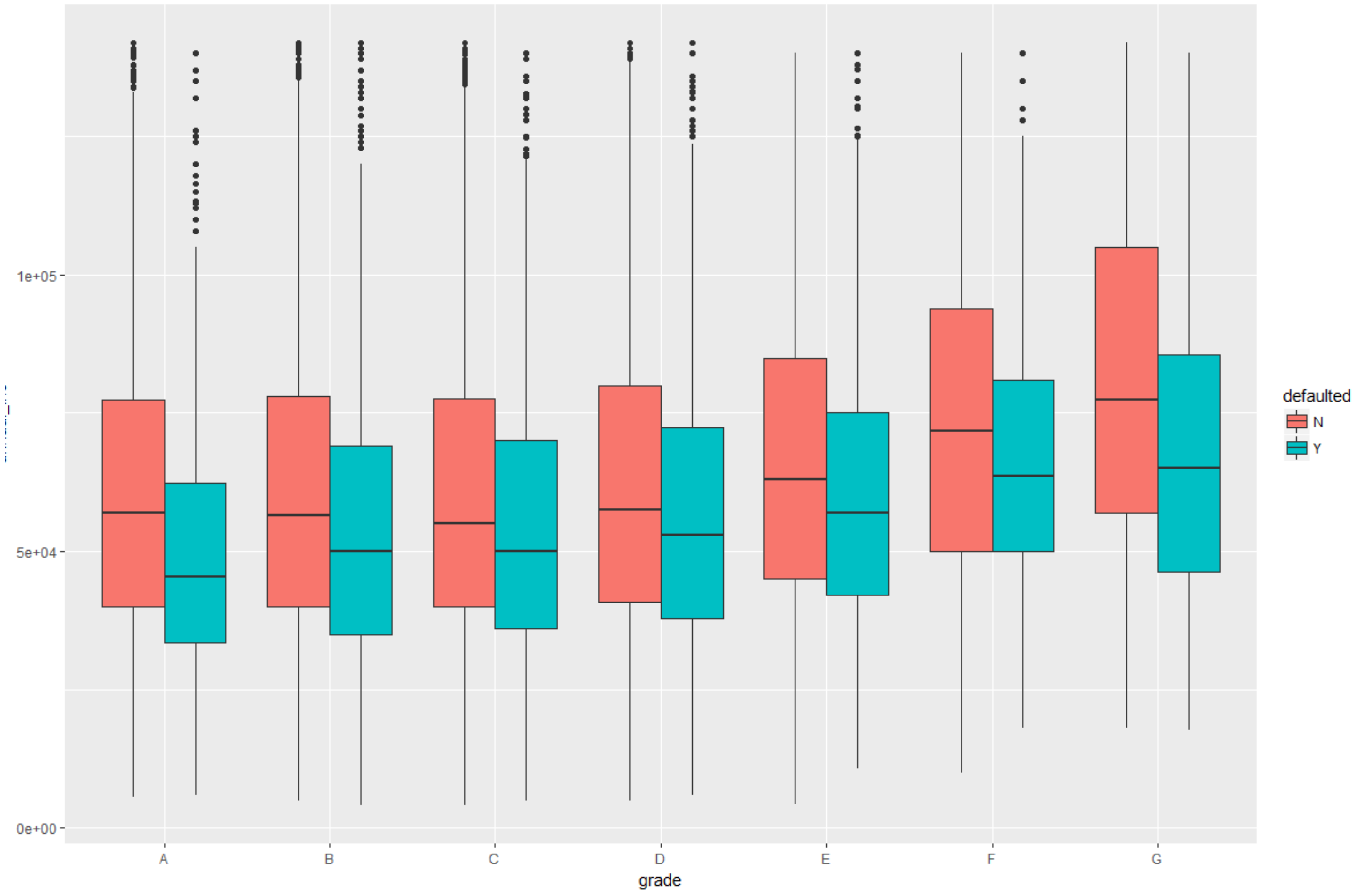
With higher number of publicly recorded bankruptcies; the chances of default also increase.

Effect of Loan to Income Ratio on Load Default



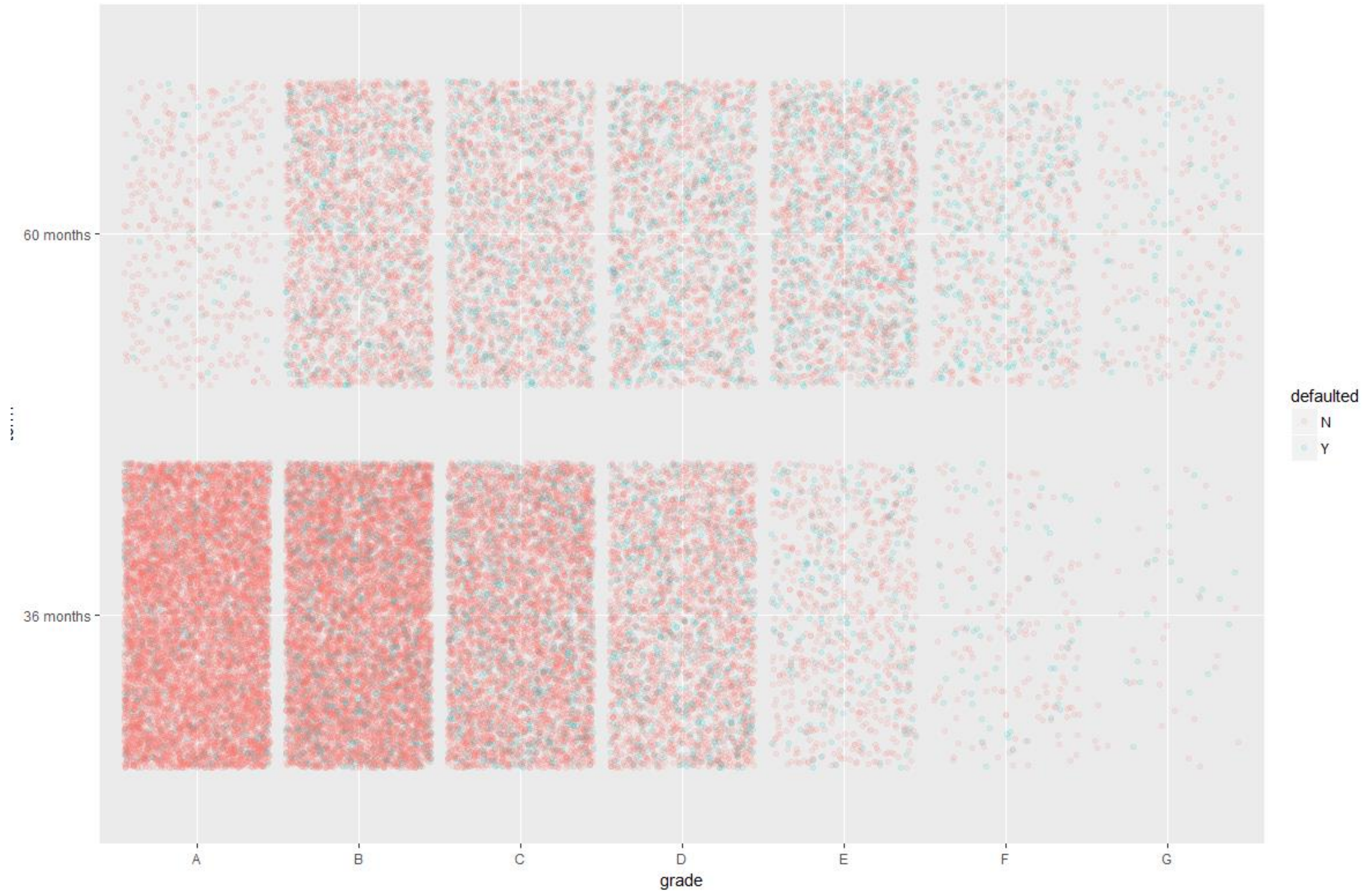
- When the loan is more and income is less; it results in loan defaults.

Income and Grade Bivariate Analysis



Defaulters have the lower income when compared to non-defaulters irrespective of grades

Term and Grade Bivariate Analysis



Borrowers with "A" grade take very few loans with a term of 5 year; while borrowers with "G" grade get fewer loans with a term of 3 year

Conclusions

Following attributes are strong indicators of loan default:

Loan Attributes

- Higher Interest Rate
- Term of 60-months
- Lower Grade / Sub-grade
- Purpose (Small Businesses, Renewable Energy, Education etc)
- Higher Loan Amount
- Higher Unfunded Amount

Consumer Attributes

- Lower Annual Income
- Non-disclosed Years of Employment
- Higher Debt-to-income Ratio
- Higher Revolving Line Utilization Rate
- Presence of Public Records
- Presence of Publicly Recorded Bankruptcies
- Number of inquiries in past 6 months
- State Listed by Borrower should be analyzed
- Higher Loan amount to Income ratio

Recommendations

- Reduce lending to (or the amount lent to) borrowers more likely to default
- Risk-based pricing: Charging higher interest rates to borrowers more likely to default
- Diversification: Diversify the borrower pool by lending to borrowers with different purposes, physical location, employment history etc.
- Covenants: Such as the borrower periodically reporting their financial condition
- Purchasing credit insurance
- Having in place adequate loan review mechanisms
- Allocating risk capital for credit risk
- Hedging based on collaterals