Lending Club Case Study

Submitted by:
Mohan Reddy
Shalaka Ghosh

Problem Statement

We work for a **consumer finance company** which specialises in lending various types of loans to urban customers. When the company receives a loan application, the company makes a decision for loan approval based on the applicant’s profile. Two **types of risks** are associated with the bank’s decision:

- If the applicant is **likely to repay the loan**, then not approving the loan results in a **loss of business** to the company
- If the applicant is **not likely to repay the loan,** i.e. he/she is likely to default, then approving the loan may lead to a **financial loss** for the company

Attributes definition

Column Name	Description
loan_amnt	Amount applied by the borrower
funded_amnt_inv	Amount given by the investor
annual_inc	annual income of the applicant
term	tenure of the loan
int_rate	rate of interest for the loan
grade	category assigned by the club
issue_d	date when the loan is issued

Attributes definition contd..

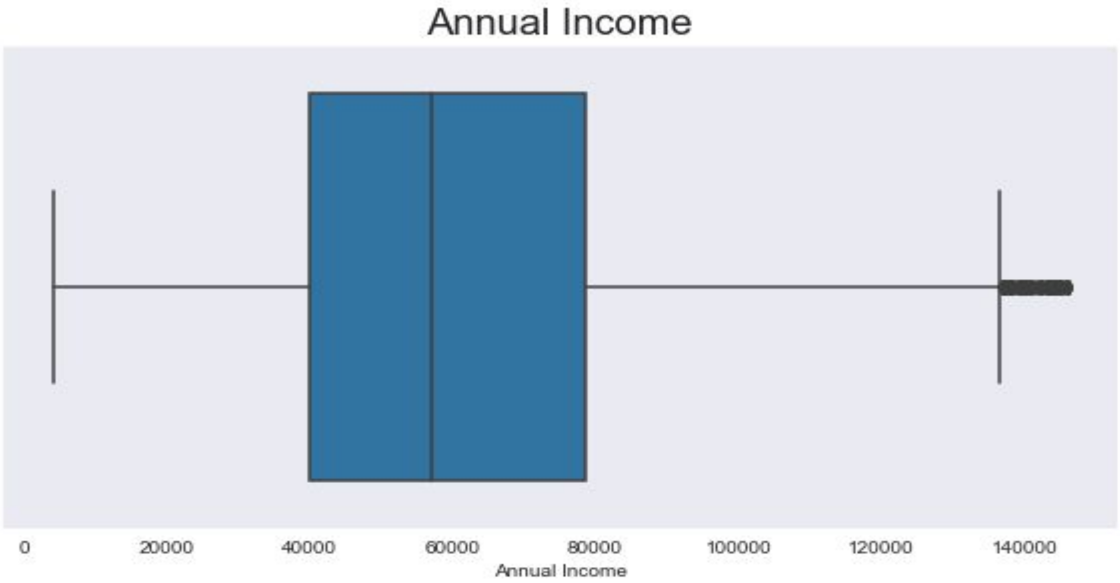
Column Name	Description
loan_status	status of the loan - possible values are Charged Off, Fully Paid
purpose	the reason given by the borrower in the loan request
emp_length	the experience of the applicant
home_ownership	status provided by the borrower, possible values are: RENT, OWN, MORTGAGE, OTHER
dti	the ratio of the total debt paid monthly divided by the annual income
verification_status	status if the income is verified or not

Annual Income

The annual income helps us determine whether or not a person will be eligible for repayment of the loan incurred.

Found many outliers in the income column data.

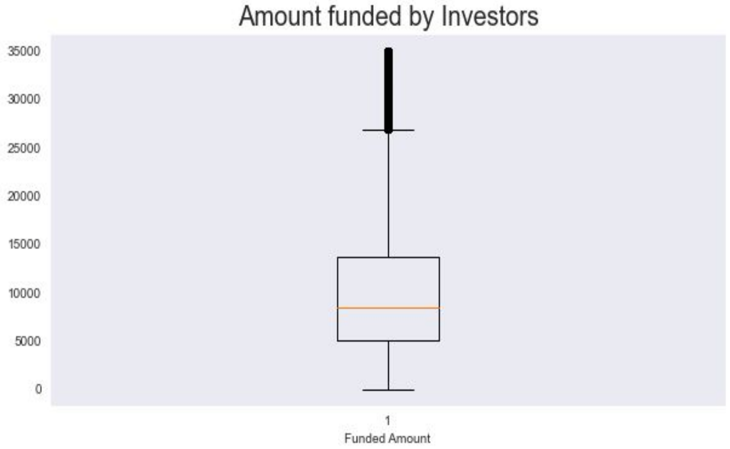
Dropping them using IQR method (values less than Q1- 1.5*IQR and values more than Q3+15*IQR)

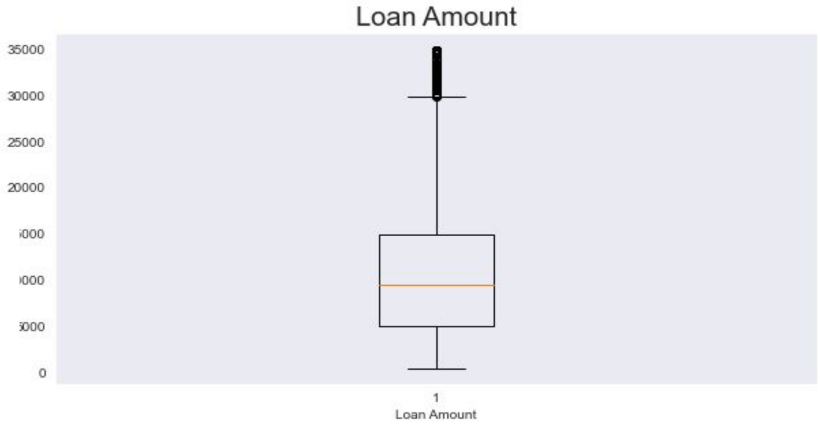


Funded Amount by Inv

Loan Amount

Most of the borrowers applied for loan amount between 5000 and 15000

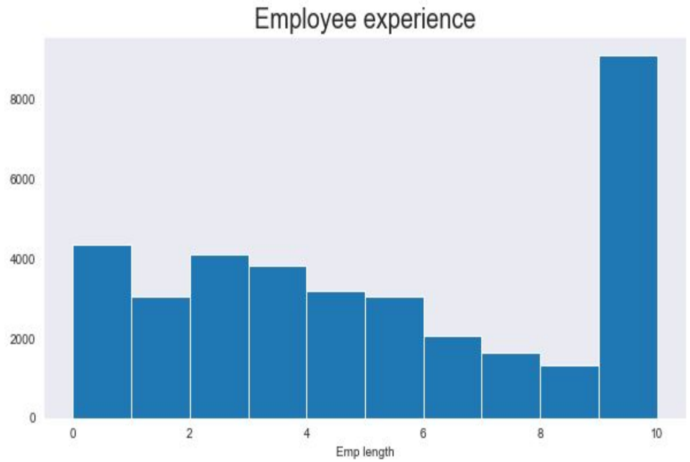




Most of the borrowers received the loan amount between 5000 and 15000 There is a small change in the average amount applied and received by the borrower

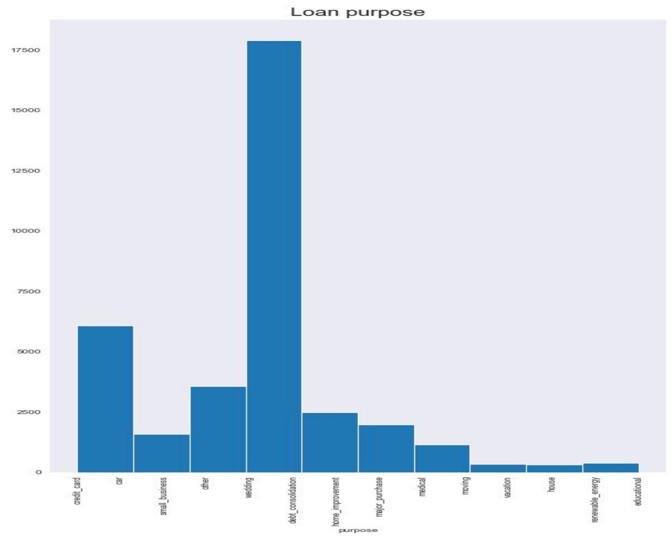
Employee Experience

Most of the loan applicants have 10+ years of experience



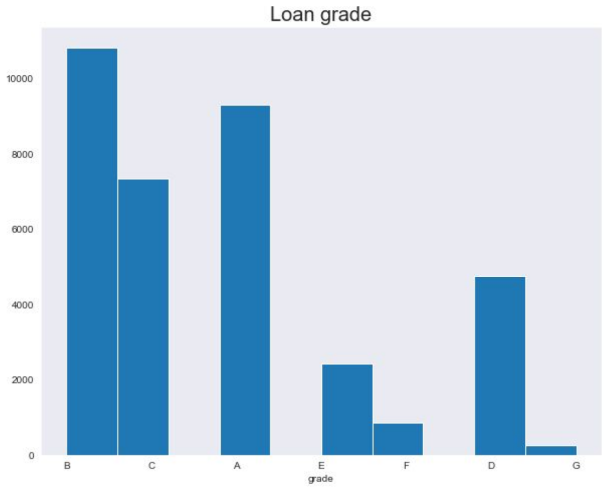
Loan Purpose

Most of the loan applicants are applying with the debt consolidation as the purpose.



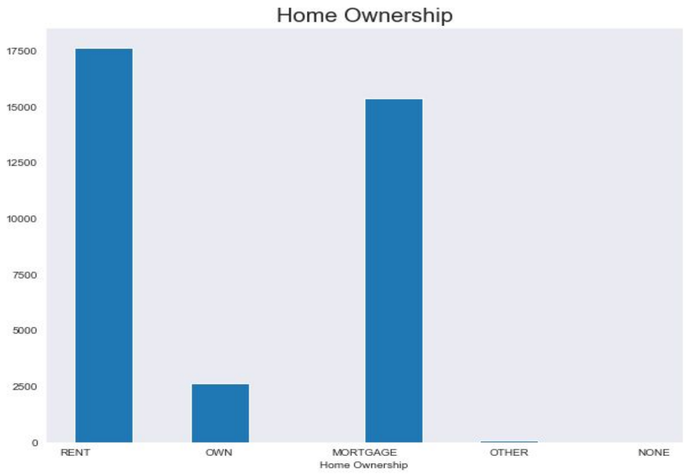
Grade

Most of the applicants are in A,B and C grades



Home Ownership

The borrowers having RENT and MORTGAGE as home ownership are the top applicants

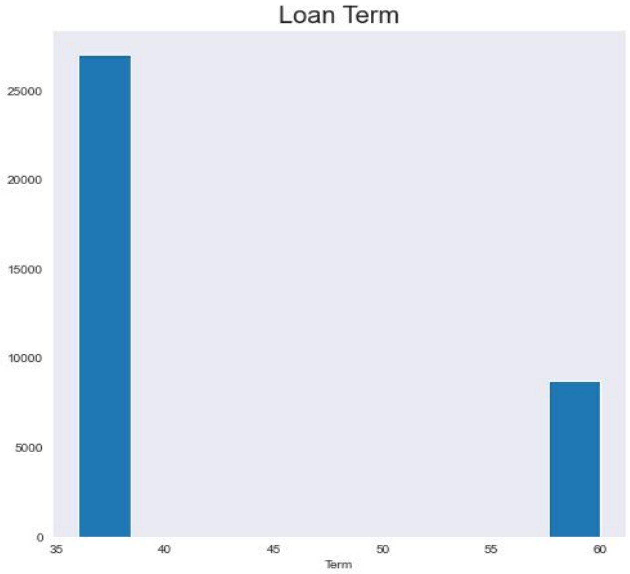


Verification Status

Most of the applicants' income is not verified by LC

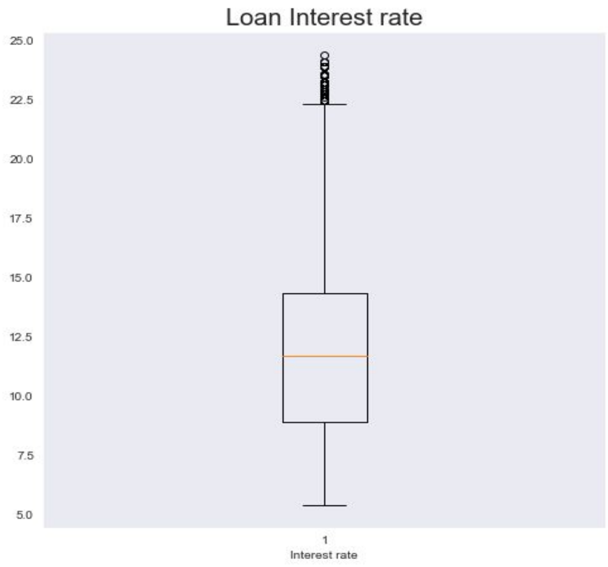
Loan Term

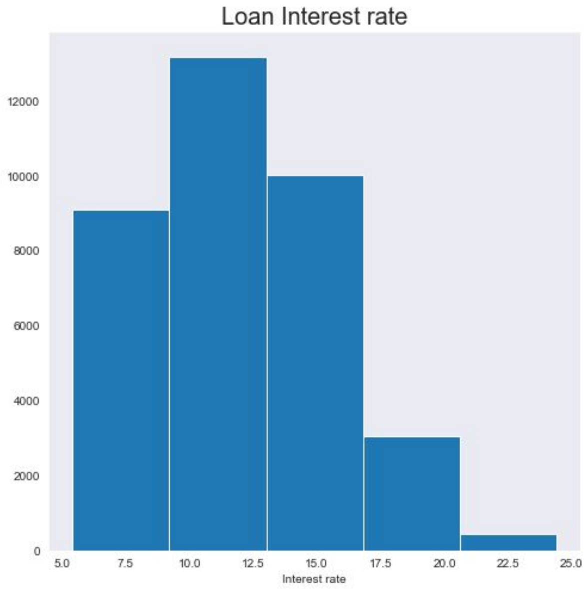
Most of the loans are having 36 months as loan term



Loan Interest Rate

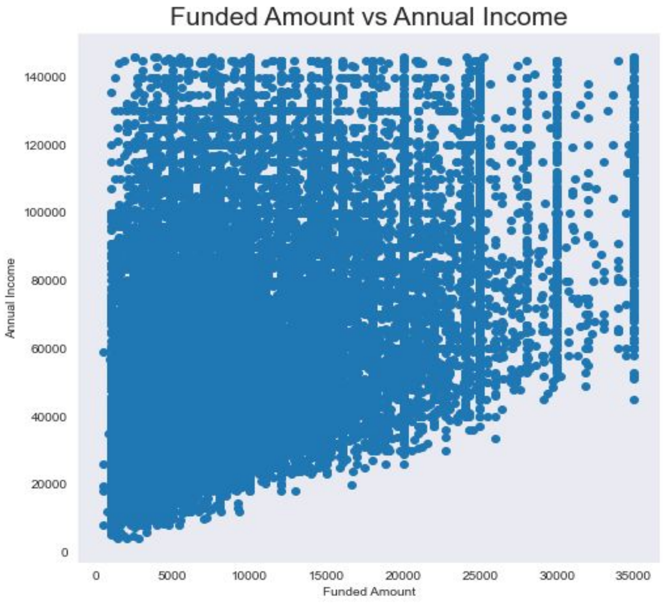
The average interest rate falls in between 9 and 14





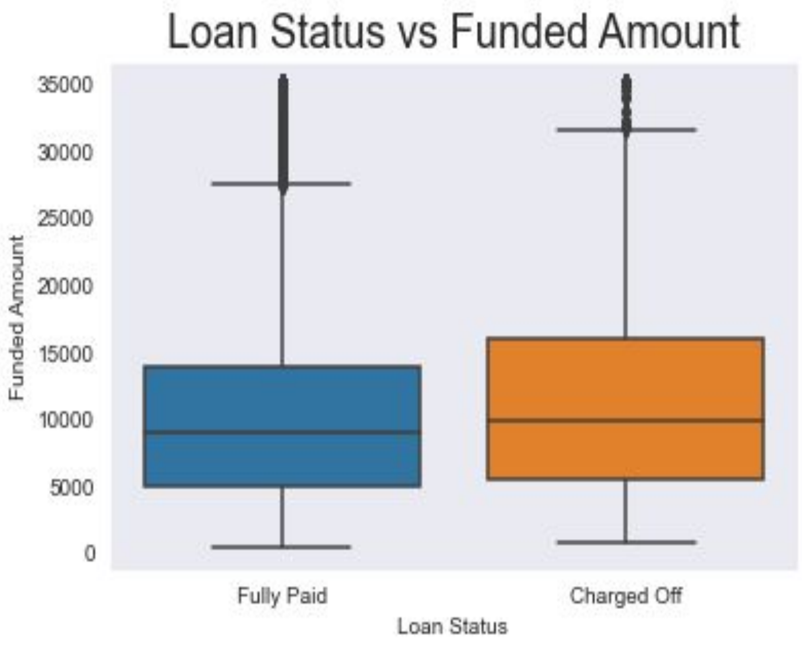
Funded amount vs. Annual income

The loan amount/funded amount increases with the increase in the annual income



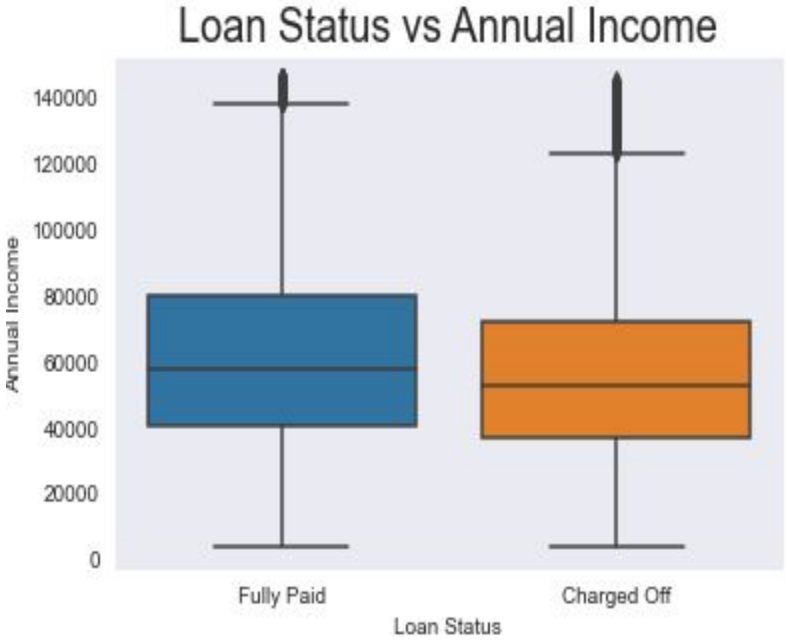
Loan status vs. Funded amount

The average funded amount for charged off applicants is more than that of Fully paid applicants



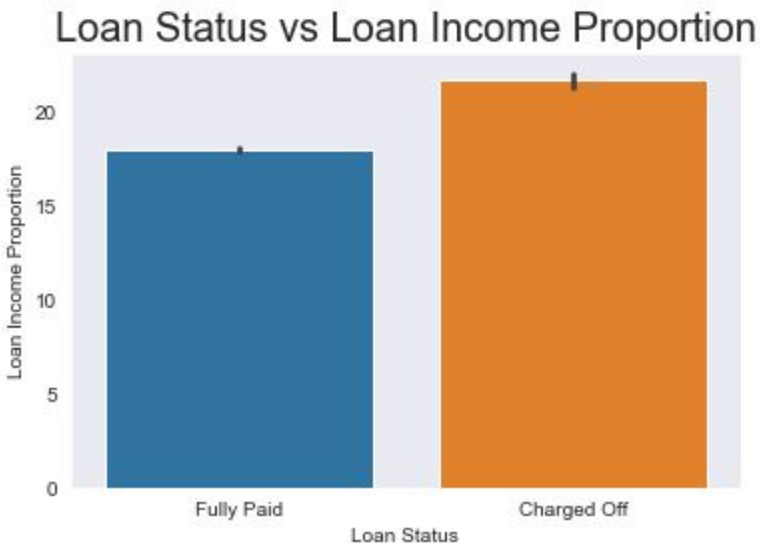
Loan status vs Annual income

The average annual income for charged off applicants is less than that of Fully paid applicants



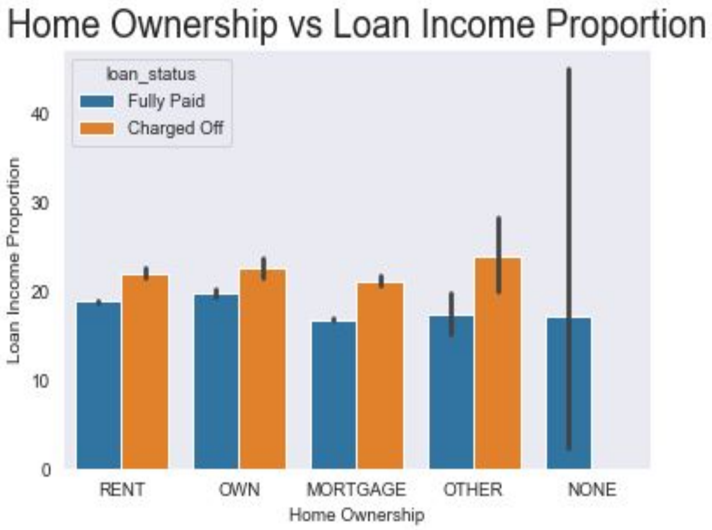
Loan status vs loan income proportion

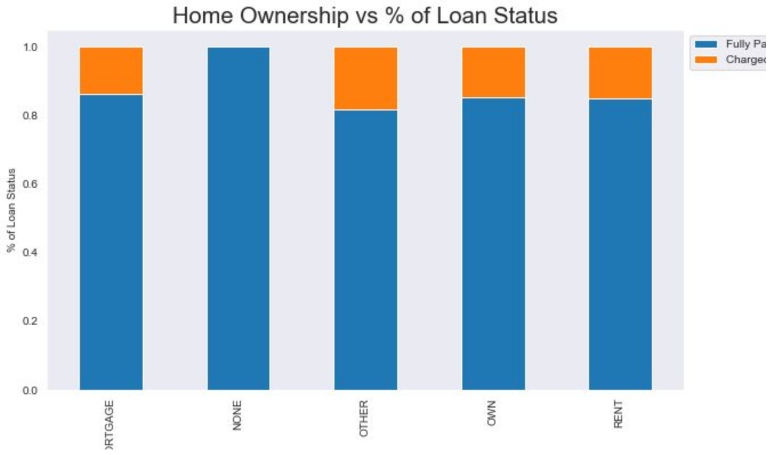
The chances of defaulting are more when the funded amount to income proportion is more.



Home Ownership vs Loan Status/Loan Income Proportion

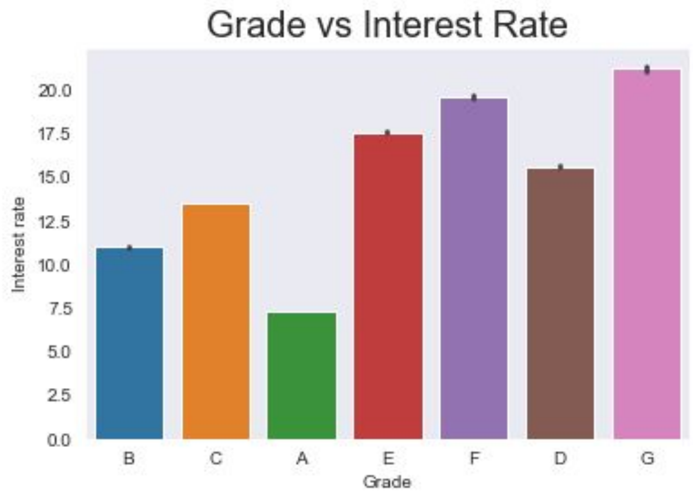
20% of the applicants are likely to default more in any category of home ownership except NONE category.





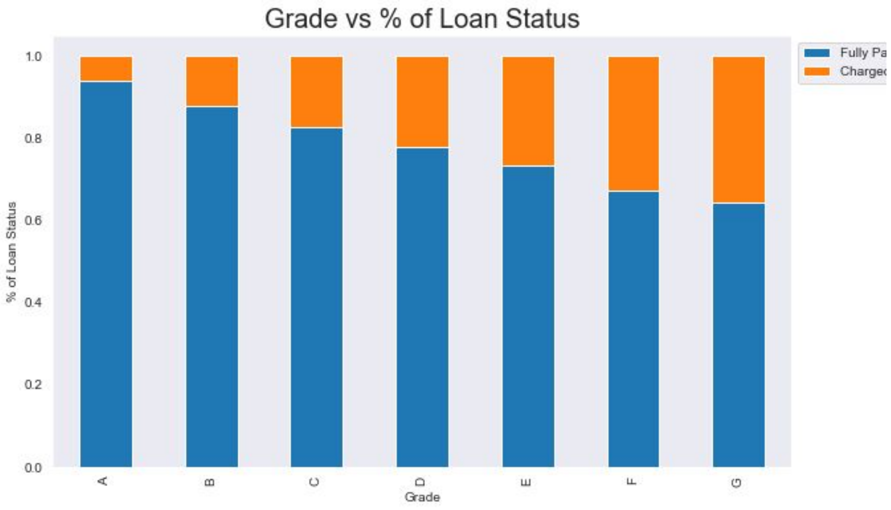
Grade vs. Interest rate

The interest rate is higher for grades F and G so the loans within these grades are risky



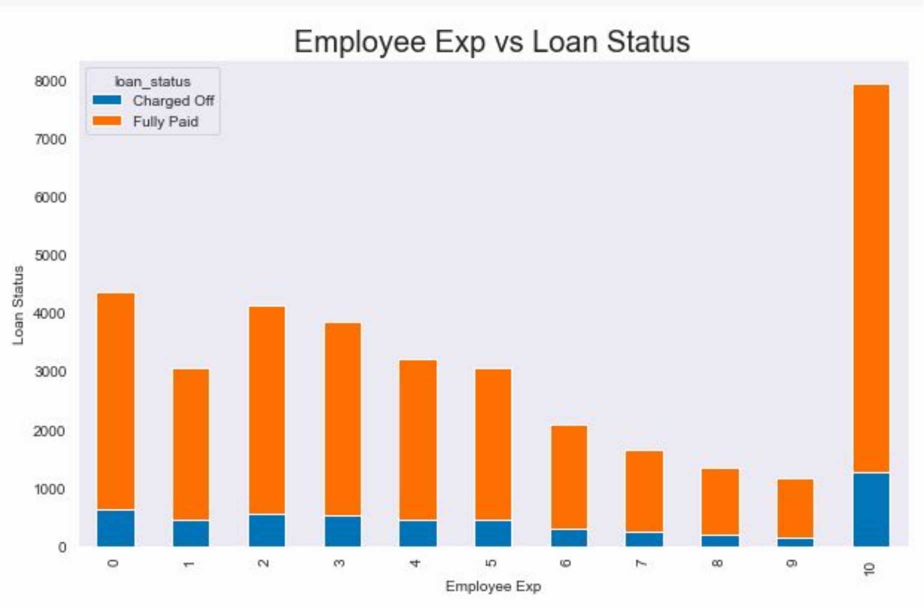
Grade vs. loan status

The chances are more to default with the loans in the grades like E,F and G grades



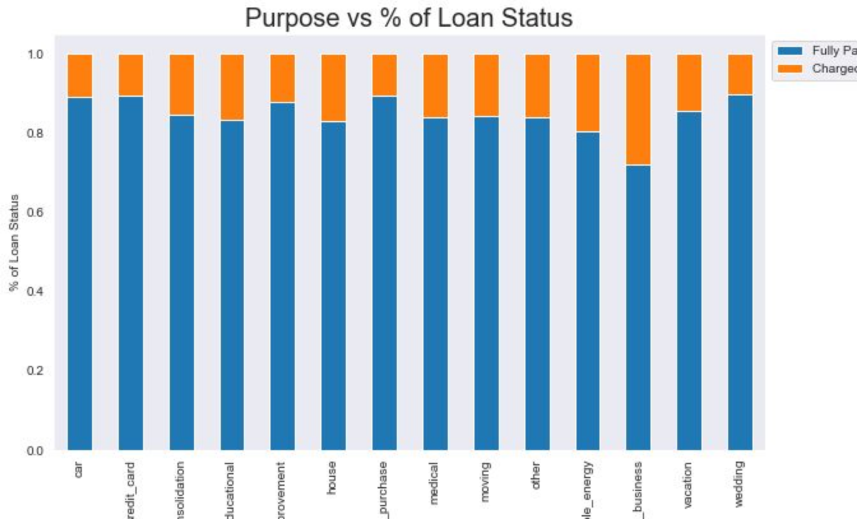
Employee experience vs Loan status

There are higher chances of an employee with 10+ years of experience when compared with the experience of other employees



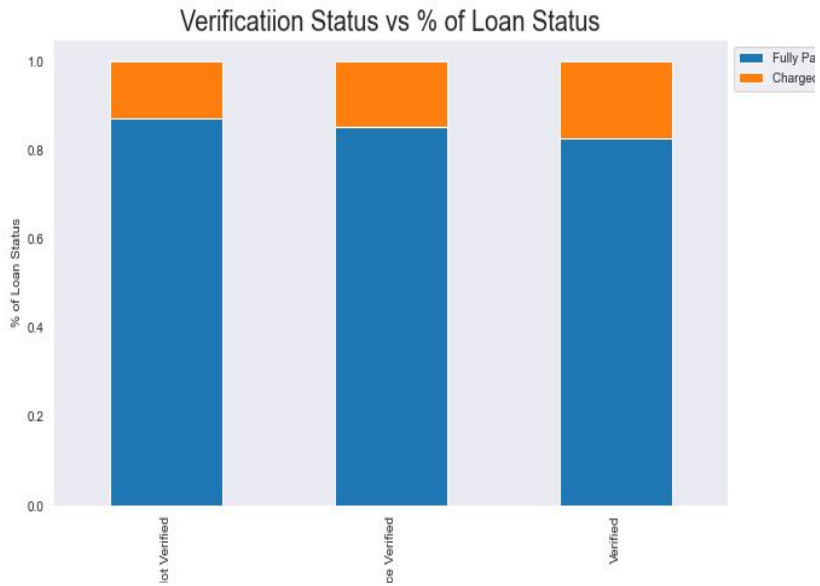
Purpose of loan vs. status

The applicants applying for small business are having good chances to get defaulted.



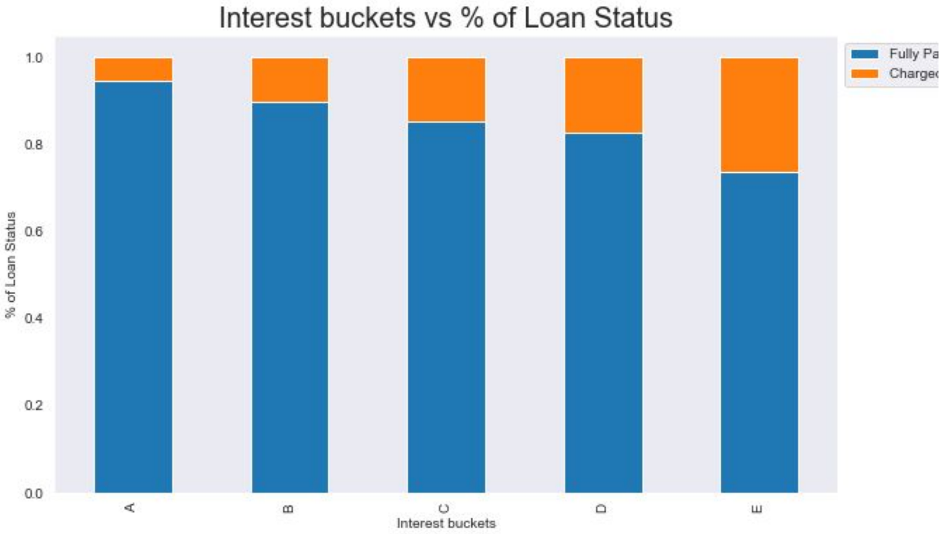
Verification status vs. % of loan status

10-15% of the applicants are likely to default in any category of the verification status



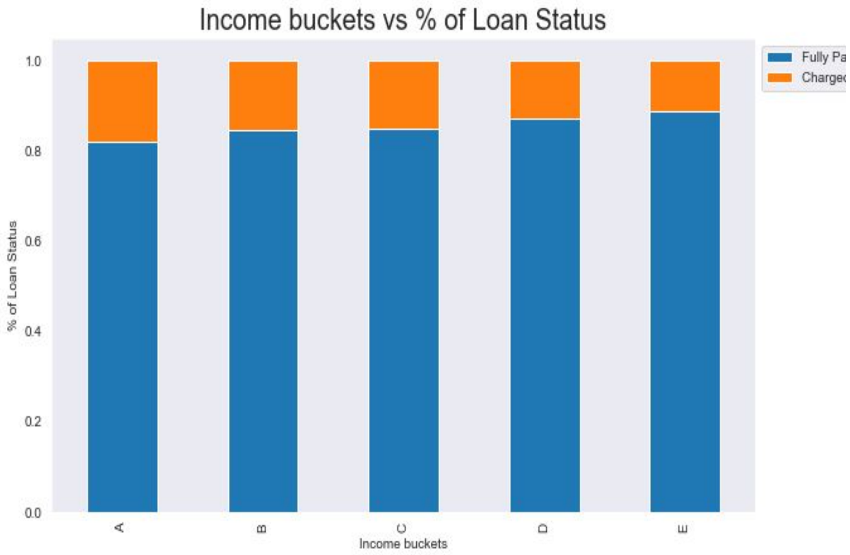
Interest bucket vs. % of loan status

The interest data is divided into buckets(A<B<C<D<E). The higher the interest rates, the higher the chance for defaulting



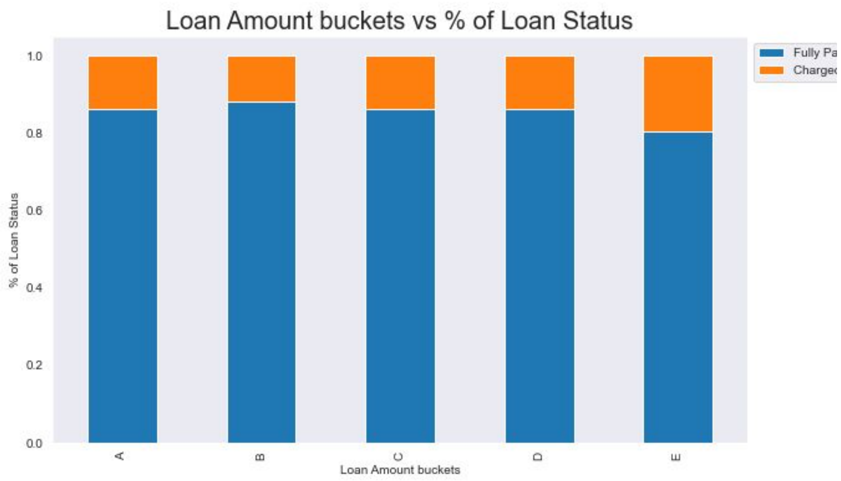
Income bucket vs. % of loan status

The income data is divided into buckets(A<B<C<D<E). The lower income applicants are likely to default more than that of higher income applicants



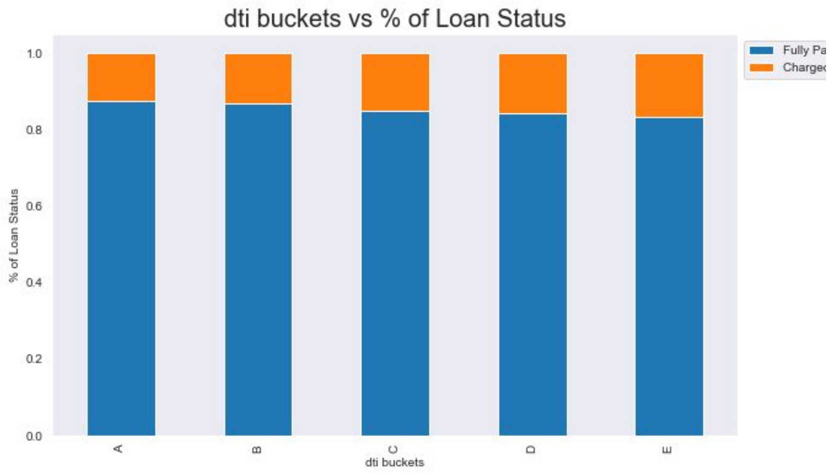
Loan amount bucket vs. % of loan status

The loan amount/funded amount data is divided into buckets(A<B<C<D<E). The applicants with higher funded amount are likely to default more than that of applicants with lower funded amount



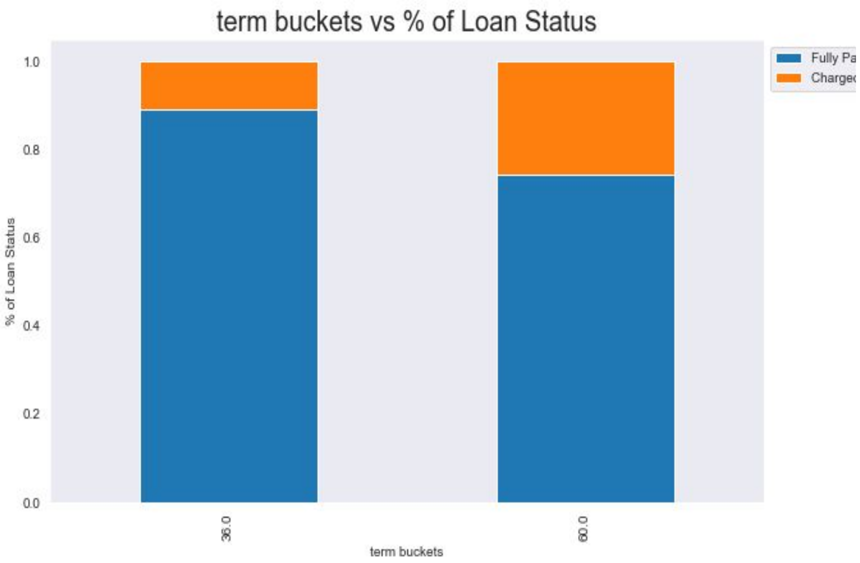
dti buckets vs. % of loan status

The dti data is divided into buckets(A<B<C<D<E). The applicants with higher dti are likely to default more than that of applicants with lower dti



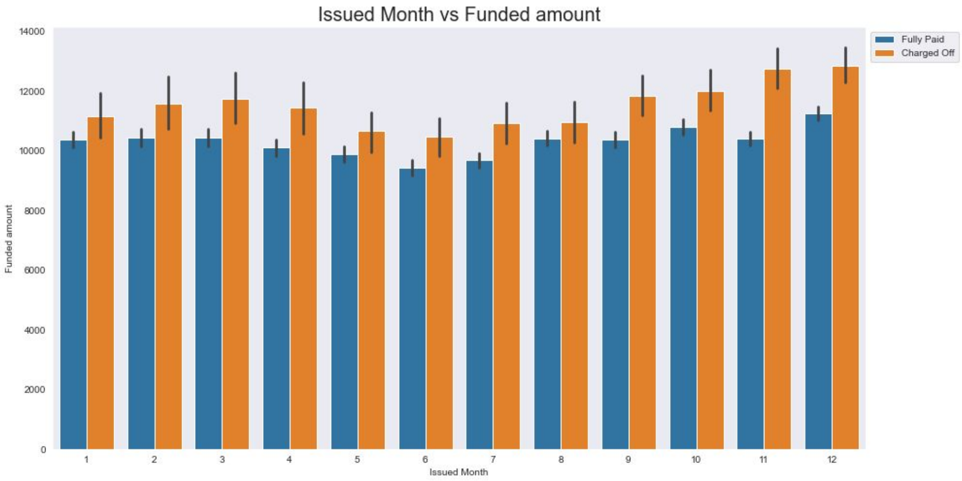
Term buckets vs. % of loan status

The term data is divided into buckets(A<B<C<D<E). The loans with 60 months term are likely to default more than that of loans with 36 months term



Issued Month vs Loan Amount

The loans issued in Nov and Dec are having higher chances of going default



Recommendations

The following are the recommendations for both the scenarios:

Likely repay the loan when	Likely default the loan when
Loan amount to income proportion is less than 20%	Loan amount to income proportion is more than 20%
The applicants who are given the loans of grade A	The applicants who are given the loans with grade F and G
The loans are given to the less experience applicants	The experience is more
The loans are given for the purpose of car, credit card or wedding	The loans are given for the purpose of small business

Thank you!