
Landing Club Case Study

Submission

EPGP ML C46 October 2022 Batch

Shailesh Kadre

6 December 2022

Problem Statement

This case study is for a consumer finance company which specializes in offering online loans to customers in urban segment for personal, business and medical procedure domains. The operations are very efficient and fast due to a good interface and company offers low interest rates to their customers.

- The main factor to perform this study is to find out and create a knowledge base of the the dependent variables which cause default. This study will be a supporting document for the company risk management and loan portfolio.

Business Goals

In this case study, the all customer data base is available for a duration of five years. Following are the key objectives to be answered from this case study:

- Clear identification of default customer with the help of study of dependent variables
- Creation of knowledgebase for the future risk assessment and customer portfolio.

Analysis Approach

Step 1: Data Exploration and Cleaning

- Drop columns with null values, all random values or single category value
- Convert values to proper int, float, date representations

Step 2: Univariate Analysis

- Drop columns with null values, all random values or single category value
- Convert values to proper int, float, date representations

Step 3: Segmented Univariate Analysis

- Analyze variables against segments of other variables
- Create derived variables

Step 4: Bivariate Analysis

- Do correlation analysis Check how two variables affect each other or a third variable
- Plot Heat Maps to Analyze correlations

Step 5: Important Factors & Recommendations

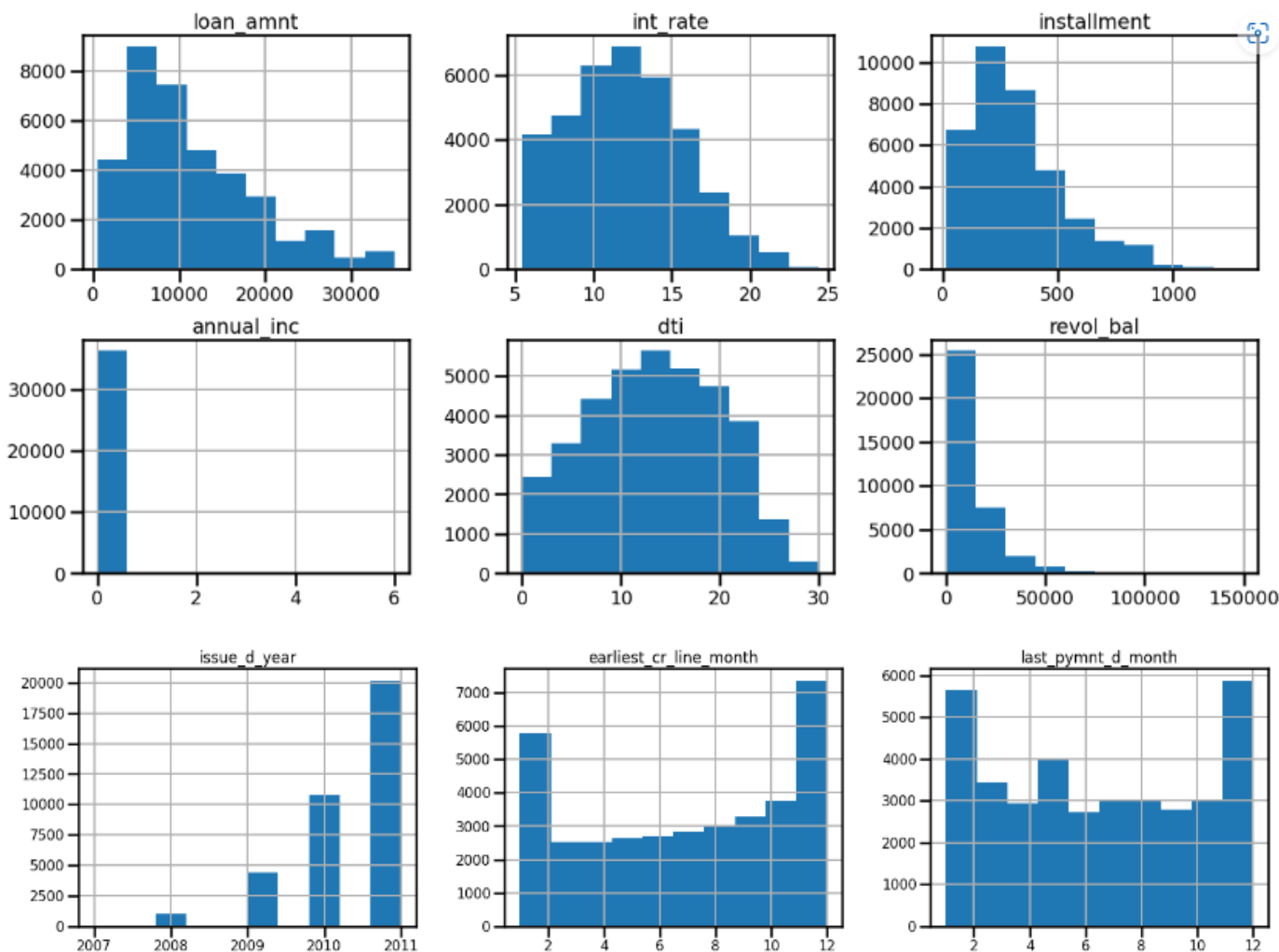
- Publish insights and observations

Data Exploration and Cleaning

- Many columns have all (39717) null values by the study of the null count sum, we can drop the columns who have more than 20000 total null values
- Dropping all the rows with null values as their number is small. Similarly drop the columns with zero values and same characters/ Strings
- Cleaning of the columns by removing special characters and strings which creates issue for the mathematical operations
- Correct the column data types as per domain requirements such as integer, floats, categorical variables etc.
- Drop the columns which do not any value to the practical analysis
- Correct the date format and split it into months and year by adding additional columns

By following above steps, the number of columns were dropped significantly and the total data was compressed to 91.7% without loss of any critical information to the business

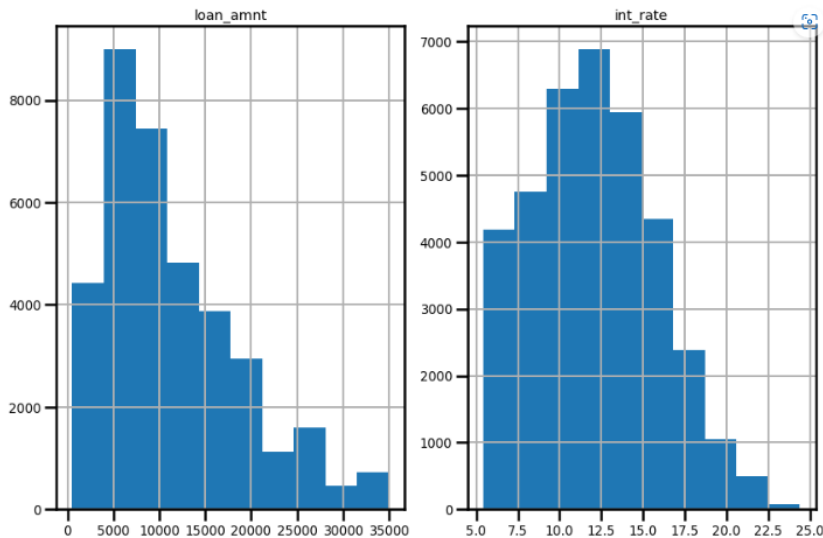
Univariate Analysis



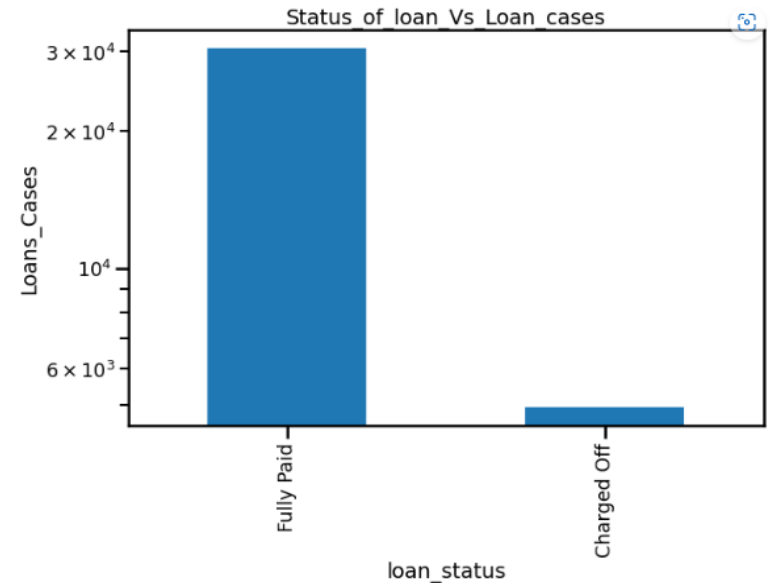
Univariate Analysis- Key Observations

- There is a huge number of outliers in last payment amount and remaining outstanding principle portion column
- Most installments are about USD 300 with 12% interest rate while the the 50 percentile of loan is of 10000 USD
- Most DTI ratio is 14
- Most revolving credit balance is 10000 USD
- Most Revolving line utilization rate is around 50
- The Most principle received is 8000 USD.
- Most of the employment length is about 5 years
- 50th percentile of total number of credit lines held by customer credit is about 20
- Mostly the loan was funded in July month .
- Mostly in July month, Borrower's credit line was opened
- Min year 2013, mostly last month payment was received
- Loan enquiries and outliers have large range.

Segmented Univariate Analysis

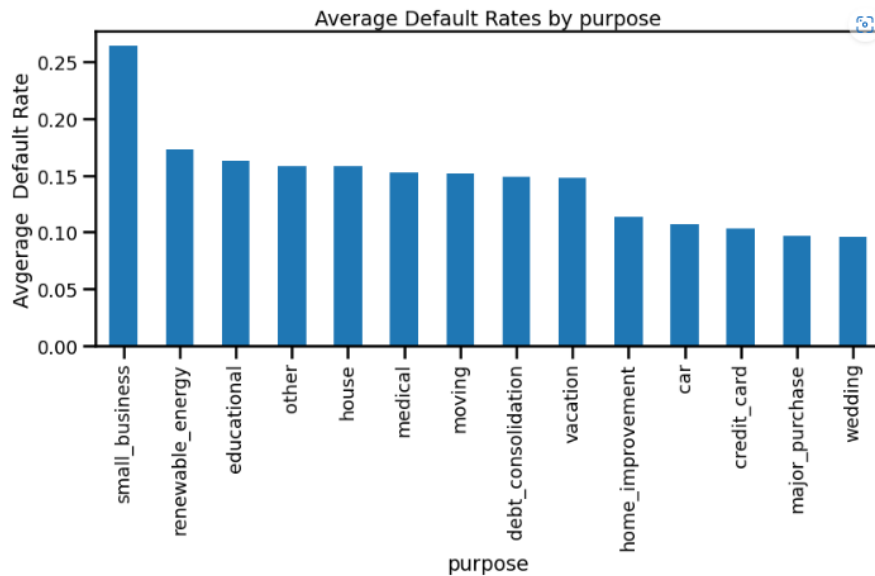


There is a huge amount of people between 5000 to 10000 USD with interest rate of 10 to 12%

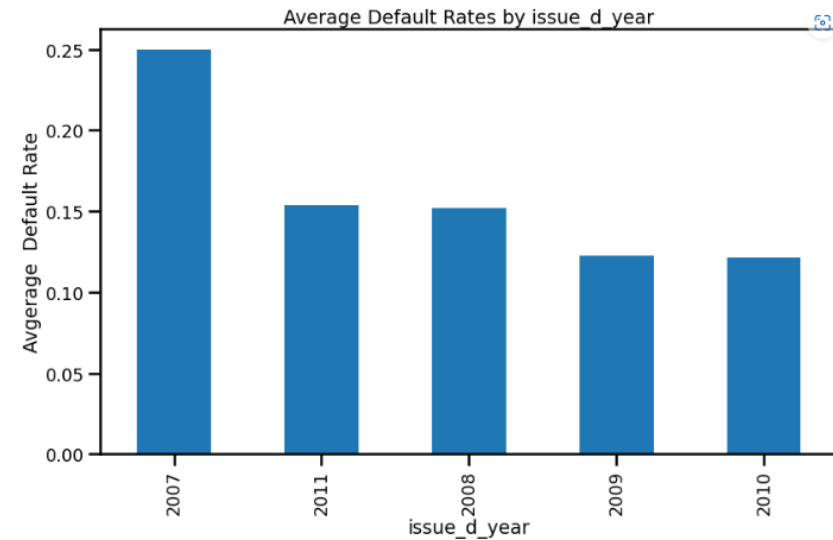


The loan status shows that the 60% of the loans amount is charged off and company is making 15% profit on fully paid

Segmented Univariate Analysis

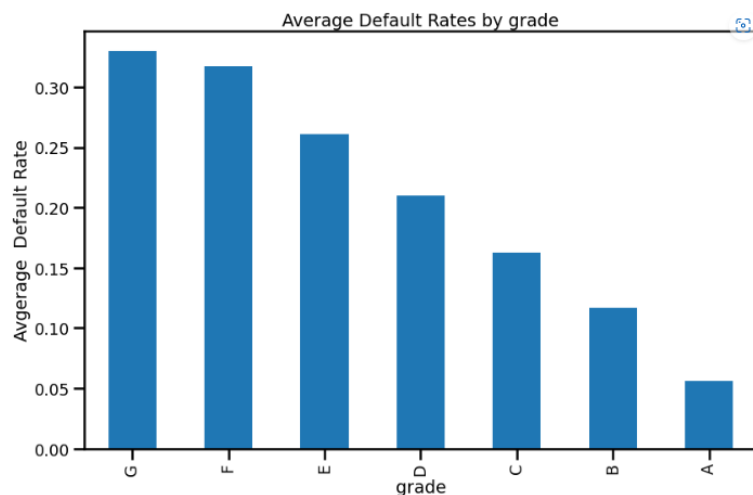


- This should be paid due attention by the business
- All others are in the range of 15 to 10 %. Loans taken for weddings and major purchases seems to have least defaulters

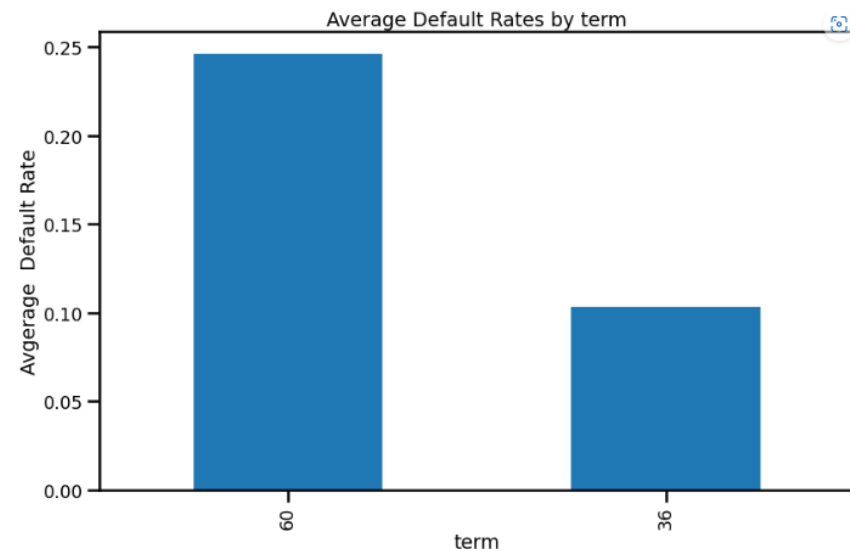


- In the year 2007 there were maximum loan default cases. The obvious reason for this can be global recession but this factor needs further investigation

Segmented Univariate Analysis

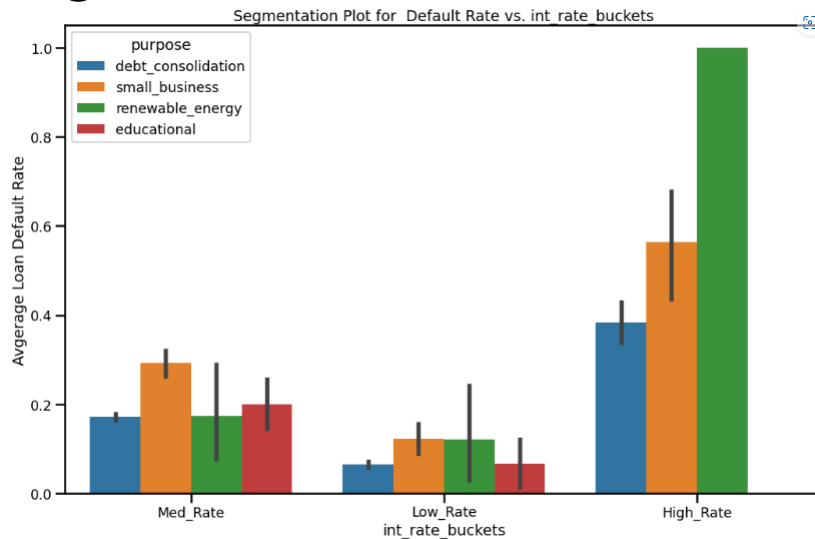


- Grades G and F contribute close to 30% of average default grades which are maximum. More attention from business is required to address this issue
- Grades A and B are minimum defaulter and contribute close between 5% to 10% of average default grades

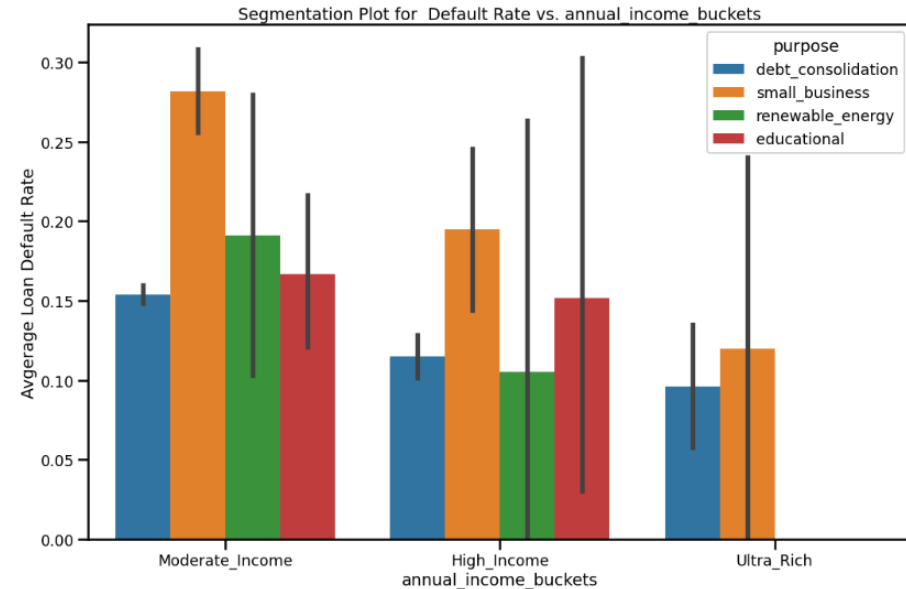


- 36 months loan term (short term) borrowers have minimum average default rate close to 10%
- 60 month loan term(long term) borrowers have major issue and average default rate close to 250%

Segmented Univariate Analysis

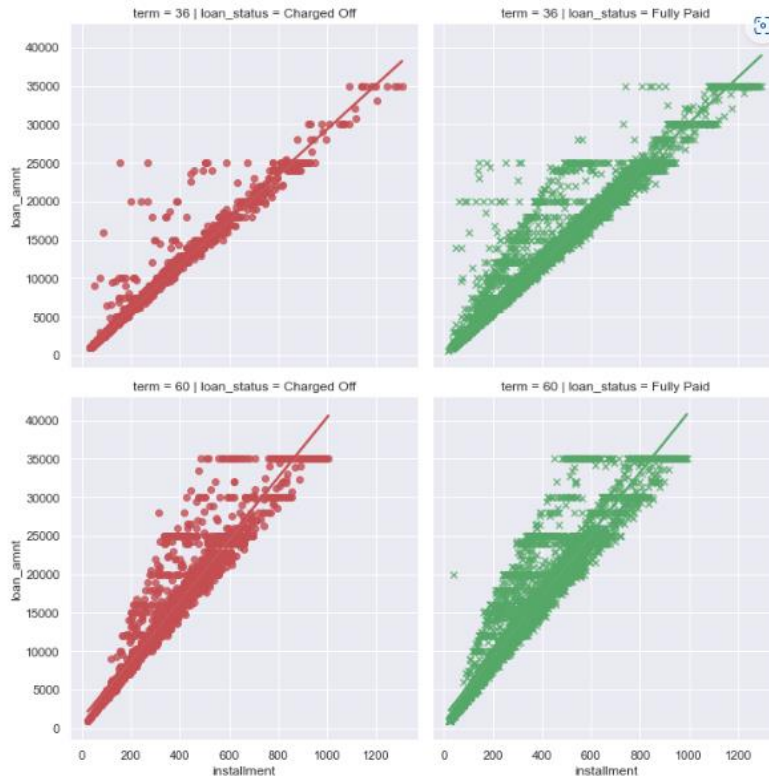


- For High interest rate buckets, Renewable Energy is the largest defaulter
- For Medium rate, Small Business is the highest defaulter _ For low interest rate, the number of defaulters are least

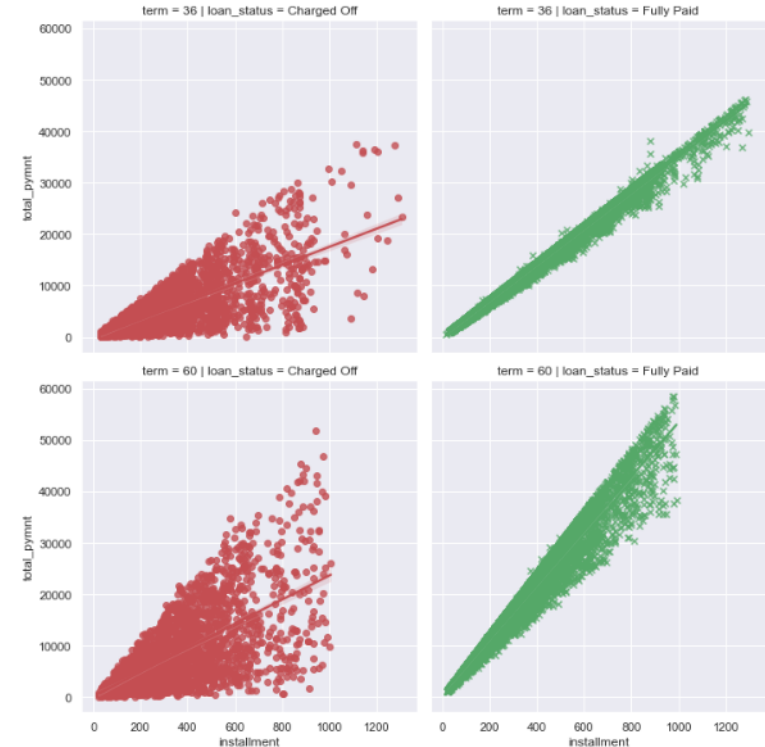


- For moderate income group , average loan default rate is highest for small business
- For Ultra rich income group also small business is the highest defaulter.

Bivariate Analysis- Scatter Plots

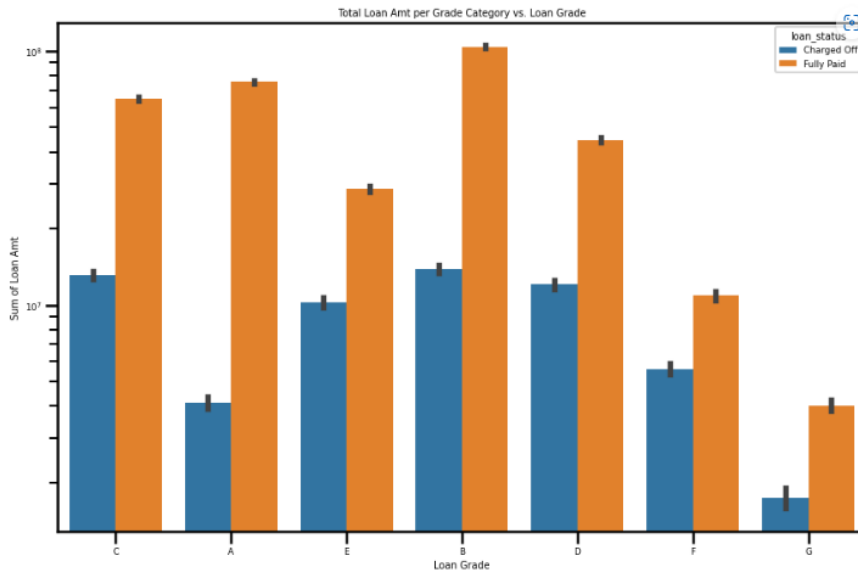


- Correlation matrix shows that relation between loan amount and installment is 0.93
- This average Ratio of fully paid for 60 months is more than 36 months

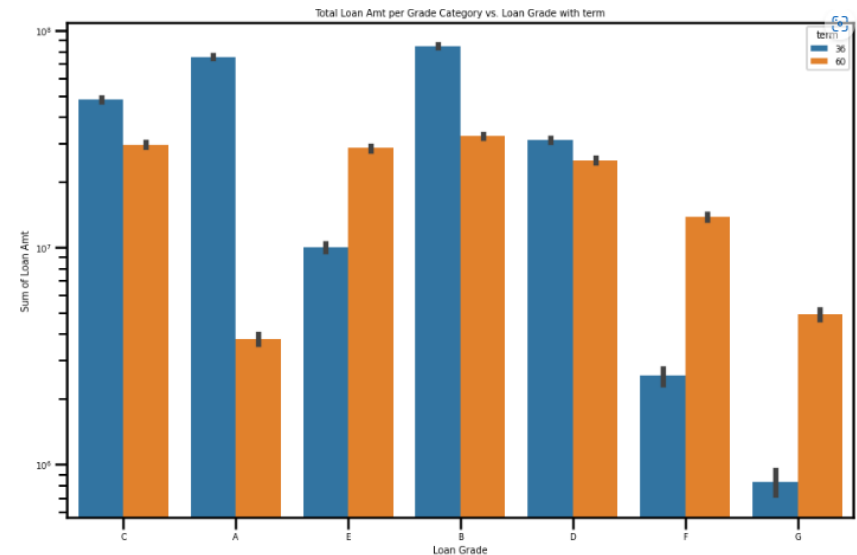


- Correlation matrix shows that relation between total loan payment and installment is 0.83
- This average Ratio of fully paid for 60 months is more than 36 months

Bivariate Analysis- Scatter Plots



- Highest Fully paid loans are in the category of A, B and C
- The grade G has highest risk for the business



- For the Grade B, the 36 term loan is highest. For grade B and C, the 60 months term loan is highest for sum of loan amount

Analyst's Finding and Recommendation

Driving factors behind loan default are as follows:

- **Annual Income** – Salary < 0.1 million USD has a higher risk of loan defaulter. Less annual income more default.
- **Loan Interest Rate** – Interest rate greater than 14% results into higher default rates. High interest rate, higher the risk of the default
- **Applicant Grade** – Grade A, low default, Grade G more default rate
- **Loan Term** – 36 months term have low risk and 60 months have high risk
- **Purpose of Loan** – Small Business pose the highest risk of default (25%), Renewable Energy(17%) and Education (15%)

The Company Portfolio and risk assessment should be considered based on the above factors

Major Findings

- The loan status shows that the 60% of the loans amount is charged_off and company is making 15% profit on fully paid
- Grades G and F contribute close to 30% of average default grades which are maximum. More attention from business is required to address this issue. Grades A and B are minimum defaulter and contribute close between 5% to 10% of average default grades
- 36 months loan term (short term) borrowers have minimum average default rate close to 10%. 60 month loan term(long term) borrowers have major issue and average default rate close to 250%
- Small business borrowers are most defaulter which is close to 25%. This should be paid due attention by the business. All others are in the range of 15 to 10 %. Loans taken for weddings and major purchases seems to have least defaulters
- Most defaulted Loans are those issued in December (15%).Loans issued in Feb ad March are defaulted lease (0 to 12%)
- Moderate income (income < 0.1 M USD) are the most defaulter (14%). Next higher defaulter category is High Income group (Income between 100000USD and 1000000USD)

Major Findings

- It is evident that borrowers with high dti (> 20) are most defaulter. Borrowers with low dti (< 10) are less defaulters. The defaulter range varies from 12% to 16%
- Clients with interest rate $< 10\%$ are safest and have only 5% default rate. As expected, clients with interest rate $> 20\%$ are the biggest risk to company and they contribute to 35% in the overall default list. Business should focus on moderate interest rate bucket which is in between these two categories so that they can be converted to low risk business
- Clients with medium loan amount (between 5000 to 15000 USD) and high interest rates ($> 20\%$) are 40% of the loan defaulter. Clients with high loan amount (> 15000 USD) and low interest rates ($< 10\%$) are lowest defaulters
- Walmart, UPS and AT&T are the top 3 loan defaulter with 30% and 20% default rates. USAF and Wells Fargo are last in top 10 defaulter list
- In 36 months term loan, the highest default rate is 20% for small business, while lowest is 10% for dept_ consideration. For 60 months term loan, the educational is highest defaulter (43%), followed by small business is next higher (40%)

Recommendation for the company

The company should focus on following pointers while giving out loans to lenders

- **High value loans** approval should be reduced
- **Long term loans** (60 months) loan approval should be reduced
- Loan for **person/ enterprices with less than 1 lakhs Salary** approval should be reduced
- Loan with a **high dti ratio approval** should be reduced
- Loans for **Higher Application Grade (F and G)** approval should be reduced
- Number of approvals where purpose is **small business** approval should be reduced