GRAMENER CASE STUDY

Group Name: Abraca-Data

1. VIKUL AGGARWAL
2. VENKITACHALAM RAMASWAMY
3. ANKIT SAHU

Background – Lending Club Case Study

Lending Club is one of the largest peer-to-peer marketplace which aims to connect borrowers with lenders. The borrowers apply through an online platform where they are assigned a score internally. Based on the score the lenders decide 1.) whether to lend and 2.) the terms of loan like interest rate, tenure, monthly instalment, etc.

Popular products: Debt Consolidation loans**,** Credit card loans, House loans, Car Loans, etc.

Business Objective:

To identify variables which are strong indicators of default and potentially use the insights in approval / rejection decision making.

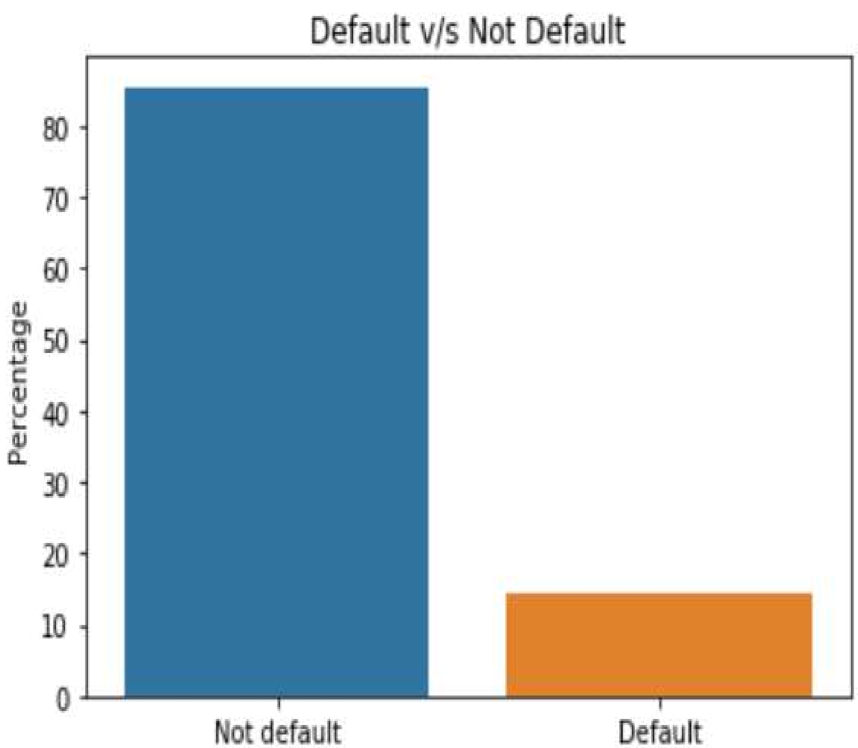
DATA UNDERSTANDING

Types of variables:

1. Variables related to customer's demographics and application
2. Variables related to the loan characteristics
3. Variables related to the customer's behavioral characteristic (after they were granted a loan).

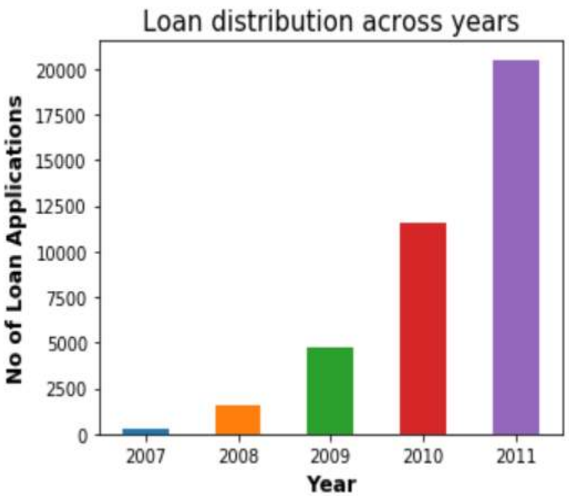
Customer Demographics & Application Variables	Loan Characteristics	Behavioral Characteristics
Employee Title	Loan amount	Earliest Credit line
Employee Length	Interest Rate	Recoveries
Annual Income	Tenure of loan	Last Payment Date
Home Ownership	Monthly Installment	Revolving Balance
Purpose	Funded amount	Last Payment amount

Overall Default Rate

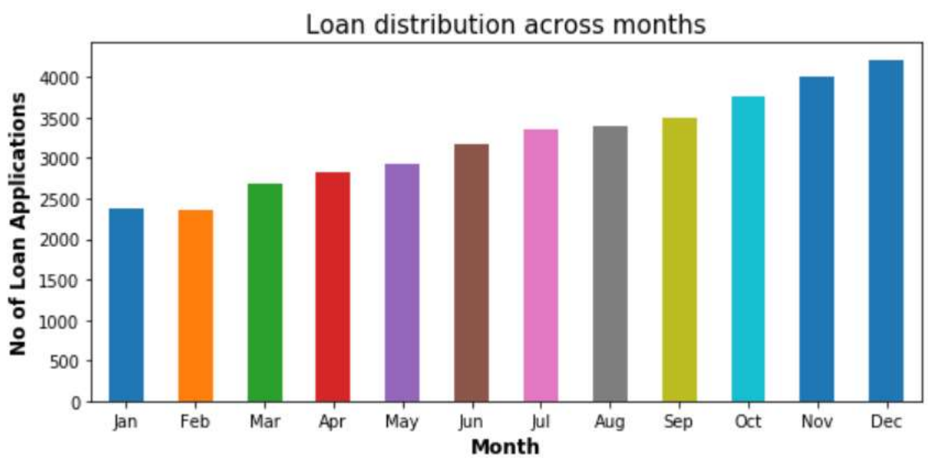


The Overall default rate comes to be ~14.6%

Loan Distribution across Years and Months

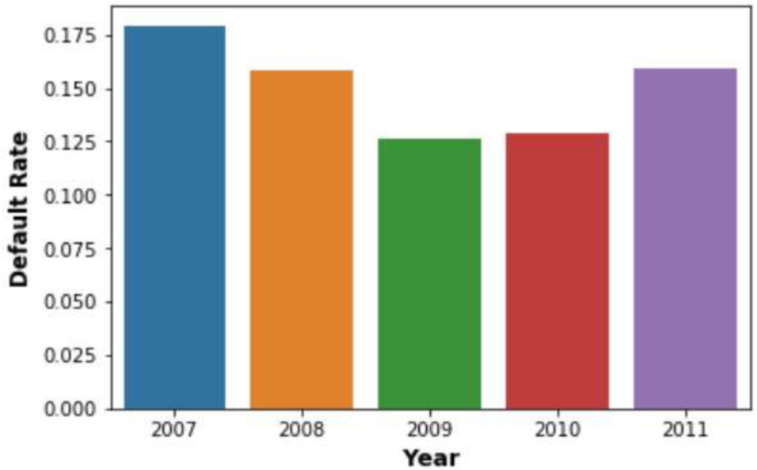


We can see a phenomenal rise in the loan distribution from 2007 to 2011.

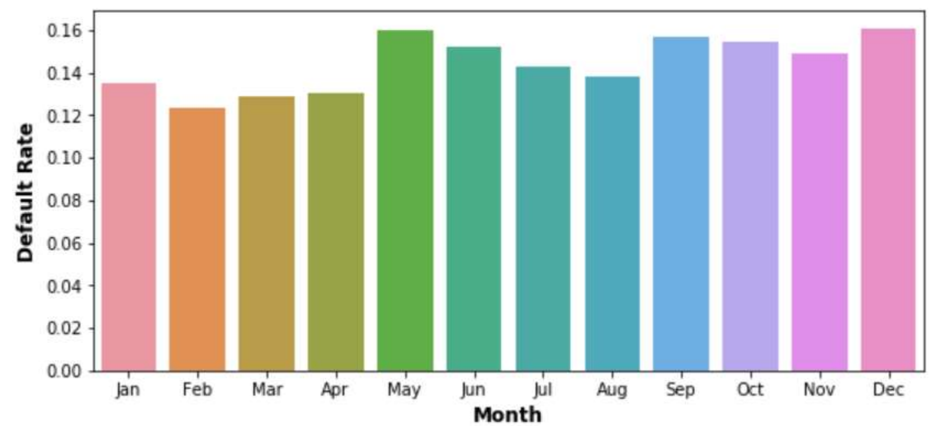


The number of loan applications are more in the second part of the year.

Default Rates across Years and Months

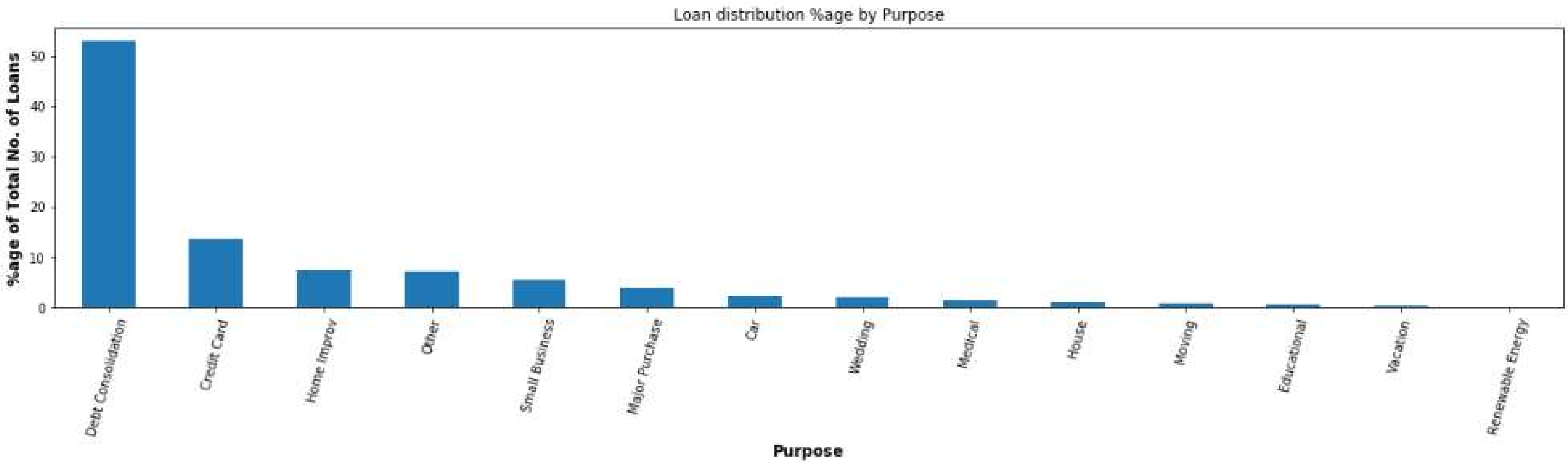


There is a substantial increase in default rates in the year 2011.



The default rates across the months are more or less the same.

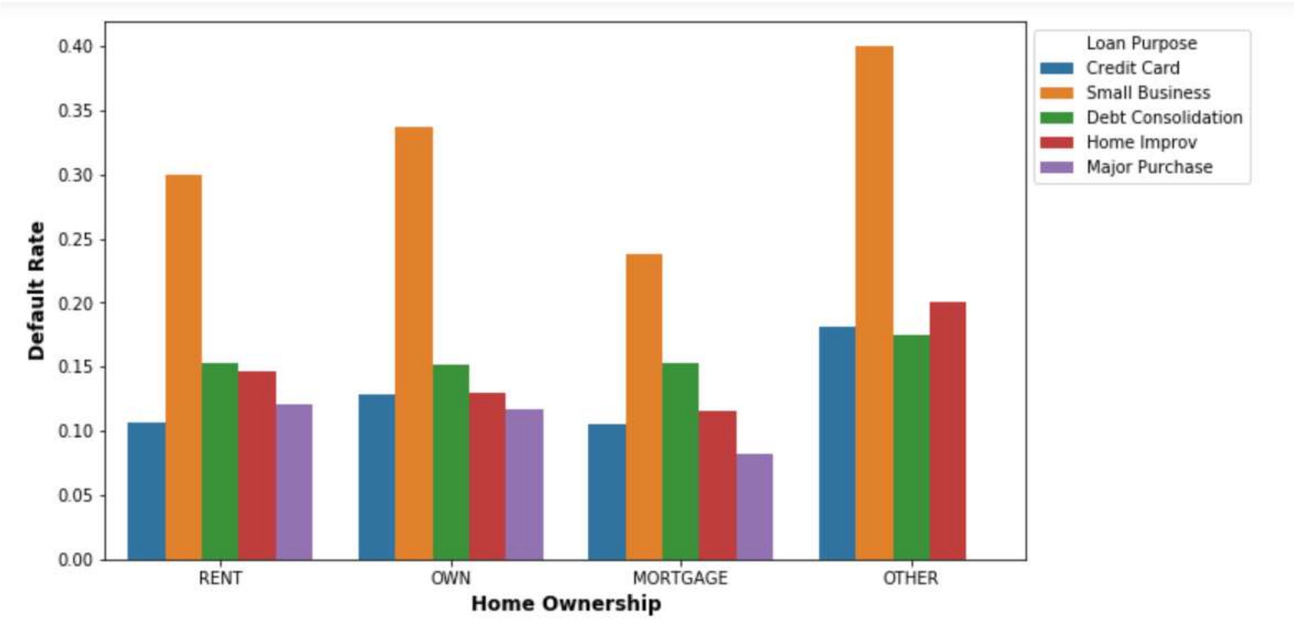
Loan Distribution by Purpose



The top 5 contributors to revenue are Debt Consolidation, Credit Card, Home Improvement, Small Business and Major Purchase (Combined ~84%).*

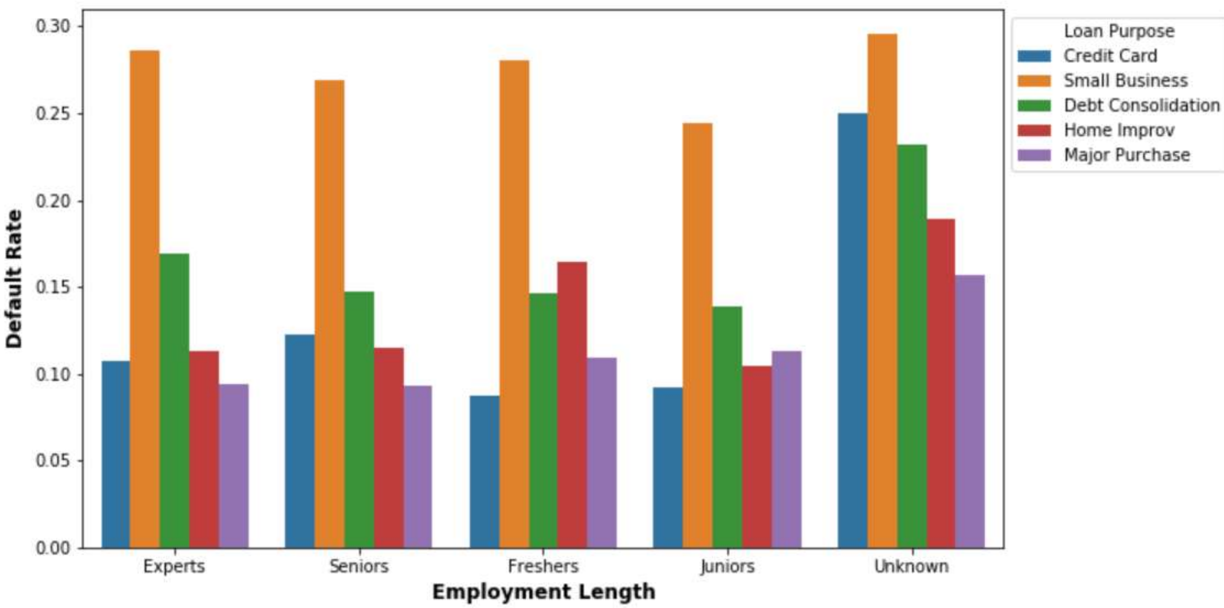
* Not considering Other

Analysis for Customer Demographic variables: Home Ownership



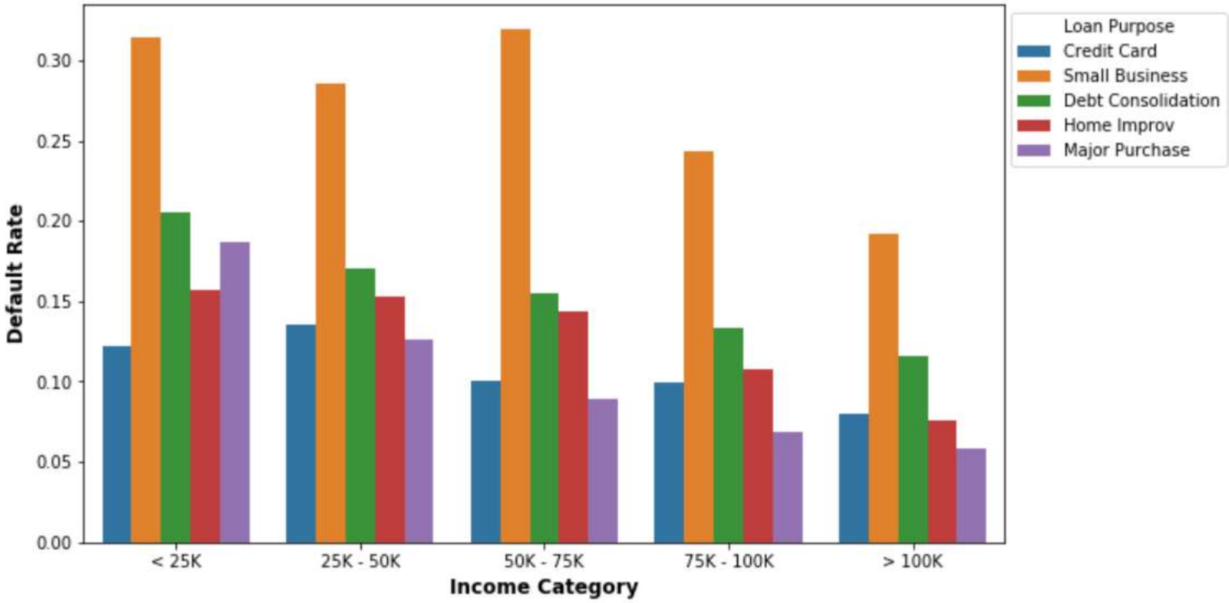
Small Business loans default the most, followed by Debt Consolidation in most of the home ownership segments

Analysis for Customer Demographic variables : Employment Length



We can see here that Small Business loans have the most default rate followed by Debt Consolidation except freshers, where Home Improvement loans have the second most default rate

Analysis for Customer Demographic variables: Income Category

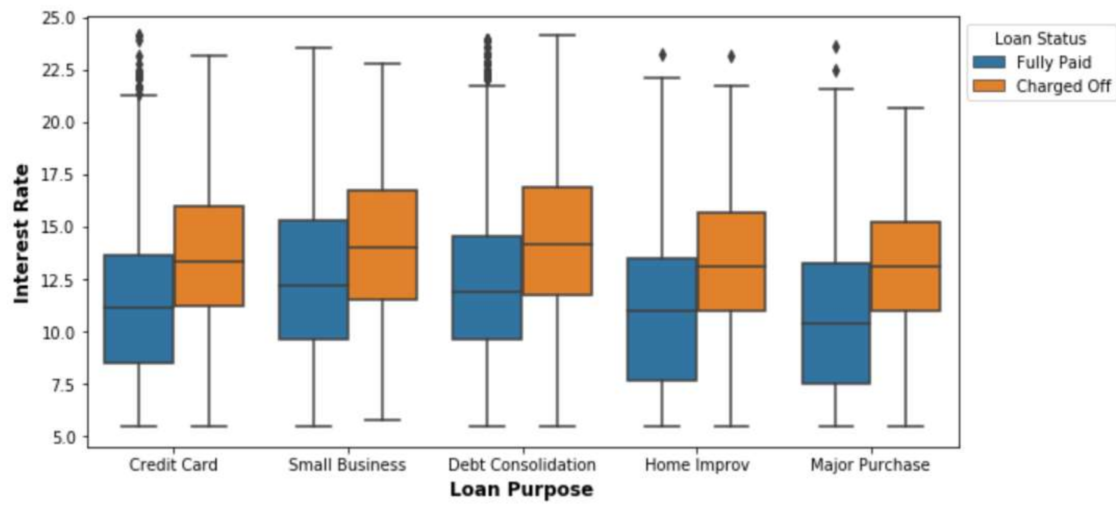


We can see that that Small Business has the highest default rate followed by Debt Consolidation in all income categories.

It can be seen that the trend is similar for all customer demographics. Debt Consolidation and Credit Card are the most popular loans. Small Business and Debt Consolidation have the highest default rate.

Analysis for Loan Characteristics: Interest Rate

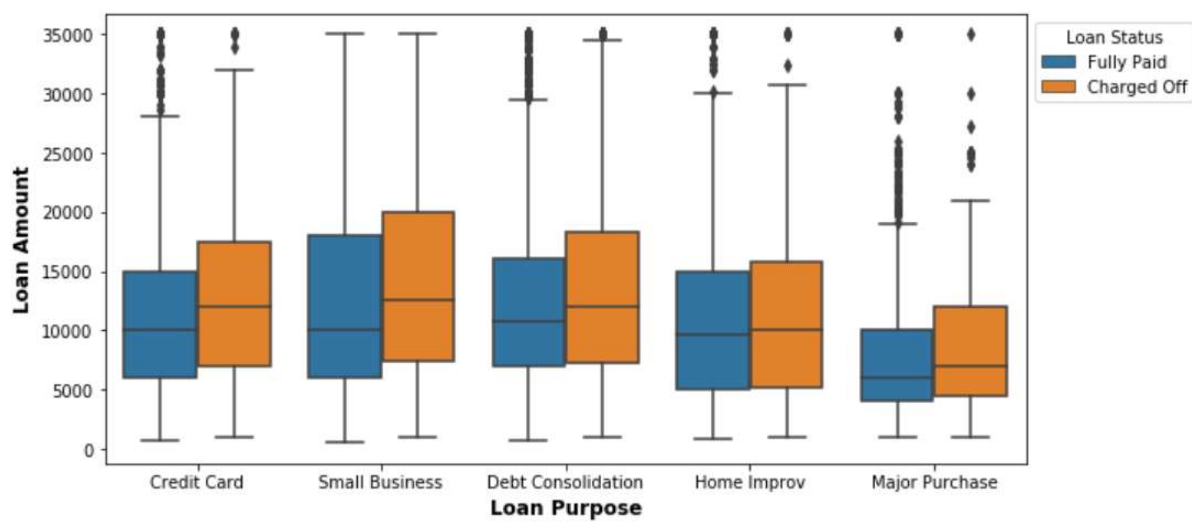
It can be seen that loans which default tend who have higher interest rate than loans that don't.



We can see that above observation holds true across purposes, that loans that default tend to have higher interest rates than those that don't.

Analysis for Loan Characteristics: Loan Amount

We can see that as the loan amount goes up, the changes of default increases.

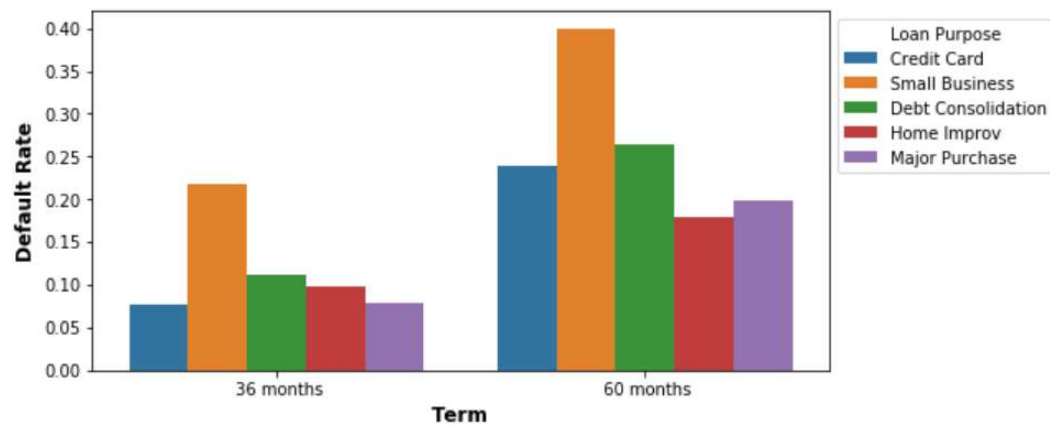


Here we can see that for the purposes which have a higher default rate (Small business, debt consolidation and credit card), the median borrowed amount is higher in case of the defaulted loans.

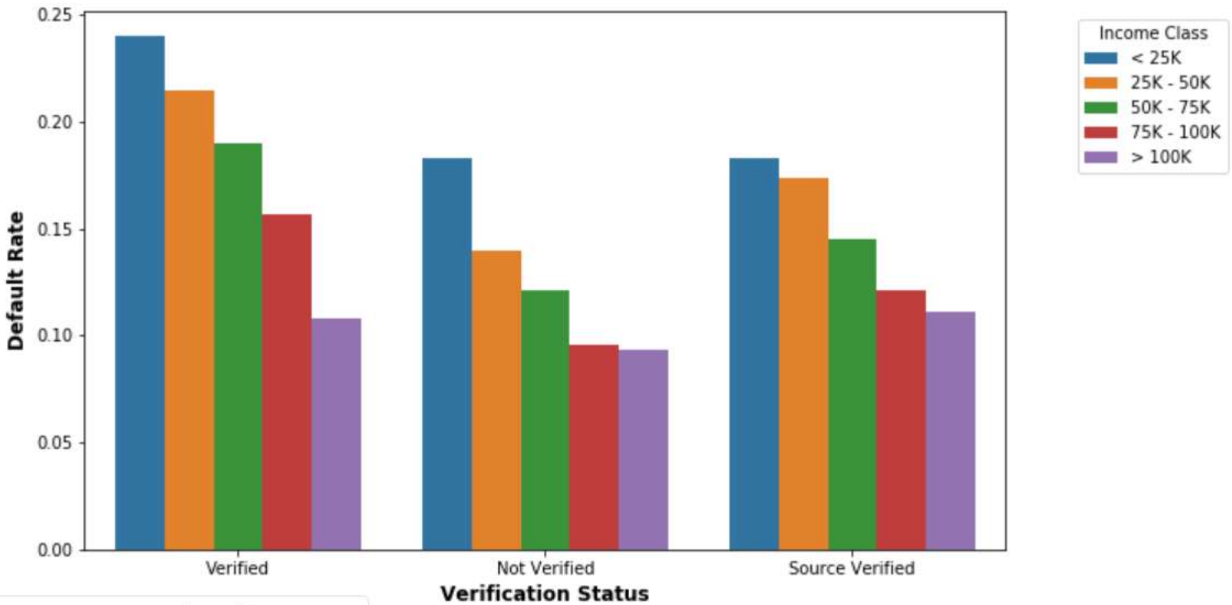
Analysis for Loan Characteristics: Loan Term

We can see that the number of defaulters is almost the same for both the term periods. However there's a noticeable difference when it comes to fully paid loans.
This indicates a higher likelihood of default on loans with the longer term period of 60 months.

The observation made is consistent across purposes, the default rate is high for loans with a higher term.



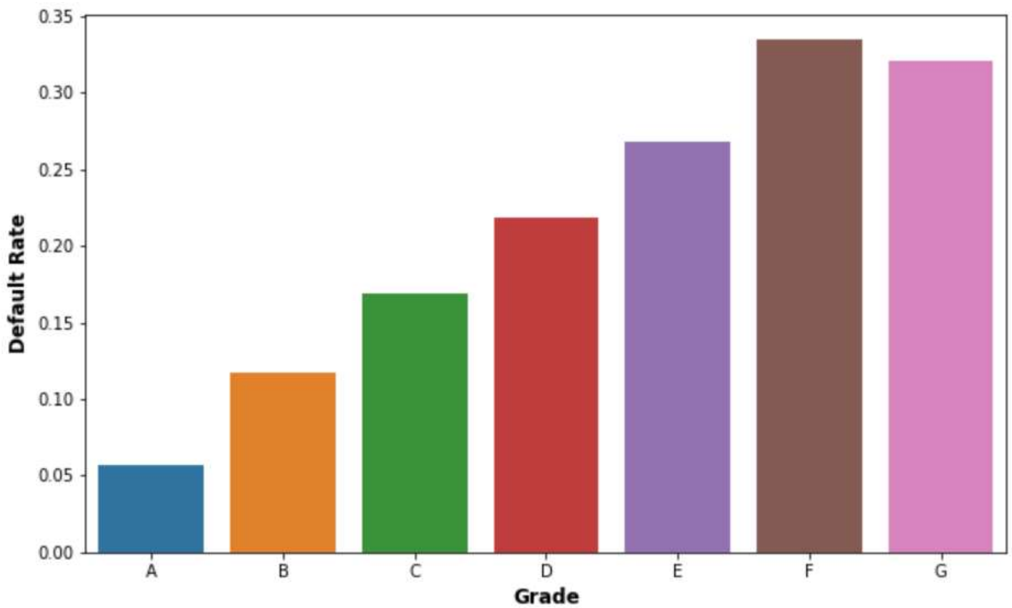
Analysis for Loan Characteristics: Verification Status across Income Categories



This is a striking observation that for all verification statuses, lower income groups have a higher default rate than higher income groups.

Also, borrowers who're not verified have a larger contribution to the business and maintain a lower default rate as compared to verified and source-verified customers.

Analysis for Loan Characteristics: Grades

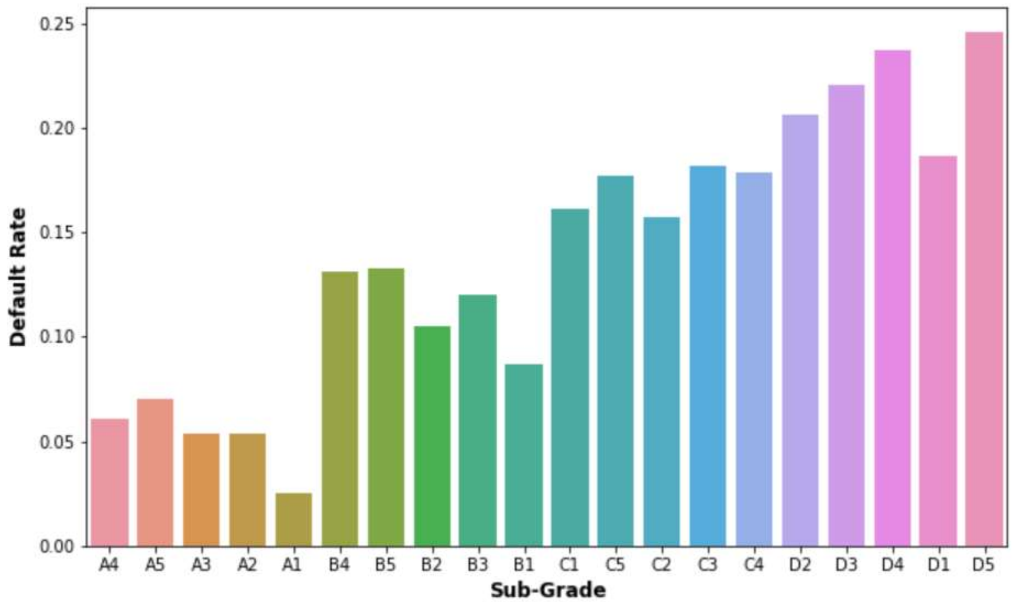


We can see here that the although the default rate for grade A is low (less than 5%), default rate for B(~12%), C(~17%) and D(~21%) is high.

A, B, C & D account for ~80% of the business.

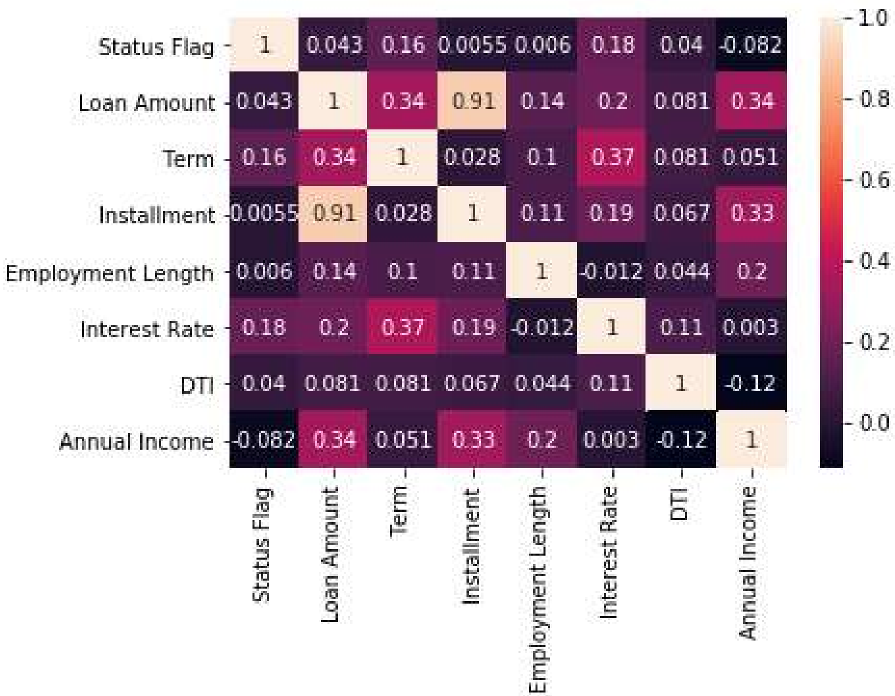
A small improvement in grades A, B, C and D w.r.t. default rate can significantly decrease the credit loss.

Analysis for Loan Characteristics: Sub-Grades



We can see here that the default rates for the sub-grades are in line with order of the default rate for the grades.

Correlation Study



The correlation heat-map tells the following about defaulters:
1. Longer term period most likely comes with a higher interest rate. It also suggests higher loan amount and installment. These factors show a higher likelihood of defaulting.

2. Income is negatively correlated, although the value is small. This is in-line with lower default rates observed in high income groups and vice versa.

3. Larger installment is not very strongly correlated with defaulting. However when coupled with high interest rate and loan amount it becomes significant.

4. It also shows that high income borrowers taking small low interest loans are less likely to default.

Inferences

Strong areas of focus to predict the risk of default

- **Purpose** - Debt Consolidation and credit card loans account for more than 65% of the business and have high default rates.
- **Income Group** - Low income groups have high default rate for Debt Consolidation and Home Improv loans. Small Business loans have high default rate across all income categories.
- **Interest Rate** - Loans with high interest default more as compared to low interest rate across all identified loan purposes.
- **Loan Term** - Loans with higher term tend to default more across all identified loan purposes.
- **Verification Status** - One strange observation made is that Verified loans tend to default more than Not-Verified loans.
- **Grade** - Grades A, B, C and D contribute the most to the revenue and have high default-rates.