Gramener Case Study

Group:

1. Akash Dusane
2. Chandana Joshi
3. Yogitha Goswami
4. Karan Ramchandani

Gramener Case Study Analysis

Objective : Reduce credit loss for the company by analysis of the data

This case study is for identifying risky loan applications using Exploratory Data Analysis.

Analyze the sample loan data to find the driving factors for loan defaults, ascertain the risk factors before approving the loan.

This will help the company to reduce the risk of loan defaults and to avoid credit loss.

Data Analysis Approach

Step 1 :Understanding the Data:

We have a dataset of the loan applications which have already been processed. The data consists of attributes of the loan and payments made and also attributes of the customer (borrower).

There are several factors which can affect the loan payment and the data has been given for some important factors like if *loan status*, *borrower has other debts*, what is *purpose of the loan*, *income of the borrower*, if the *source of income* of the borrower has been verified.

We need to further drill down and clean the data for better understanding of the data.

Step 2: Data cleaning:

Remove columns/ attributes having NA/ NULL values.

Remove duplicates

Remove special characters and convert to valid formats for measures and dates.

Data Analysis Approach

Step 3 : Adding derived columns where required

Added **income category** column to classify the borrowers based on income.

Added **region** column to classify if defaulters are from specific region based on state.

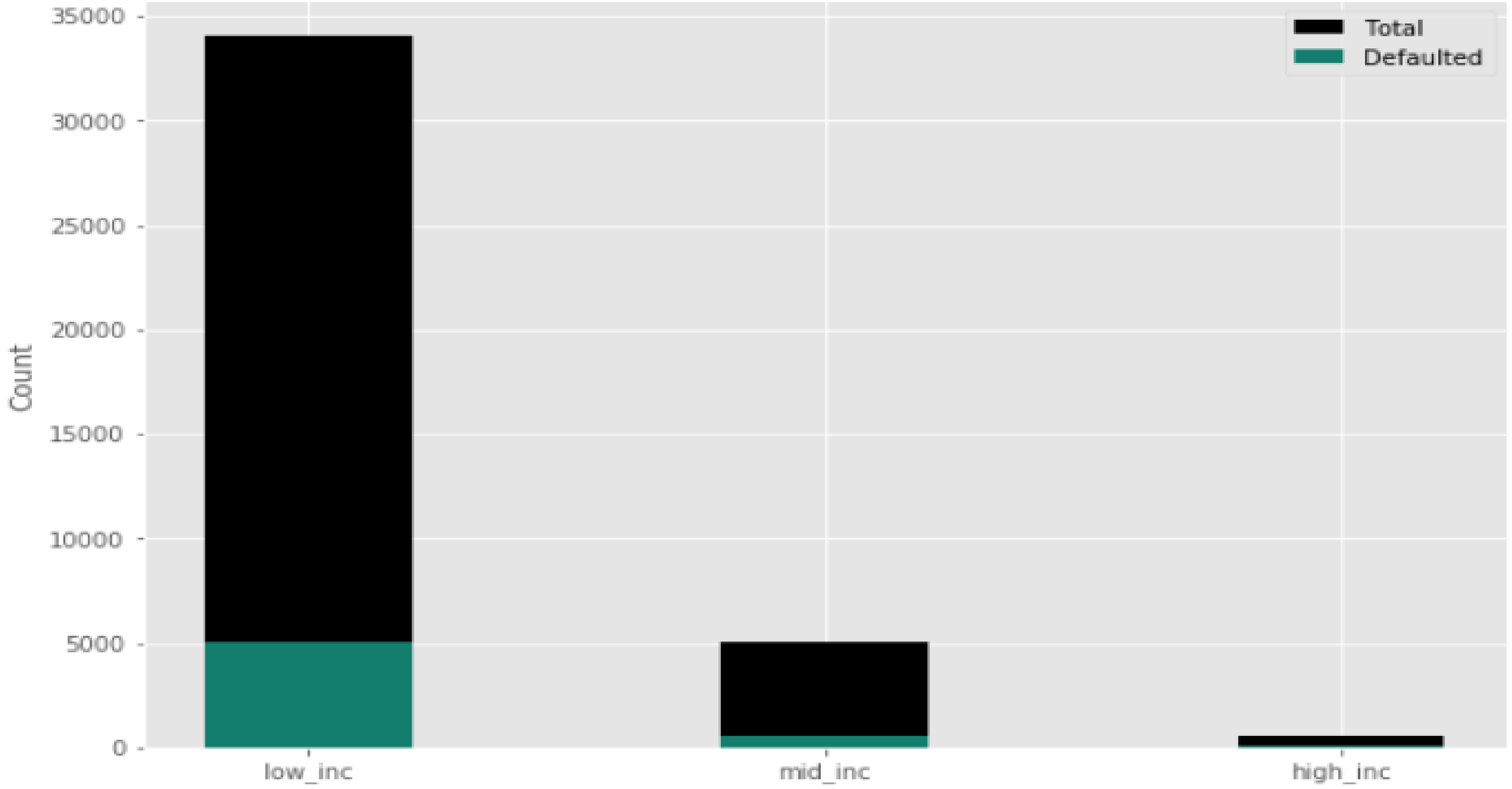
Added **Defaulter** column as binary to indicate if the particular borrower defaulted or not.

Added Loan **Issue year and month** to show the increasing loan applications every year.

Step 4: Analysis

Here we take different attributes and try and see if there is a significant trend that we can see with which we can directly infer some driving factors for loan defaults.

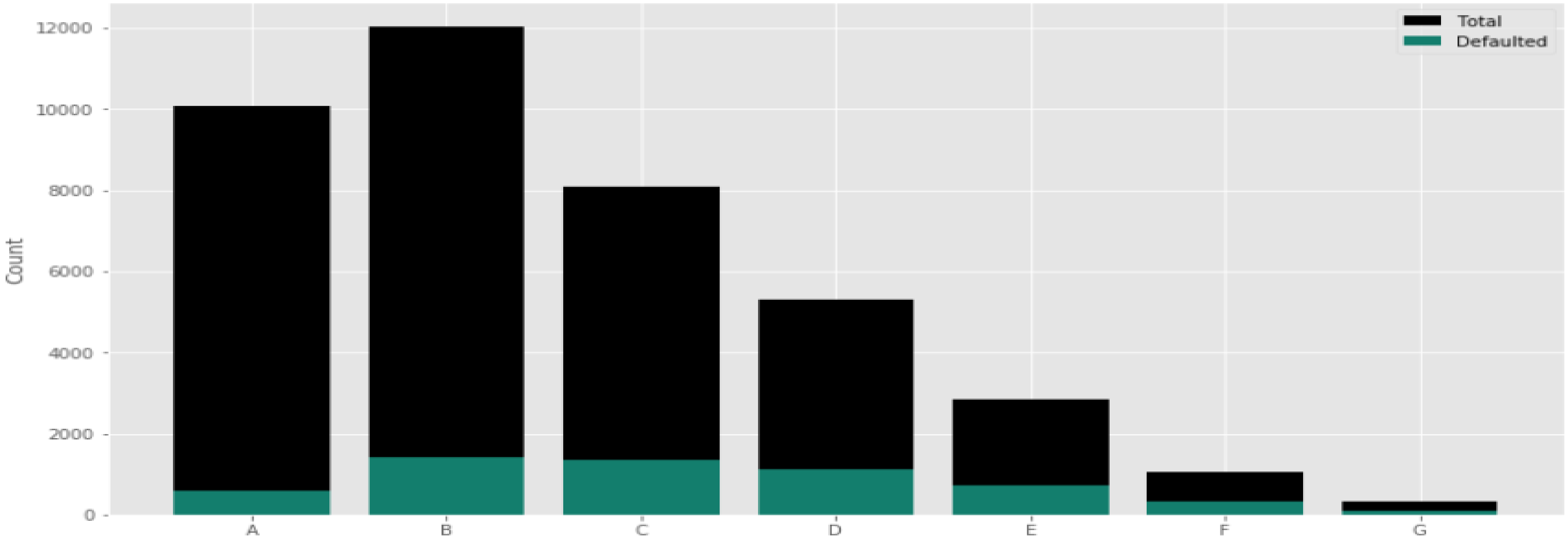
Analysis by Income Category



Default ratio is number of defaulters to the total number of borrowers. Default ratio for each income category:

low_income: 14.8%   mid_income: 10.43%   high_income: 10.24%

Analysis by Grades (assigned by LC)

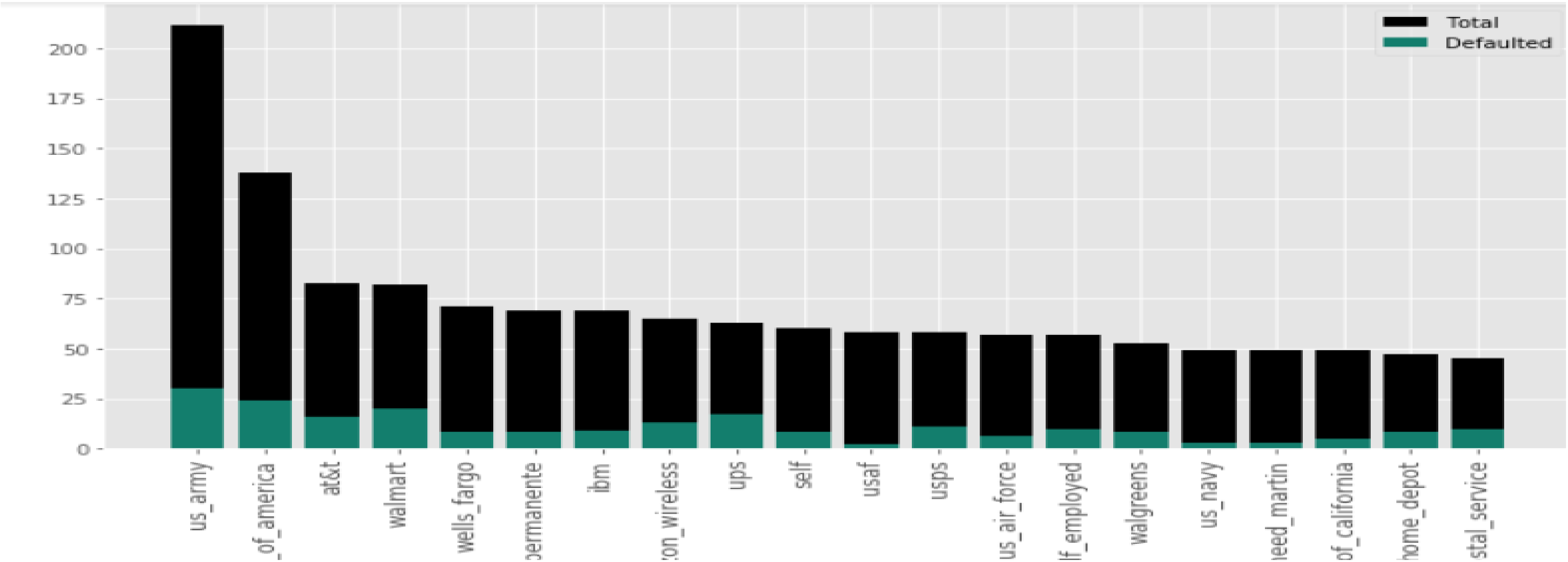


Default ratio for each grade:

A: 5.97%   B: 11.86%   C: 16.63%   D: 21.07%   E: 25.16%   F: 30.41%

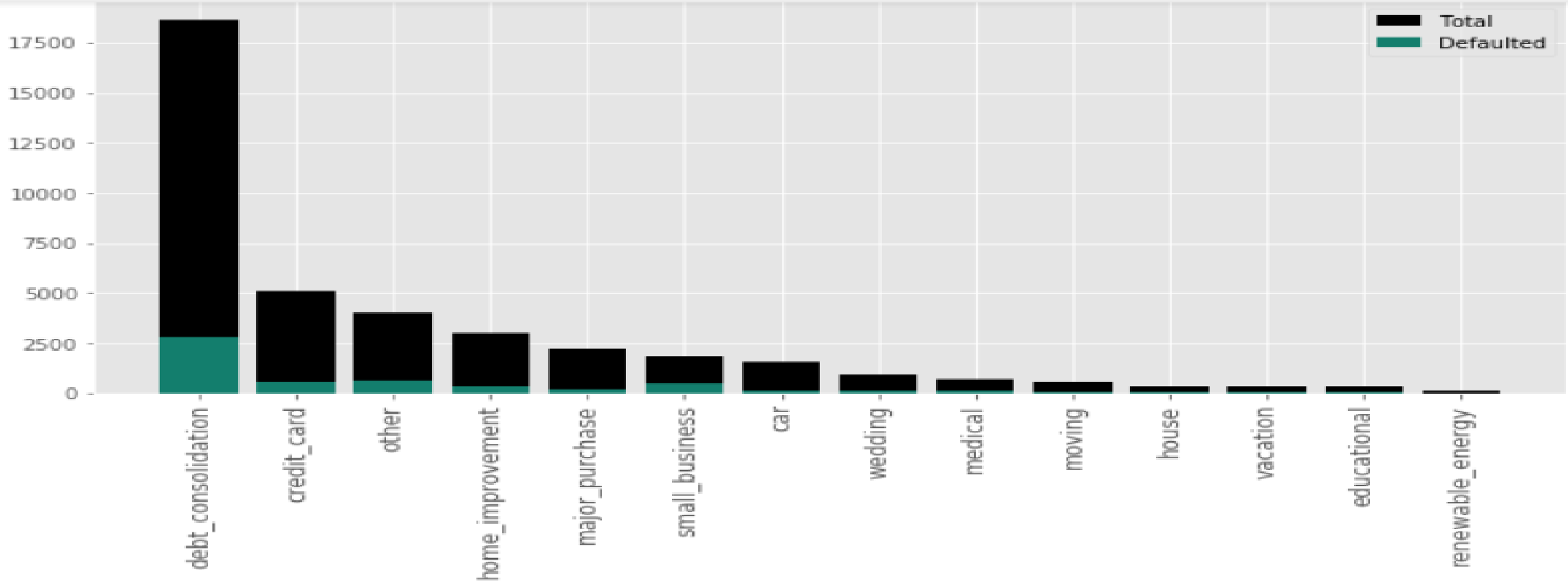
G: 31.96%

Analysis by employment



- Most number of borrowers are employed at the US Army followed by Bank Of America, but highest default ratios are from employees of:
- UPS: 27%
- WALMART: 24.4%
- US POSTAL: 22.22%

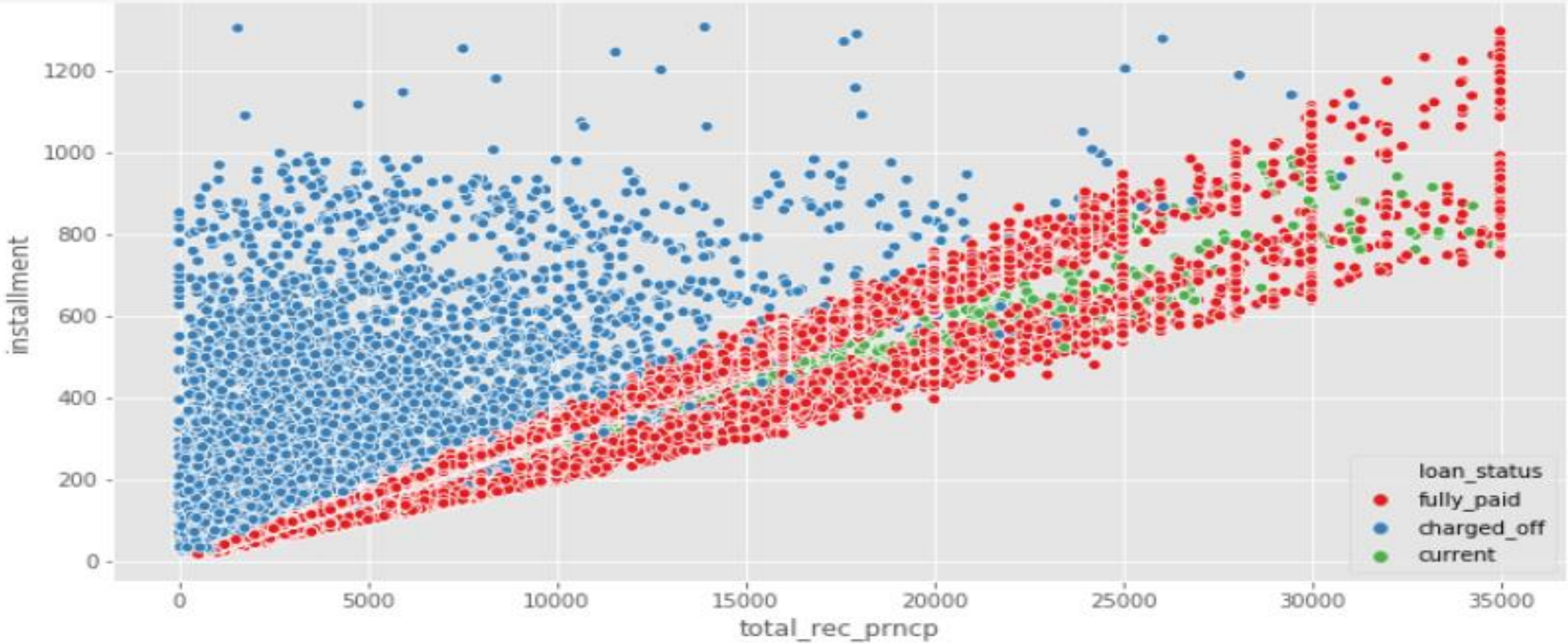
Analysis by purpose of loan



Debt consolidation is the most common reason to take a loan and makes up for 46.93% of client base.

Highest default ratio exist for small_buisness (26%) followed by renewable_energy (18.5%)

Bivariate Analysis



Shows correlation between installments and reoccurring principle.

Summary

Below factors are significant driving factors for loan defaults. Considering them carefully while approving loans in future will reduce credit losses.

1.Grade:

Grade and rate of default are heavily inversely correlated. The highest default ratios are present for the three worst grades namely G: 31.96% F: 30.41%   E: 25.16%

Recommendation: Offering loan to grades above E is considered highly risky, take caution while approving loans for grades E,F,G. These customers have a tendency to default and hence been given these grades.

2.Interest Rate:

There's a 20% positive correlation between interest rate and tendency to default. Interest rate is heavily correlated with grade as better the grade, lesser the interest rate. Therefore borrowers without a good grade (A, B) will receive higher interest rates that decrease their chance of paying of their loan. (Thus increase chance of default.)

Recommendation: Reducing interest rates for lower grade applicants, can potentially enable fringe borrowers to more regularly pay monthly installments.

3.Total Recurring Principal:

There is a strong negative correlation (34%) between Total Recurring Principal and tendency to default. This effect can be seen that defaulters usually have higher installment charges and higher total payments for an increasing value of TRP.

Recommendation: Since interest rate and Principle has correlation with grades, Offering loan to grades above F is considered highly risky, take caution while approving loans for grades F and above.

Summary

4.Purpose of Loan:

47% of borrowers stated 'Debt Consolidation' as purpose of loan and make up the majority of borrowers. The purpose with the highest default ratio is 'Small Business' (26%) followed by 'Renewable Energy' (18%).

Recommendation: Small business is a risk factor as the business may fail and customers end up defaulting on loans as already they would have incurred losses. Avoid offering loans for small business and renewable energy purposes.

5.Income:

Borrowers from the 'Low Income' category make up for 85.68% of the client base of borrowers, which is an overwhelming majority. 15% of these tend to default their loan. Probability to default is 4% more in case of low income borrowers compared to mid and high income. 28.5% of all low income, bad grade (E, F, G) borrowers tend to default.

Recommendation: Examine the grade before approving loans for *low income category,* this might significantly help in reducing the credit losses.