



## **LENDING CLUB CASE STUDY**

# **SUBMISSION**

**Presented By:**  
**Ishita Aggarwal**



## Background of Lending Club

---

### Background

Lending club is the largest peer-to-peer marketplace connecting borrowers with lenders. Borrowers apply through an online platform where they are assigned an internal score. Lenders will decide,

- 1) whether to lend or not
- 2) the terms of loan such as interest rate, monthly instalment, tenure etc.

Some popular products are credit card loans, debt consolidation loans, house loans, car loans etc.

### Business Objective

To identify variables which are strong indicators of defaulters and potentially use the insights in approval / rejection for their decision making.

---



## Data Exploration and Understanding

There are 111 attributes in Data set. So we have to segregate the attributes and group into different categories before starting our analysis.

Types of attributes:

1. Customer attributes
2. Loan attributes

### Customer attributes

Employment Length

Employment Title

Annual income

Zip code

Description

### Loan attributes

Grade

Loan Amount

Interest Rate

Purpose

Funded amount

Funded amount investment

Home Ownership



## Data Cleaning and Manipulation

---

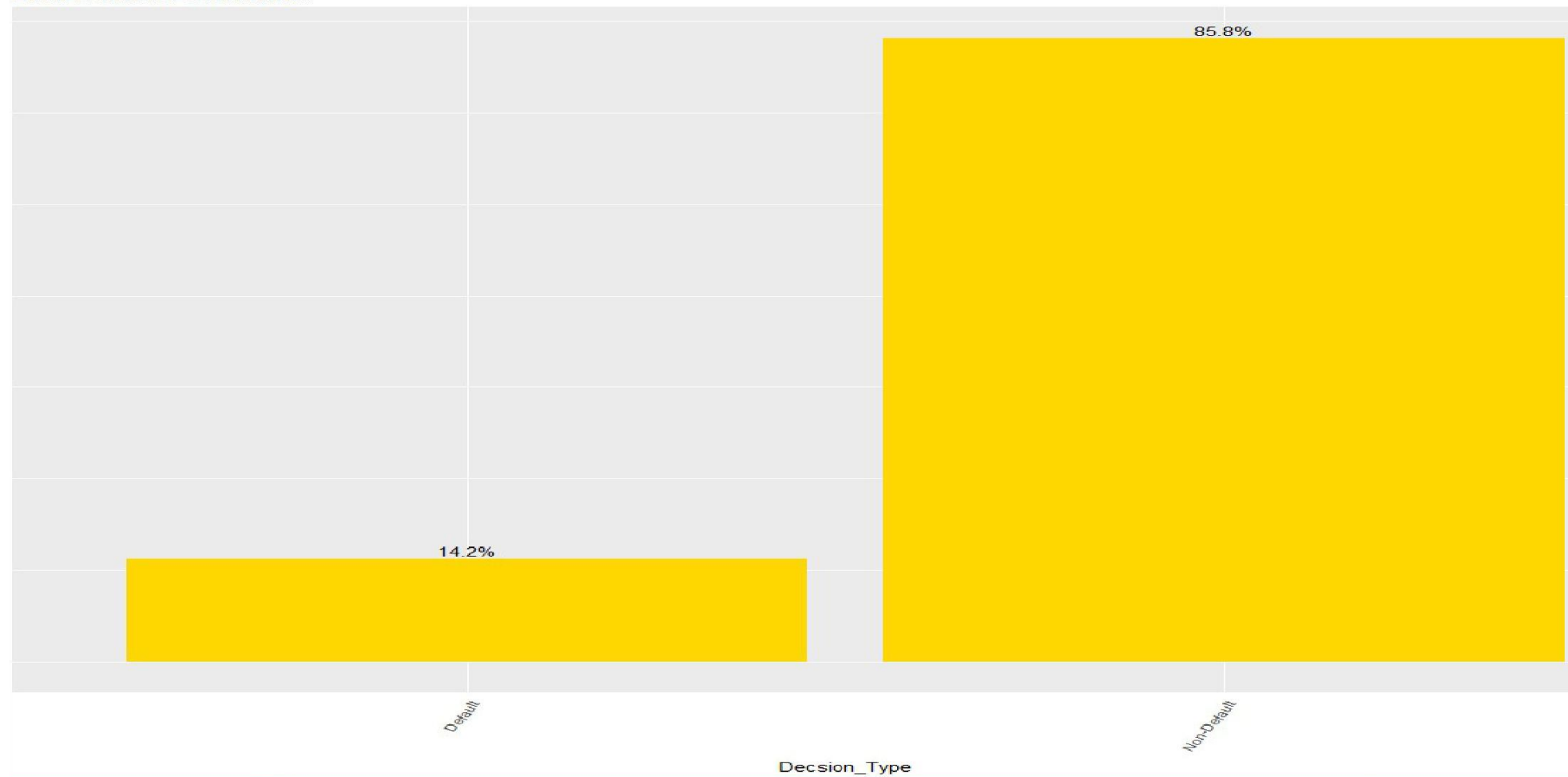
### Missing and NA values:

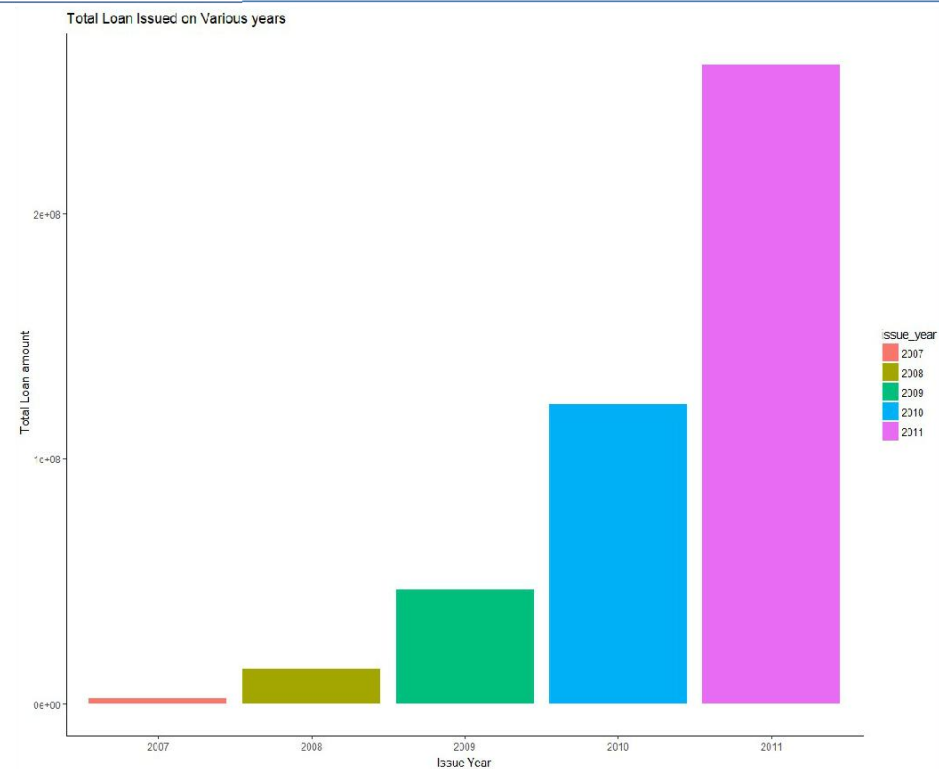
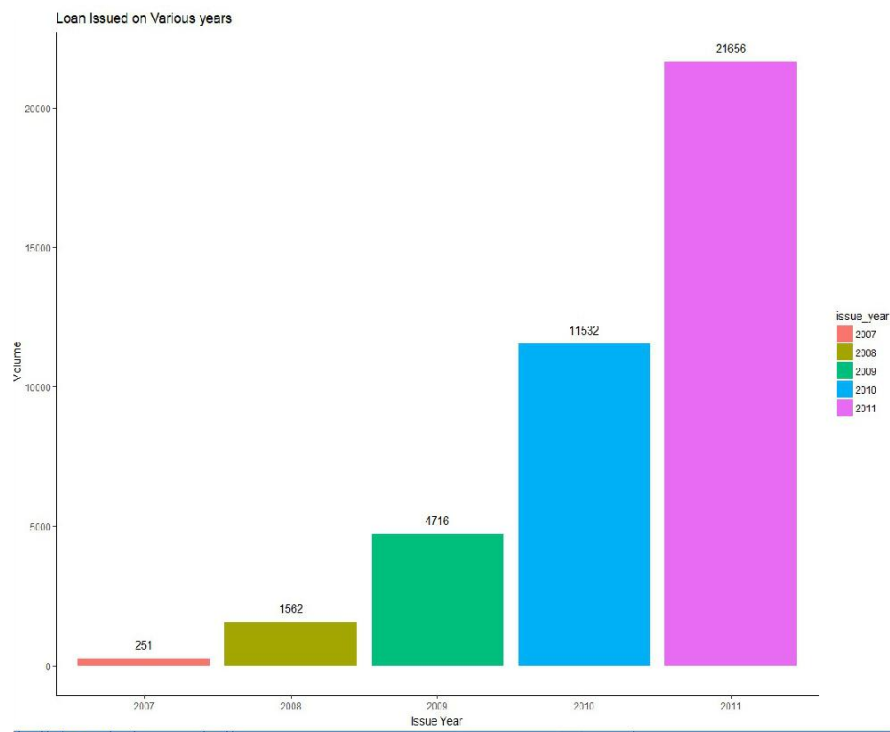
1. There are 56 attributes with all observations as NA has been removed.
2. There are 11 attributes which contain same values across all observations and they are removed.
3. There are some columns which are imputed by 0's for NA Values.
4. There are some columns where NA values has been imputed with median value.

### Data cleaning & others:

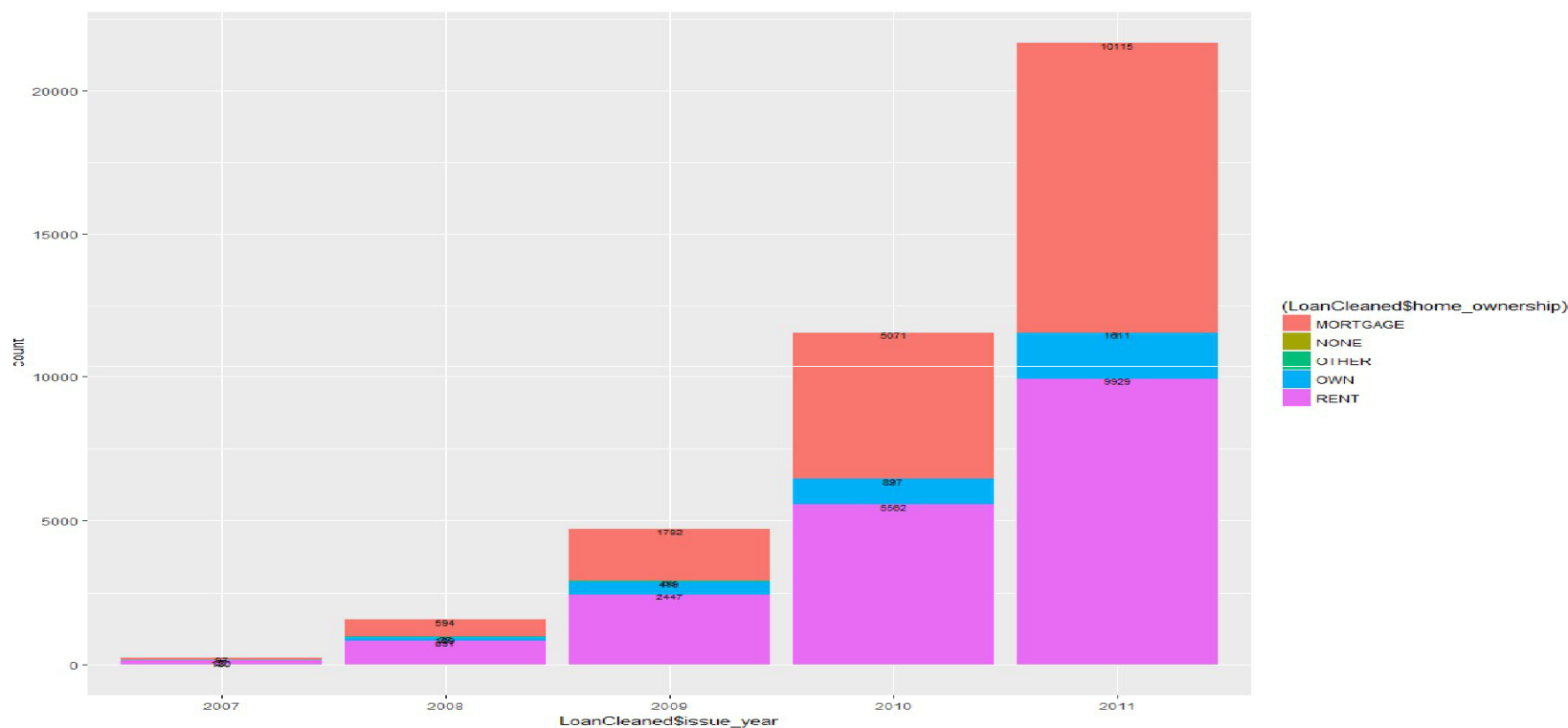
1. Some attributes like "interest rate" where "%" sign is removed.
  2. There are attributes like term(tenure) where the word "months" got removed.
  3. Converted few attributes into proper date format.
  4. Converted few categorical attributes as factors.
-

Loan Defaulter Distribution





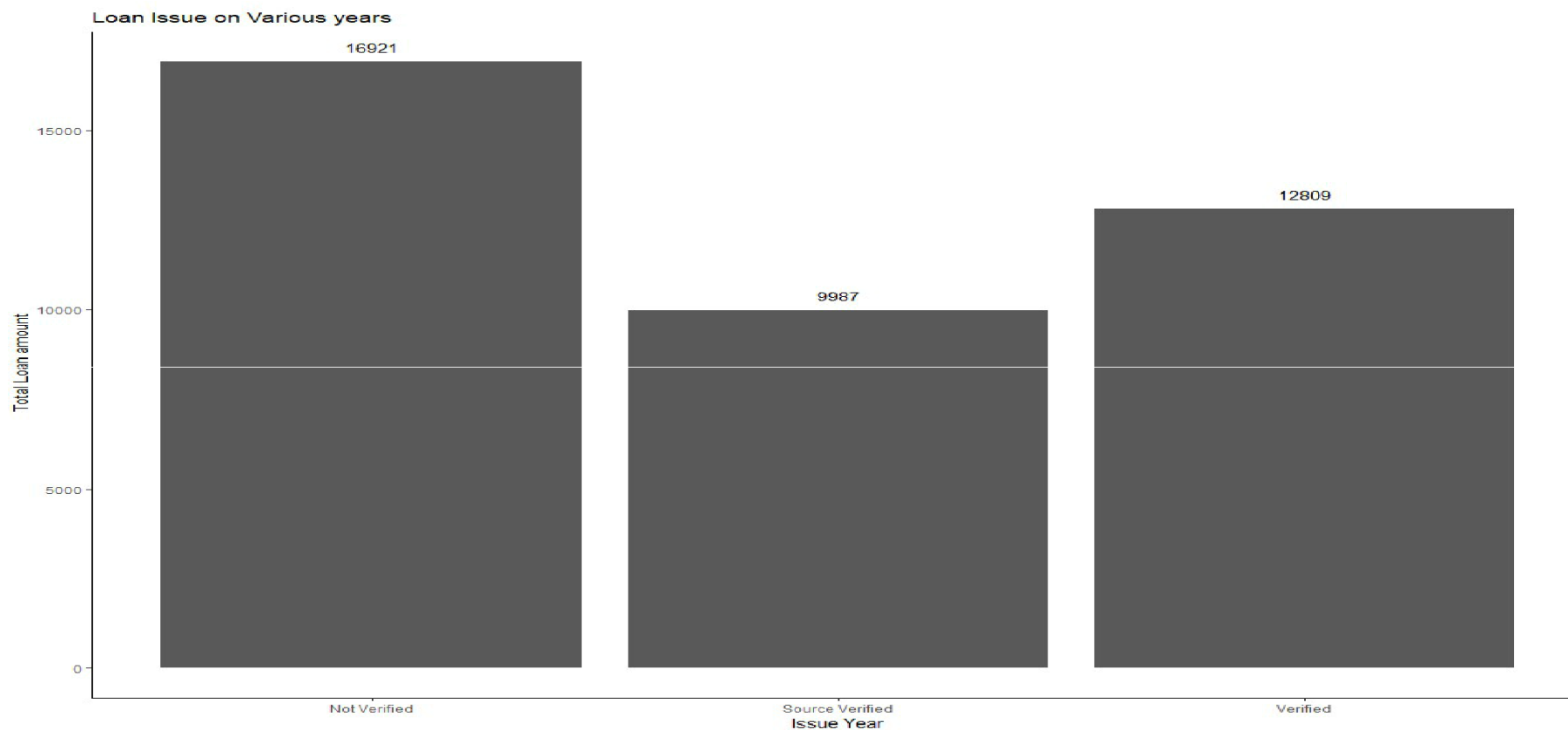
**It is evident that lending club started getting more customers and their business increased in 2011.**



People who have got home ownership of “Mortgage” and “Rent” approaching lending club got increasing over the years from 2008-2011.



## Univariate analysis on Verification Status



The Number of “Not verified” customer base is huge in comparison with other verification status. So in order to increase their competition in #market , they are taking risk by issuing lending to not verified profiles



Purposes of Loans



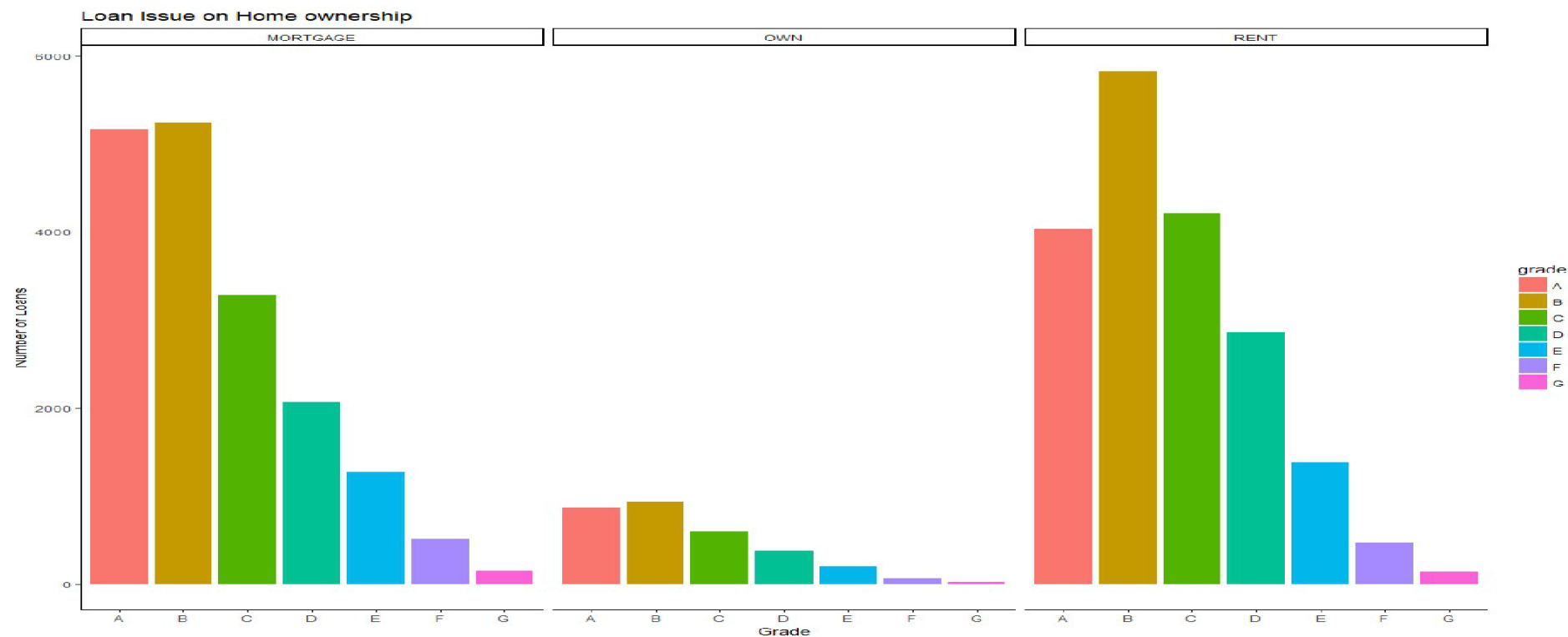
From the treemap we can easily say the top 3 reasons of customer taking loans.

1. Debt Consolidation
2. Credit Card
3. Other





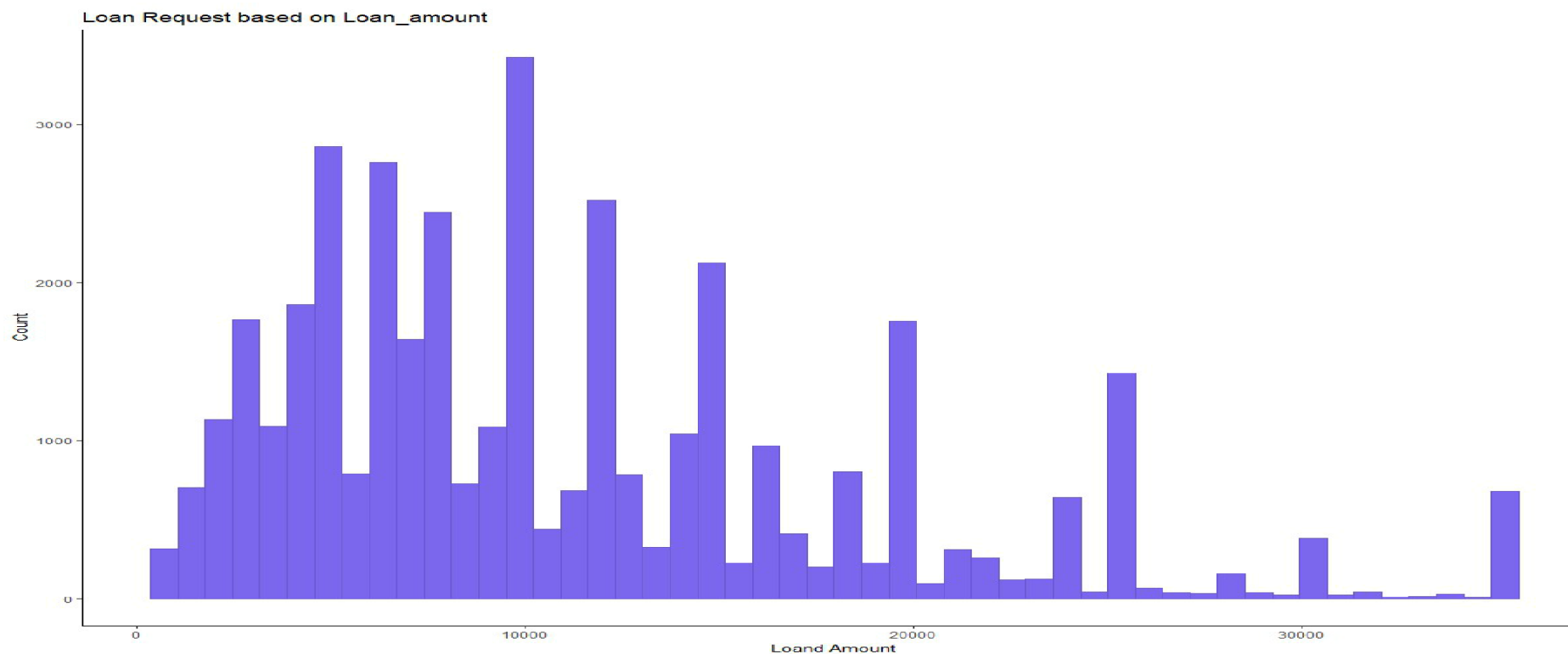
# Univariate analysis of Grade based on Home Ownership



Customers who are from B grade and belongs to the home ownership category of Rent and mortgage are approaching more in numbers to lending club. This is a key data driven metrics for Lending Club.

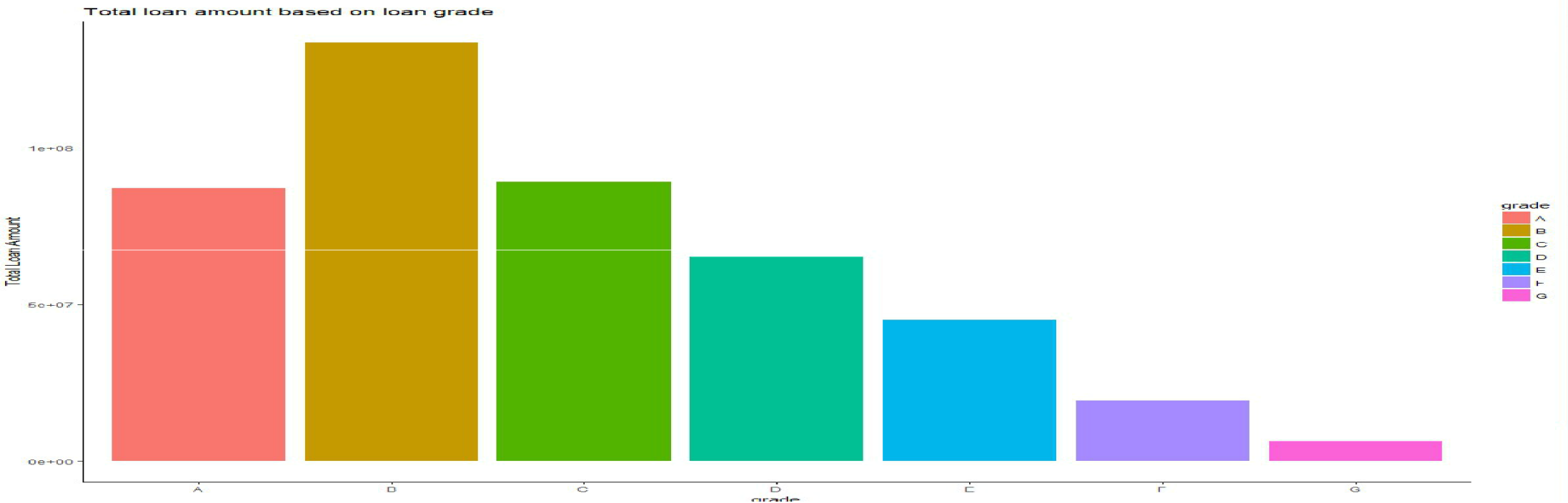


## Loan Request based on Loan amount



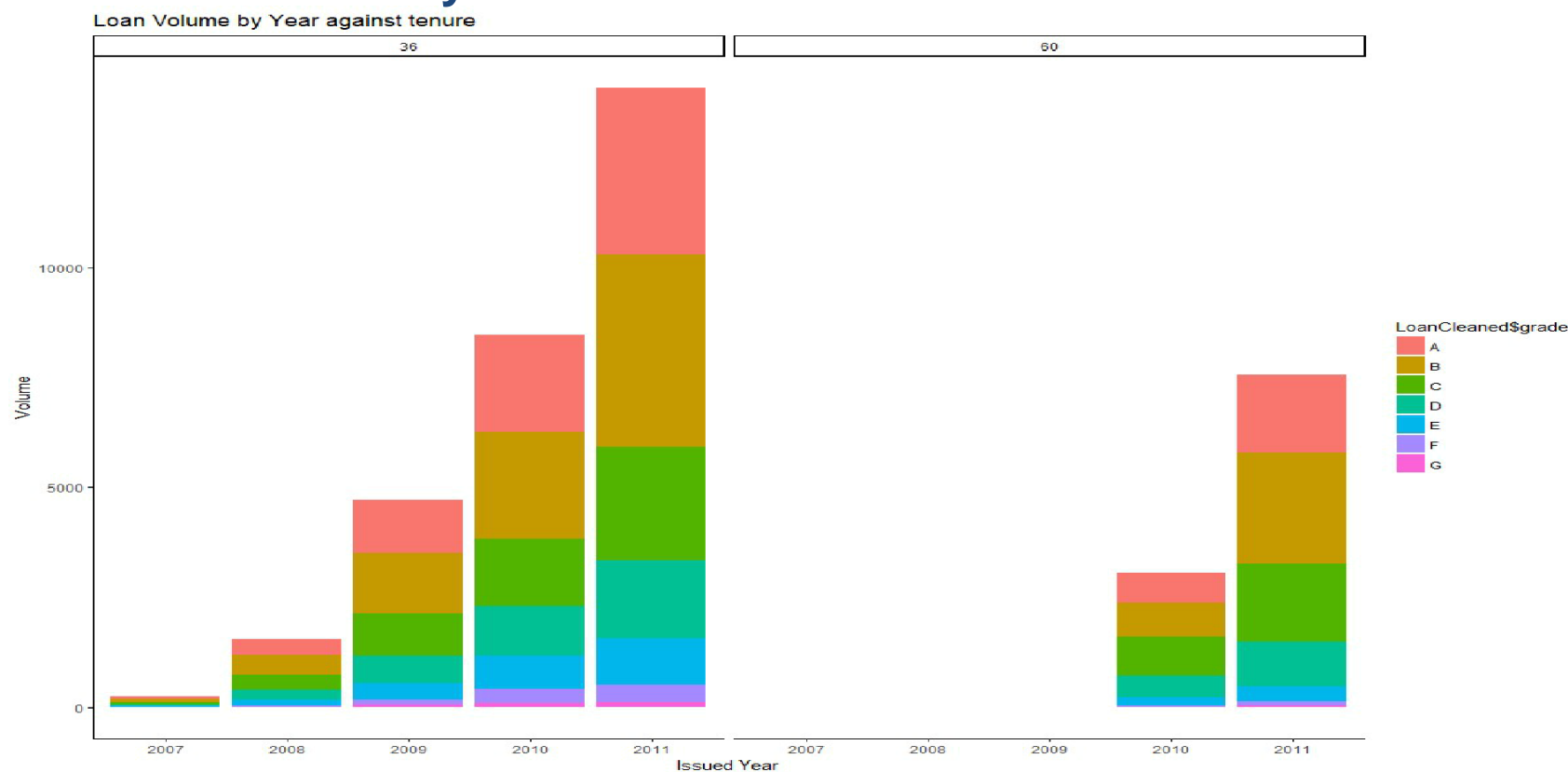
From the histogram we can say that more number of Loan request for Loan amount of 9000 to 12000 dollars

# Analysis on Loan Grade



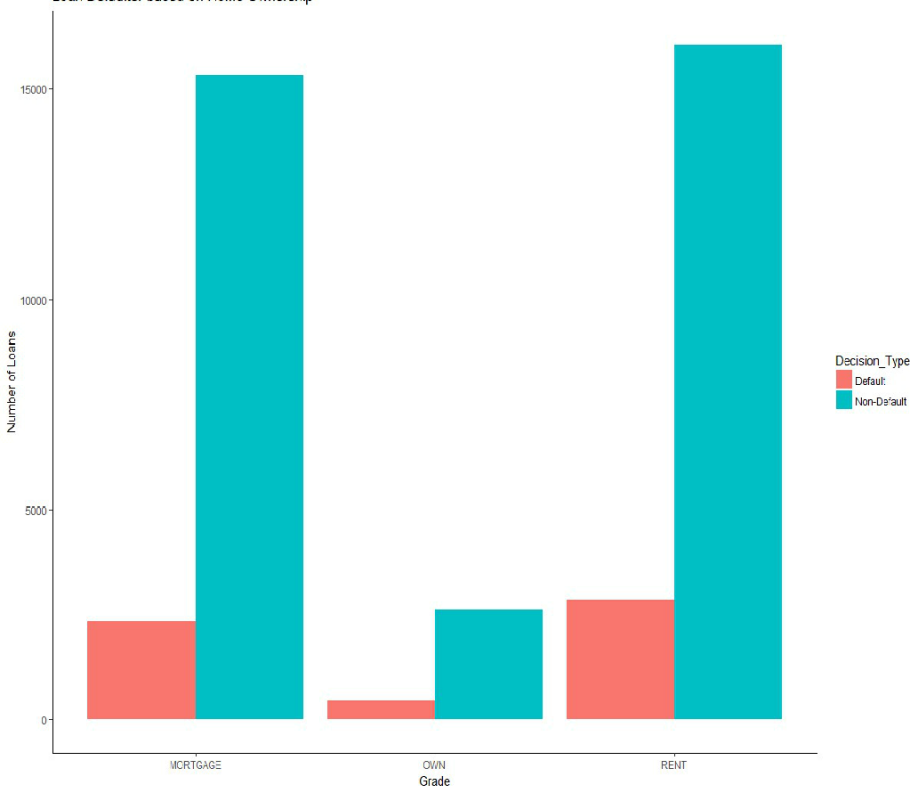
**Clearly customers who are belongs to grade B are more in demand of Loan from Lending Club , So it is a data driven metric for Lending club.**

## Loan Volume analysis for Grade based on Tenure



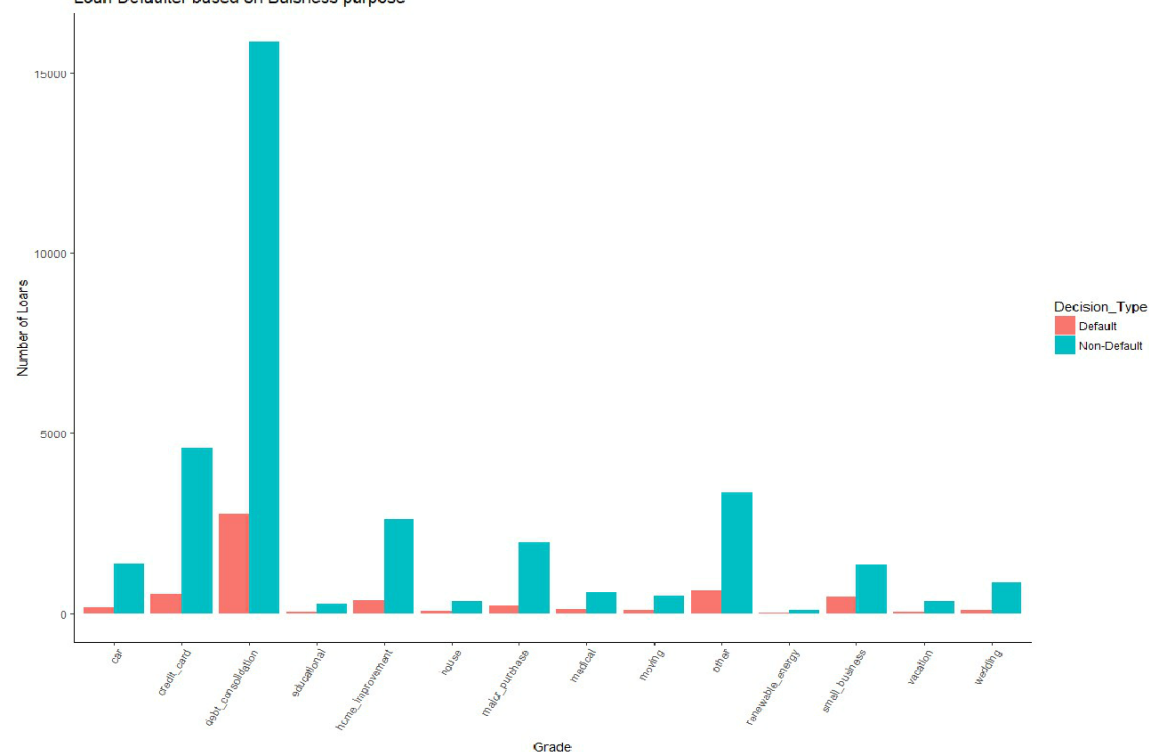
Clearly there is a increase demand for a short tenure of 36 months than 60 months across all grades.

Loan Defaulter based on Home Ownership



- Defaulters are more under home ownership of Rent and Mortgage

Loan Defaulter based on Business purpose

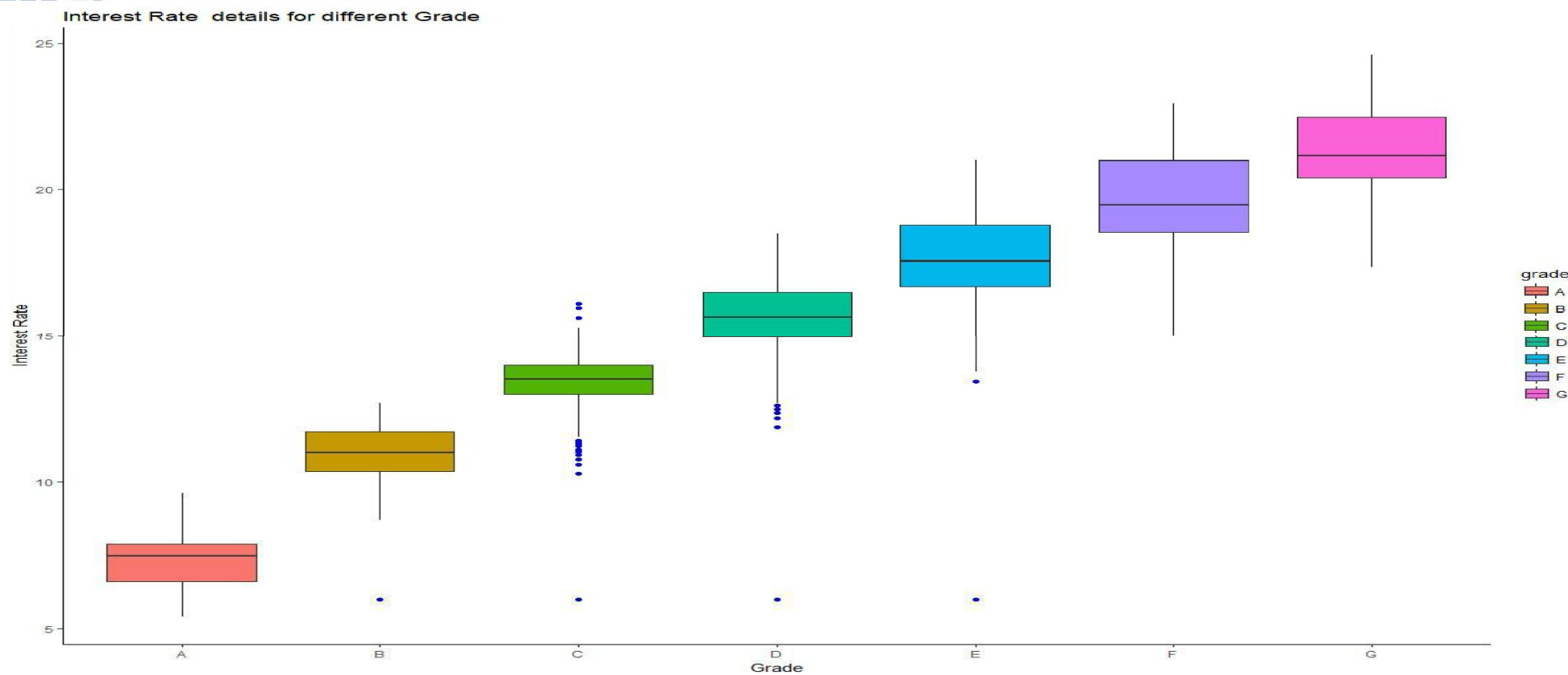


Defaulters are more under business purpose of debt consolidation, credit cards and other



## Interest Rate Analysis based on grades

UpGrad

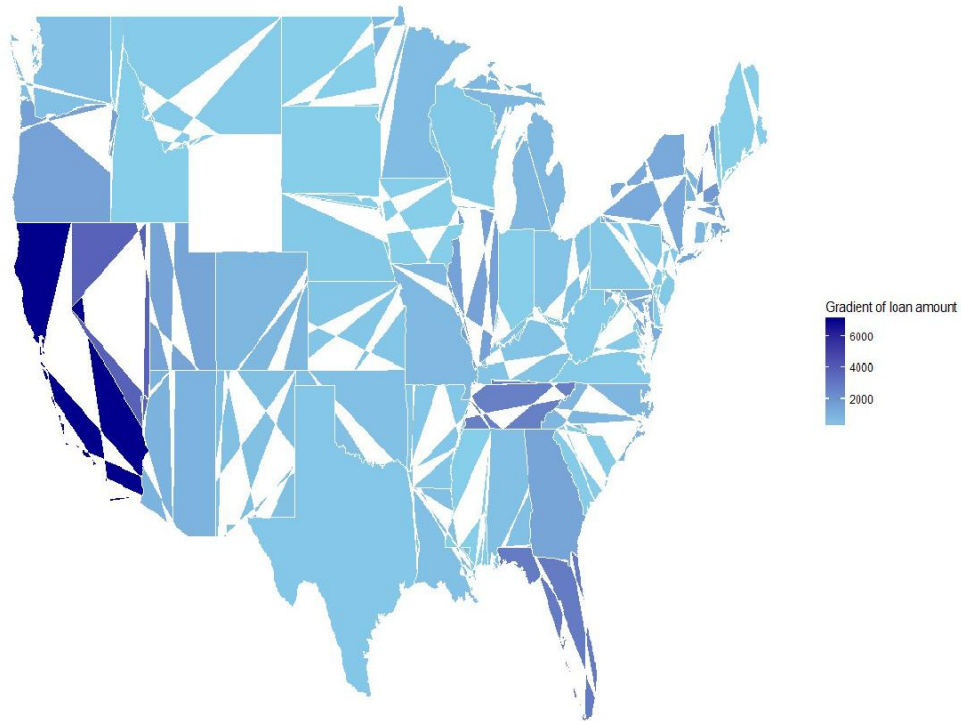


Clearly interest rate are very less for Customer who belongs to A, B and C

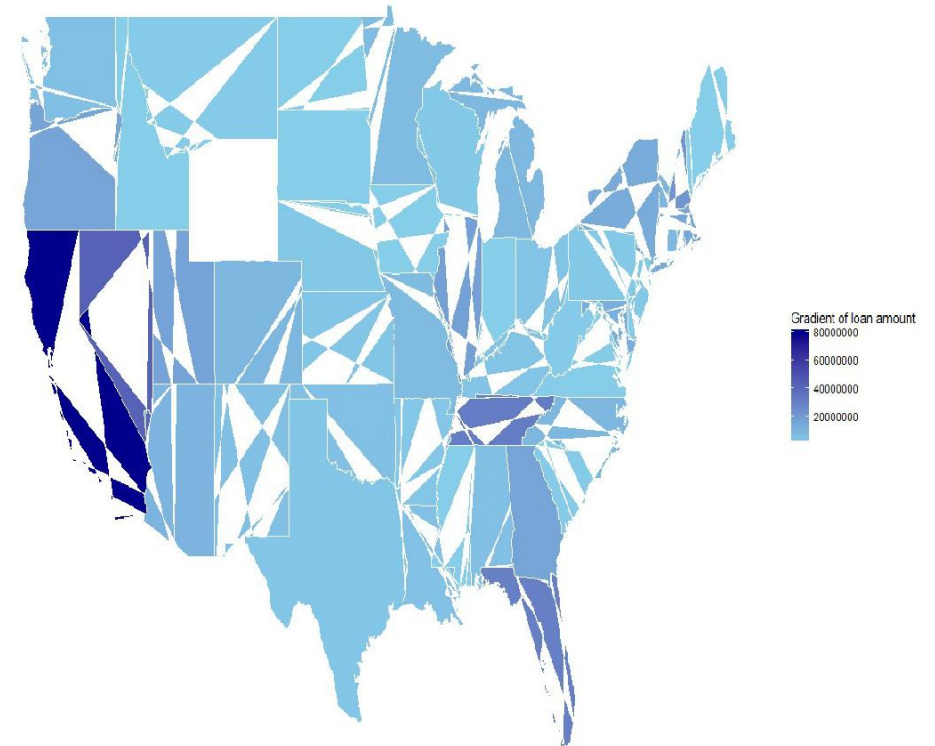


## Loan Volume and Total Loan amount analysis based on State

Heat Map of loan amount volume in all states



Heat Map of total loan amount in all states



**From both the maps, it is clear that loan amount and loan volume lending is high in California**

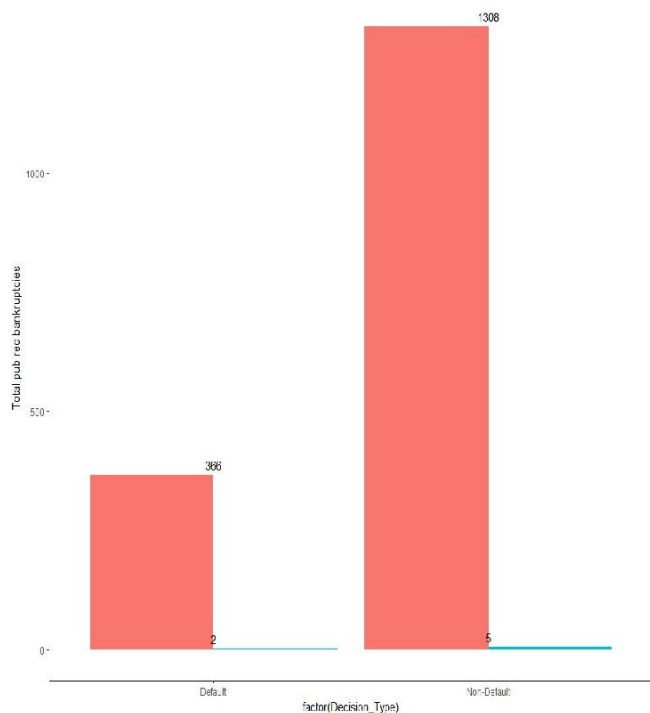




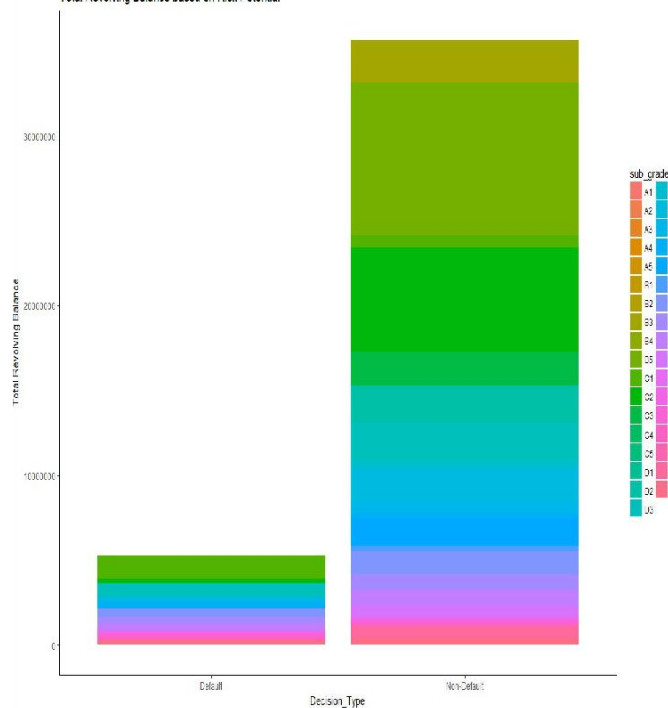
# Defaulters Analysis Based on Debt income, public record bankruptcy and revolving balance

UpGrad

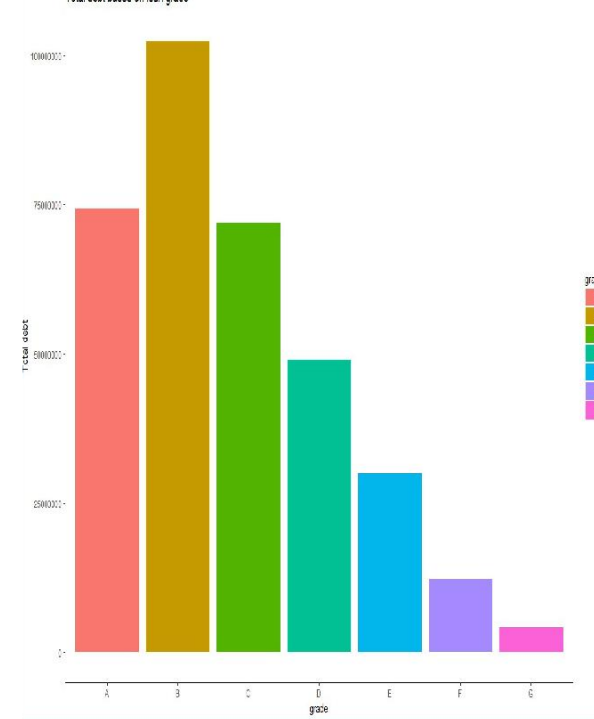
Total pub rec bankruptcies based on Decision Type



Total Revolving Balance based on Risk Potential

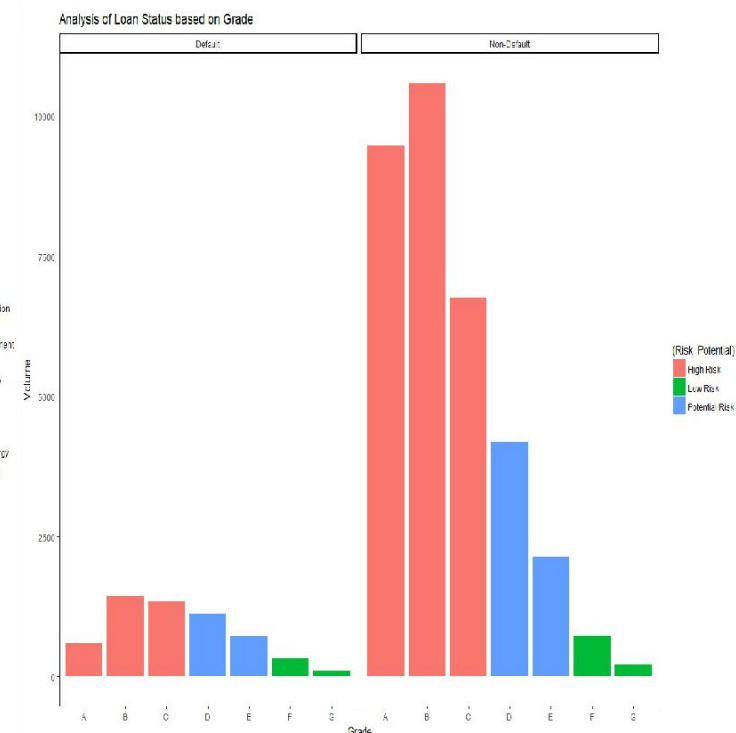
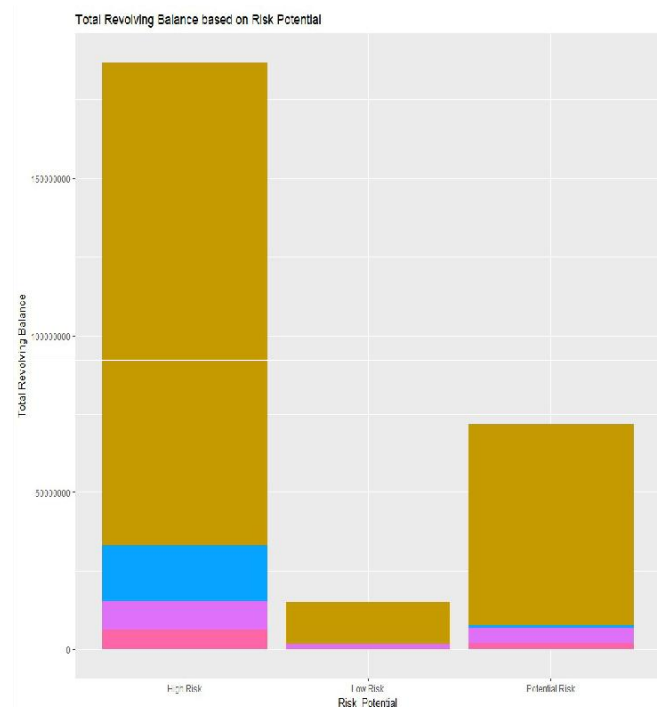
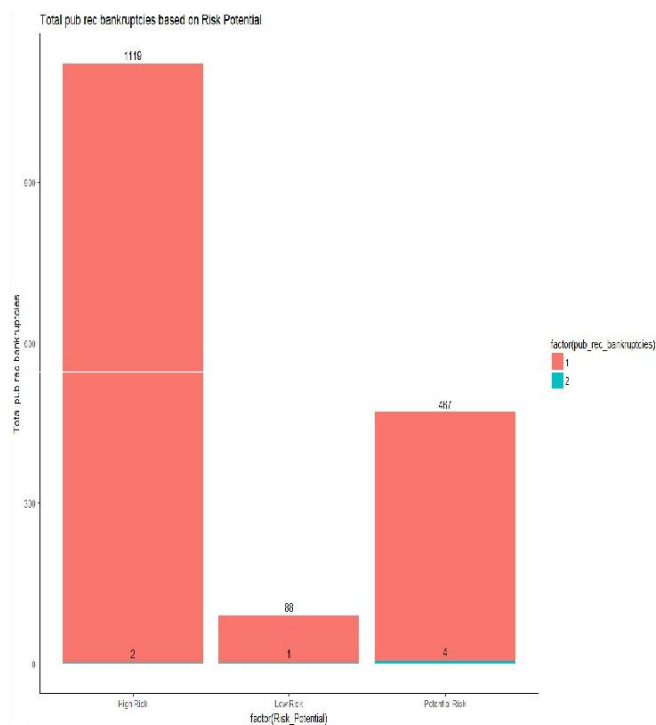


Total debt based on loan grade



Debt income is more for people belong to B grade .

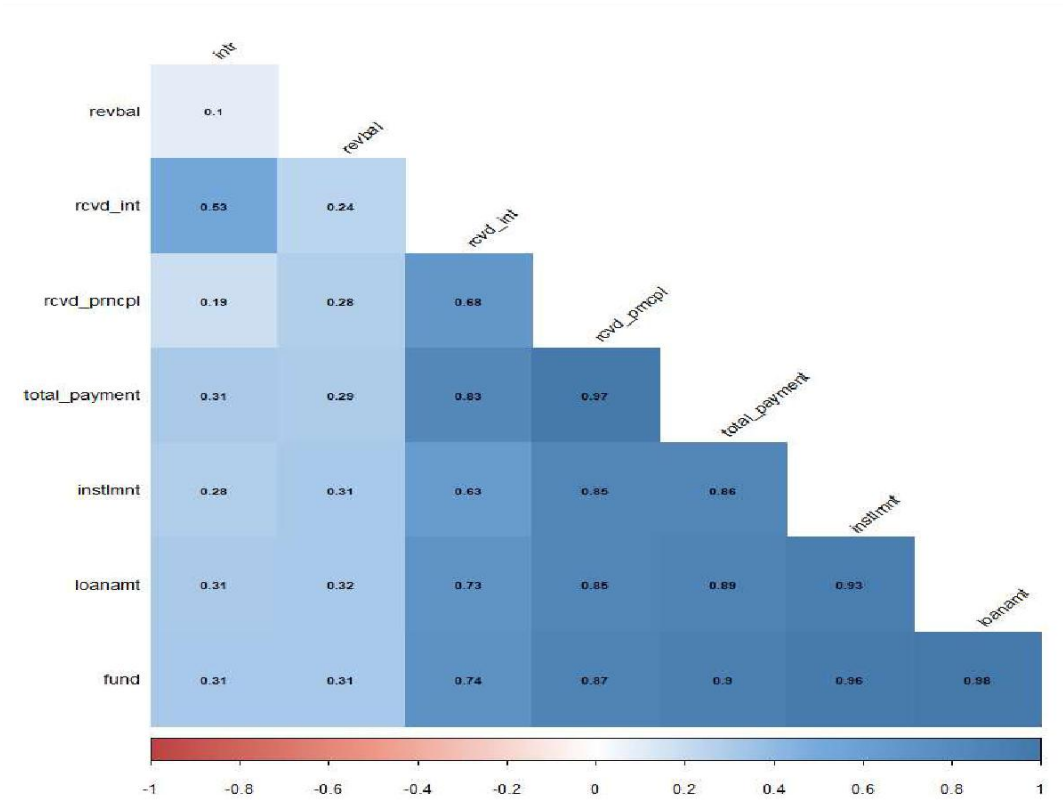
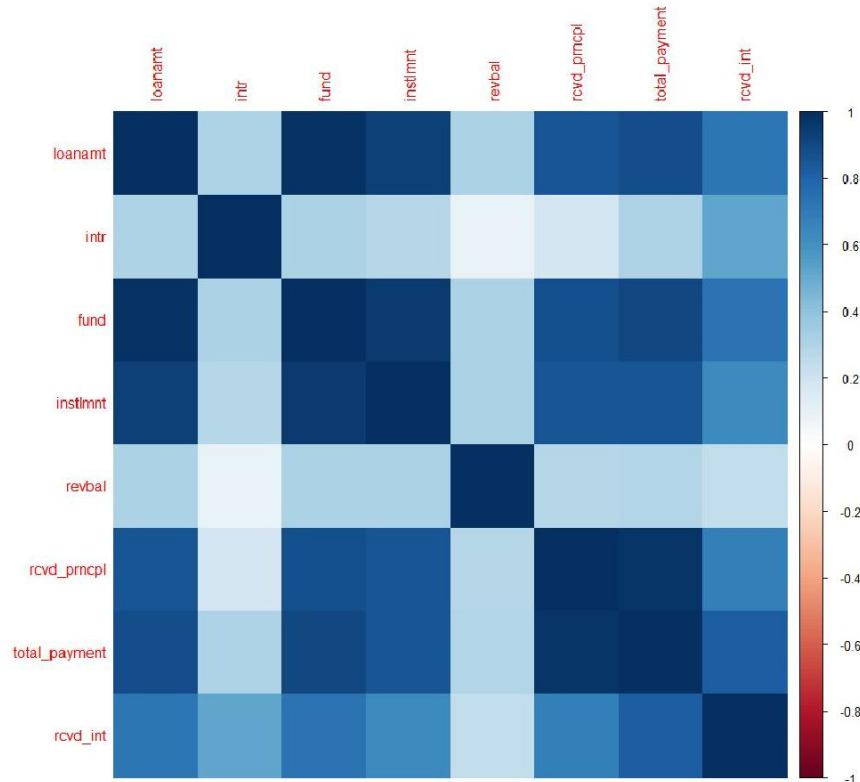
# Risk Potential Analysis based on revolving bal, Grade and bankruptcy records



These plots tells risk category of their customer which is categorized as High Risk, Low risk and potential Risk.

1. B & C grade have got High risk to default.

2. Total revolving balance amount is high for people who are taking loans for debt consolidation got high risk.



**We see that there is a positive correlation loan amount and fund\_amt.**

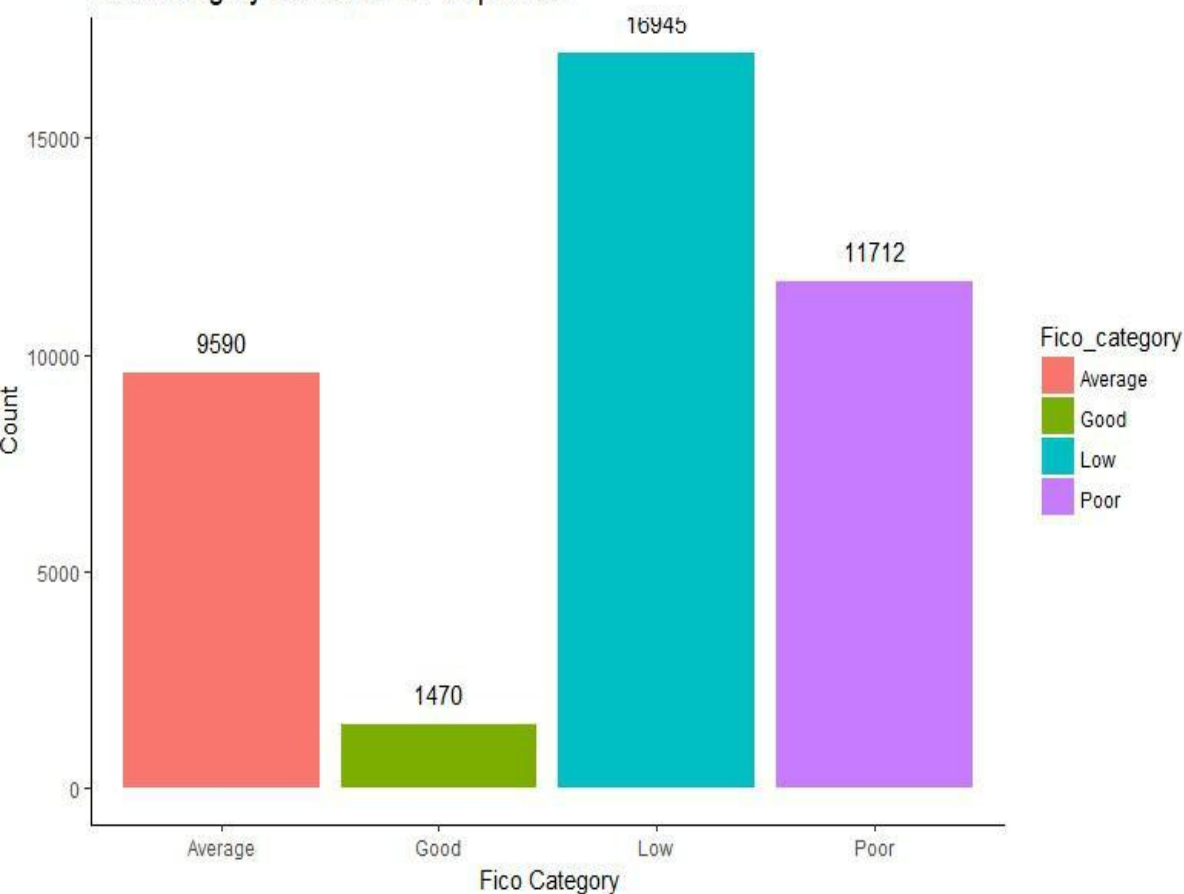
**Similarly we see a positive correlation between loan\_amnt and installment.**

**And we see a positive correlation between total payment and reveived principal**



## Business Driven Metric: Fico Category

Fico Category Derivation for all profiles



Fico Category is derived metric which is calculate based on FICO Score.

Fico score Range	Fico_Category
250 – 499	Poor
500 – 579	Low
580 – 618	Average
620 -- 679	High
680 – 800	Excellent

It is evident the lending club has to focus on Fico\_category of profiles with Low & Poor to avoid credit loss



# Conclusions

UpGrad

## Summary:

we have a clear introduction to the growth of LC's main business (i.e lending). It's obvious that LC is currently a fast-growing but the increasing of volume and amount of loans is erratic.

## Conclusion:

1. Lending Clubs business grows significantly since the start of 2010 and 2011.
2. The largest markets of LC is California.
3. Debt consolidation and credit card are most popular reasons for borrowing from LC from Treemap.
- 4 .LC is not that efficient in pricing the interest rate with different terms.
5. LC Customers who have home ownership of 'MORTGAGE' and 'RENT' have much more demands of borrowing money than customers in 'OWN'
6. Also the High Risk Customer who may default are those who take home ownership of "Mortgage" and Rent compared to others.
7. It is also clearly evident Lending Club issued loans to large borrowers whose profiles have not been verified fully. This shows a clear process gap to evaluate their borrowers and this has to be fixed immediately.
8. From FICO category we can say that Lending Club got a lot of profile who got Low and Poor FICO scores. This is a high risk area and they have to come with business plan to recover their money from them.



**Thank You**