Lending Club Group Case Study

FINAL SUBMISSION

Group Members:

Pavan Kumar Doddi (Group Facilitator)

Angad Singh

Abstract

Objective:

The overall objective is to identify the predictors that result in loan default. Lending Club specializes in lending various types of loans to urban customers. It is pertinent to find out the pattern of default for the Lending Club as, if it doesn't lend money to a good customer it is a business loss and if it lends to a possible defaulter then it would be a financial loss to the company.

Analysis Approach:

Exploratory Data Analysis (EDA - univariate & bivariate) has been conducted on historical loan records by cleaning the data, analyzing the data that will predict the default possibility to identify the characteristics impacting the loan repayment.

Data Sources:

loans.csv (historical loan data from Lending Club) Findings Historical loan data that provided information of current loan, defaulted and closed loan which was cleaned up, transformed & analysis has been conducted to provide some recommendations to Lending Club

Recommendation:

Final recommendation has been given to Lending Club based on the data analysis on few important variables

Data Exploration

- Explore the historical date and find out significance of each variable

Data Cleansing

- Clean the data by removing null values from data set (column and row cleaning)

Data Modification

- Check if we need to impute the missing values also create additional columns from existing data

Data Analysis

- Perform univariate and bivariate analysis on the data set

Recommendations

- Provide actionable recommendations

Data Understanding

Types of variables

- Applicant demographics
- Loan related information
- Customer Behavioral Characteristics

Applicant Demographics

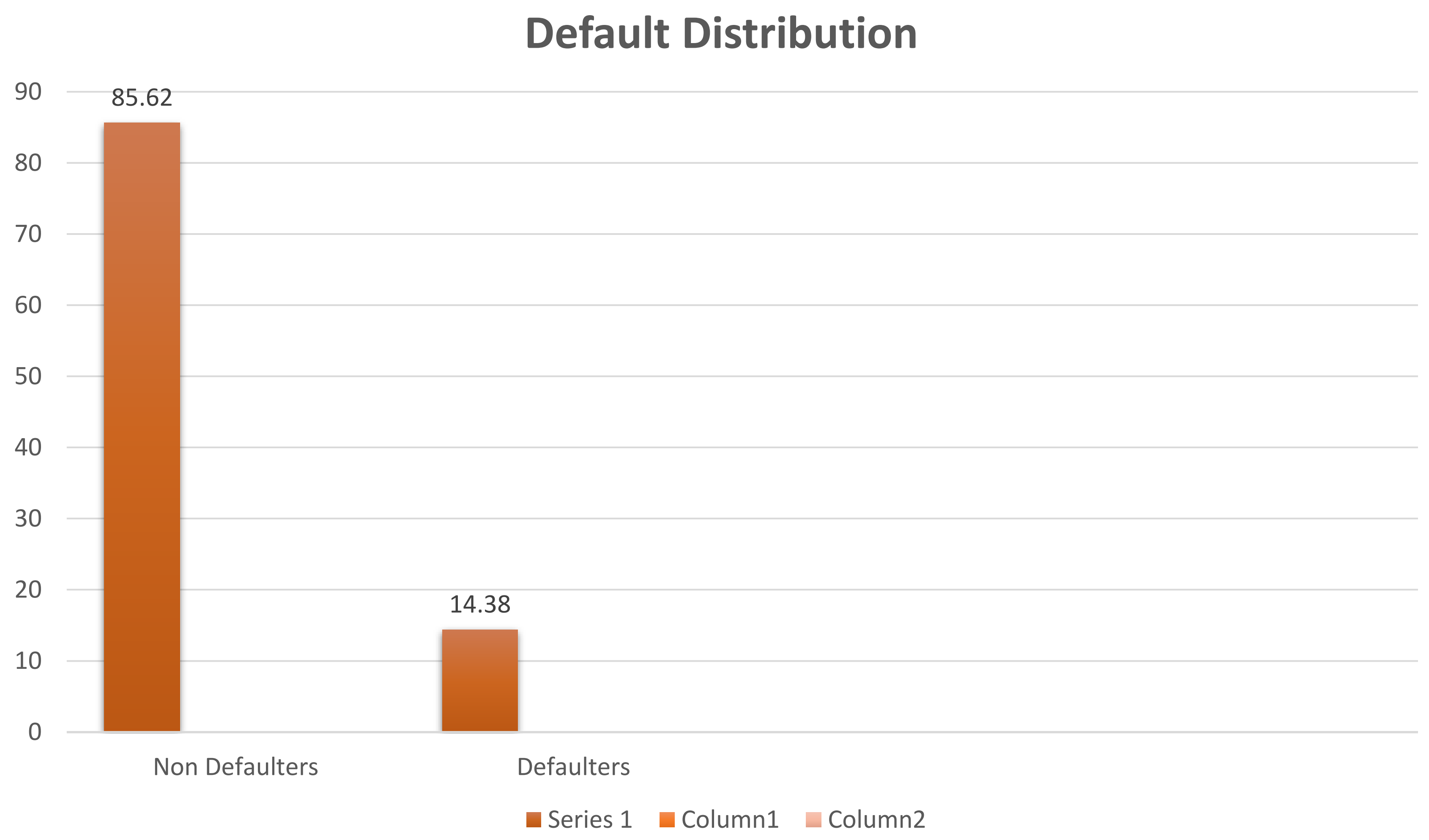
- Employment Title
- Employee Service Length
- Annual Income
- Description

Loan related information

- Loan Amount
- Funded Amount
- Funded Amount Investment
- Grade/Sub Grade
- Interest Rate
- Loan Status

Customer Behavioral Characteristics

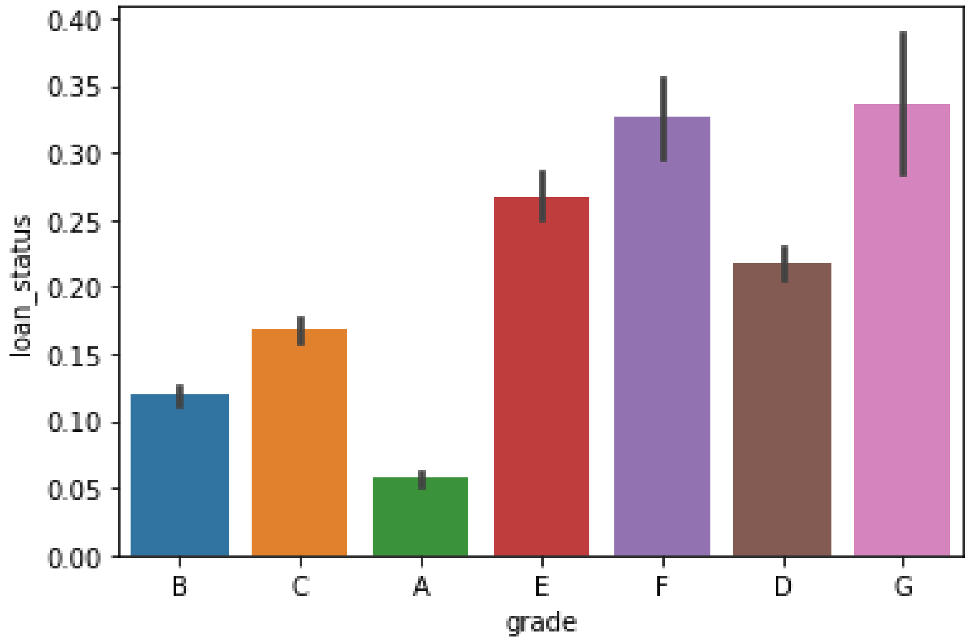
- Delinquency Year
- Loan Purpose
- Revolving Balance
- Recoveries
- Credit Line



The Default Rate is 14%

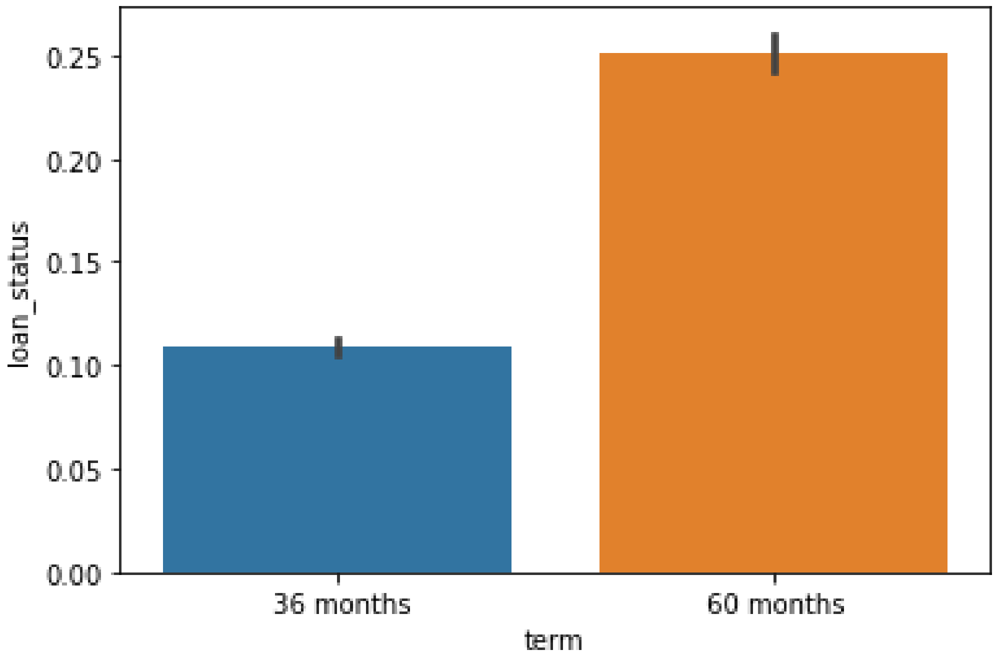
Univariate Analysis

Default by Grade



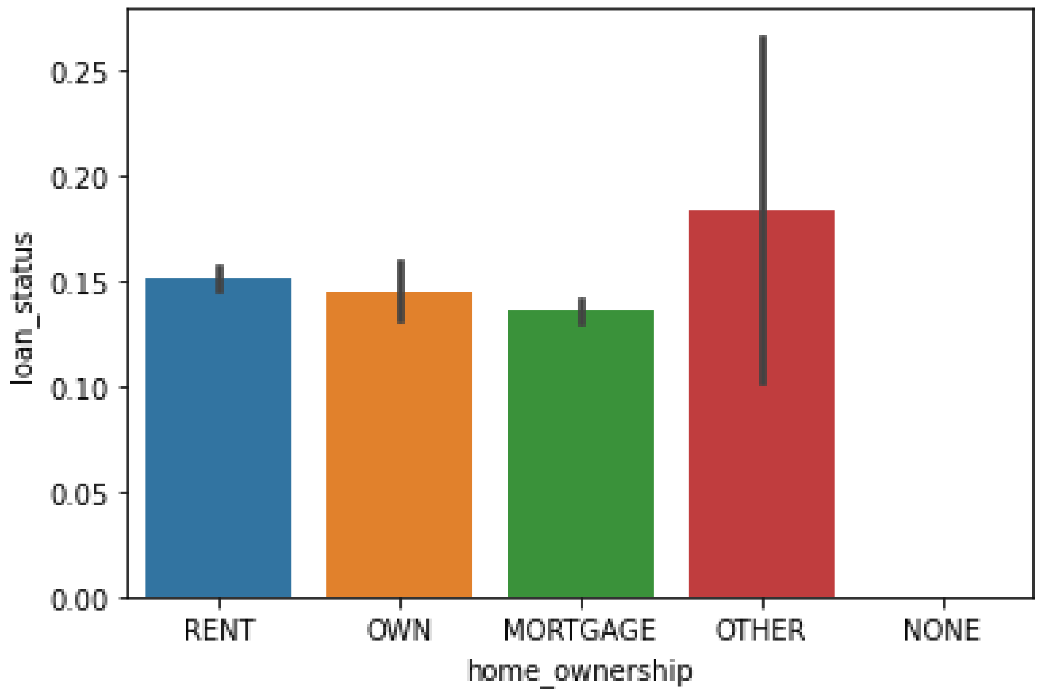
As the grade of loan goes from A to G, the default rate increases.

Default by Term



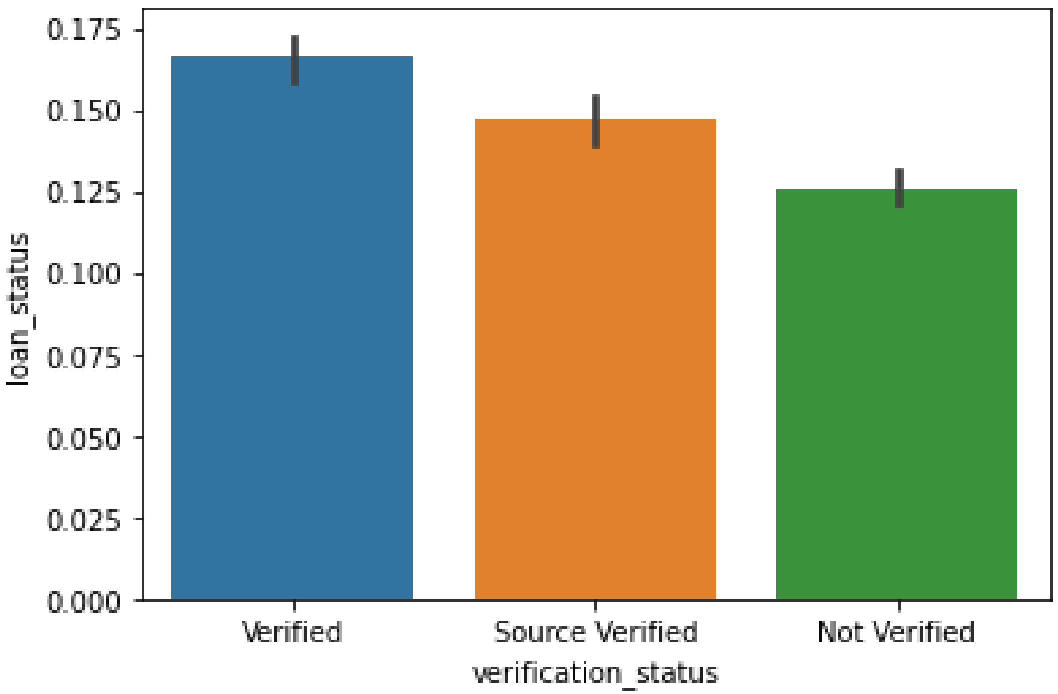
We can see that chances of default is more with longer period

Default by Home Ownership



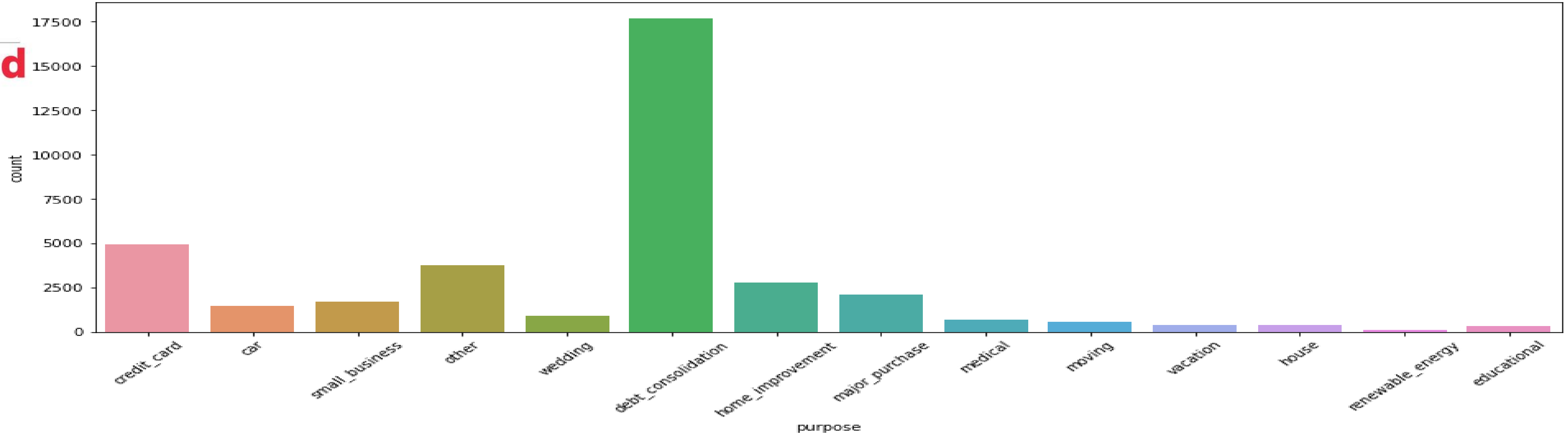
Default rate is lower when mortgage is there and high with Other.

Default by Verification Status

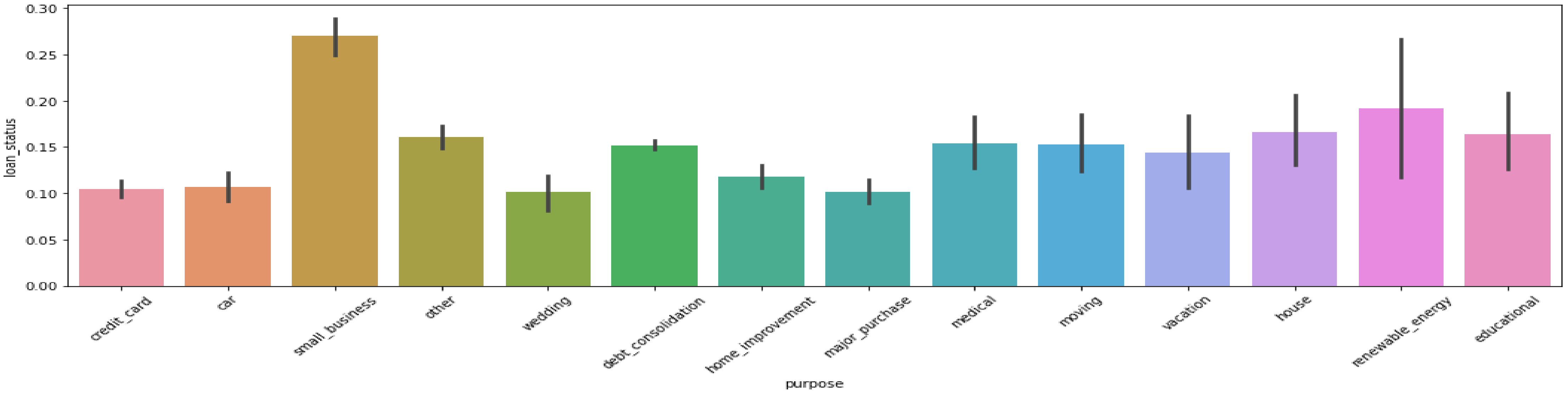


Surprisingly where the verification status is verified the default rate is very high.

No of Loan applications by Purpose

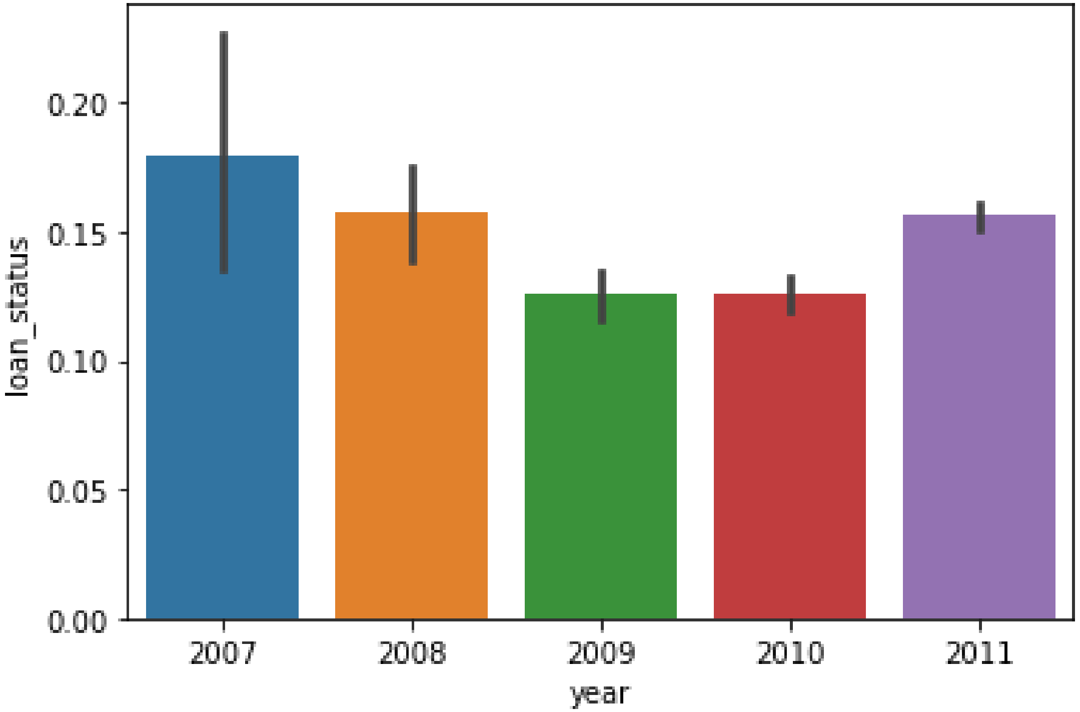


Default by Purpose



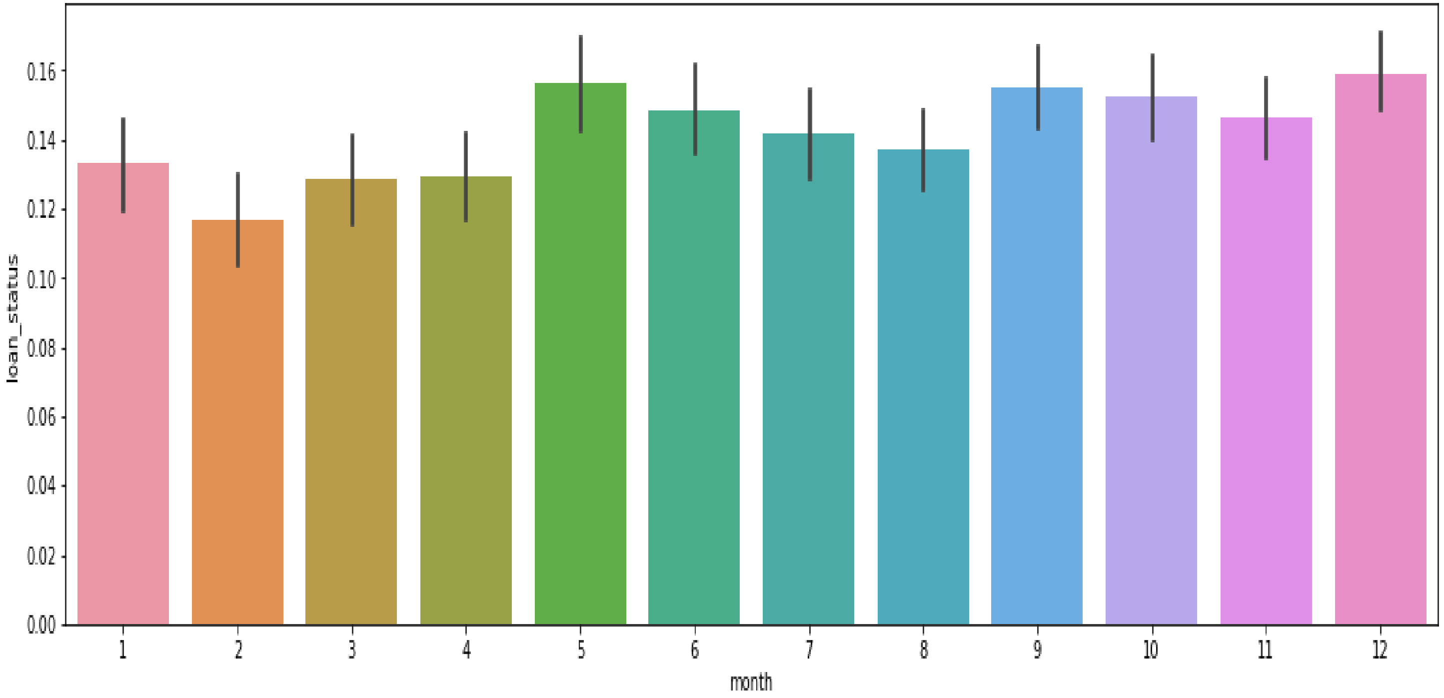
Small business loans default the most then Renewable Energy, Education and Home

Default by Year



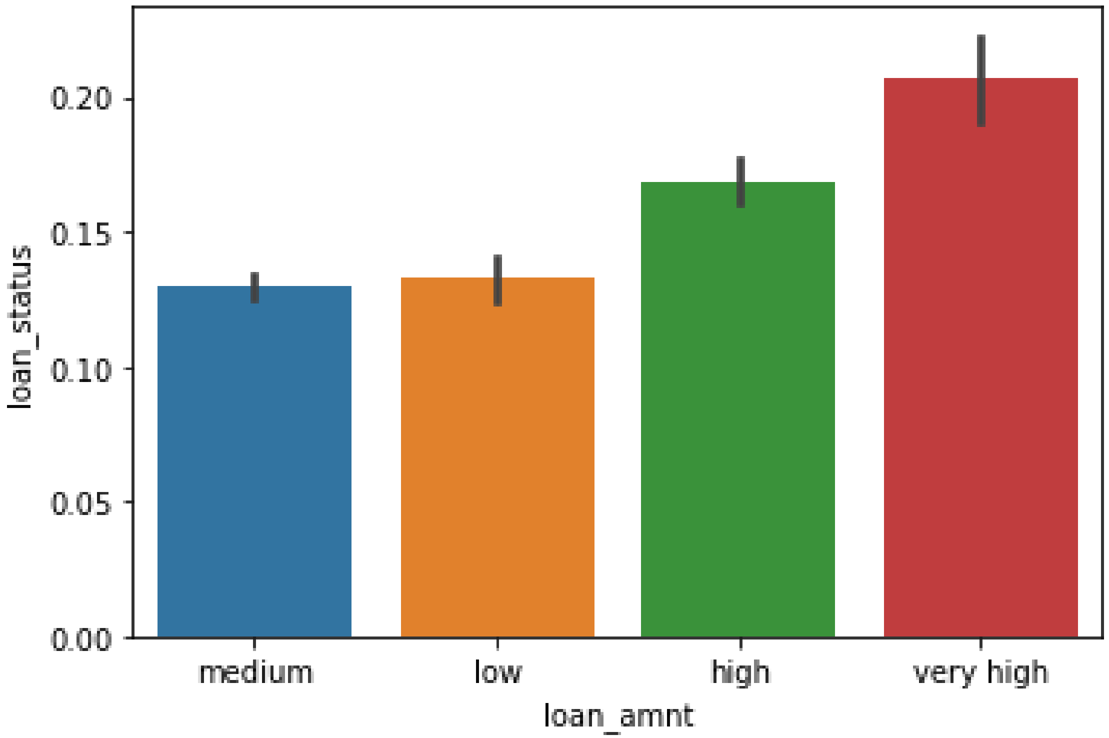
We can see that default rate has increased in 2011 unlike the trend from 2007

Default by Month



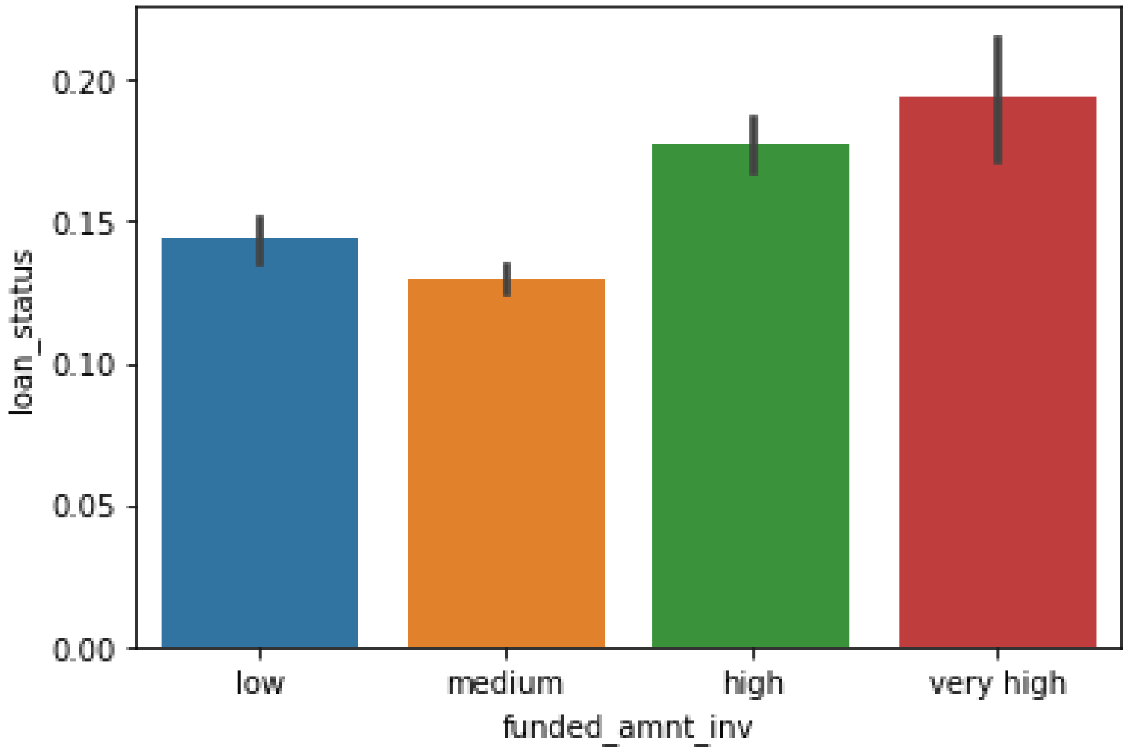
There is not much change in the default rate across months except for little change in second half of the year

Default by Loan Amount



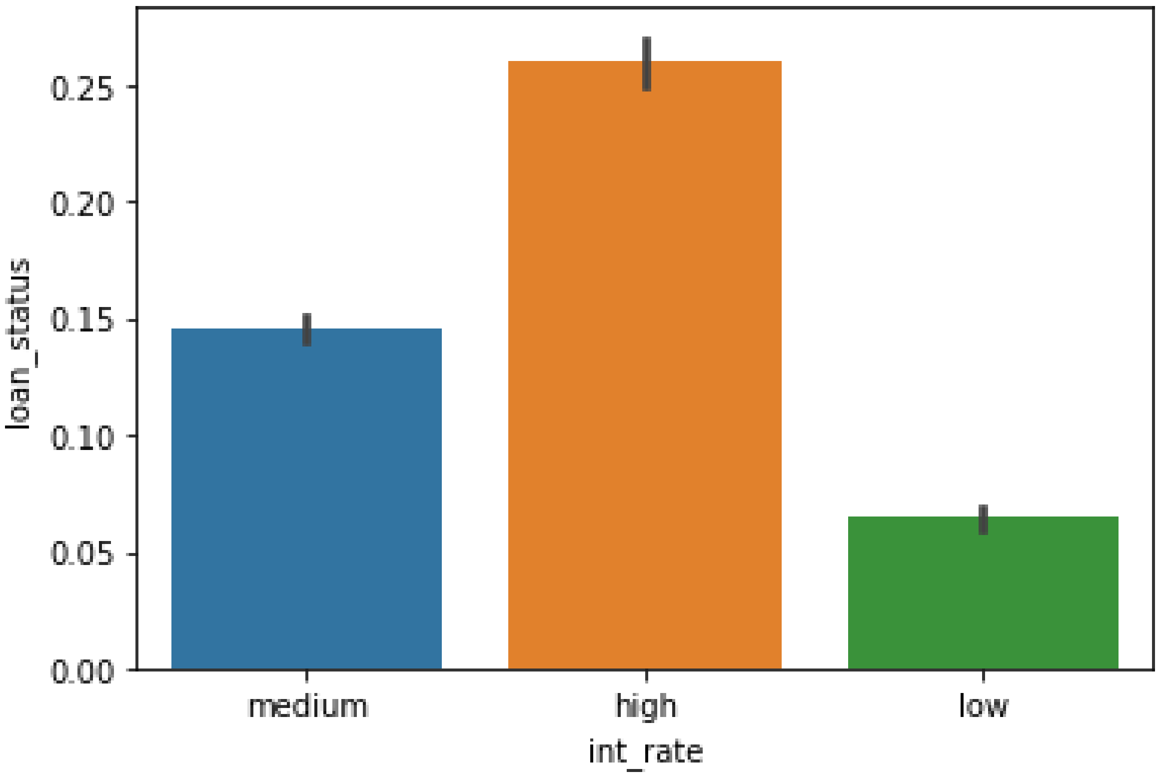
We can see that as the loan amount increased the default rate also increased considerably

Default by Funded Amount Investment



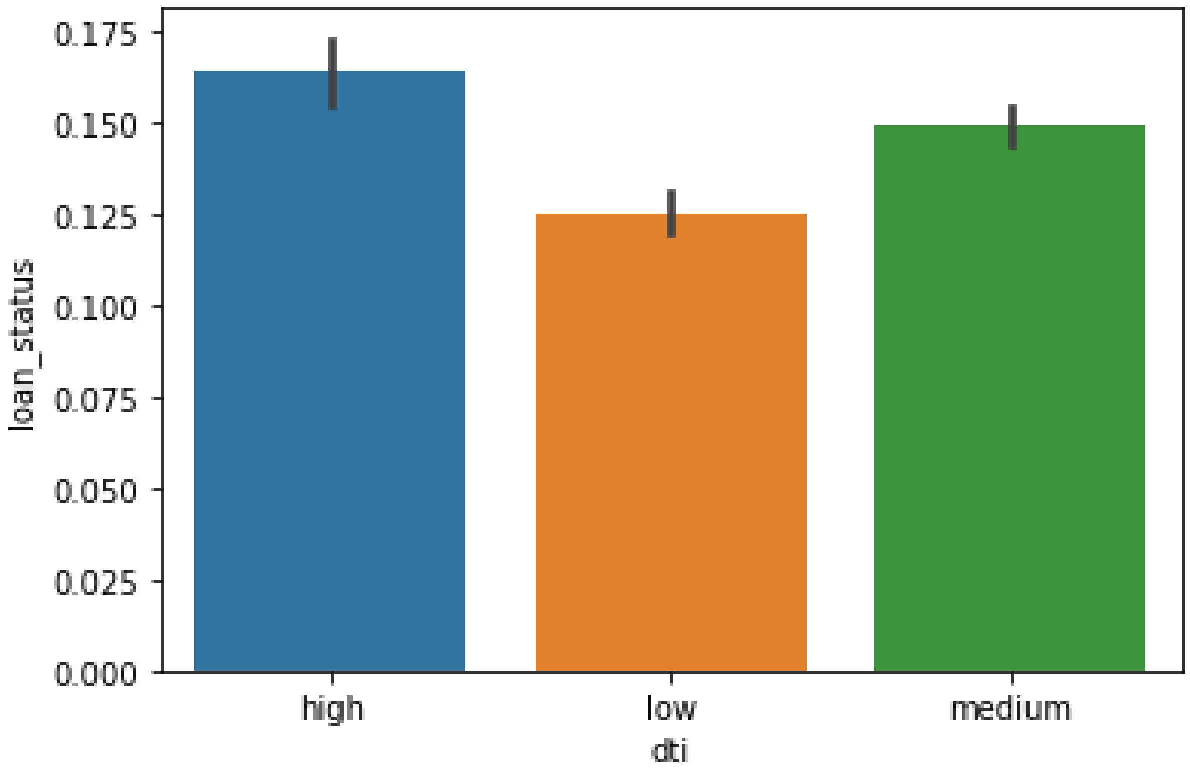
Higher the funded amount invested higher is the default rate

Default by Interest Rate



It is evident that higher the interest rate higher the chance of default

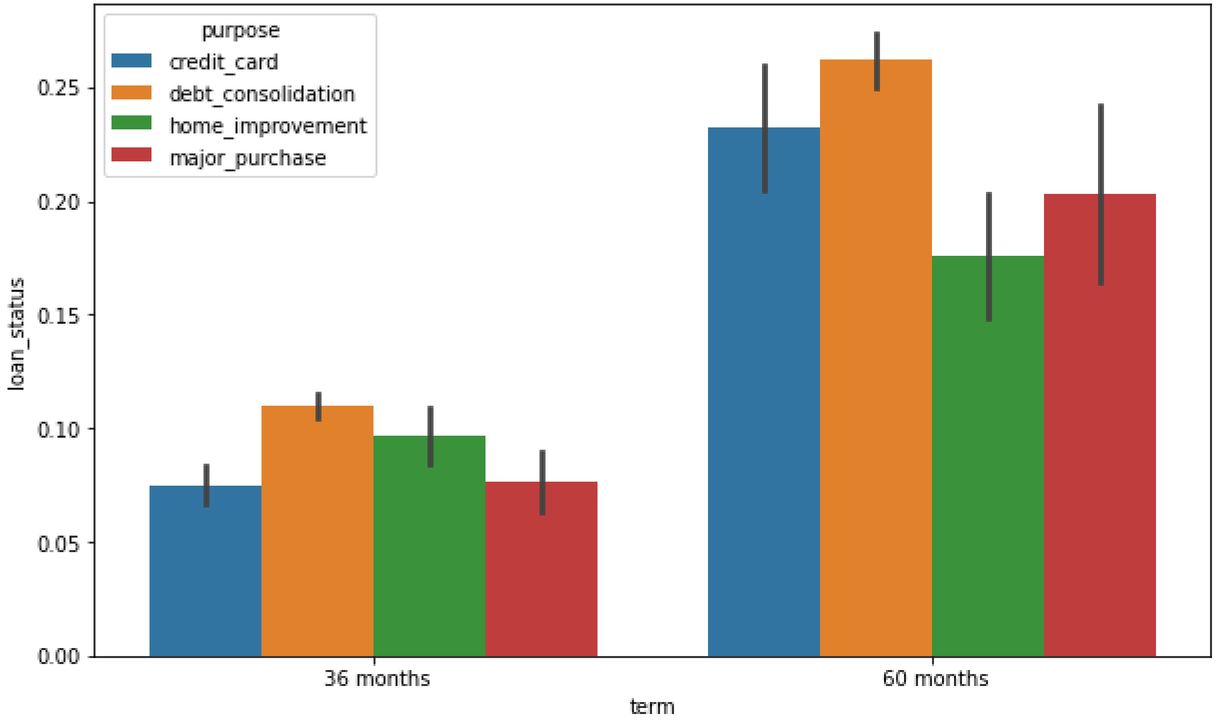
Default by Debt to Income



Wherever debt to income ratio is higher the default probability is high

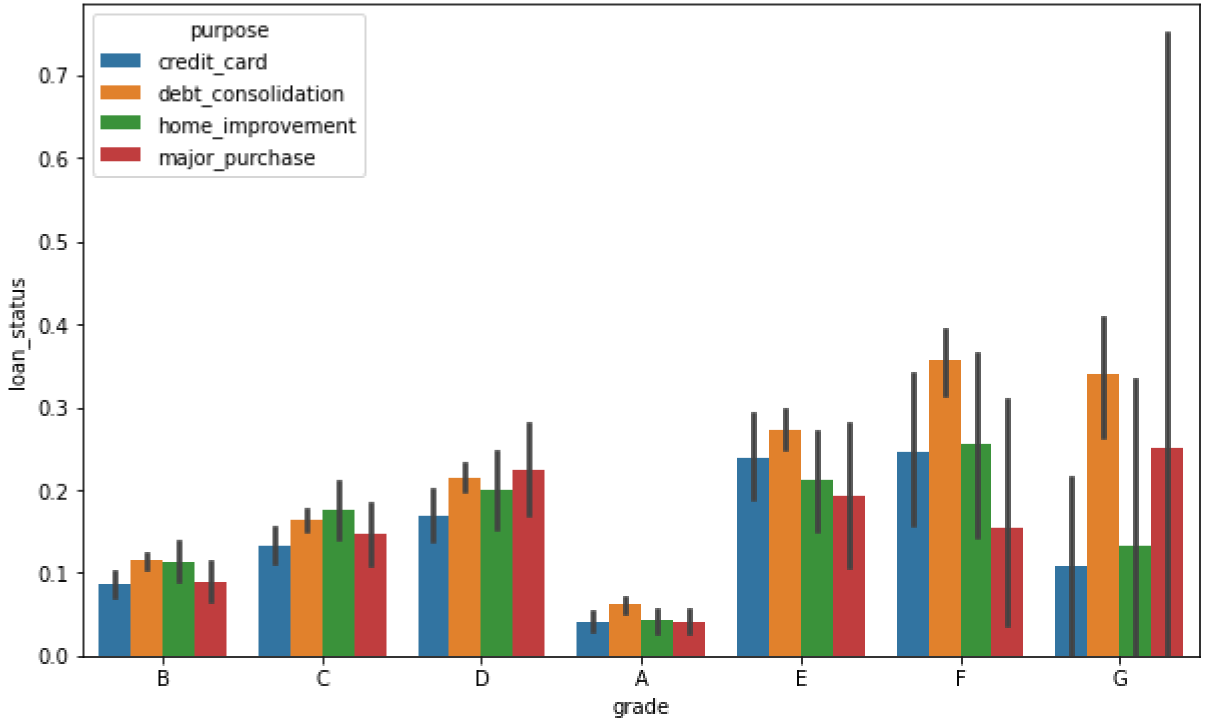
Bivariate Analysis

Default by Term across effective Variables



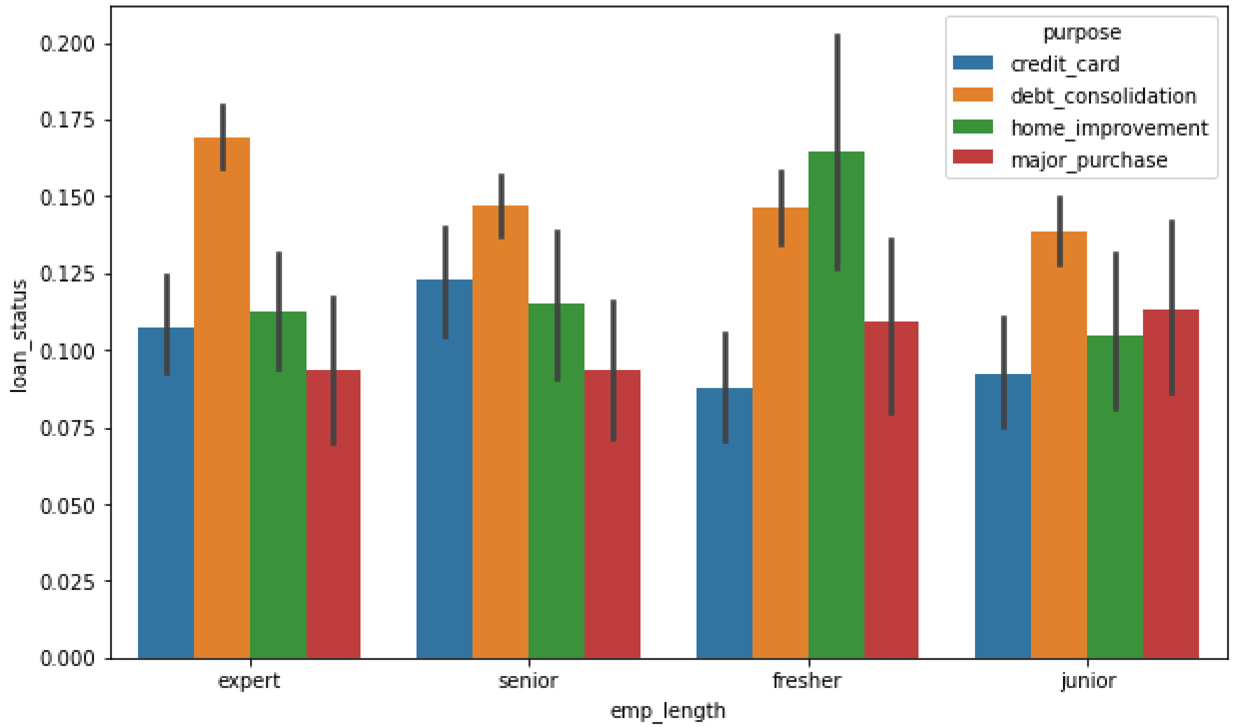
Debt Consolidation is the has the highest incidence of default when term is more where as when the term is less there is not much difference among the variable

Default by Grade across effective Variables



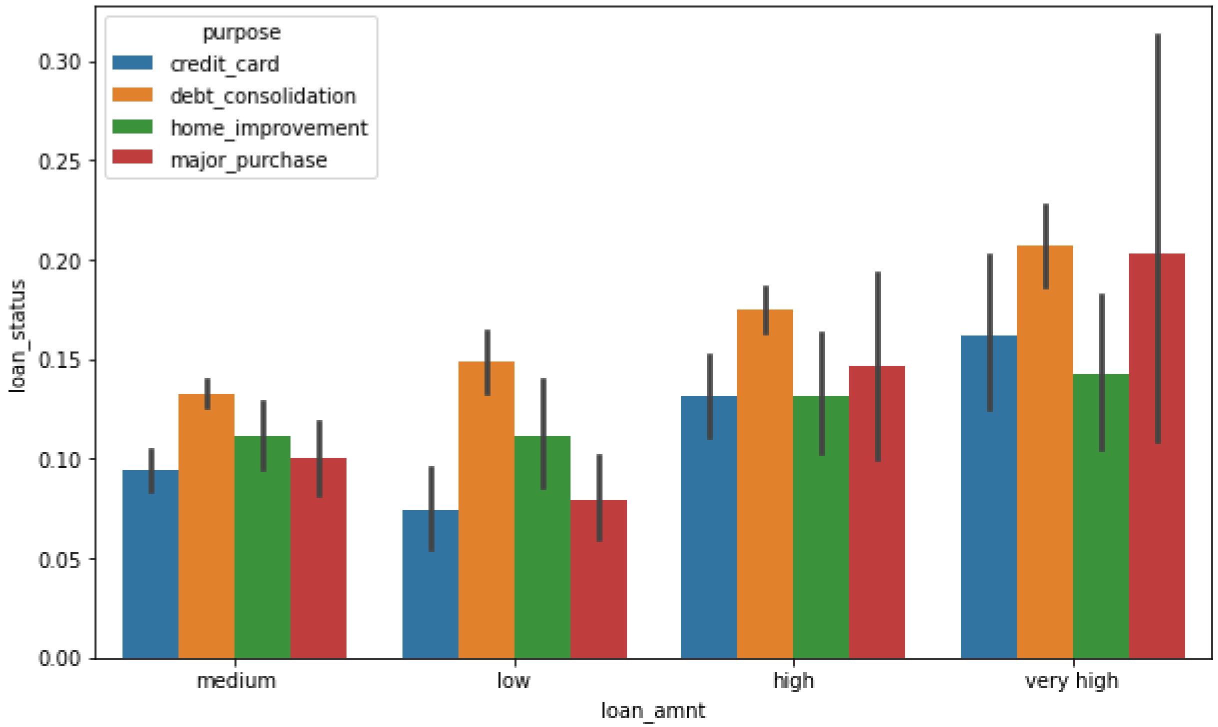
Debt Consolidation is the has the highest incidence of default when Grade is F & G, where as when the is less there is not much difference among the variable

Default by Employee Service Length across effective Variables



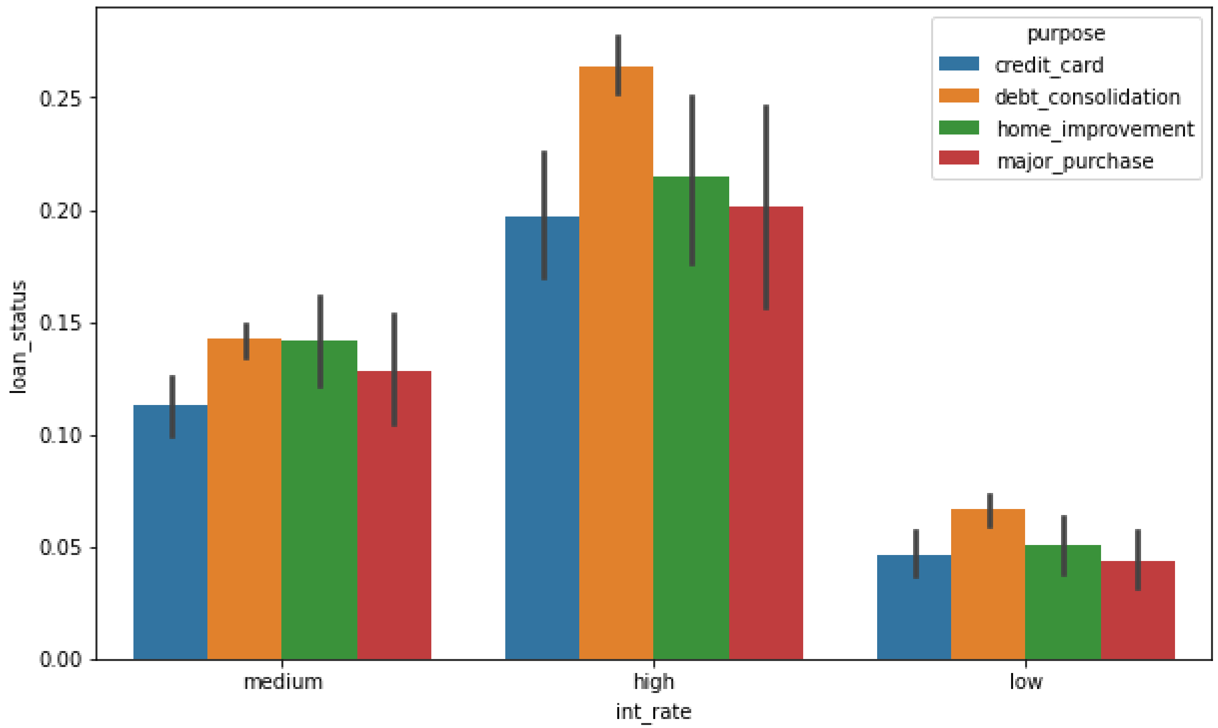
We can see that in the category of experts the loan availed for debt consolidation has higher chance of default whereas in the category of freshers home improvement has the high possibility of default

Default by Loan Amount across effective Variables



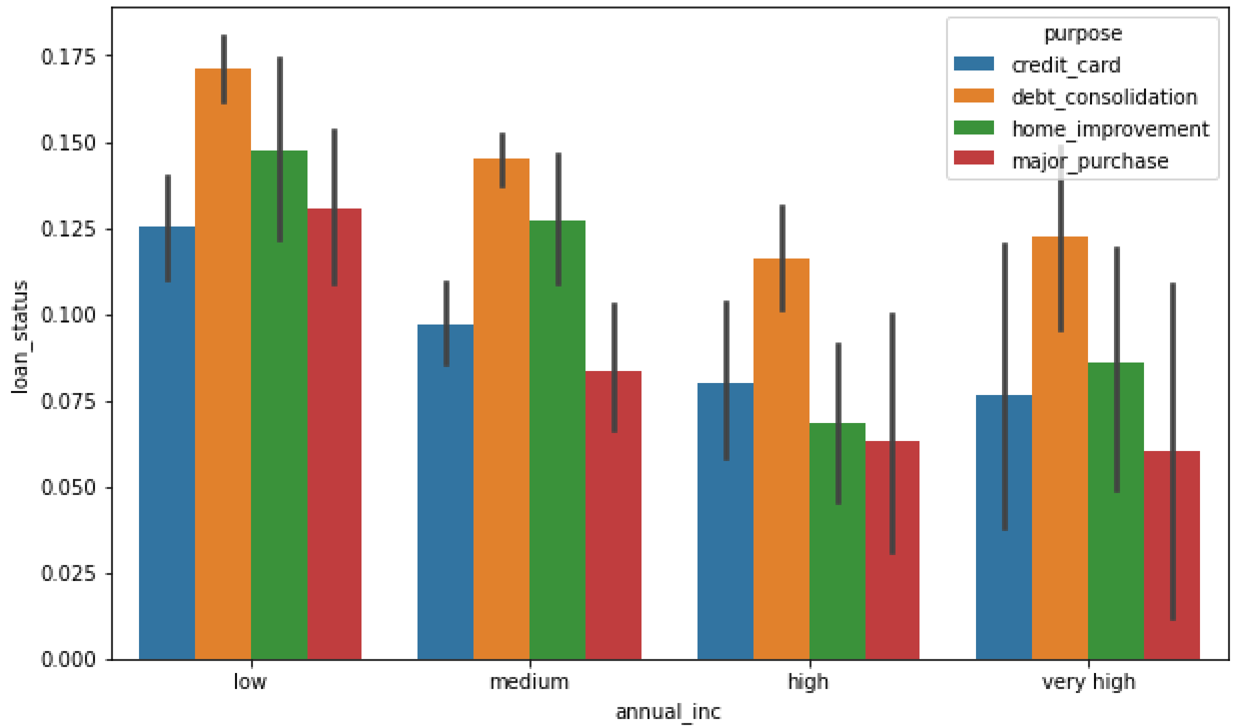
We can see that when loan availed is very high chances of default is also very high in both debt consolidation and major purchases.

Default by Interest across effective Variables



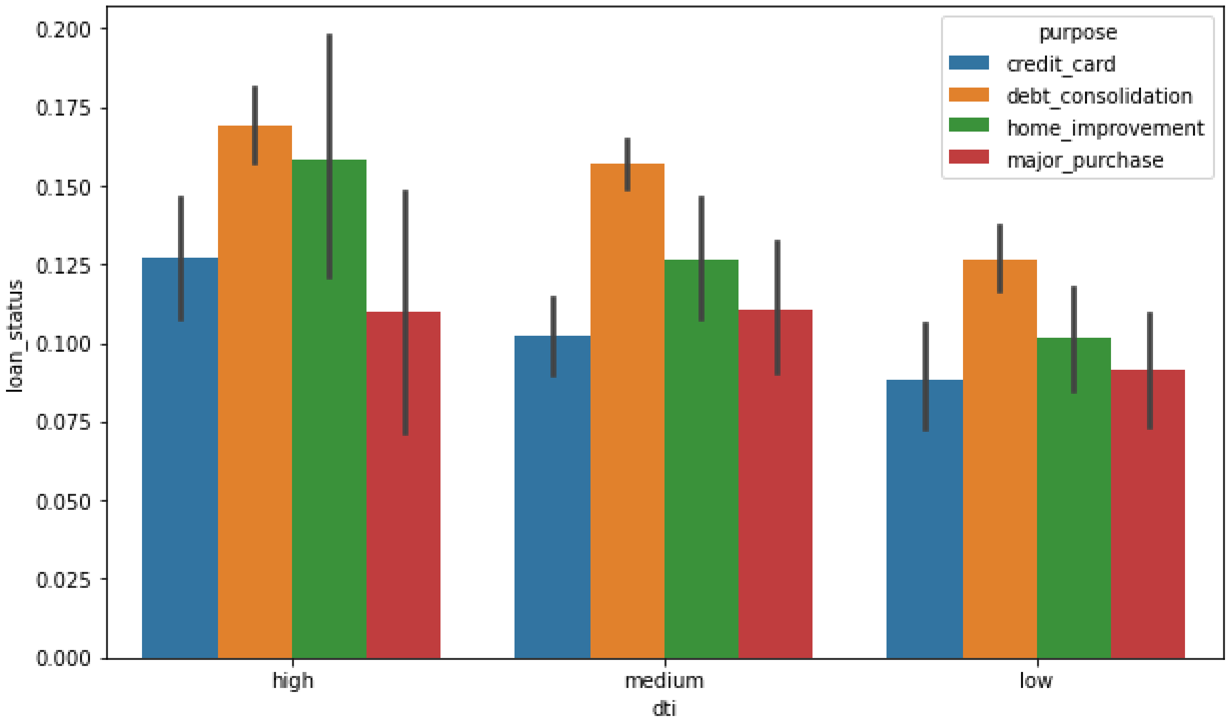
When the Interest rate is high the incidence of default is very high

Default by Annual Income across effective Variables



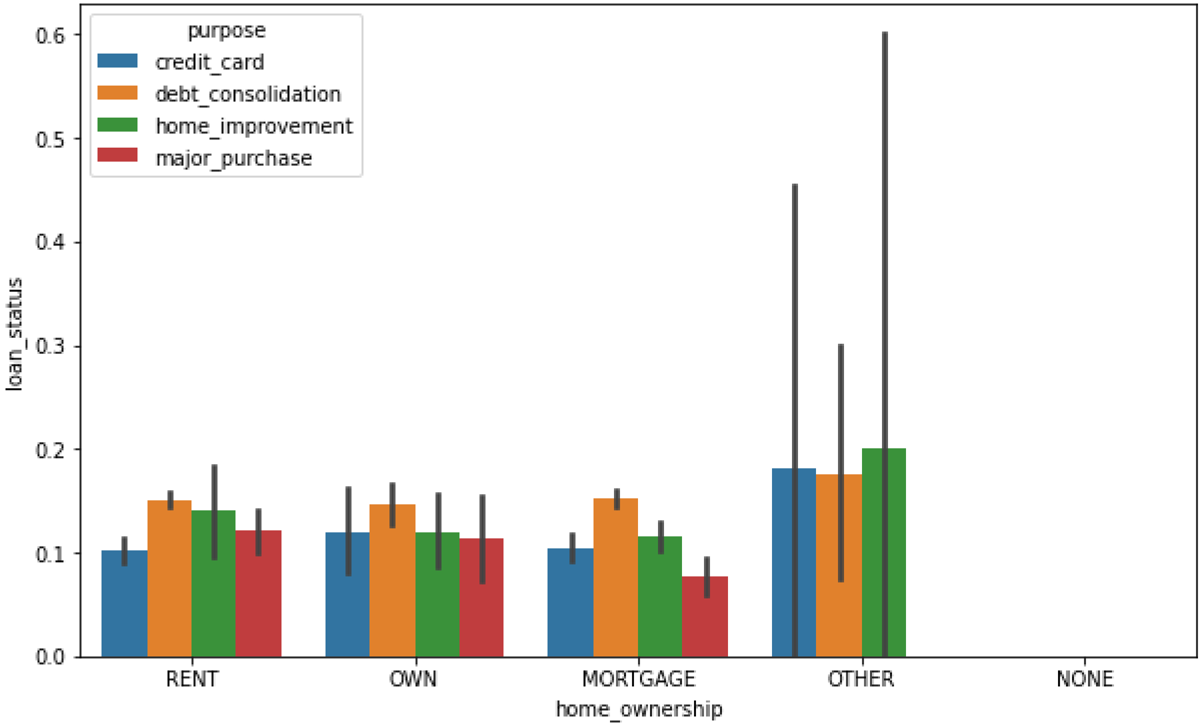
Low Income applicants are more prone to default in general and variance in the very high income group is very high in comparison to other income groups.

Default by DTI across effective Variables



The default rate is very high where debt to income ration is very high especially loans taken for home improvement

Default by Home Ownership across effective Variables



The default rate is very high where Home Ownership is others and mortgage

Findings

• Loans with higher term (60 months) are defaulting the most
• Low quality loans are defaulting the most (grades D, E, F & G) • Loans approved even after source verification process resulting in default.
• Higher loan amounts and higher interest rates resulting more defaults
• Loans borrowed for small business purpose resulting in more defaults
• Borrowers with 10+ years of experience and borrowers with less than 3 years of experience are defaulting the most
• Borrowers with low income are defaulting the most.

Suggestions

STOP
• Approving high loan amounts for borrowers with less income and higher dti (debt to income)

REDUCE
• Number of approvals for the purpose of small business
• Loan amount and interest rate for loans applied by borrowers with 10+ years of experience and less than 3 years of experience

REVISIT
• Source verification process (as loans approved after source verification process are defaulting the most)

Thank You