Gramener Case Study(EDA)

Analysis

Group Name:

1. Saurabh Singh
2. Saurabh Dongare
3. Shantanu Kanchan
4. Ravi Rathod

Problem statement

• A **consumer finance company** specializing in lending various types of loans to urban customers. Customers apply for loans online and when the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile.

• Now the company decides whether to grant them loans or not using internal profiling and also they decide the stakes at which it can be provided like tenure, installments, interest rates, etc.

• Considering the risk of non repayments they need some matrices which could specify and help them reduce the risk. They provide various products like credit card loans, housing loans, car loans, etc.

• Using EDA to understand how **consumer attributes** and **loan attributes** influence the tendency of default.

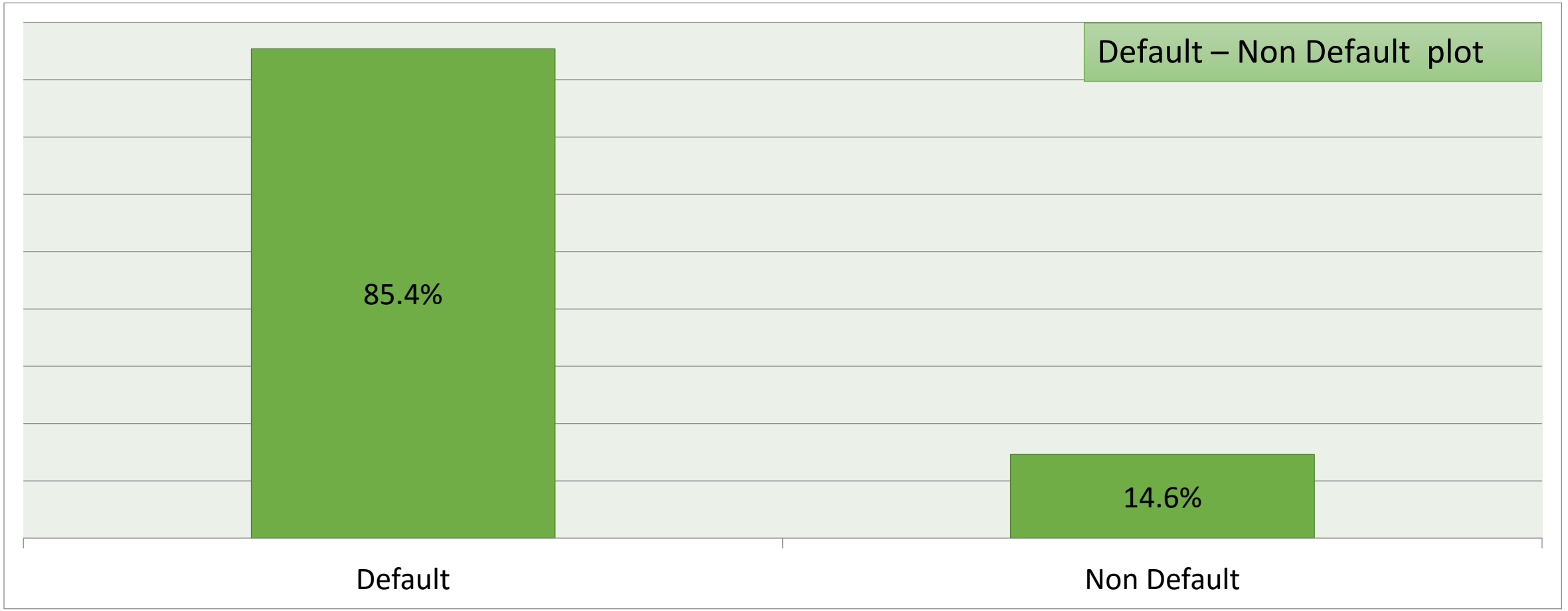
Business Objectives:

To identify the driving factors i.e. the variables which are strong indicators of loan default which the company can utilize this knowledge for its portfolio and risk assessment.

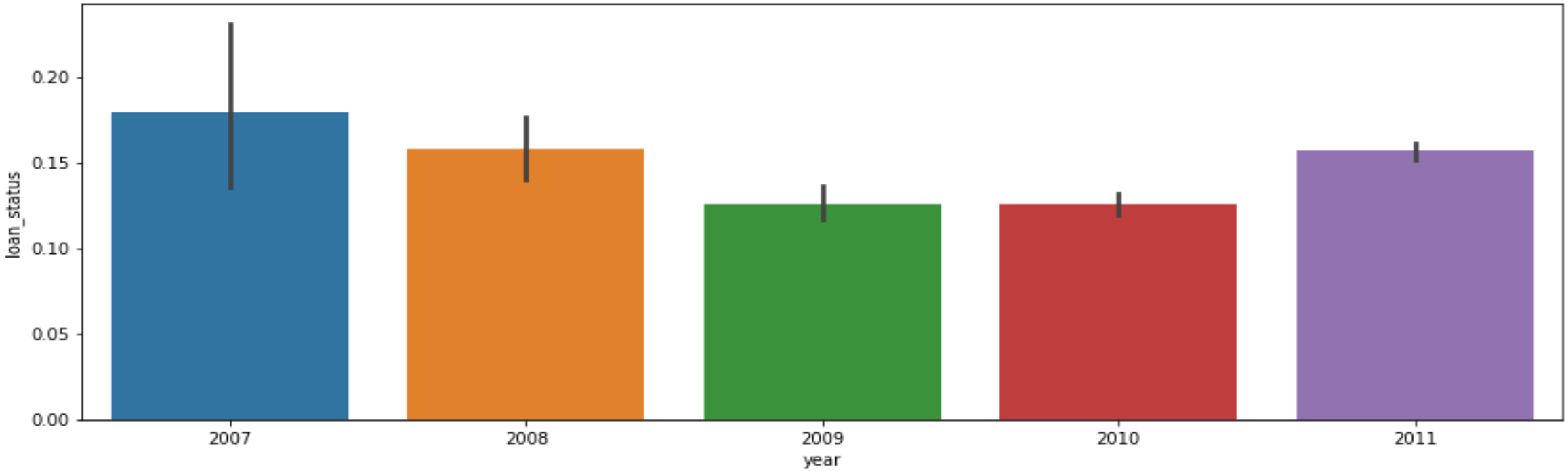
Analysis Approach

- Data Understanding (Understanding Problem statement, Columns provided and data dictionary)
- Data Cleaning (Cleaning missing values, removing redundant columns, etc)
- Data Analysis (Univairiate Analysis, Segmented Univariate Analysis, Bivariate Analysis, etc)
- Recommendation/Results of Analysis

Overall default rate – 14%



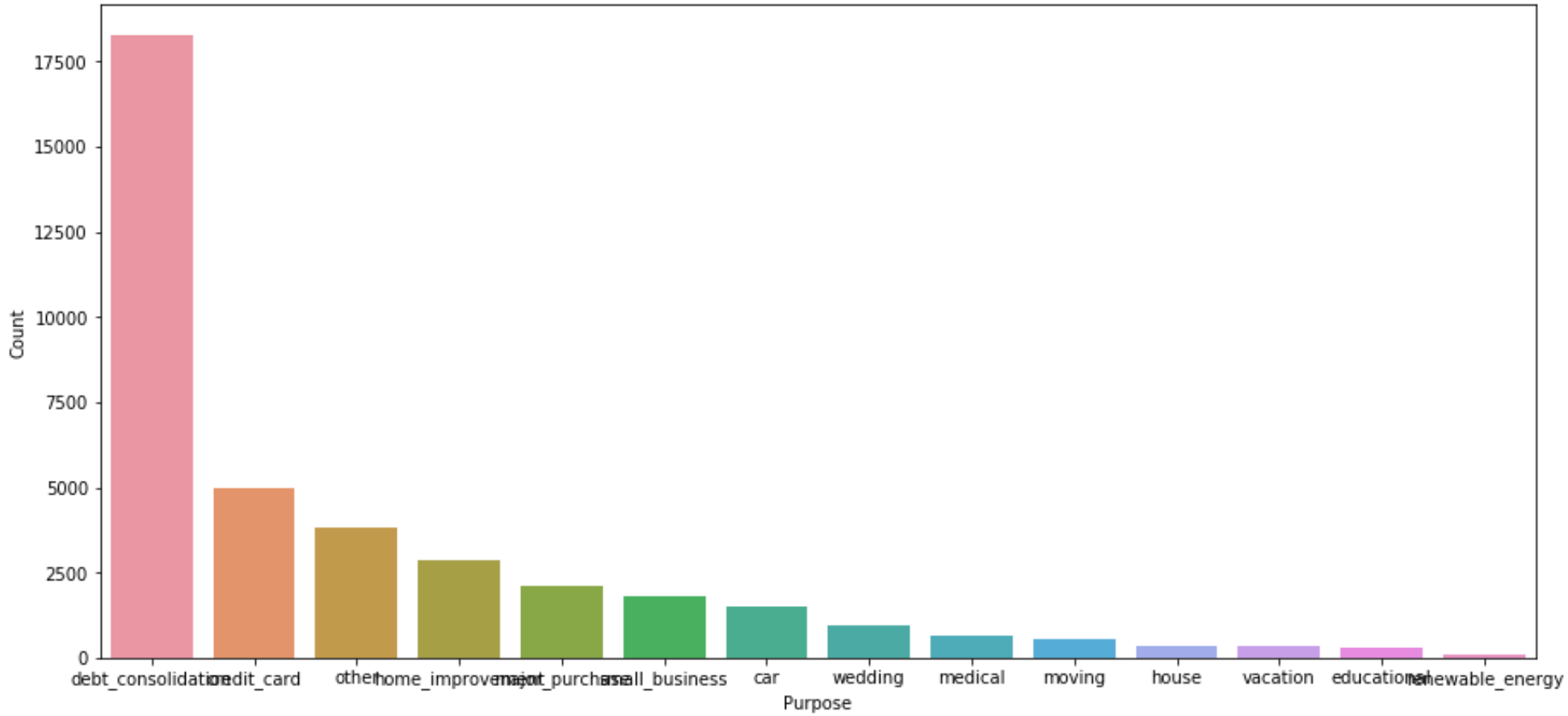
Time Frame of the data 2007-2011



Here we can see that the default rate decrease from 2008 till 2010 and suddenly increased in 2011.

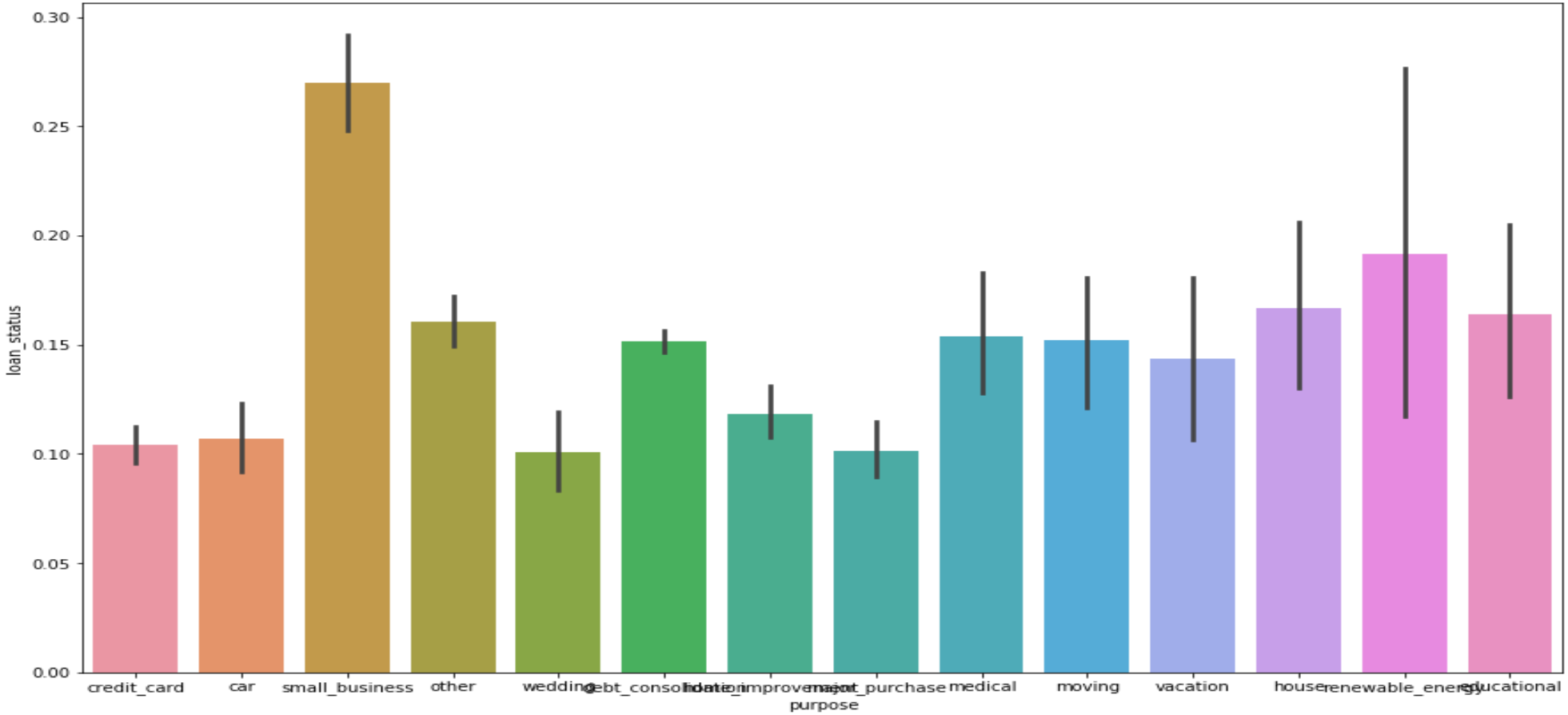
Loan Purpose Distribution : Debt Consolidation

•Debt consolidation (46.8%) is the most popular loans followed by credit card loans(13%) and home improvement loans(7.5%)

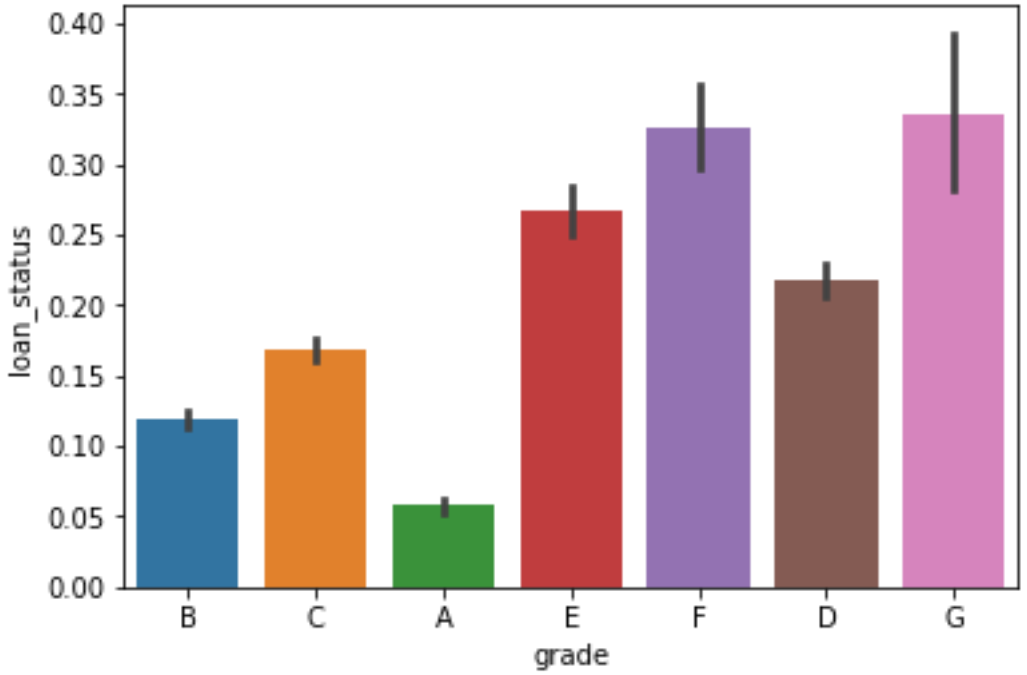


Default Rates By Loan Purpose

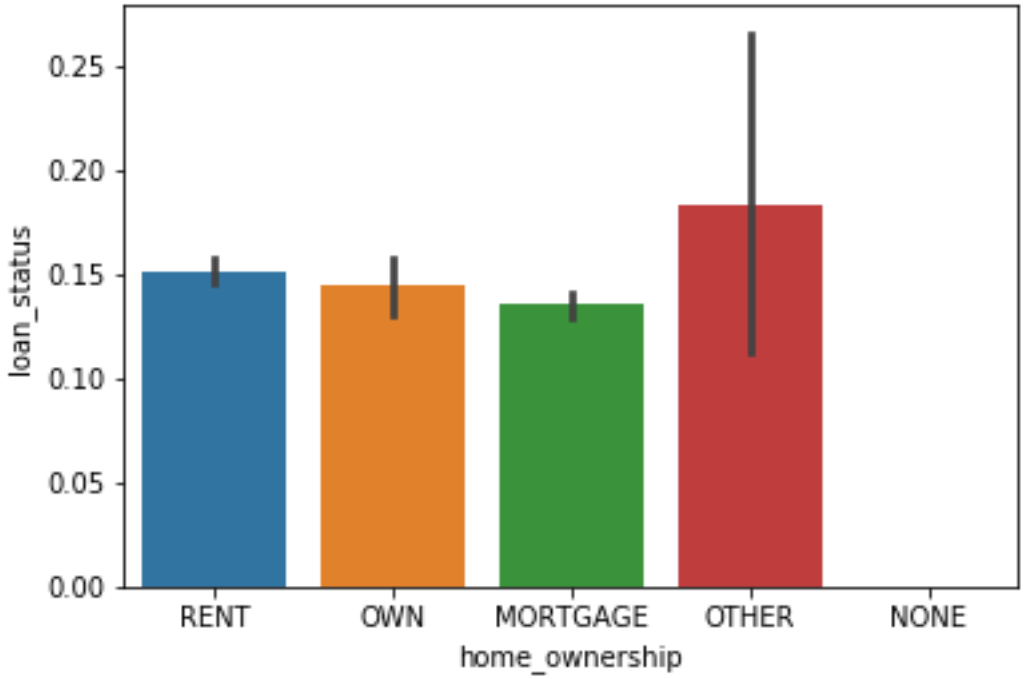
Most defaults are from small business(27%), house(17%), Educational loans(16%) as shown below.



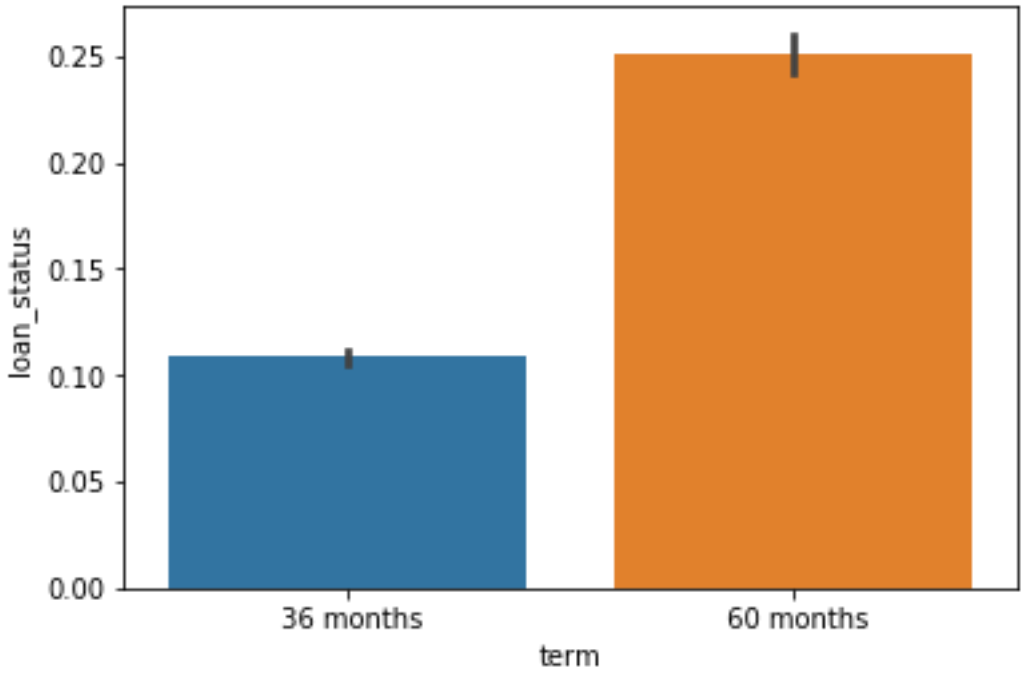
Default by grades



Default by homeownership

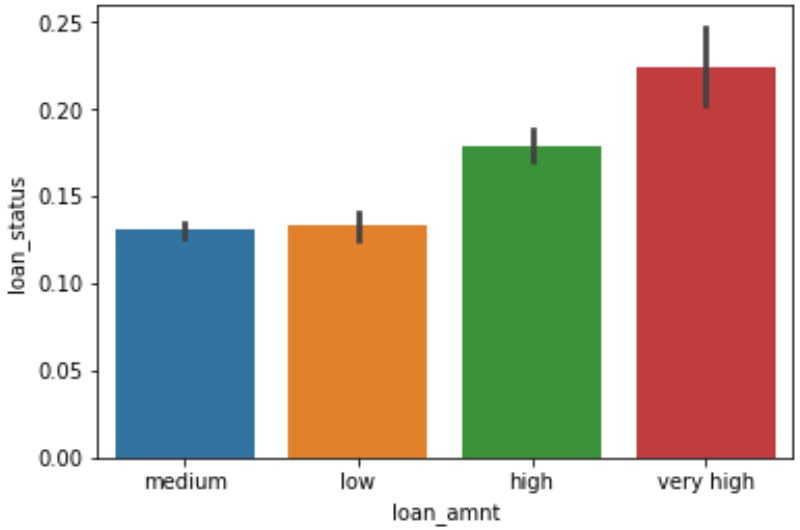


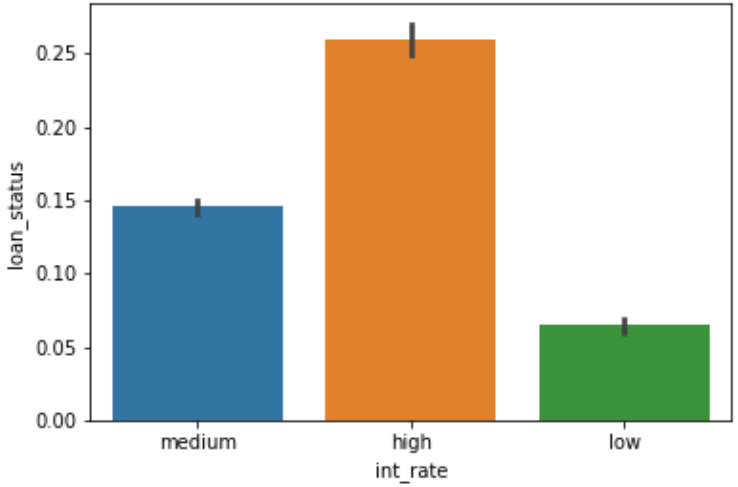
Default by term

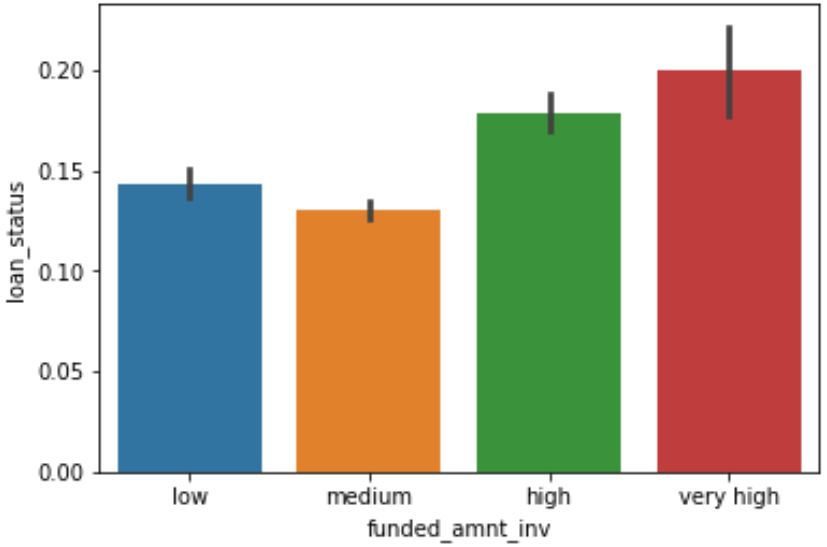


Default by loan amount, interest rates and amount invest

High loan amount, interest rates and amount invested leads to more default

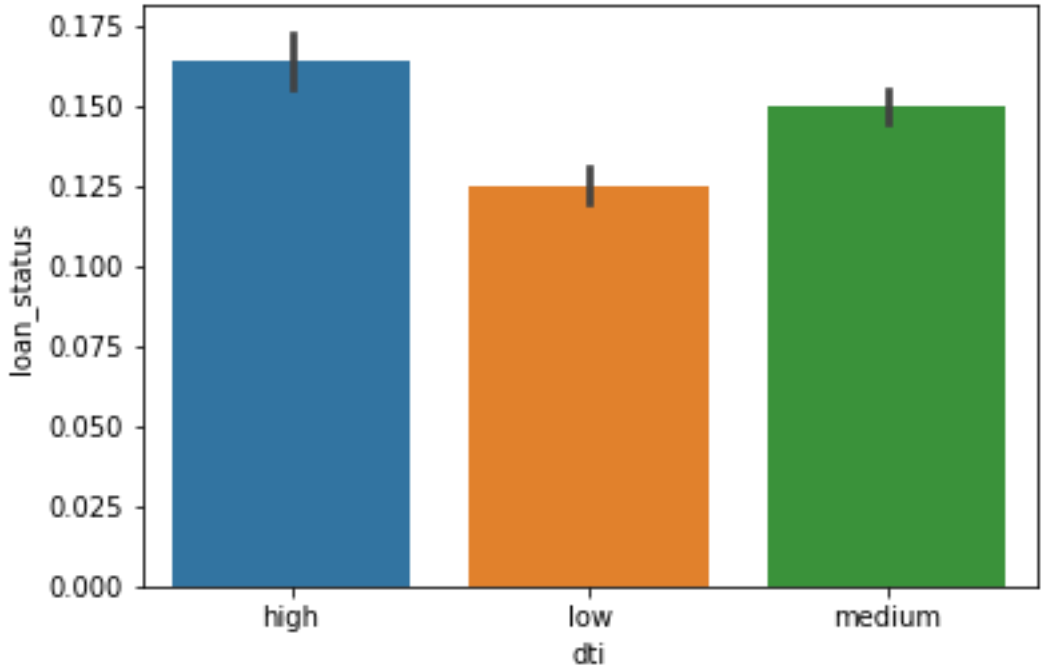






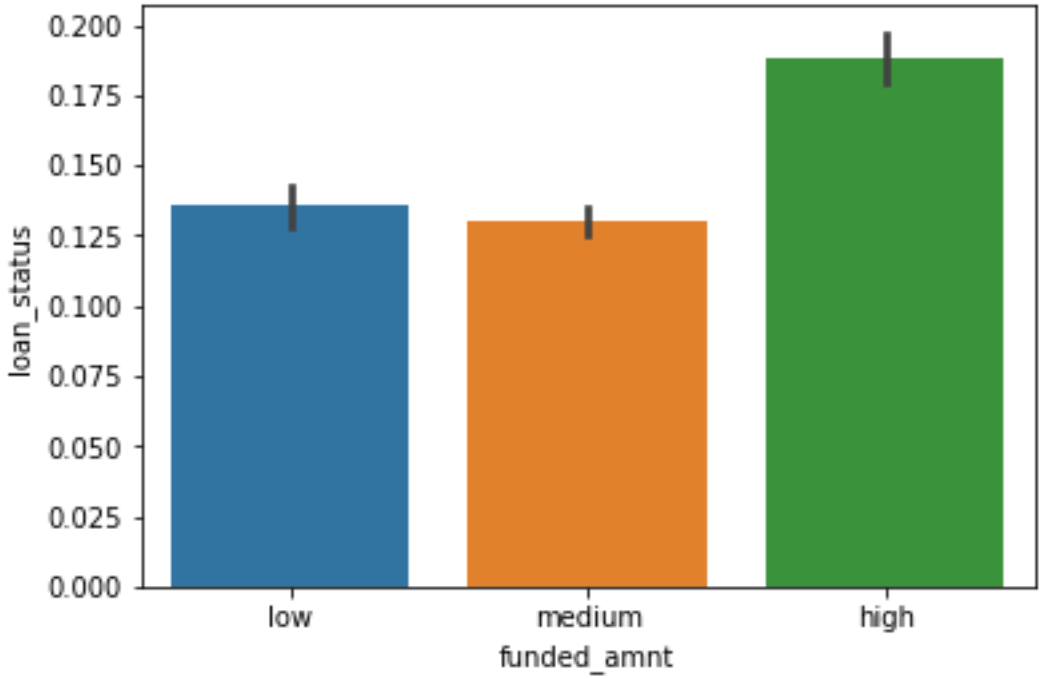
Default by dti

High dti leads to more default



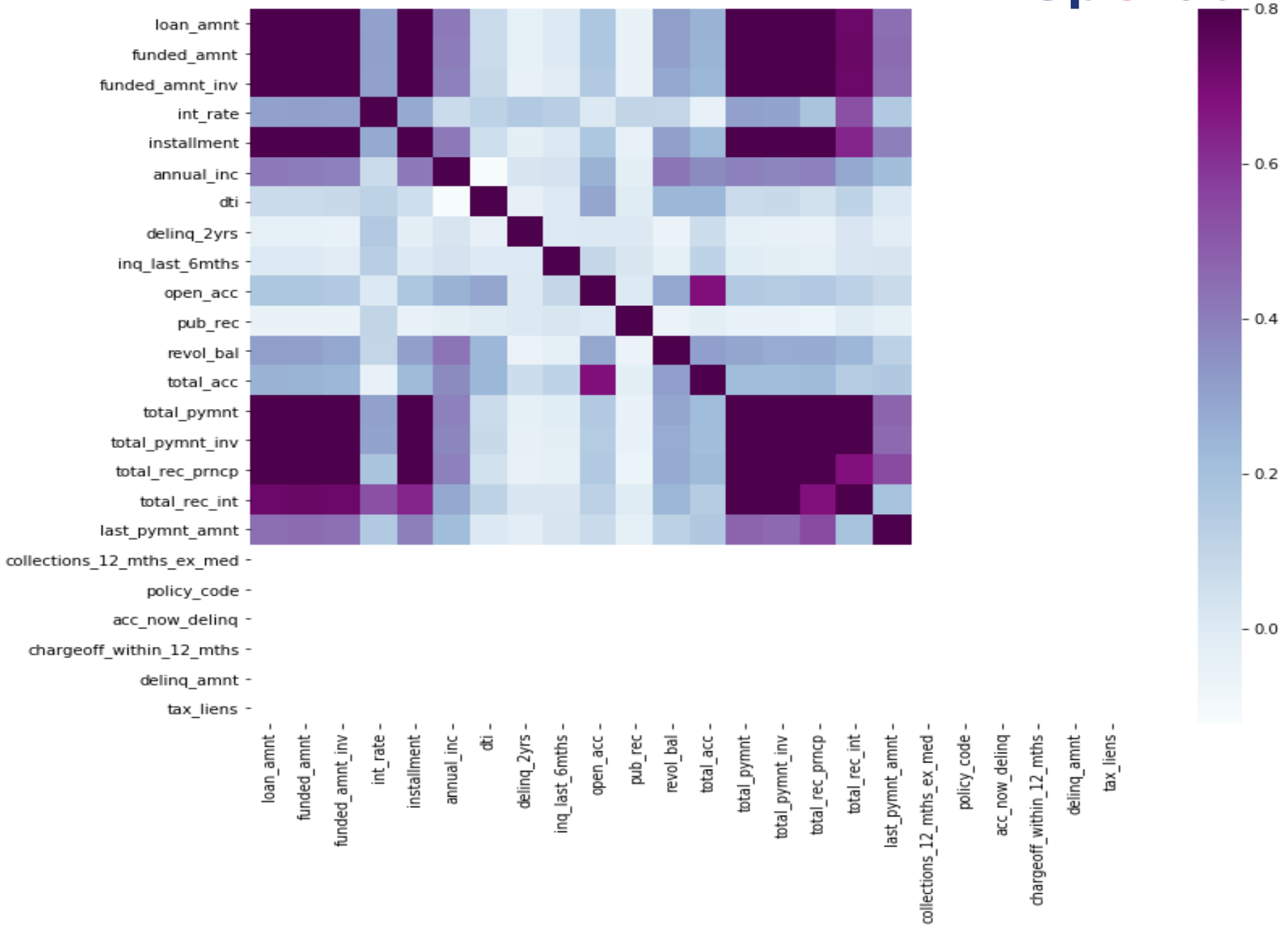
Default by funded amount

High funded amount leads to more default



Correlation Matrix

Variables like: loan amount, funded amount inv, total payment inv, total rec prncp i.e. the **numerical variables** all are highly correlated .



Conclusion

Hence we conclude by this analysis that :

1) Most of the loans are taken to consolidate debts from the other loans.
2) Most of the loans are taken in 2011
3) Around 70% of the loan are of the tenure of 36 months
4) The most prevalent interest rate are between 12 to 16 percent
5) Majority of the applicants stays in rents or mortgage house
6) Large number of applicants have the annual income between 500000 to 1000000