

# **Lending Club Case Study**

## **SUBMISSION**

### **Group Member:-**

- Suraj.B.S
- Vinay Singh Shekhawat (Group Facilitator)

# Introduction

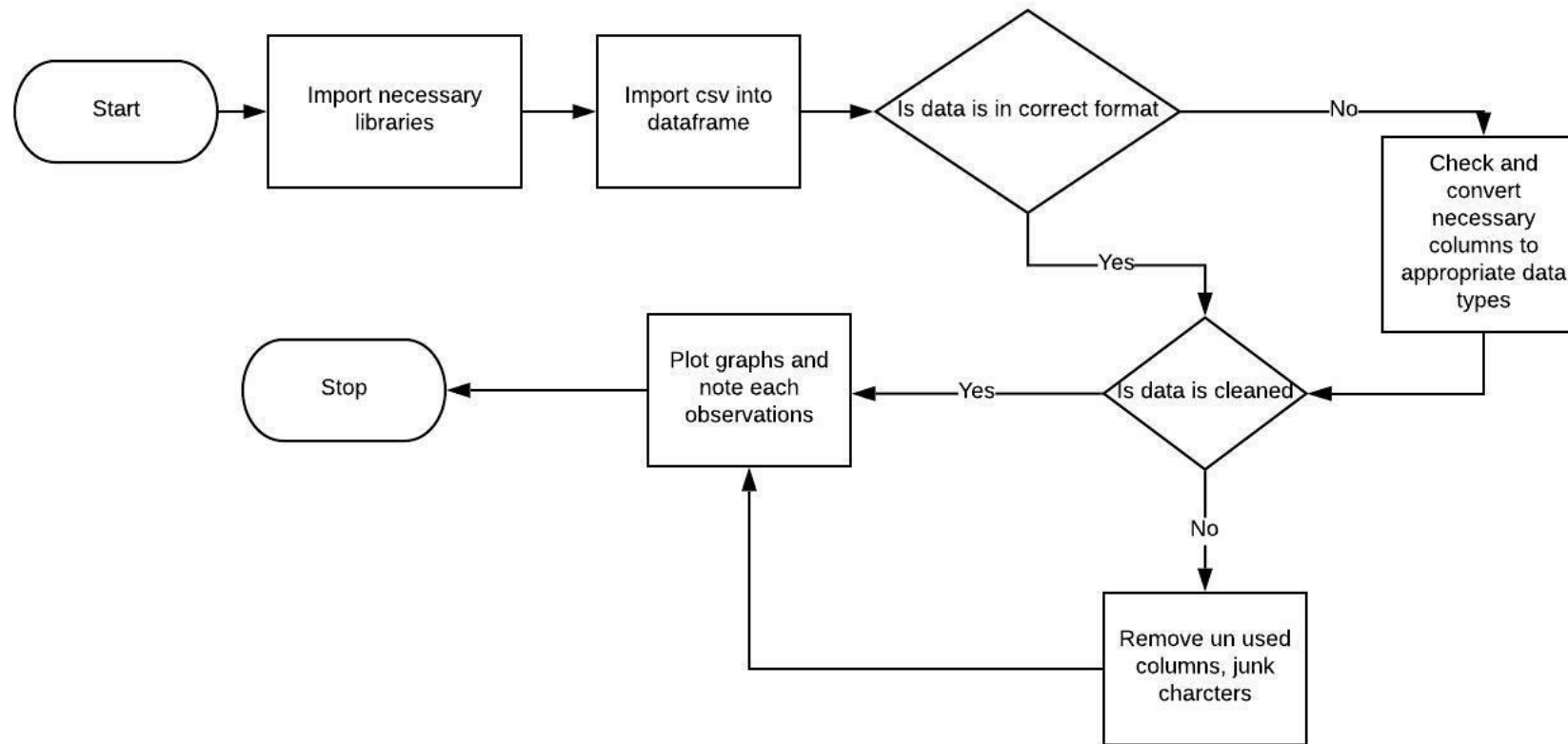
## ➤ **Company**

- Lending Club is largest online loan marketplace, Business loans, Facilitating personal loans and financing of medical procedures.
- Borrowers can easily access loans in lower interest rate through a fast online interface.

## ➤ **Problem Statement**

- The company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilise this knowledge for its portfolio and risk assessment.
- Company wants to analyse whether approving loan of certain applicant will help to boost their business or not.
- Dataset is containing information about past loan applicants, their grades and loan status. Using this information, plot graphs and see whether there exists defaulters or not.
- Use Exploratory Data Analysis.

# Flow Chart



# Methodology

Below are some steps which we were used to solve our loan application problem.

-Remove NULL values from datasets  
-Calculate NULL values percentage and remove respected rows

Understanding the data and getting knowledge of all the columns

Analysing the columns and plot the distribution of respective columns.

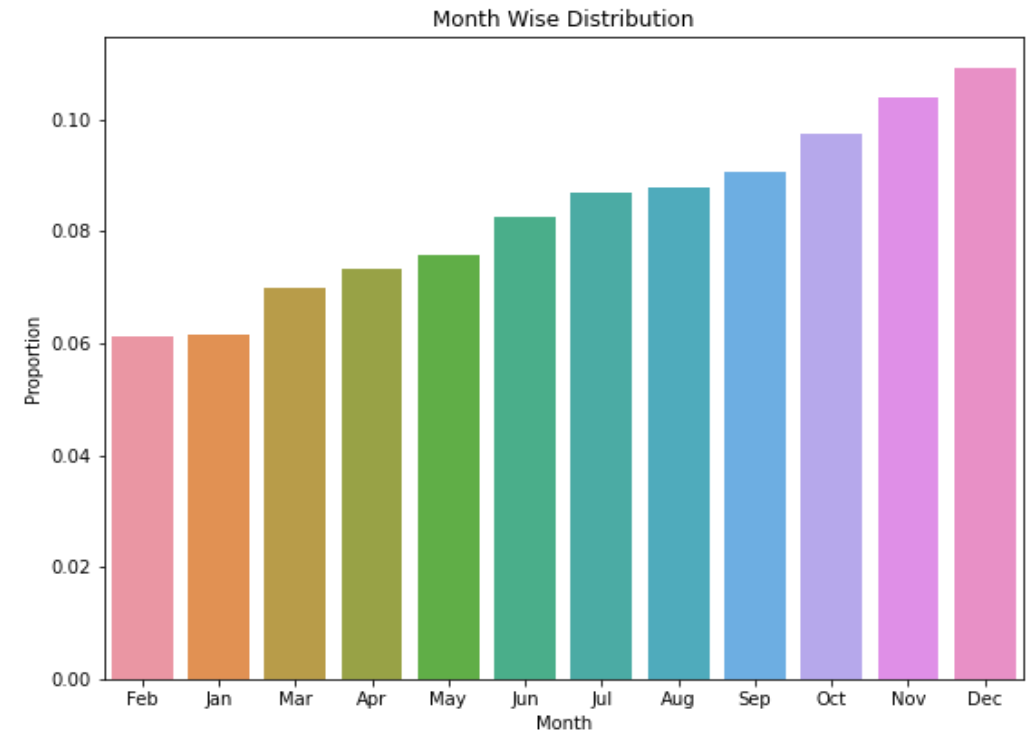
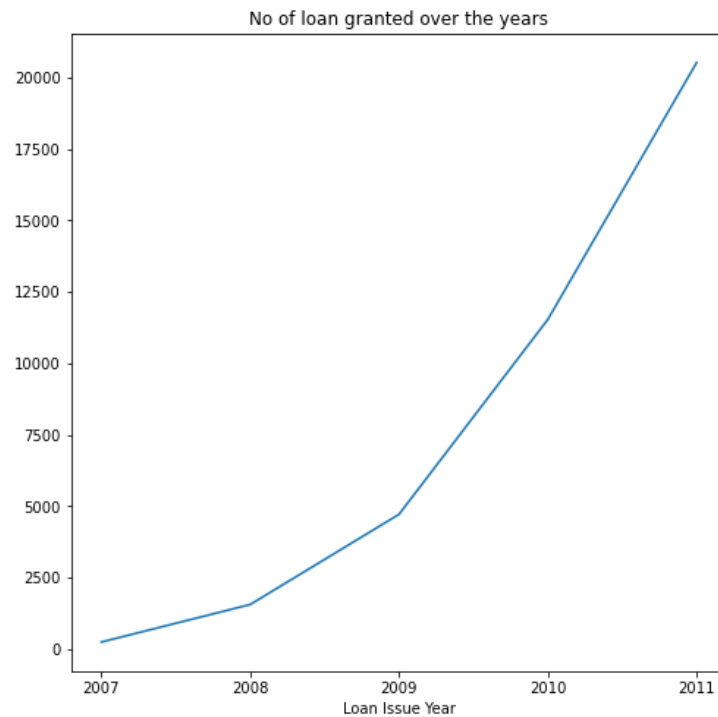
Analysing the continuous data columns with respective to categorical columns.

Analysing the two variable behaviour like term and loan status with respect to loan amount.

Publish insights and observations

# Analysis - Understanding Loans

- ✓ We found that almost every year, Lending club has managed to double its business.
- ✓ December is a month in which higher loans were distributed, where February is lowest.
- ✓ Loan distribution keeps increasing from start to end of year.

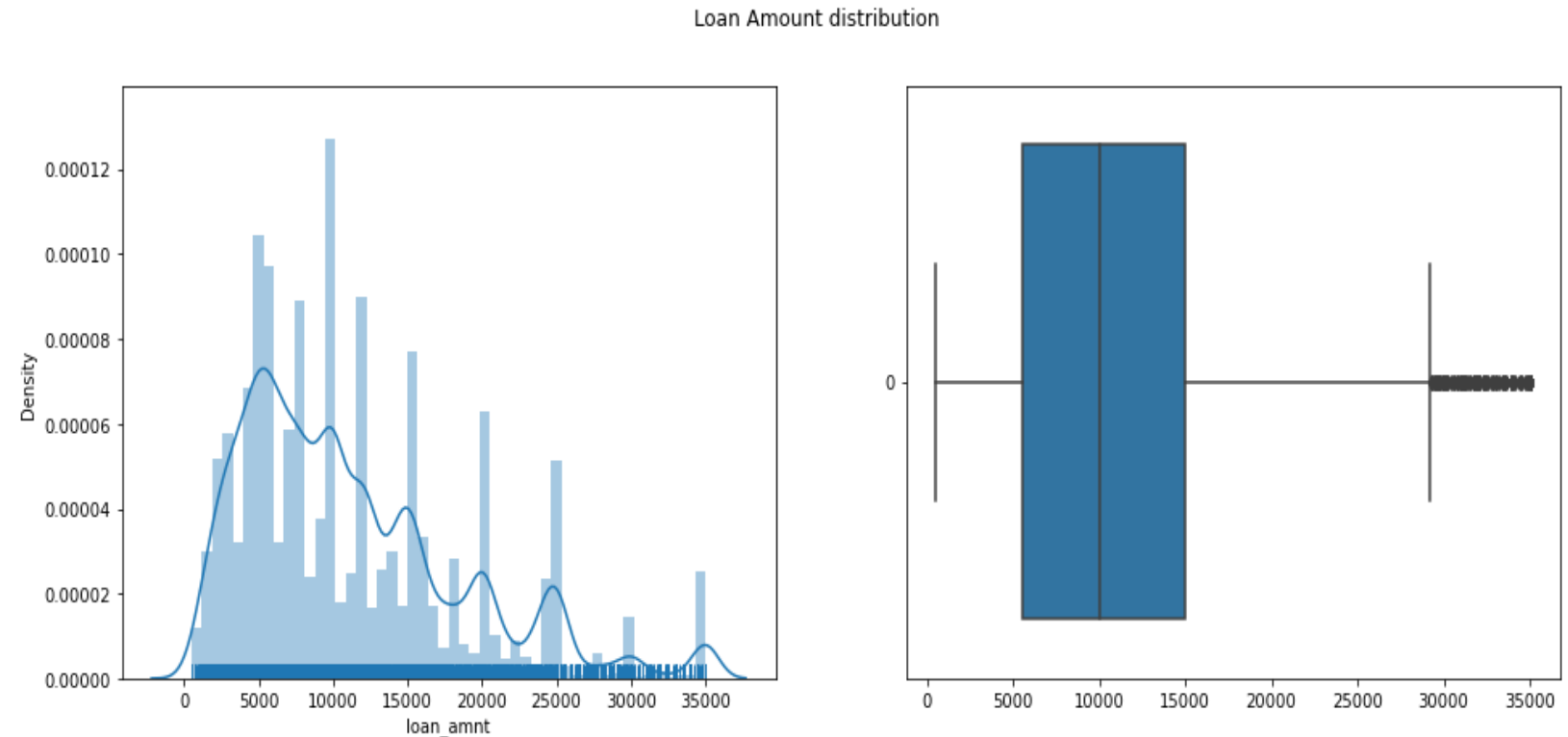


# Univariate Analysis

A Univariate Analysis is a type of analysis in which only one variable from dataset is used. Data may contain more than one variable, but we will only focus on that single variable. “Uni” means **Single**.

We have decided to look ‘loan amount’, and we found that

- 500 is minimum and 35000 is maximum loan amount which is distributed.
- Mean loan amount is 11219.44
- 50% applicants are given loan amount between 5000 to 15000
- This distribution is skewed.



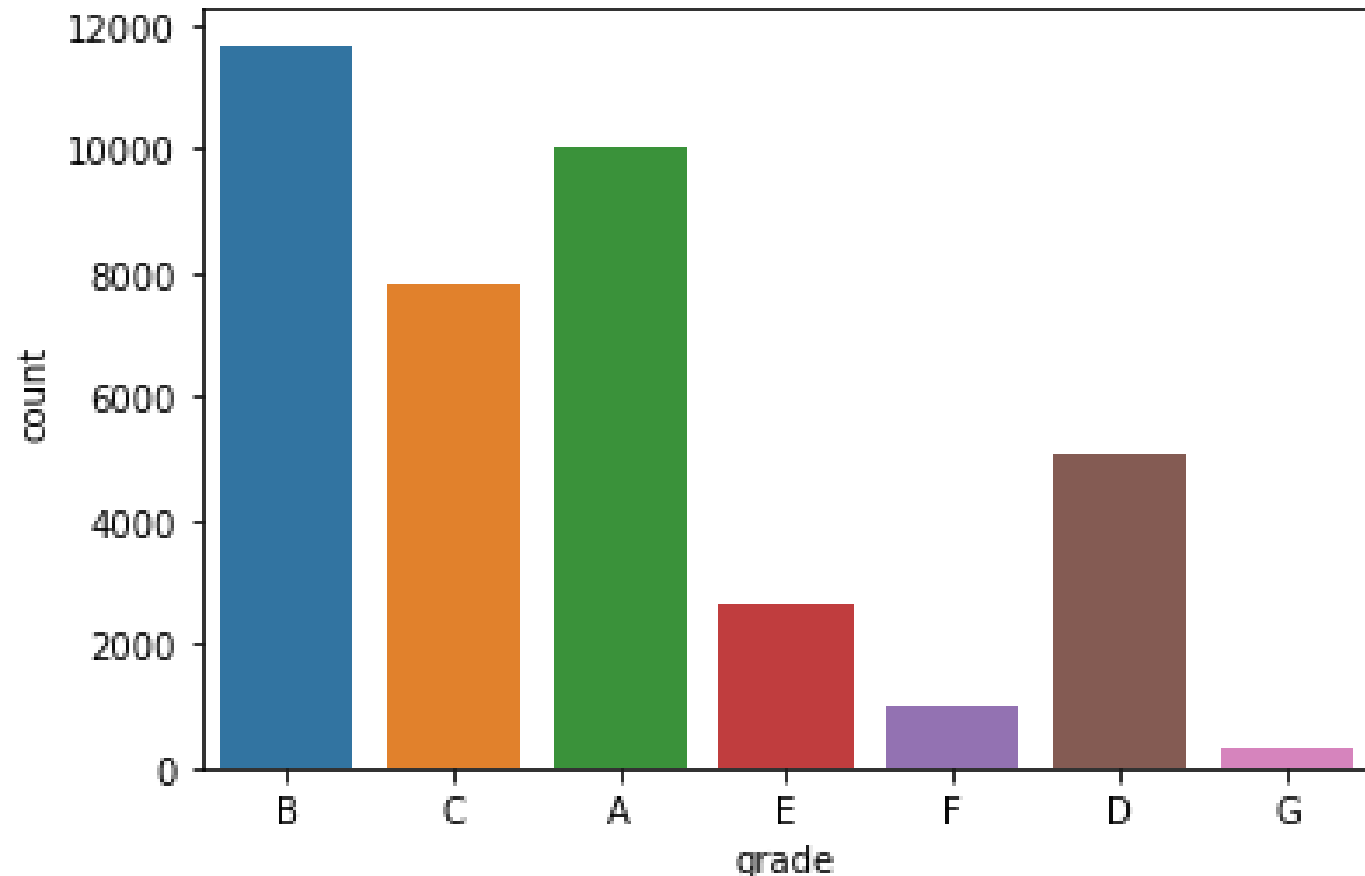
Like loan amount column, we analysed grades, because it seems an important factor while giving a loan to applicant. What do you mean by a **Grade**?

A Grade is a quality score given to a loan borrower depending his credit history. You may hear about CIBIL score. Think of it like that.

Grade **A is highest grade**, it indicates that applicant have paid all of his EMI on time, where **G** is lowest grade,

Below is our plot for grade

- B is grade hold by most loan applications.
- G is grade hold by least applications
- We can say that only in some criteria, lending club has given loan to candidates who has bad credit history.
- There is a risk, but it is low.



# Segmented Univariate Analysis

We have seen behavior of loan amount. In previous case, we just compute mean, median, 50% data distribution. We can categorize this univariate variable based on some conditions or based on existing variable.

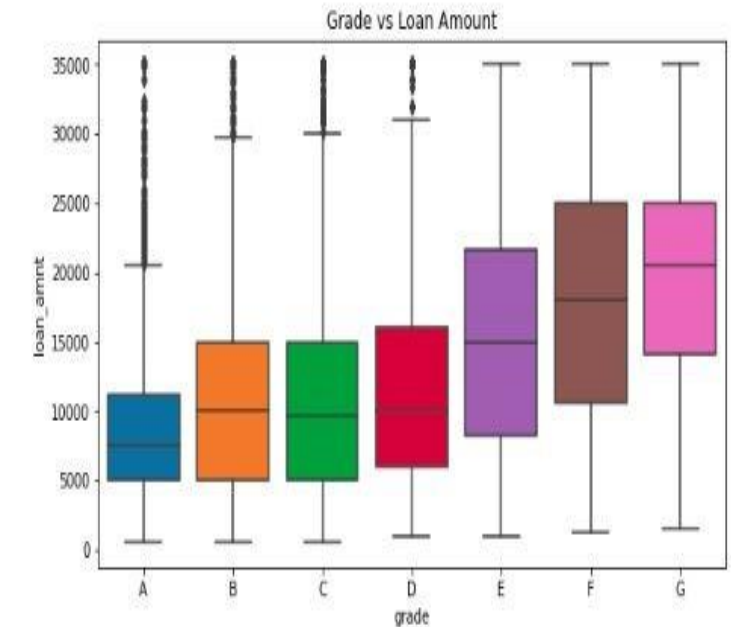
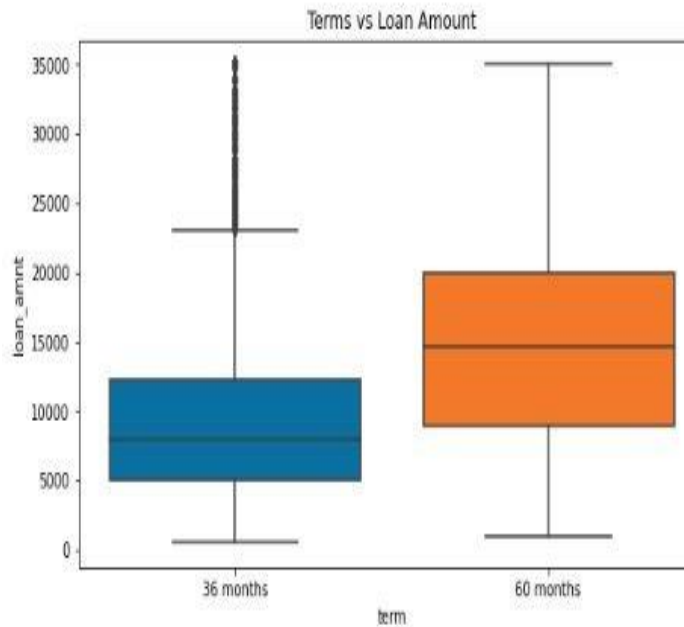
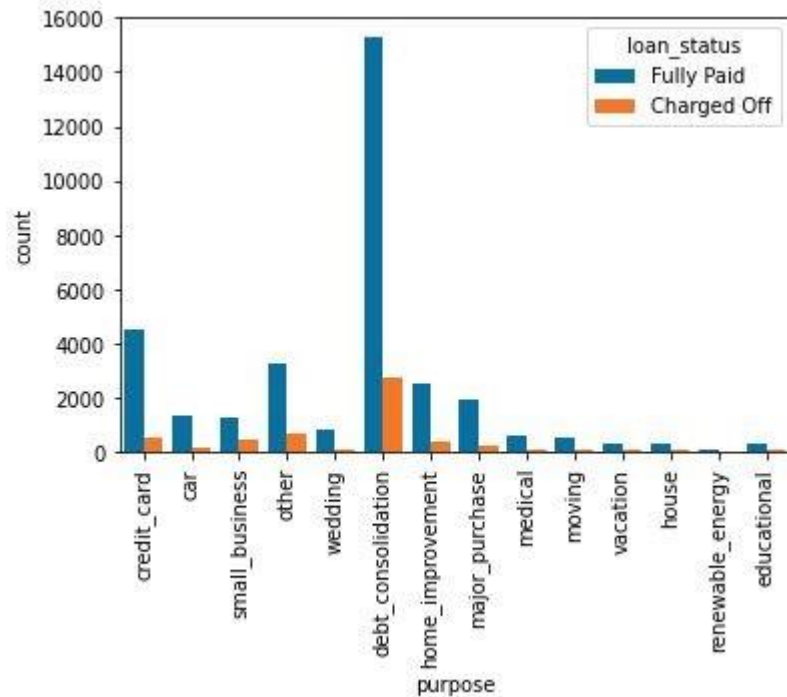
The technique of grouping dataset based on a category or mathematical function is known as **Segmented Univariate**. Below can be steps

- Take raw data
- Group by dimensions
- Summarize using a relevant metrics
- Compare the metrics along group



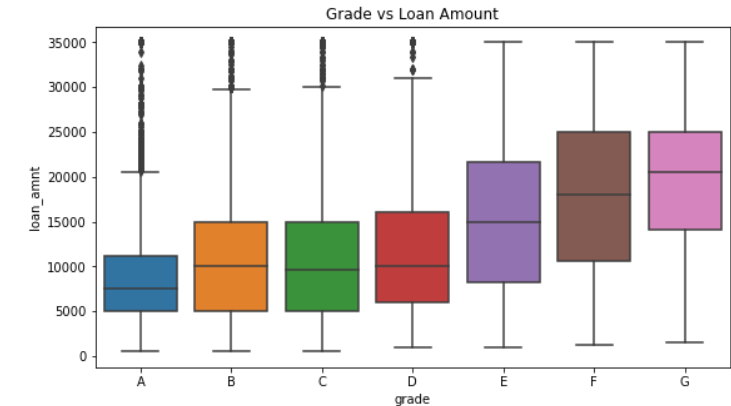
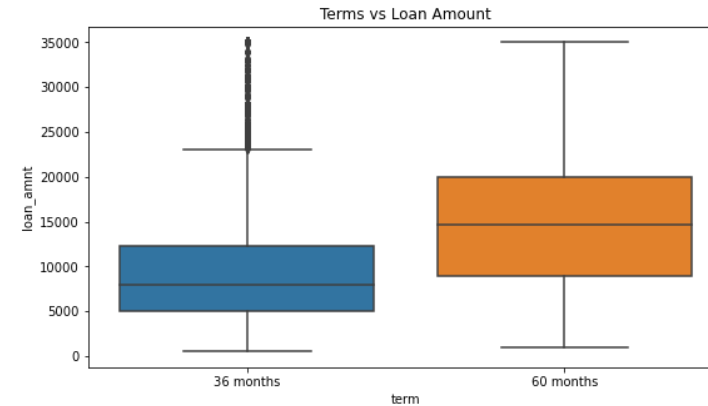
We compared loan amount in segmented way and found that –

- Majority of loan has been given for the debt consolidation purpose and has been fully paid
- Higher the loan amount, higher the tenure. Usually, high amount need some extra time to pay off entire loan.
- Lower the grade, higher the loan amount. Lower grade means candidate have lots of outstanding Emi. He need more money to pay money



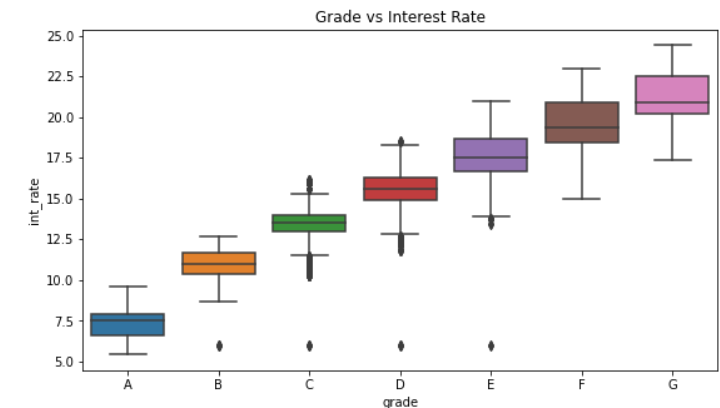
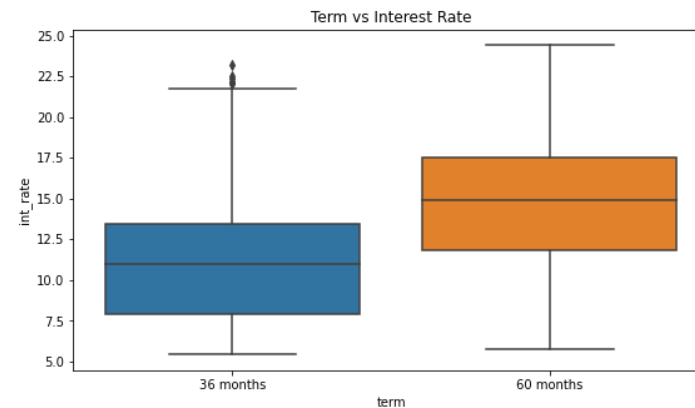
Continuing previous observations on loan amount and found that

- Higher the loan amount, higher the tenure.
- Lower the grade, higher the loan amount



We try to observe rate in this way and found that

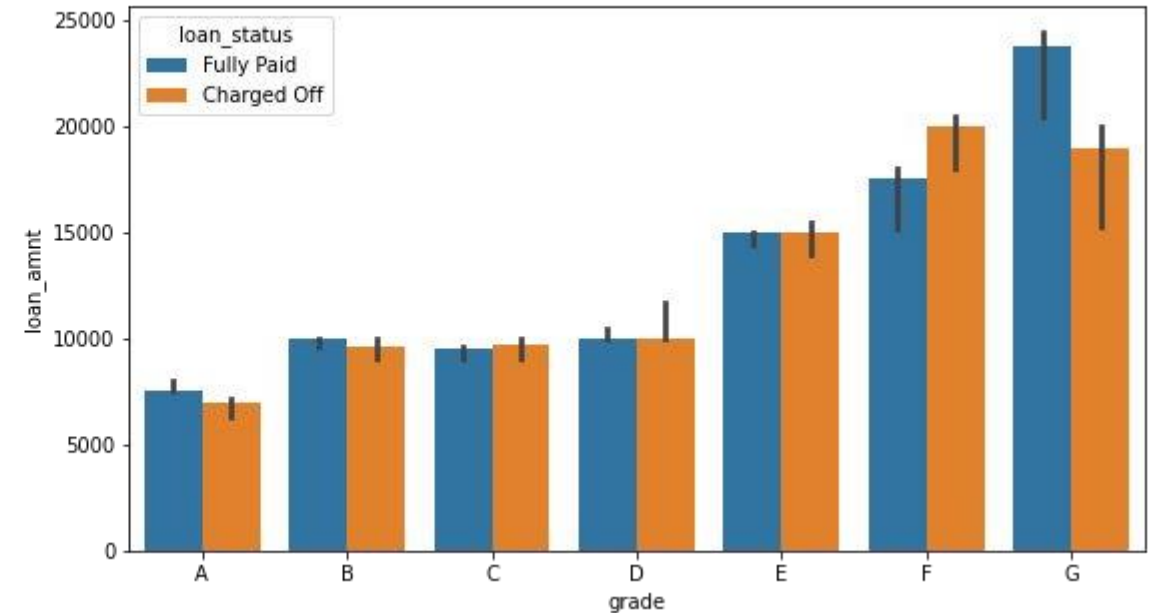
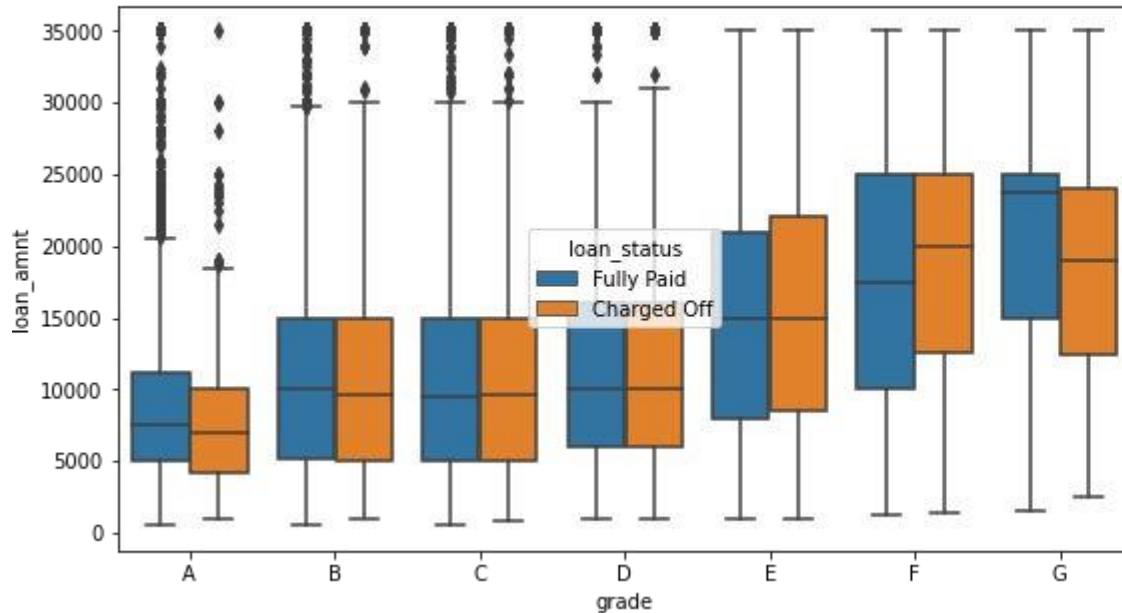
- Both tenure have different minimum rates.
- Lower the grade, higher the rates



# Analysis on the basis of grade and loan amount

- For lower grades 'F' and 'G' there are more difference between charged-off and fully paid.
- The lower grade people has taken higher amount of loans and also they are more prone to default the loan.

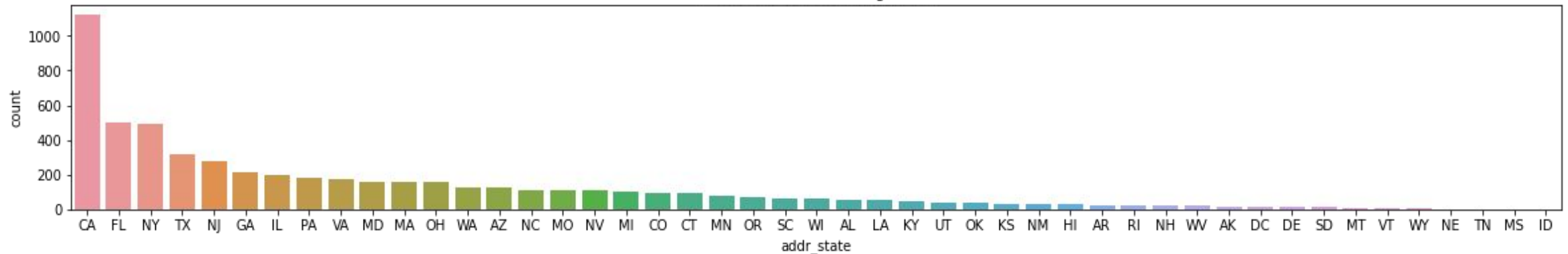
Term vs loan amount



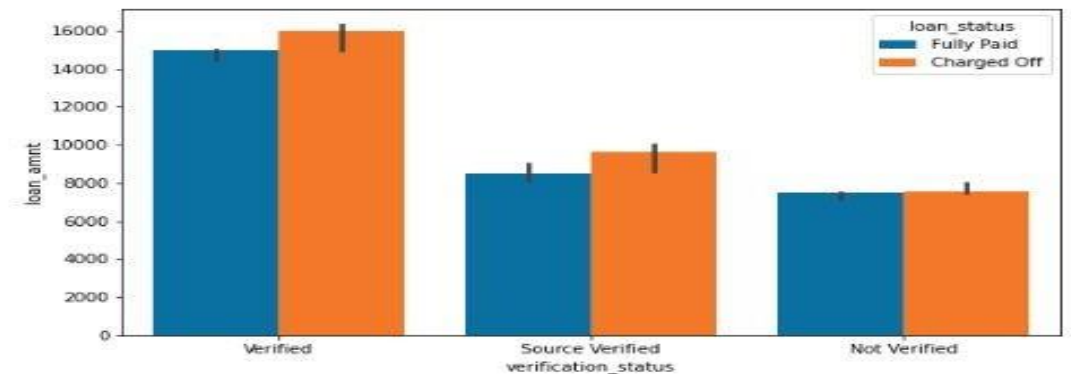
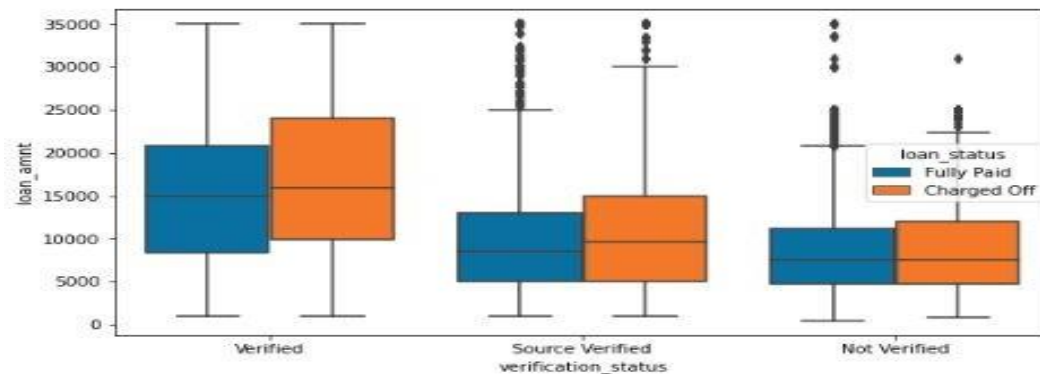
# Analysis on the basis of location & Source

- More number of borrowers defaulted in CA , FL and NY states.
- Verified loans are given more loan amounts compared to others.
- There is a slight increase in loan amount for verified and source verified loan when they are defaulted.

Address State vs Charged Off

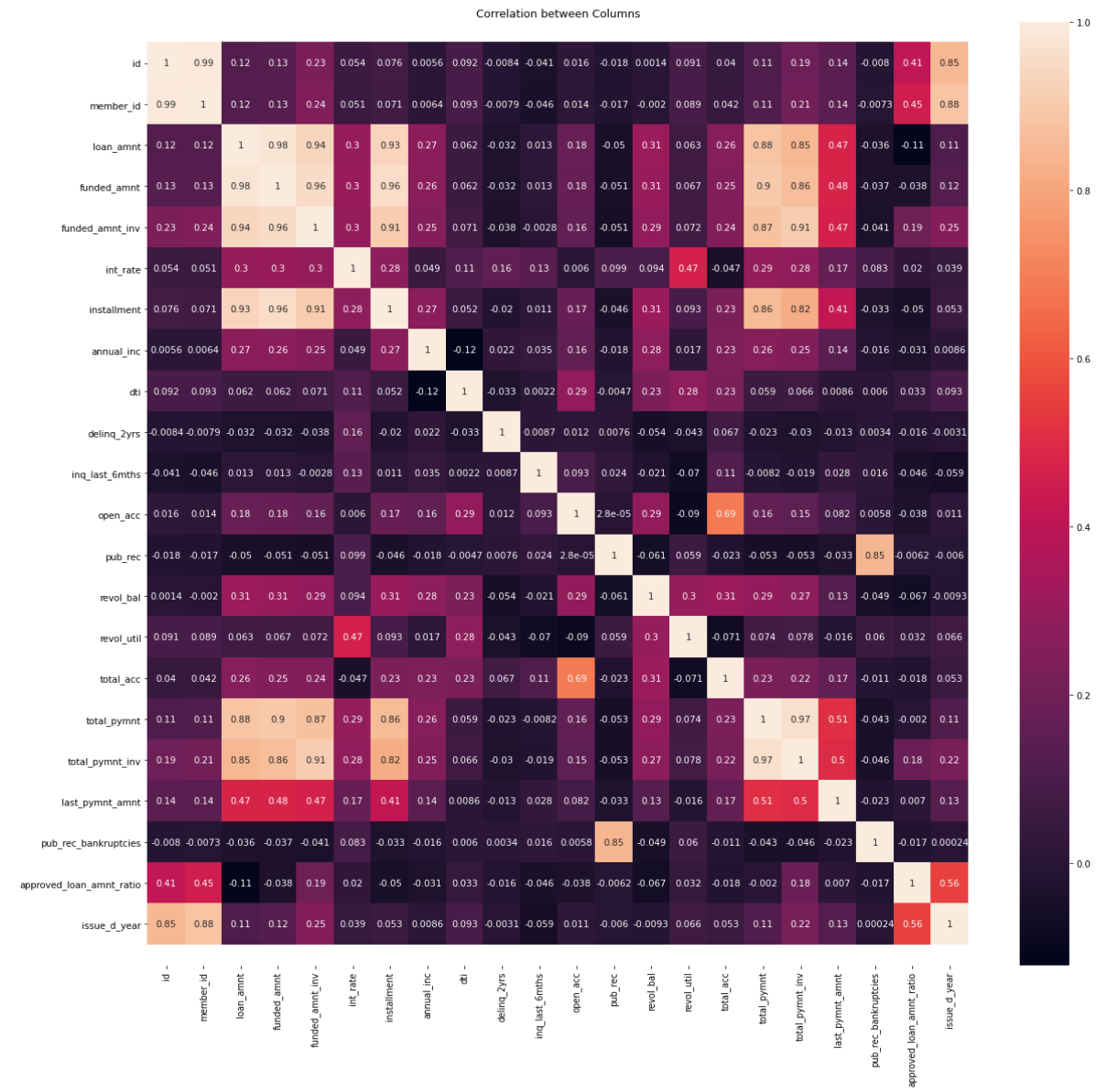
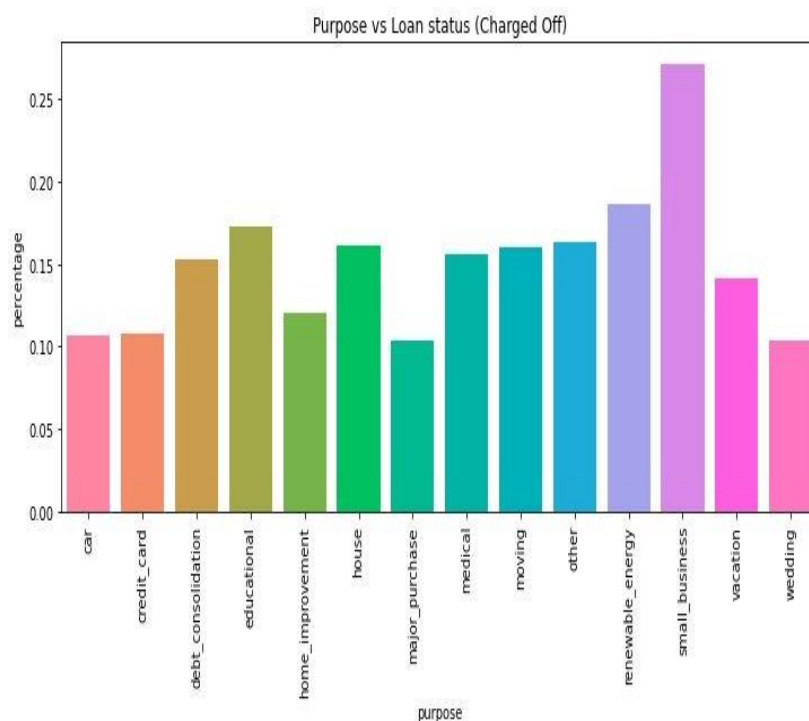


Verification Status vs loan amount



# Purpose of Loan And Correlation Analysis

- Charged-off are higher for small\_business comparatively.
- The public derogatory records column is highly (+ve) correlated with public bankruptcies records.
- Interest rates are high for people with high revolutilisation.
- For car and credit card this percentage is almost same.



# Conclusions

Based on our various univariate, segmented and bivariate analysis we reach to conclusion that

1. Lending club has managed to doubled its business by giving more loan to applications with higher grade..
2. Since low grades are more prone to default on the loan, lending club must identify the ways to lower the interest rate for them.
3. Lending club should be concerned while approving higher amount of loans to customer who have lower grades.
4. Borrowers with more work experience took more loan amounts got more default rate.
5. While approving loans for customers who are from CA, FL and NY, lending club should study more in detail about their background and thus lowering the chances of default.
6. Charged-off are higher for small\_business comparatively.
7. People with a greater number of public derogatory records are having more chance of filing a bankruptcy. Lending club should make sure there are no public derogatory records for borrower.