GRAMENER CASE STUDY ON LOANS

SUBMISSION

Prateek Pathak

Abstract

Problem Statement:

A consumer finance company specialized in lending various types of loans to urban customers, has been encountered with risky loans, which have resulted in credit or financial loss to the organization.

Objectives of the Analysis:

- Objective of the case study is to identify the driving factors behind loan defaults.
- The company wants to utilise this knowledge for its portfolio and risk assessment.

Solution Approach:

- R is used for Data cleaning from raw data sources.
- R is used to deriving new columns.
- R is used for univariate, bivariate and segmented analysis.
- R is used to plot the graphs, for better understanding of analysis.

Data Cleaning Approach

- Remove all the columns which have 'NA' or 0 values.
- Remove all the columns which have only 1 value throughout.
- Remove all the columns which are redundant for analysis.
- Converting 'Interest Rate' and 'Revolving Utilization percent' columns to Numeric.
- Derive new columns for further analysis.

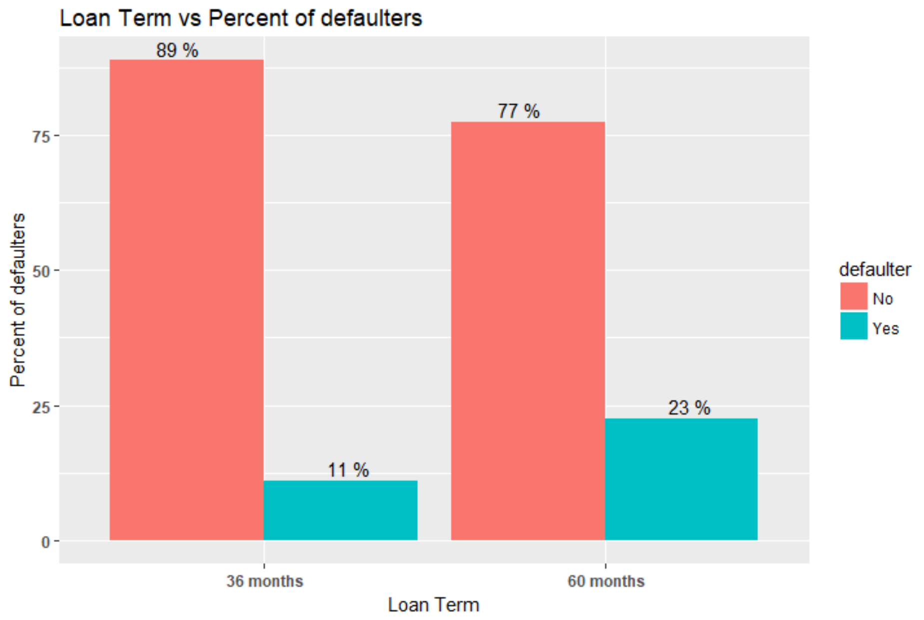
Data Description

1. Data contains **39717** unique employee loan data with **5627** charged off loans.
2. Loan is given for two term lengths **36 months** and **60 months**.
3. Employees are divided in Grade levels A to G and subgrades within grades. With A,B and C being 75% of loans.
4. Employee have experience ranging from **New (0) to more than 10+ years**. With most defaulters going to 10+ section.
5. Further employees have different home ownership like **RENT, MORTGAGE, OTHER** and **OWN**.
6. The LC data has 3 verification status for employee: **Verified, Source Verified** and **Not Verified.**
7. The loans have been issued from **2007** to **2011**. The no of loans have been doubling each year since 2007.
8. Among the various purpose stated by employees, the **LC has granted nearly 50% loans for Debt Consolidation**.
9. The maximum loans have been given in states of **CA, NY, FL and TX**.

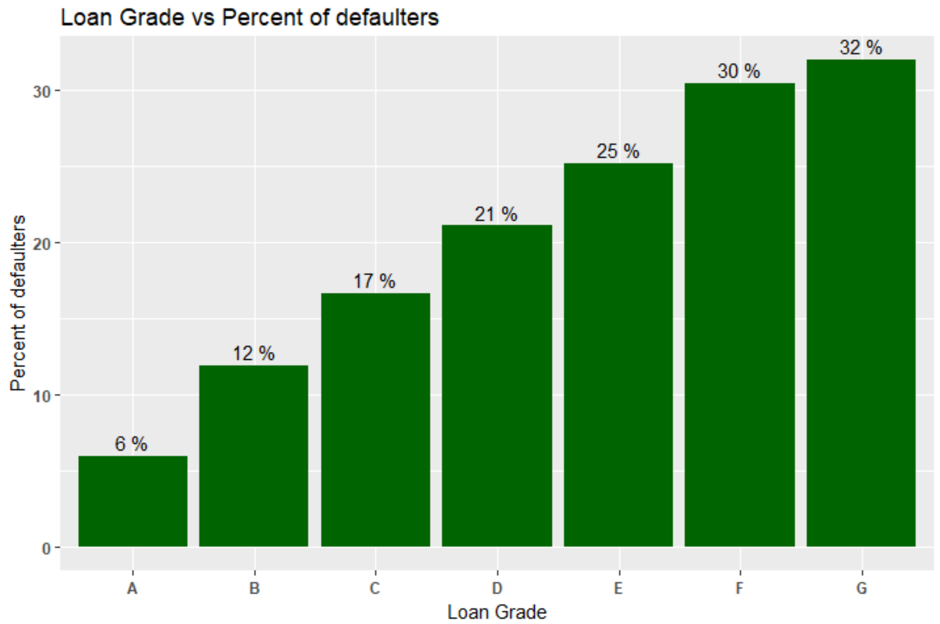
NOTE : After complete univariate and bivariate analysis following slides showcase the highlights. Rest of the plots and analysis summary can be looked up in R file.

Univariate Analysis – Exhibit 1

- Loan Term: Customers in 60 months loan term have a higher percentage of defaulting.

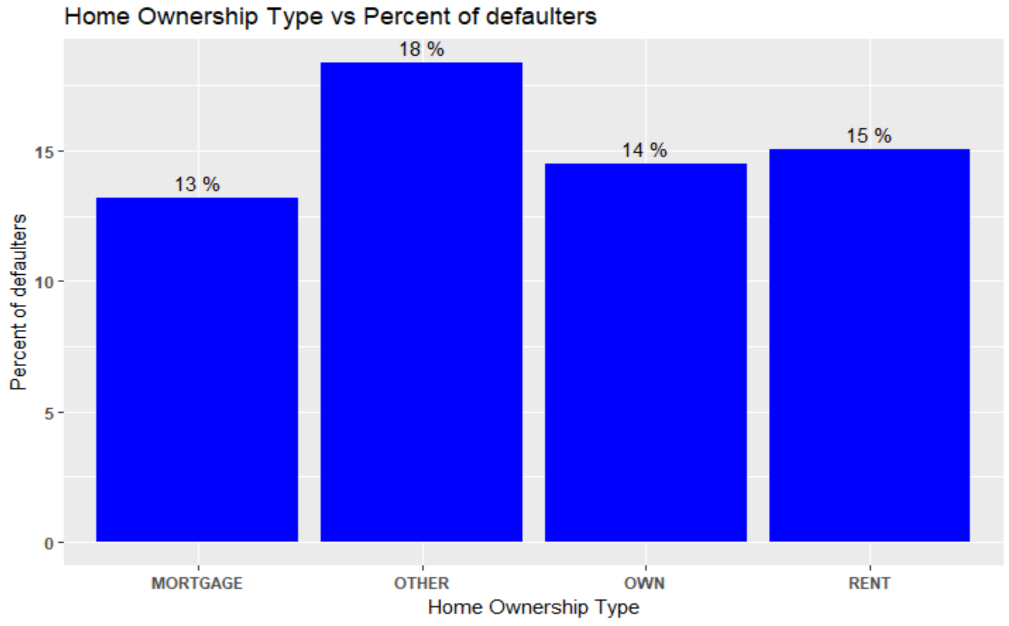


- Grades E, F, and G have the highest proportion of defaulters, making them the riskiest investment for Lending Club.

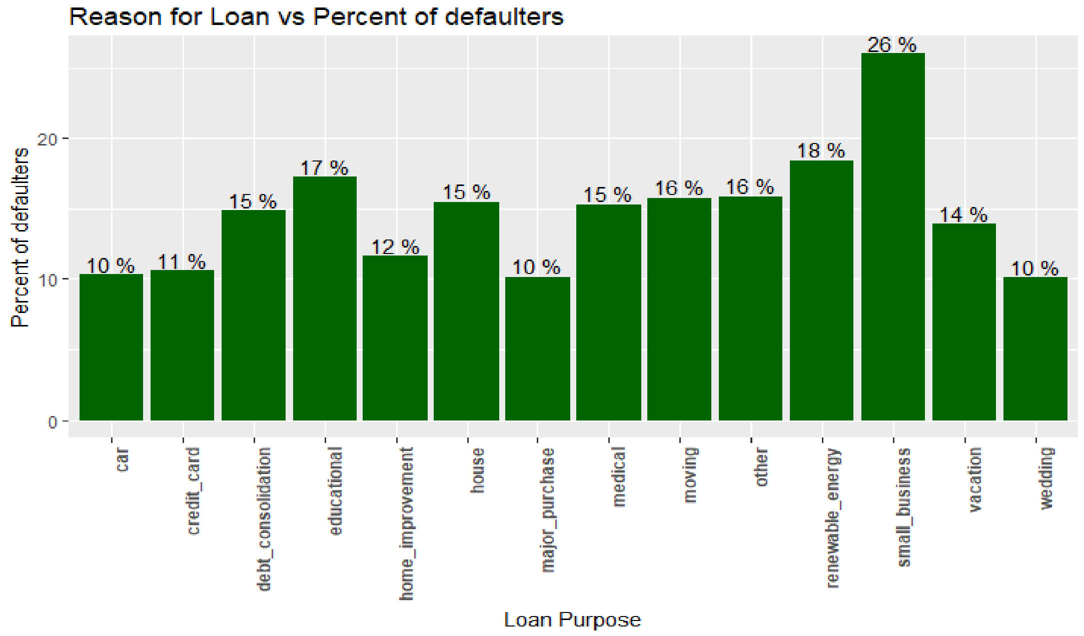


Univariate Analysis – Exhibit 2

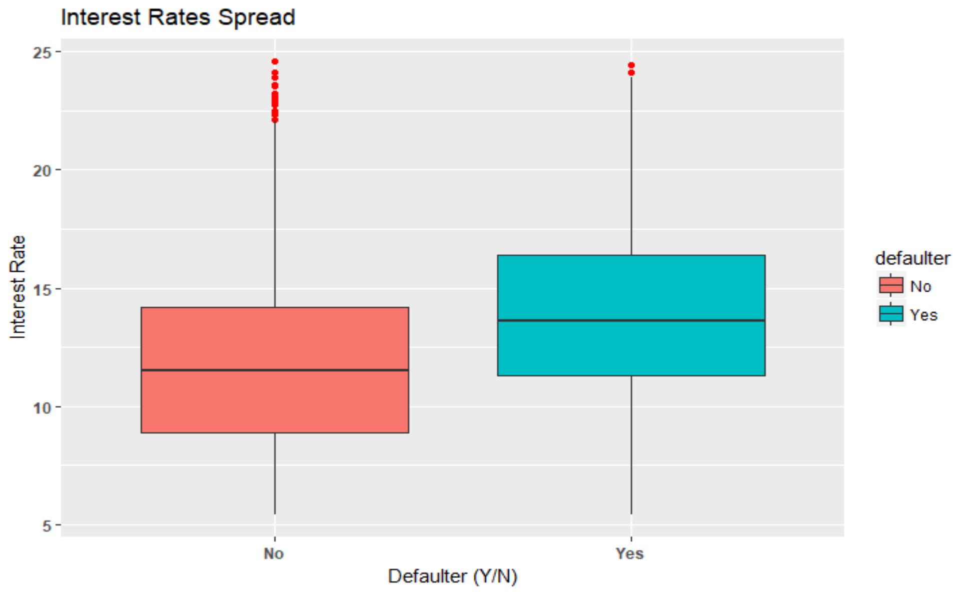
- Customers who do not have a home ownership information or have a rented house have a larger percentage of defaulting on loan.



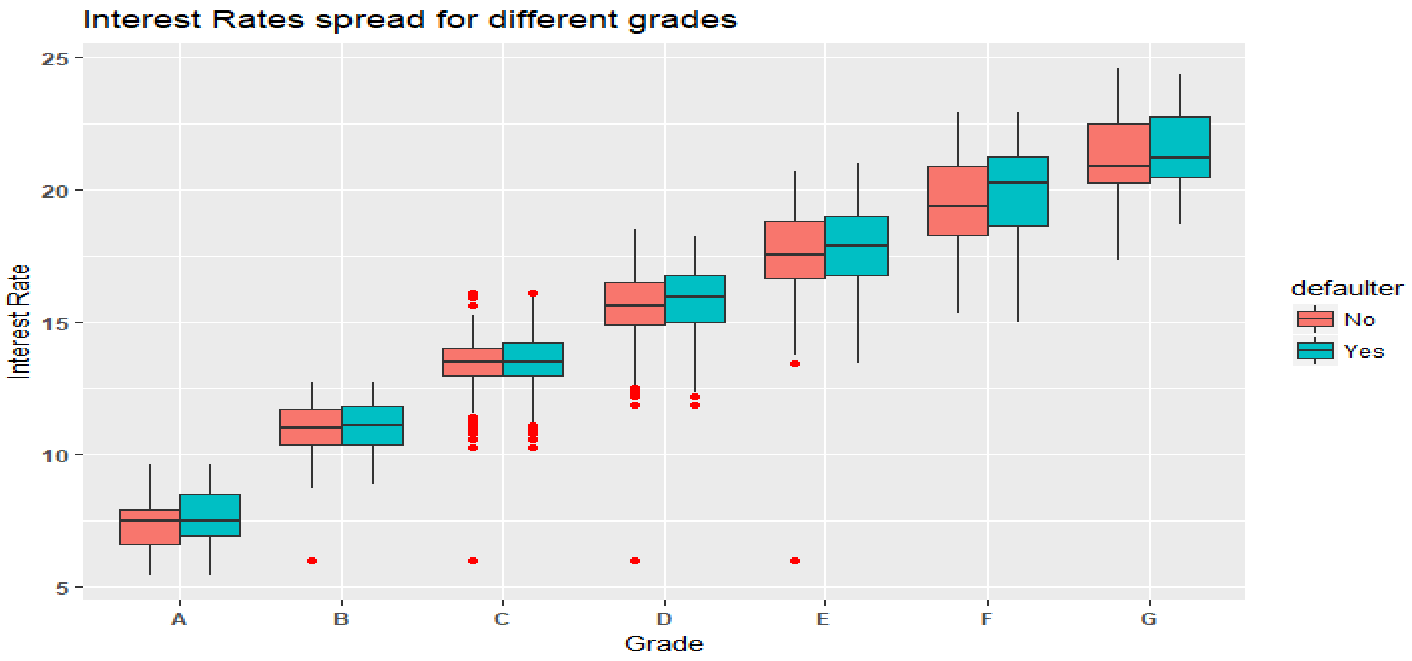
- Employees stating “Small Business” are the biggest defaulters with 26% defaulting rate.
- However the most number of defaulters are those stated “debt consolidation” and “credit card” as a purpose.



Univariate Analysis – (Continuous/Segmented)

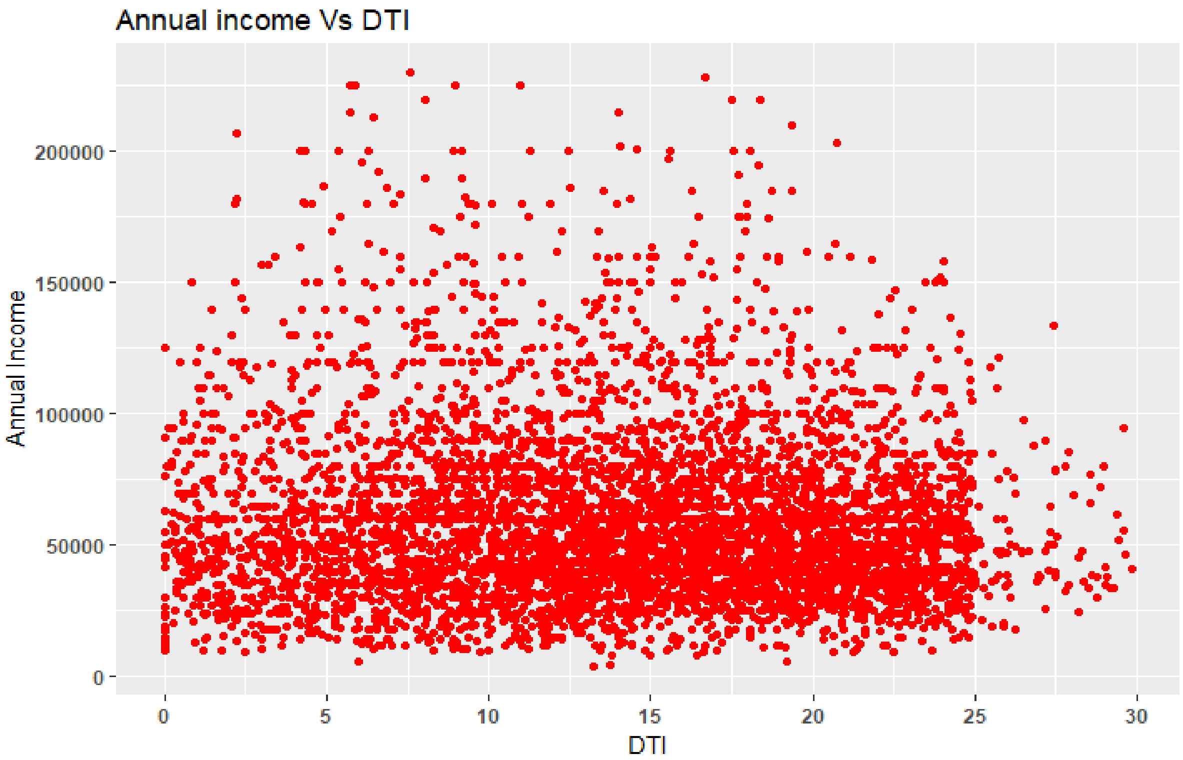


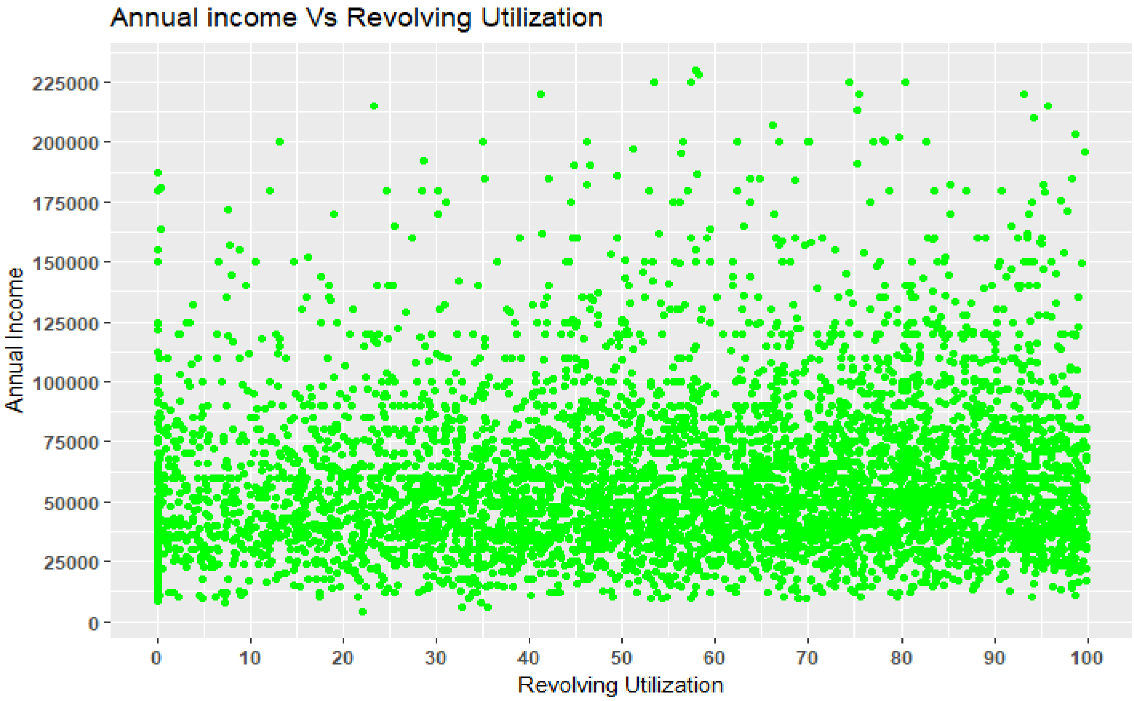
- Interest Rate: People with a higher interest rate have a higher risk of defaulting.
- This is also evident for grades E, F and G. The defaulters in these risky grades have a high median of interest rate.



Univariate Analysis - (Continuous/Segmented)

- Annual Income: In almost all cases employee with less annual income accompanied by a high debt to income ratio and revolving utilization are seen to default.

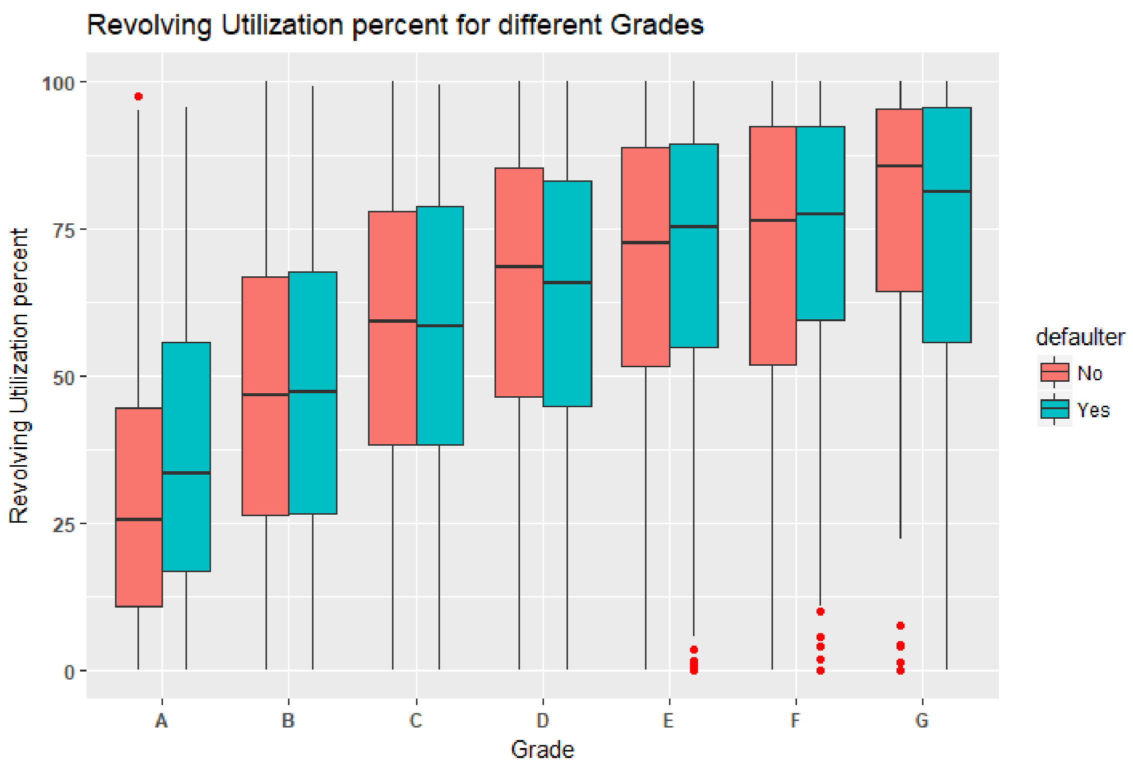


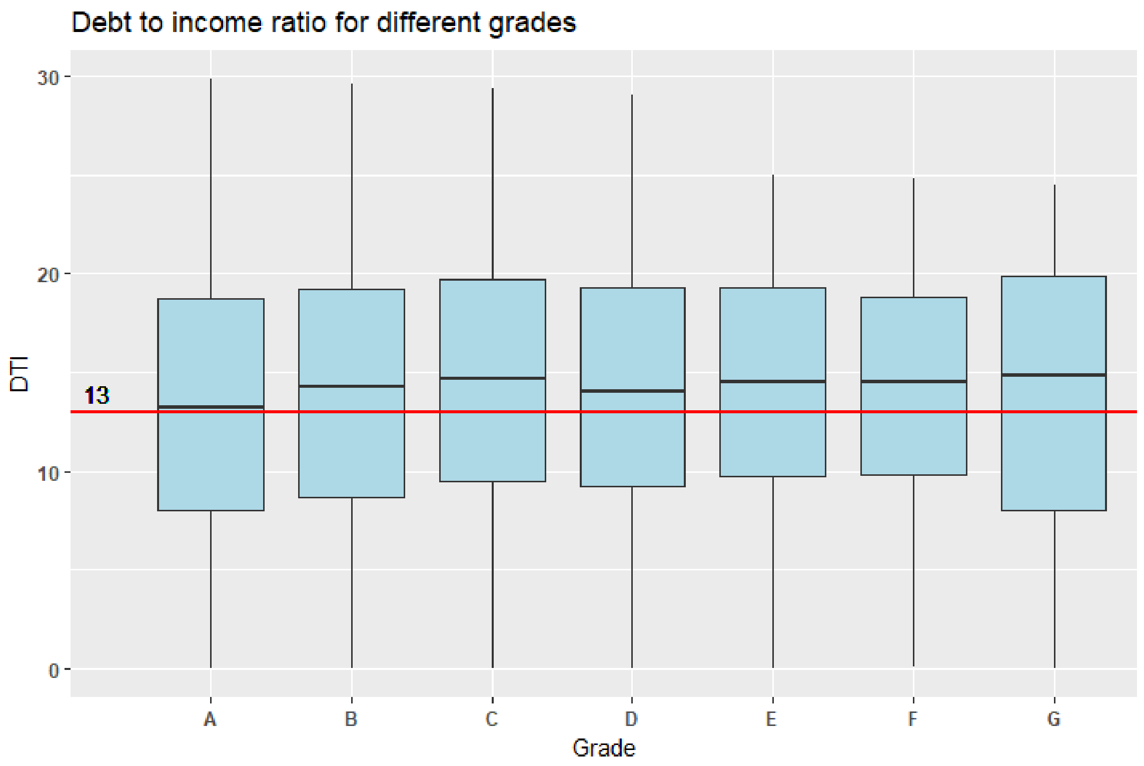


In both plots it is evident that the population of defaulters increases as the DTI or the Revolving utilization increases for low annual income group.

Univariate Analysis - (Continuous/Segmented)

- Effect Revolving Utilization on defaulters is also evident in case of riskier Grades E, F and G. In comparison to people who completed there loans, defaulters have high revolving utilization.





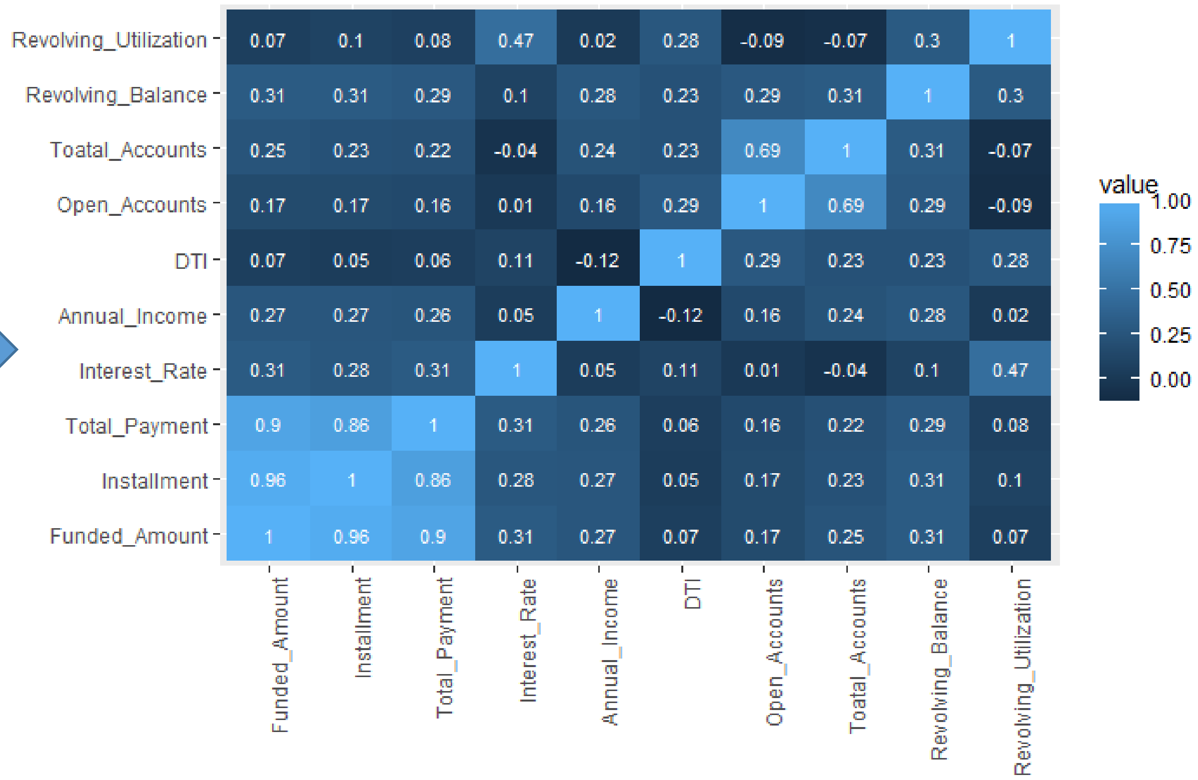
Debt to income ratio is also high in nearly all grades for defaulters.

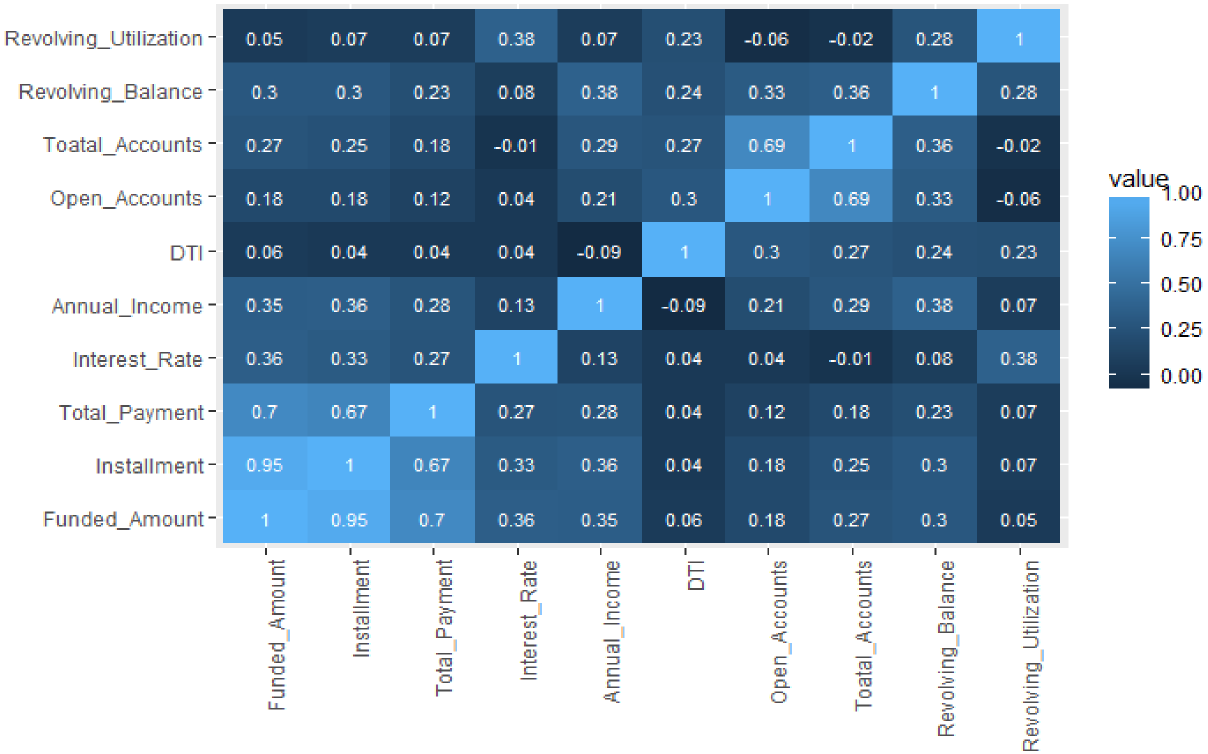
Bivariate Analysis

UpGrad

- Conclusions:

1. There is a high correlation between total payment and funded amount.
2. High correlation between funded amount and instalments.





1. However for defaulters the correlation between funded amount and total payment decreases as expected.

Conclusions Recommendations

Variables affecting Loan Default

1. Grades E, F and G are risky for giving loans. Especially when the employee has a high debt to income ratio.
2. Small Business as a purpose has a very high percent of defaulters hence becomes risk in loaning.
3. In all scenarios a combination of high DTI and a high revolving utilization along with low income points to an employee that has a high risk of defaulting.
4. Employees opting for a 60 months term have a high percent of defaulters.
5. Home ownership type OTHER also have a slightly higher percent of defaulters.