

# Lending Club Case Study

By - Vaishali Bhilave

# Problem Statement

Company main objective is to facilitate loan to borrowers through a fast online interface.

Aim of this analysis is to identify risky loan applicant, to avoid lending loans to these applicants which are the largest source of financial loss.

Applicant who refuse to pay or runs away with money owed cause largest amount of loss to the lenders. With this analysis it will help company to understand the driving factors behind loan default



# Loan Data Description

We are having data of 5 years (2007-2011) which contain information of past loan applicants .

We have to analyze factor which help to decide a person is likely to pay loan on time or not. This will be helpful for company to decide for lending loan to borrower.

Data have three types of attributes-

1. Borrower related
2. Loan related
3. Customer behaviour related



# Data Preparation

To analyze given data first we have to prepare data for analysis by -

1. Removing null value rows/columns from data
2. Removing Outliers
3. Removing unnecessary data like, Customer behaviour related data which are not required for analysis
4. Removing data where loan status is 'Current' as it is not helpful to analyze defaulted borrower
5. Simplifying data to make it easy to analyze



# Data Analysis

Let's start analyzing loan data to get responsible factors which is responsible for **Charged Off** loan

The column in provided dataset which clearly signifies if a loan is charged off or fully paid is **loan status**.

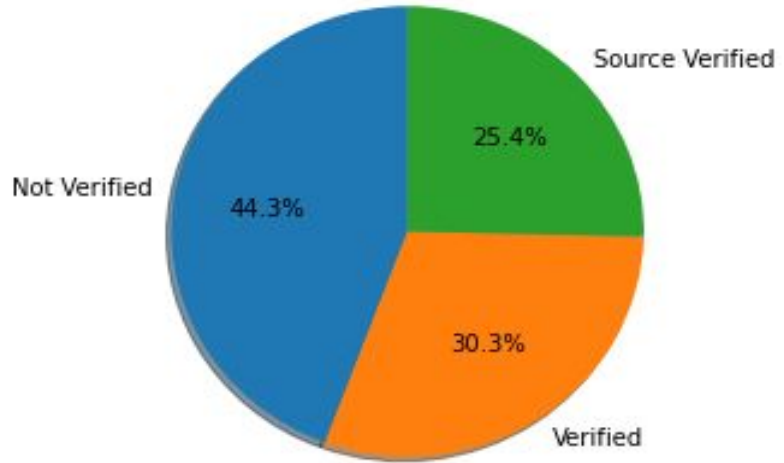
So considering **loan status** as main deciding attribute let's proceed analysis



## Analyzing verification status of borrower where loan is either fully paid or charged off

As we can see here that there are 44.3% of loan data which is “Not Verified”, which is maximum.

Other loan data verification status are Verified (30.3%) and Source verified (25.4%)

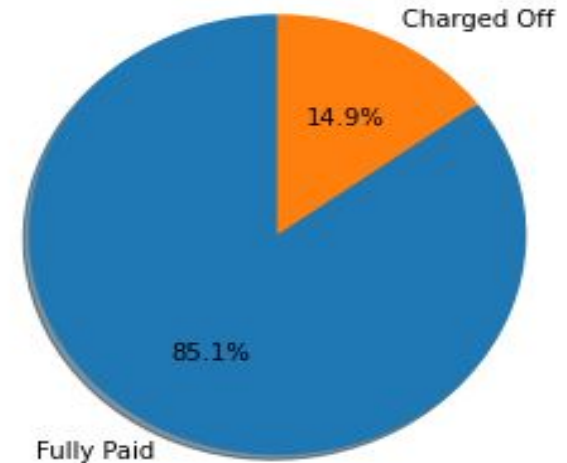


## Analyzing data where sanctioned loan is less than loan amount requested

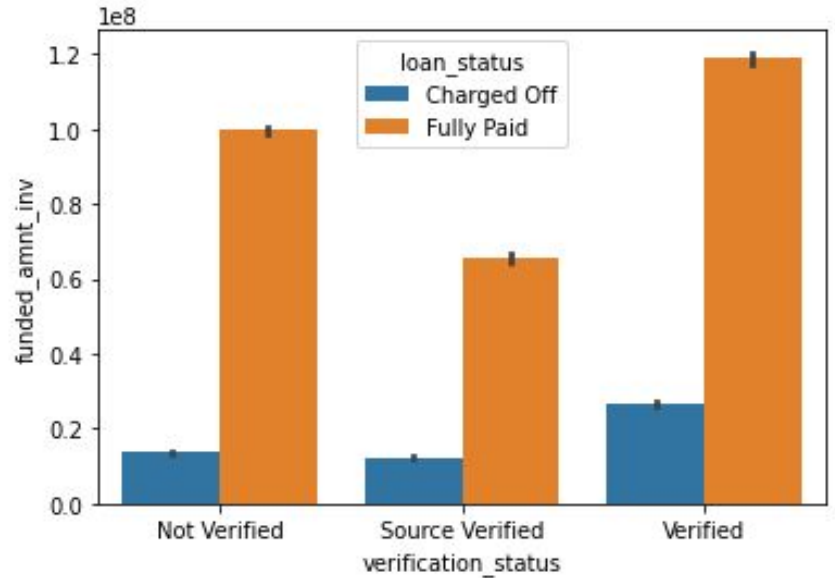
Loan data where sanction loan is less than loan amount requested , loan which is fully paid is **85.1%**

Whereas loan which is charged off is **14.9%**

Hence we can conclude that decision of lending less amount of loan than request amount was good because maximum number of loans are fully paid



## Analyze borrower verification status and loan amount sanctioned by investor w.r.t loan status



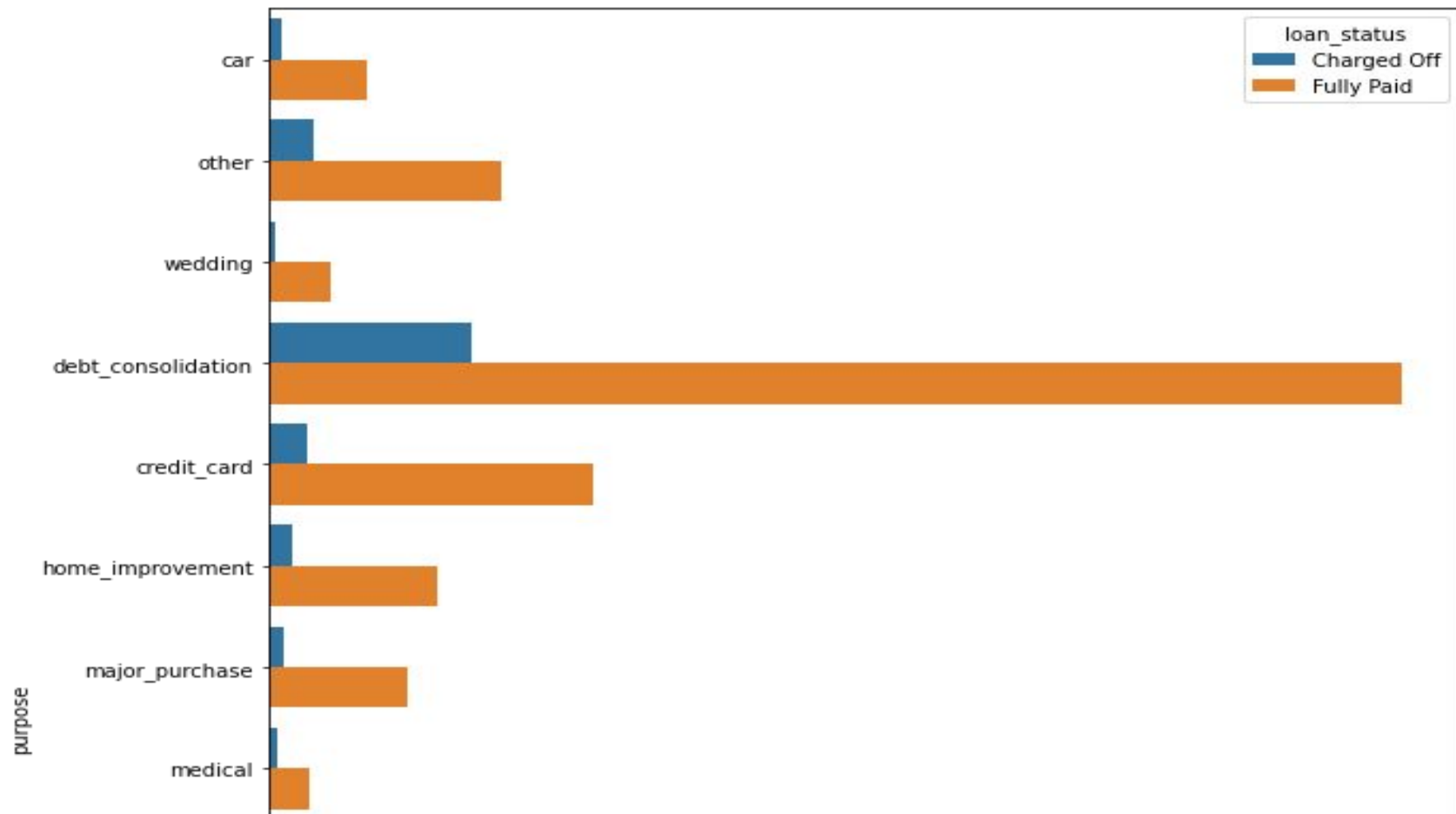
We can see here , that borrower verification is not having much impact on its loan status , As borrower which are verified are having maximum charged off loans.



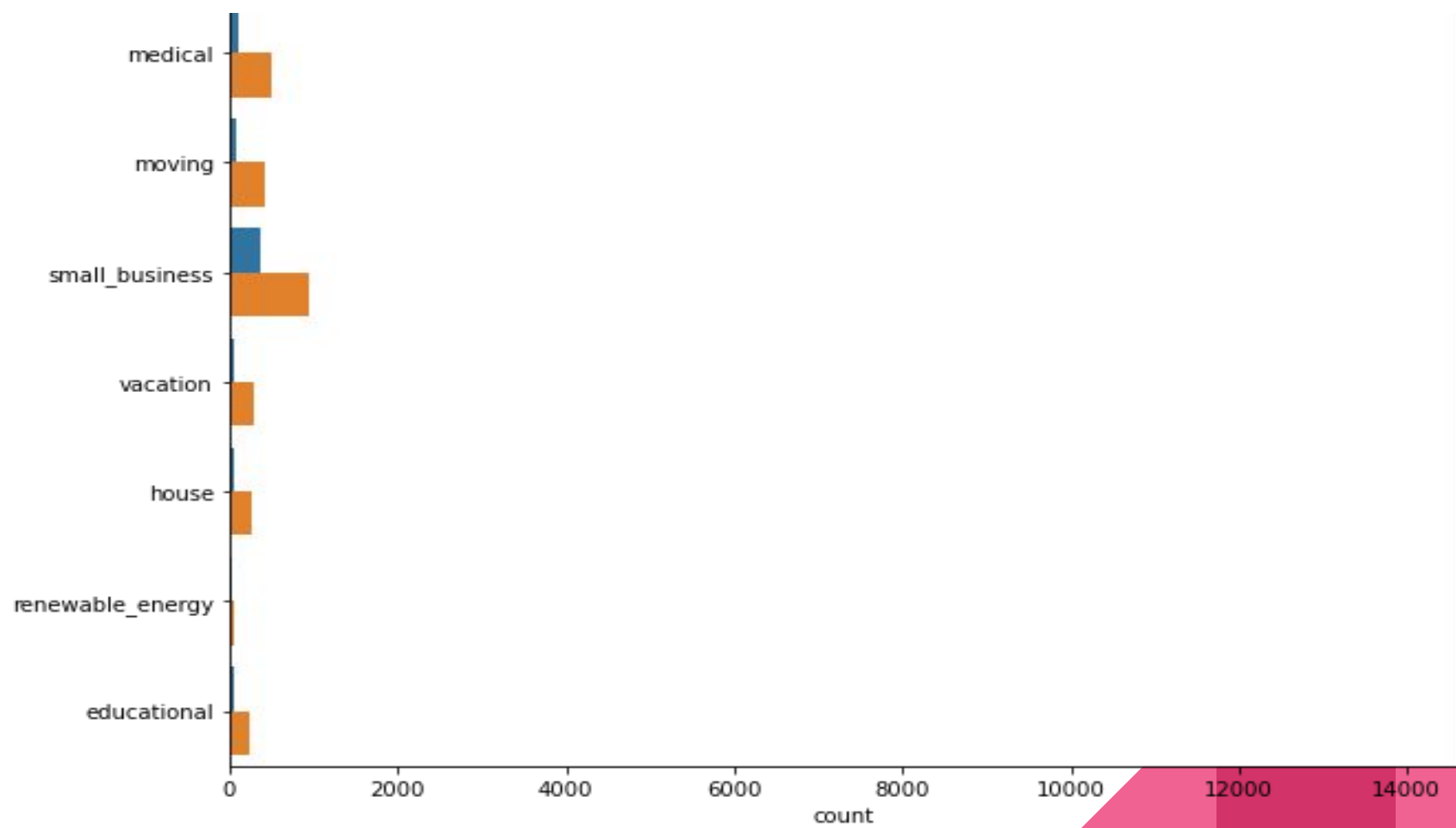
# Analyze the category provided by the borrower for the loan request

- ❖ Below plot shows that most of the loan taken for debt consolidation purpose are fully paid.
- ❖ For renewable energy purpose very less amount of loan is proposed.
- ❖ Maximum number of charged off loan is also for debt consolidation.





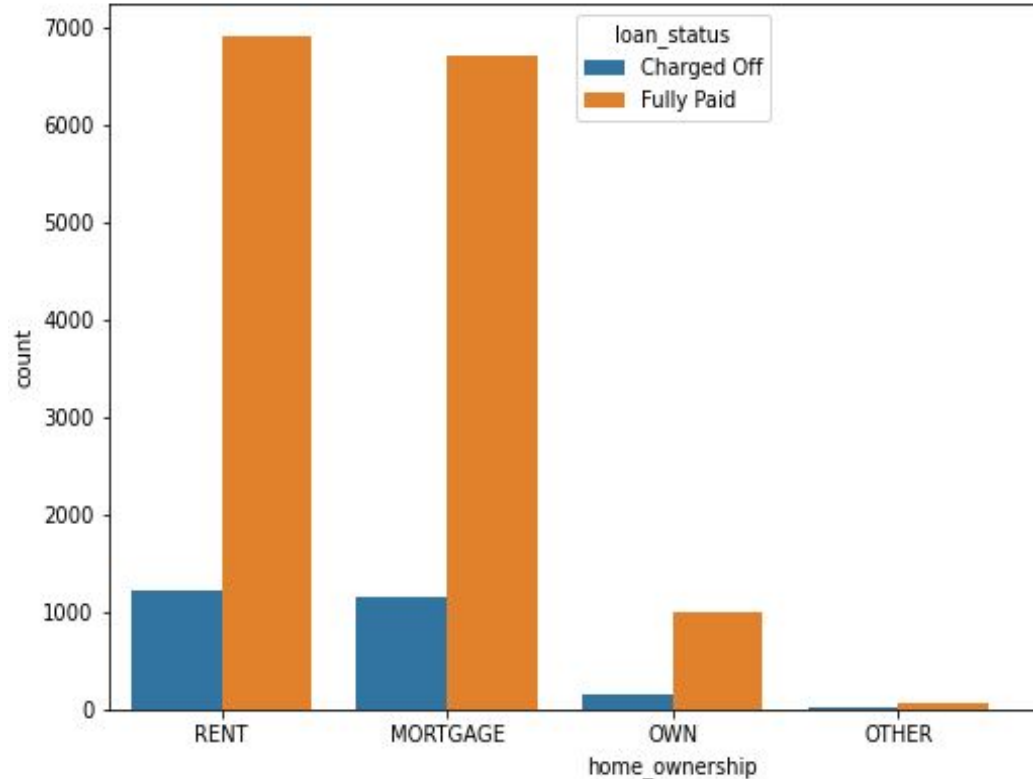
put



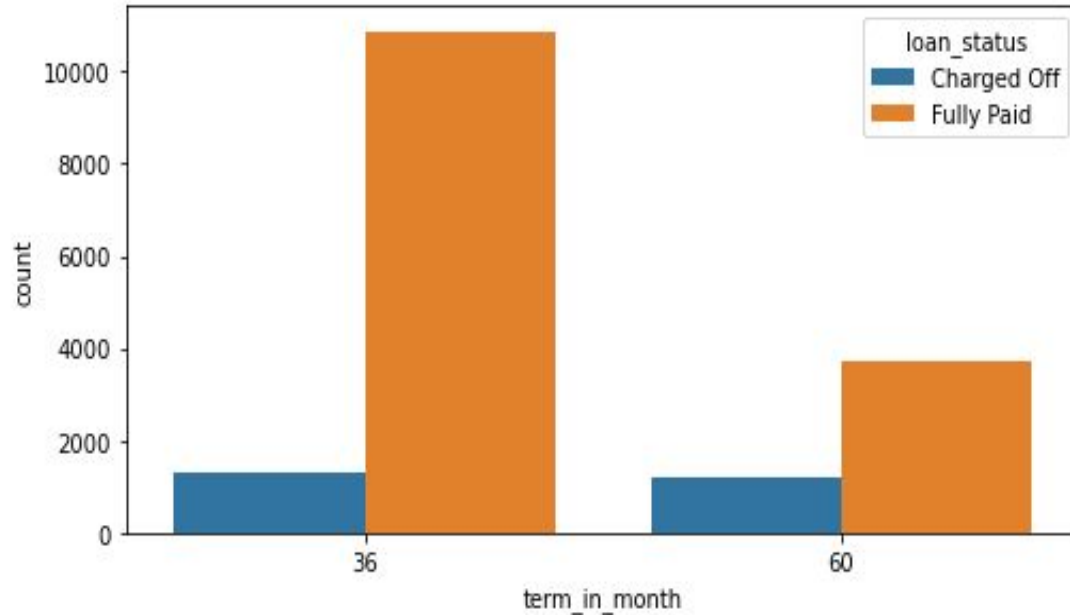
## Analysis of borrower's home ownership w.r.t loan status

This plot shows maximum number of charged off loan and fully paid loan is where borrower is living in rent

Borrower who are having their own home is having minimum number of charged off loan as compared to those who are living on rent or mortgage their home



## Analyzing number of payments on the loan wrt loan status



This plot shows that most of the loan which are having less number of payment are fully paid.

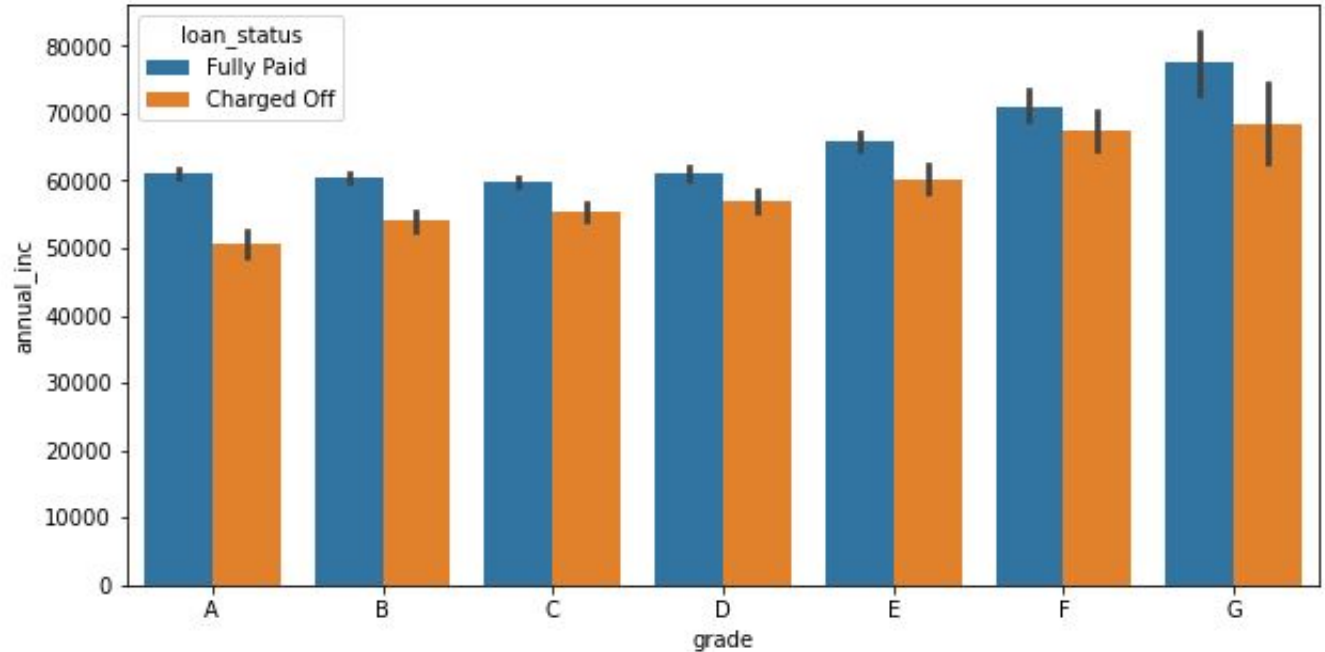
\*term\_in\_month = Number of payments for loan in month

# Analyze borrower loan grade and annual income wrt loan status

This plot shows that borrower having loan grade G have maximum charged off loan and annual income

This plot shows that borrower having loan grade A have minimum no of charged off loan and annual income

Plot shows that loan status count is deviated almost similarly with annual income



# Analyze loan interest rate and loan issued year wrt loan status

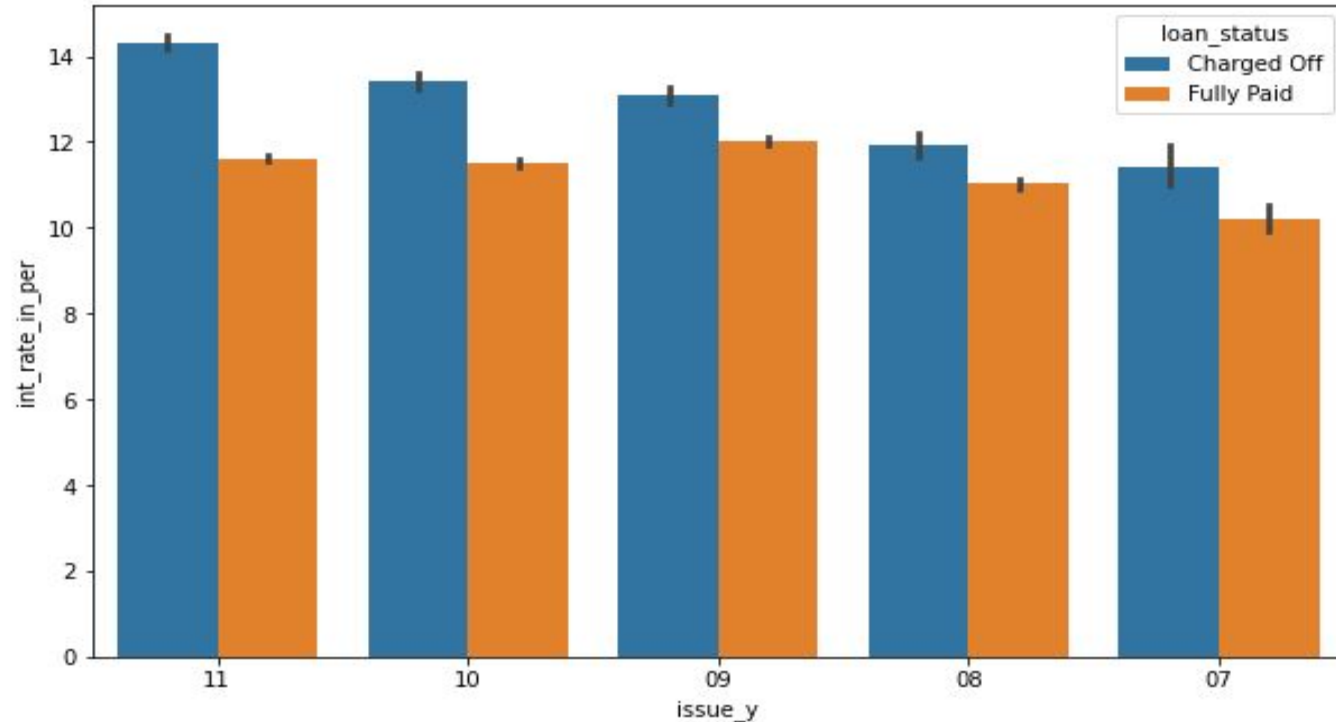
This graph shows that in year 2011 maximum number of loan is issued which are charged off

In year 2009 maximum number of loan is issued which are fully paid

Overall maximum loan issued in year 2011 and minimum in 2007

Charged off loan interest rate increased with year

\*Issue\_y = loan issued year



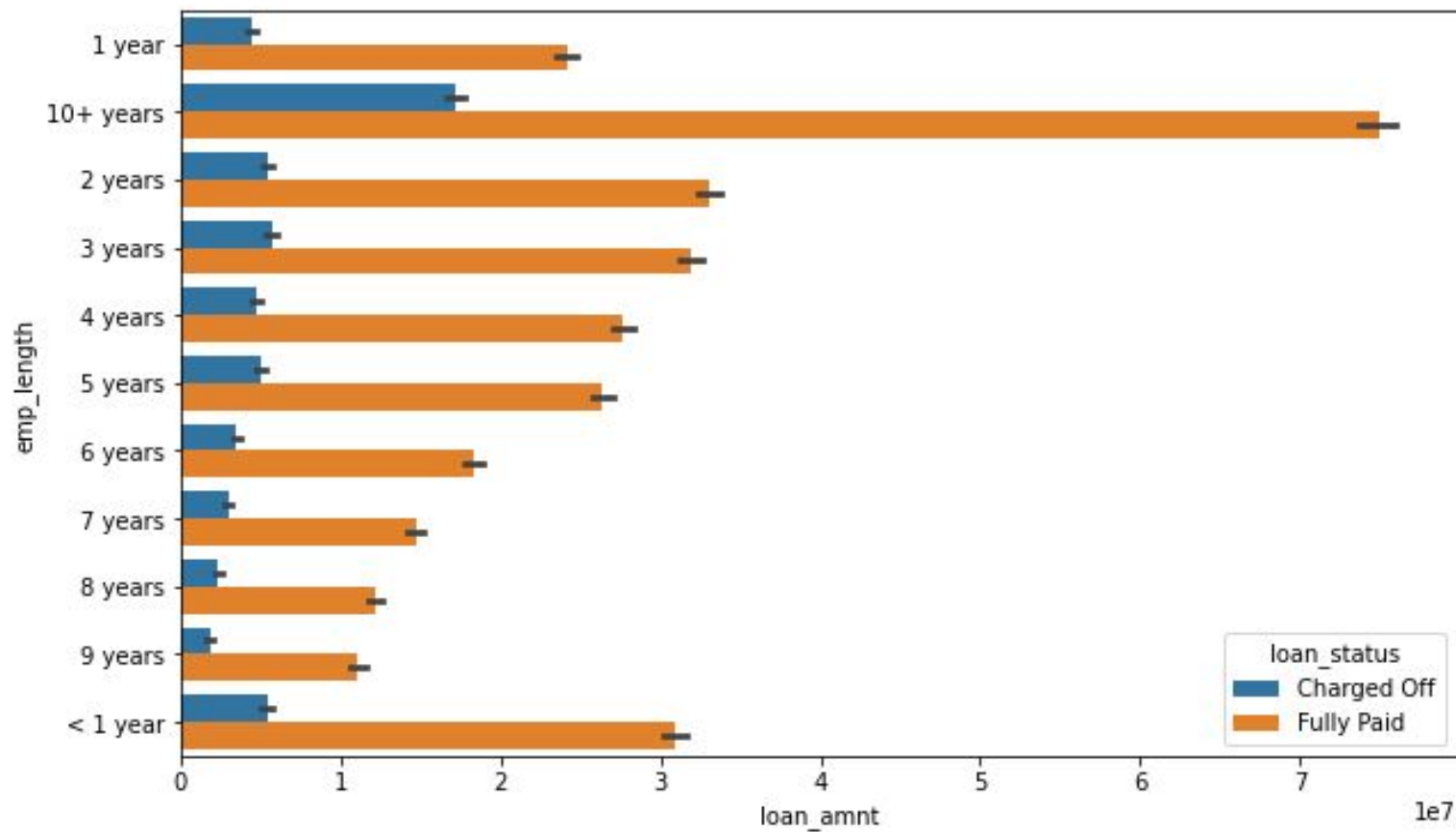
## Analyzing effect of number of years of experience of a person on loan status

Below analysis shows borrower having more than 10 years experience have maximum number of loan fully paid

It also shows charged off loan ratio is also decreasing by years of experience





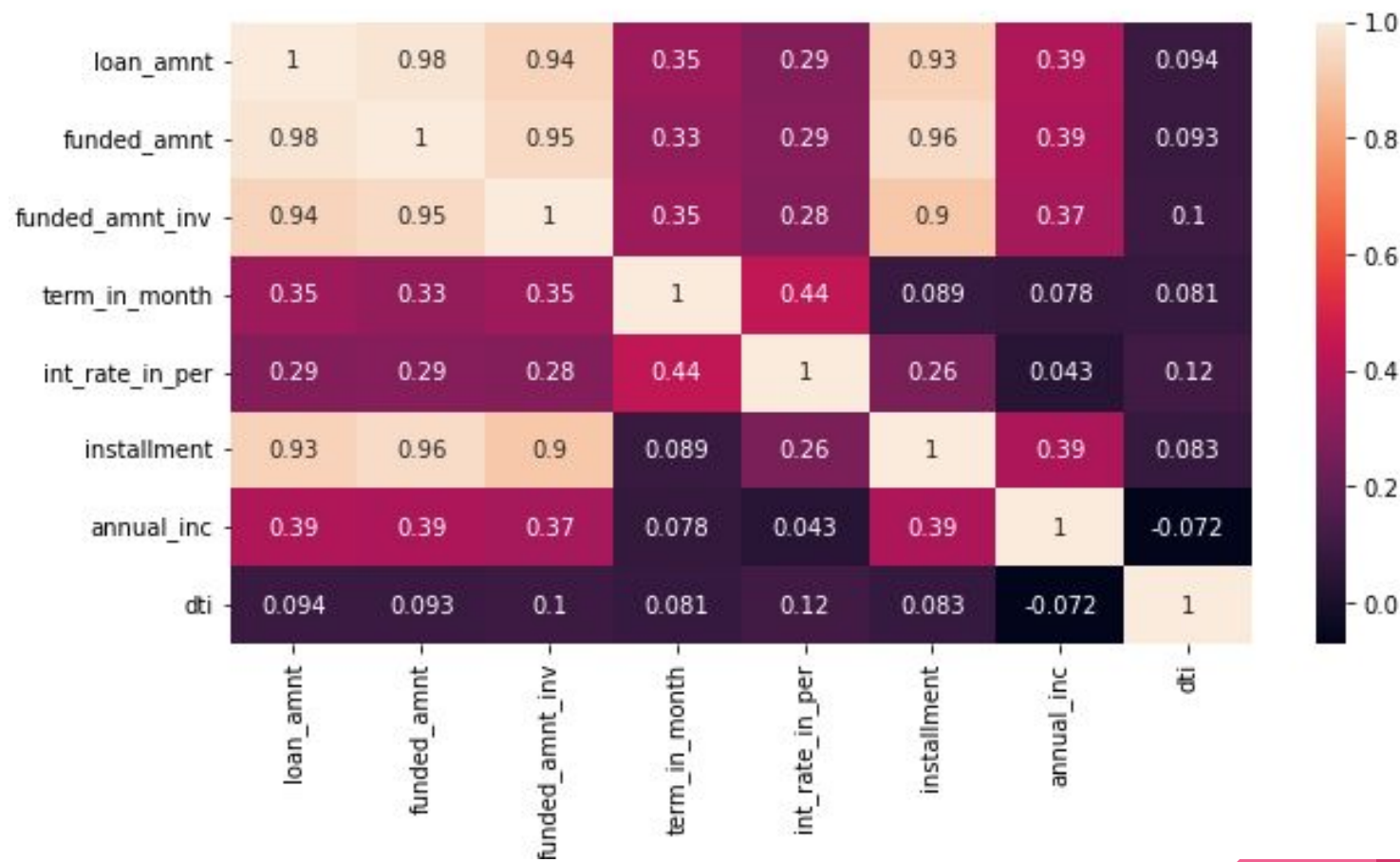


# Analyzing correlation between different categories of loan and borrower

From below Correlation Heat Map plot it is clear that -

1. Loan amount ,funded amount and amount funded by investor are highly correlated
2. Annual income and interest rate are very less correlated
3. Annual income and dti are negatively/inversely correlated





# Conclusion

On the basis of analysis we can conclude that -

1. Borrower verification is not having much impact on its loan status.
2. To avoid risk lender should sanction less amount of loan than amount request , to suspicious borrower
3. Most loan are applied for debt consolidation , so while approving loan for that pay more attention
4. Borrowers who are having their own property are much safer to lend money
5. Short term loan is much safer than long term
6. Borrower having low grade and Higher experience are much safer loan applicant.

