LENDING CLUB ASSIGNMENT

SUBMISSION

Name: Siddhika Shetty

Lending Club Case study

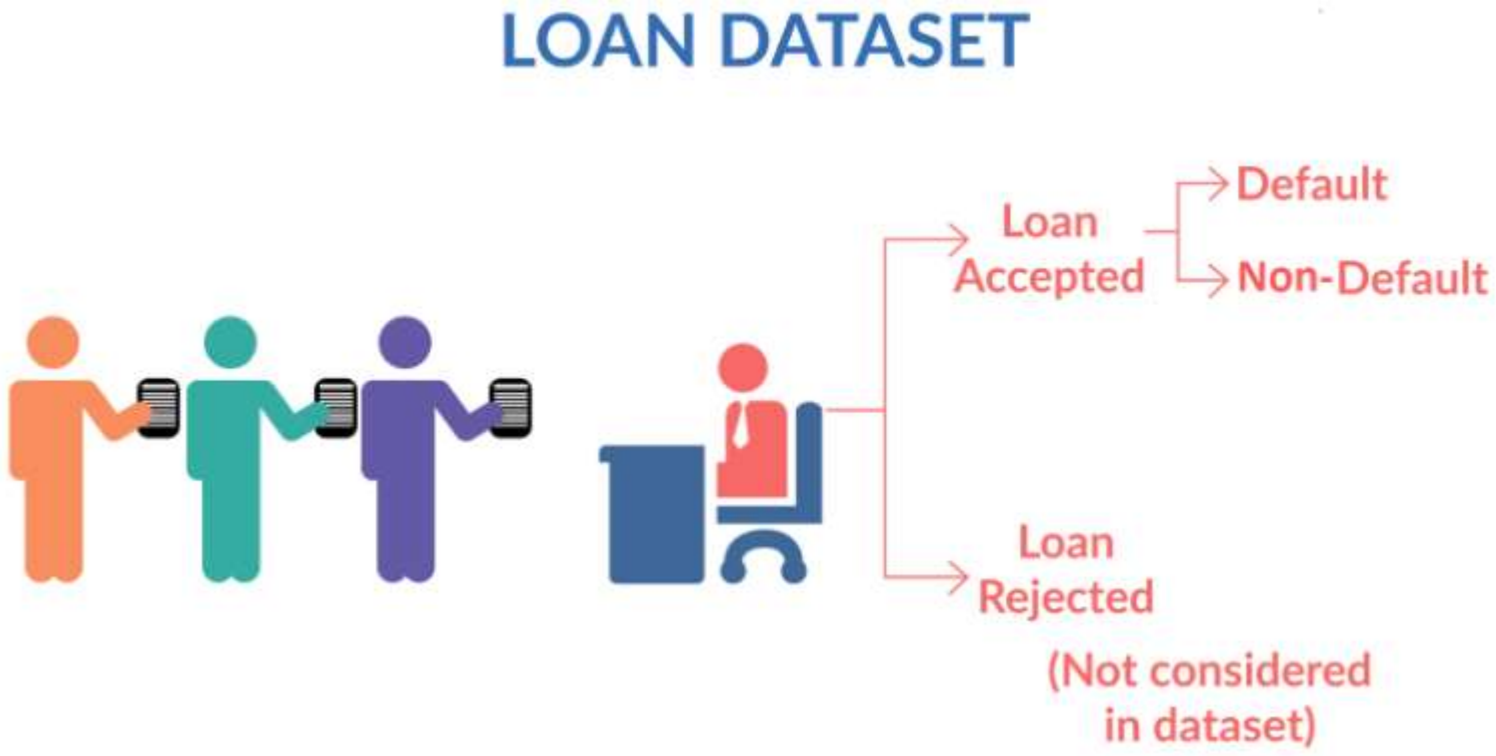
Lending club is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast online interface. .

Lending loans to ‘risky’ applicants is the largest source of financial loss (called credit loss). The credit loss is the amount of money lost by the lender when the borrower refuses to pay or runs away with the money owed.

Business objective:

- The aim is to identify patterns which indicate if a person is likely to default, which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc. Identification of such applicants using EDA is the aim of this case study.

When a person applies for a loan, there are **two types of decisions** that could be taken by the company:
- Accept Loan
- Reject Loan



Data Cleaning

- Remove columns having huge % of null values, or ones that are not required for analysis
- Remove rows not pertaining to the objective of the analysis

Univariate Analysis

- Analyze each column by plotting the distribution of each column

Segmented Univariate Analysis

- Extract insights by conducting univariate analysis on segments on data

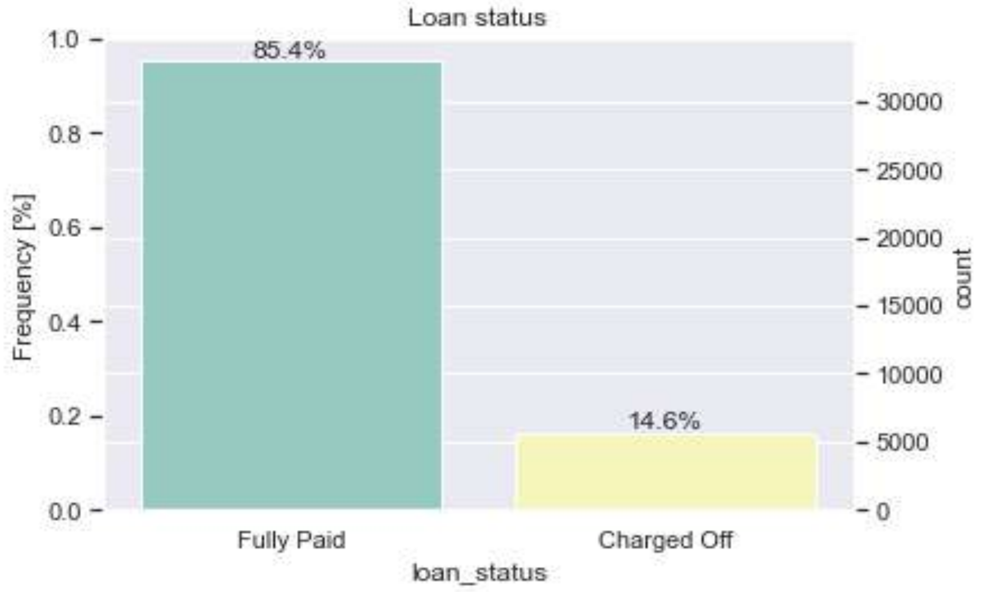
Bivariate Analysis

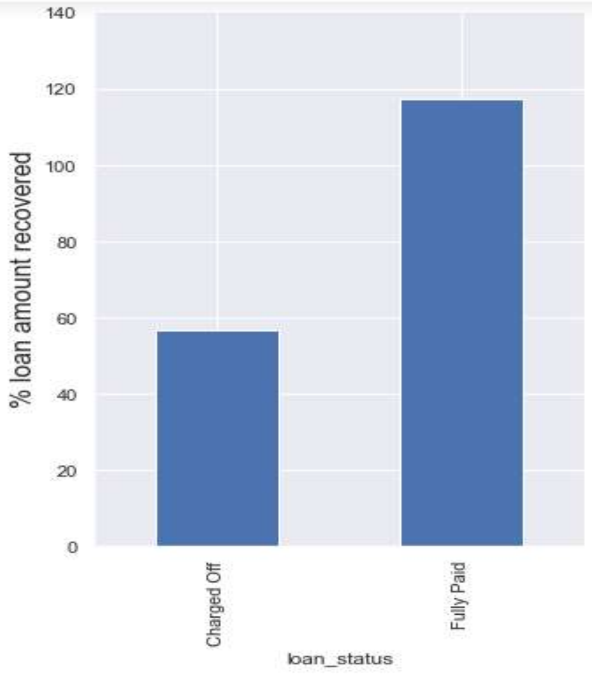
- Correlation analysis to check impact of one variable on another by analyzing plots / graphs of the concerned 2 / more columns

Recommendation

- Summarize insights and observations

Analysis

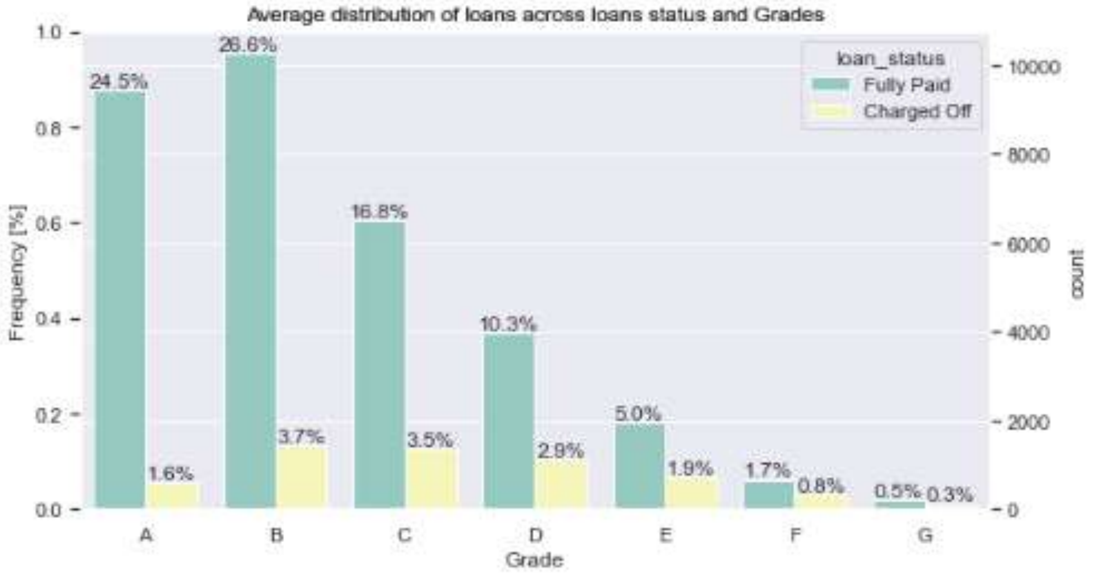


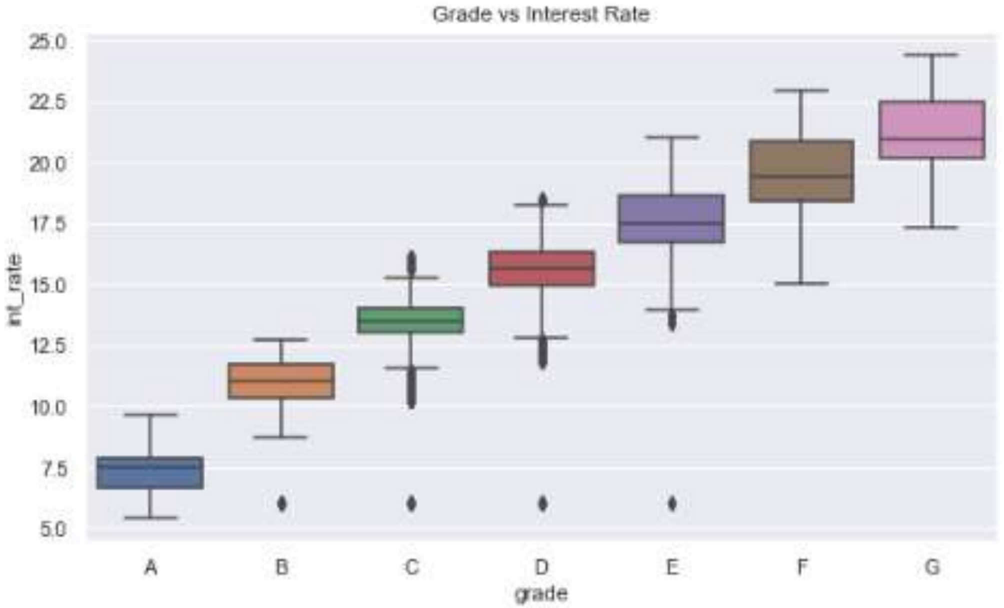


- 85.4% of the loans have status as ‘Fully Paid’ while 14.6% of the loans have turned out to be bad loans with status as charged off

- For charged off loans around 56% of the loan amount has been recovered, while for Fully paid ones more than 100% of loan amount has been recovered

Analysis- Grades



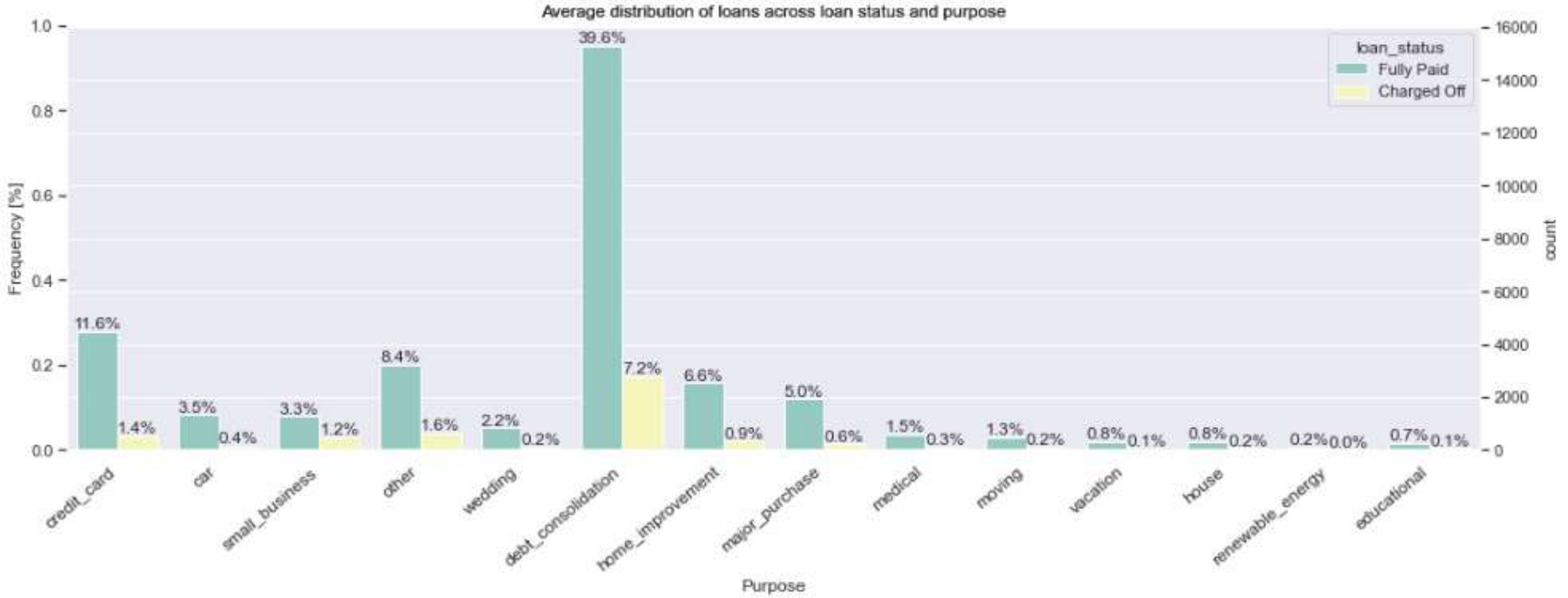


- Majority loans are of grade A & B
- As grades decrease (considering grades A to be higher than B) the number of loans granted decrease and so does the number of bad loans

For lower grades the interest rate is higher

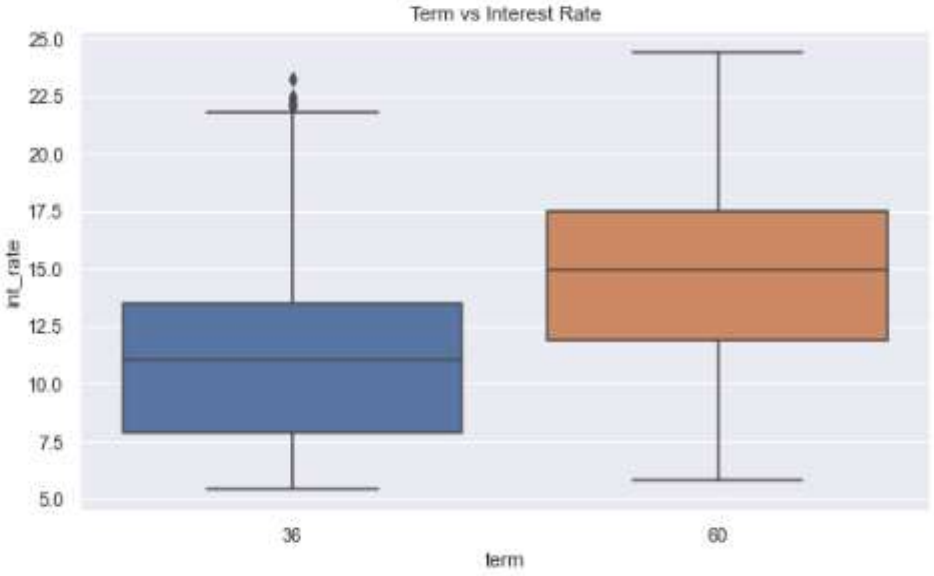
Overall grading system is working fine in identifying bad loans (Lower the grades – higher the probability of it turning into a bad loan)

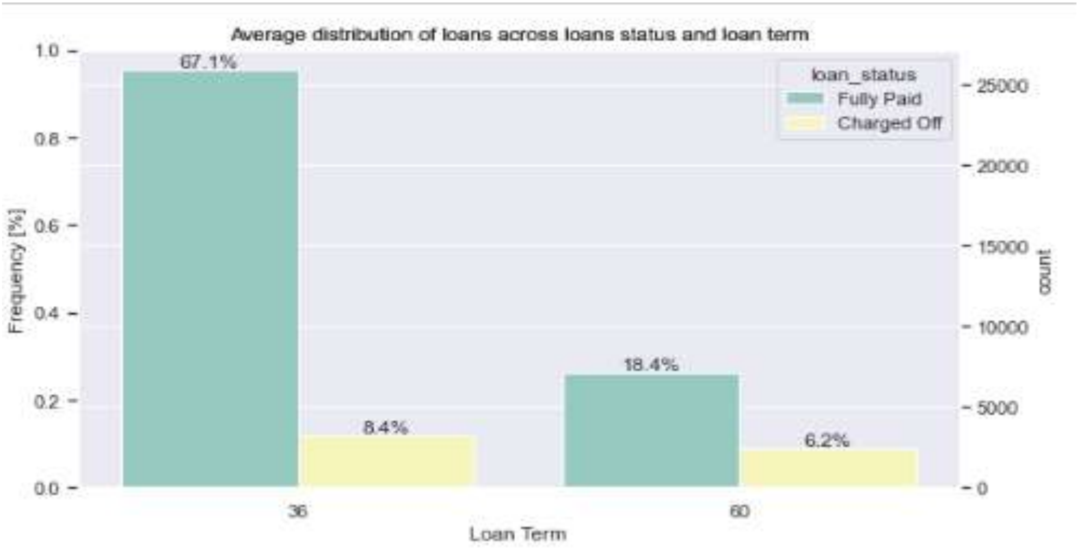
Analysis - Purpose



- Almost 47% of the loans have been taken for the purpose of debt consolidation
- Debt reconsolidation has the highest % of loans that have defaulted

Analysis - Term

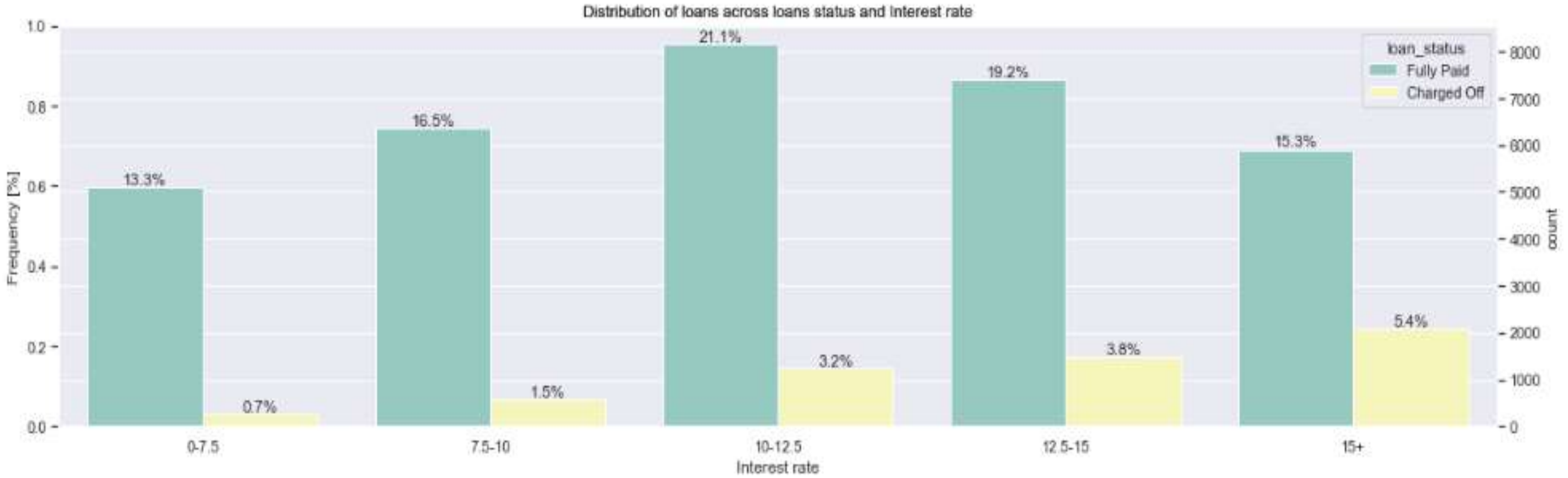


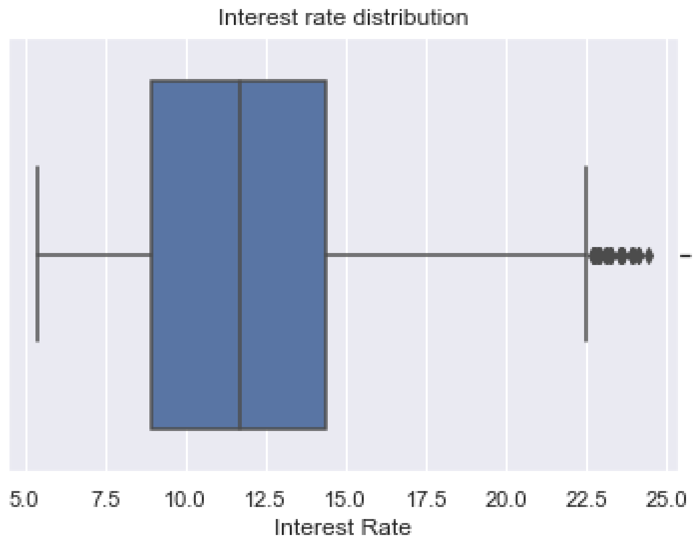


- Loans with longer term have higher interest rates

- Majority loans have lower term of 36 months and looking at ratio of fully paid and charged off loans, we can conclude that proportion of charged off loans is higher for term of 60 months

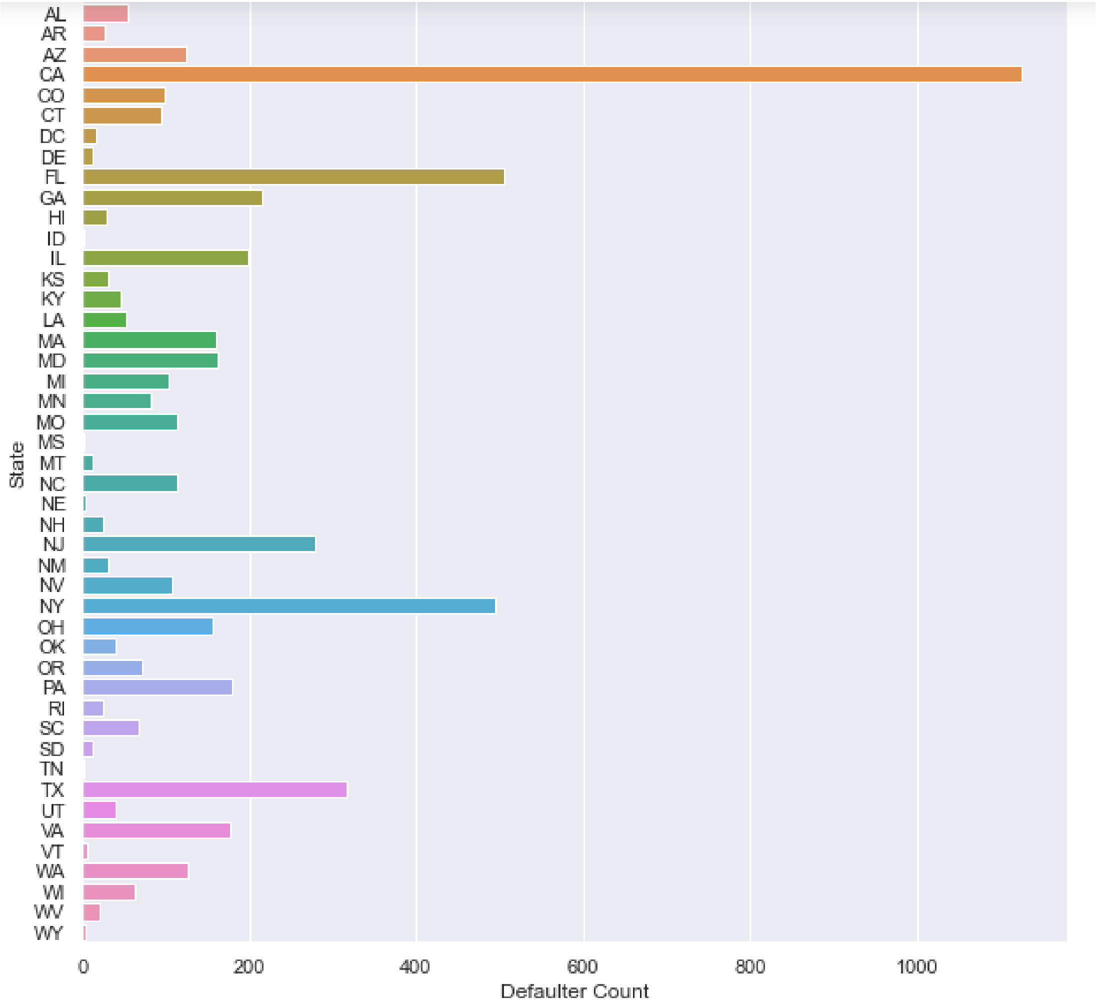
Analysis – Loan status & interest rate





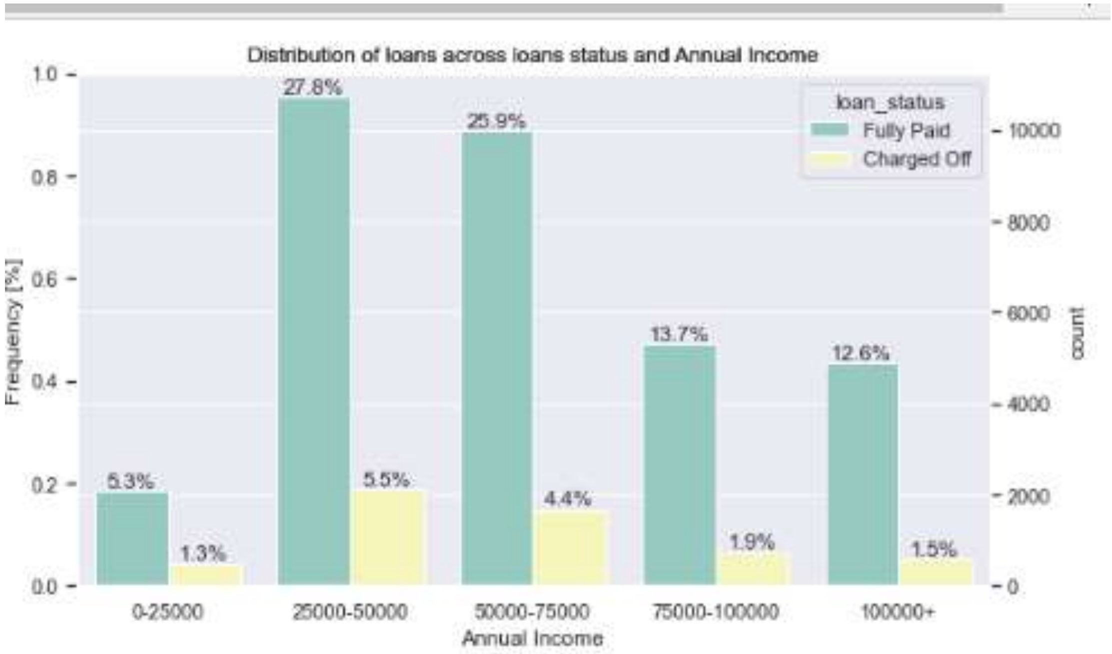
- On an average most of the interest rates are between 9 and 15%. While in some cases customers took interest rate as high as 22.5%
- As interest rate increases percentage of bad loans also increases

Analysis - State



- California has the highest number of loan defaulters
- It is followed by Florida, New York and Texas in terms of defaulter count

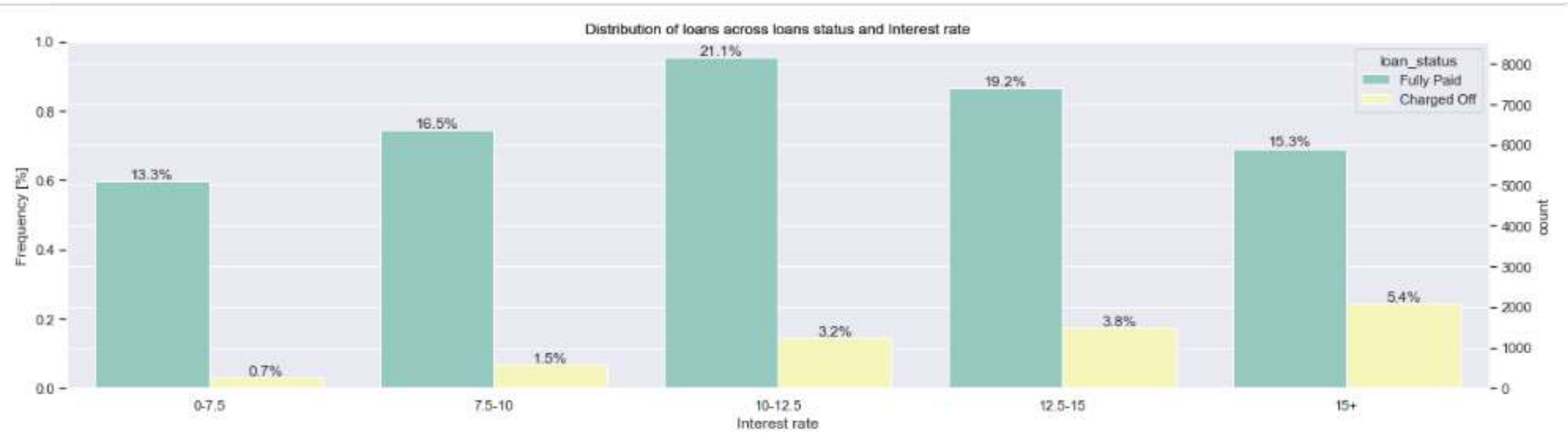
Analysis – Annual Income



- With the exception of 0-25000, the percentage of charged off loans increases as annual income decreases.

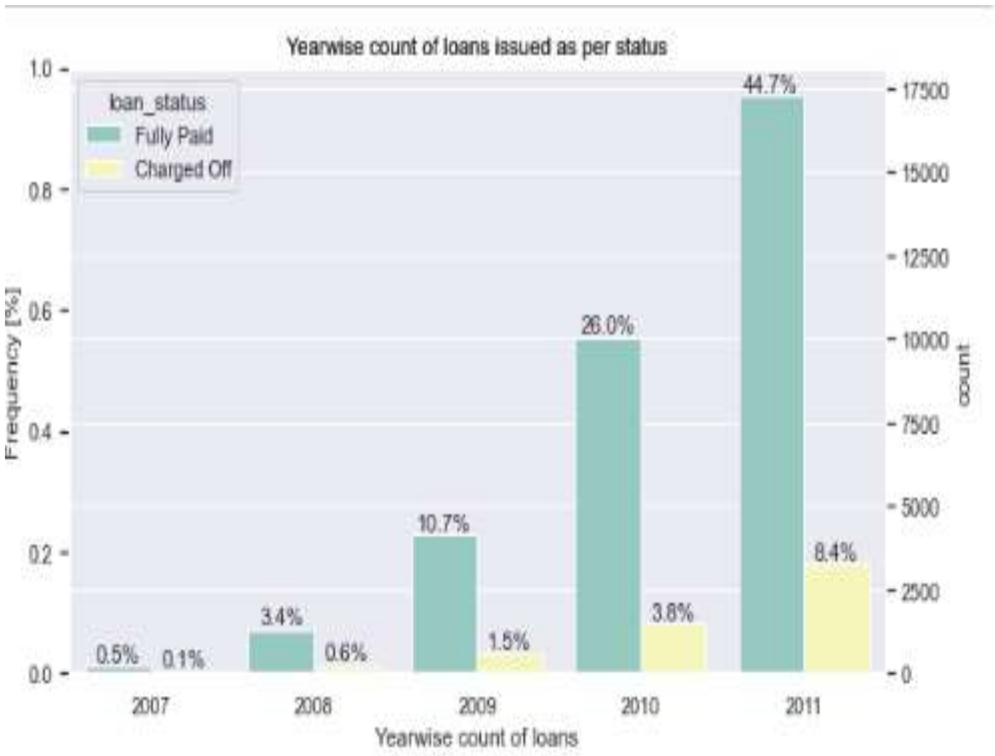
- Noticeable point is that people in lower income group 0-25000 and in highest income group 100000+ both have approximately same defaulter percentage

Analysis – Interest Rate



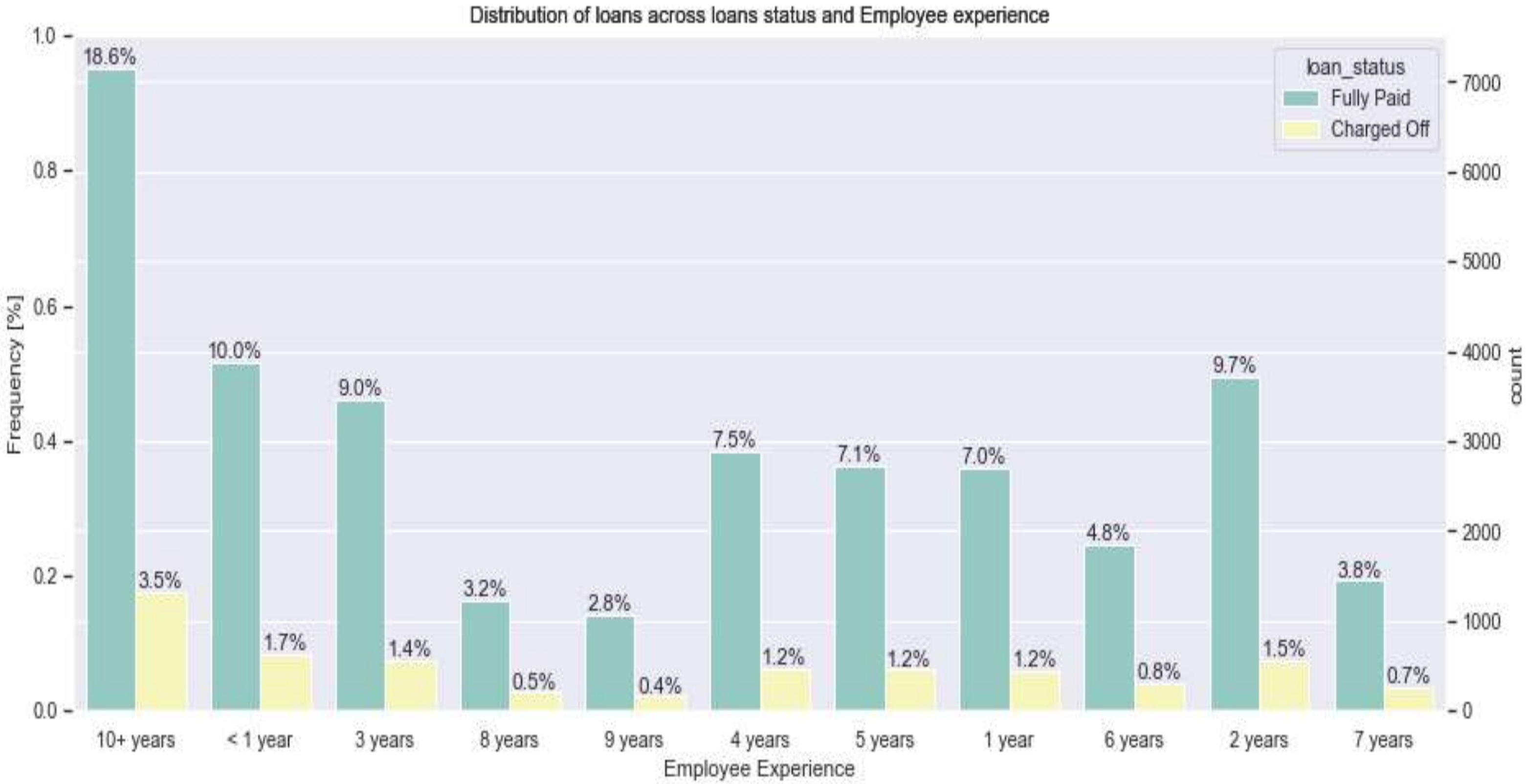
- As interest rate increases number of bad loans also increase

Analysis – Loan Issue year



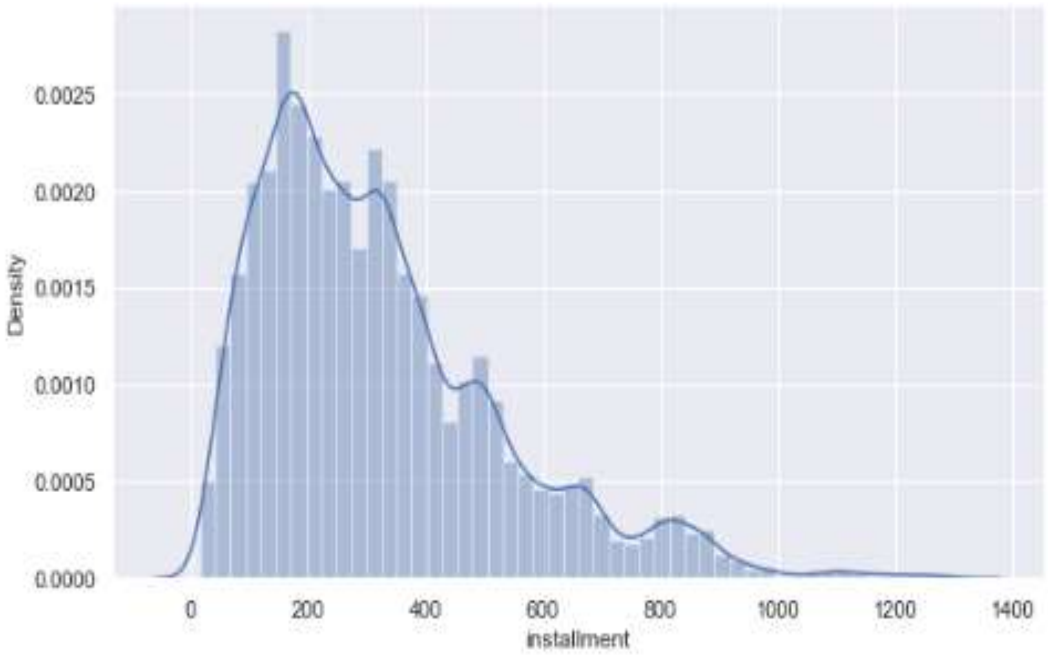
- With every new year number of loans has increased
- Number of defaulters has also increased with every passing year

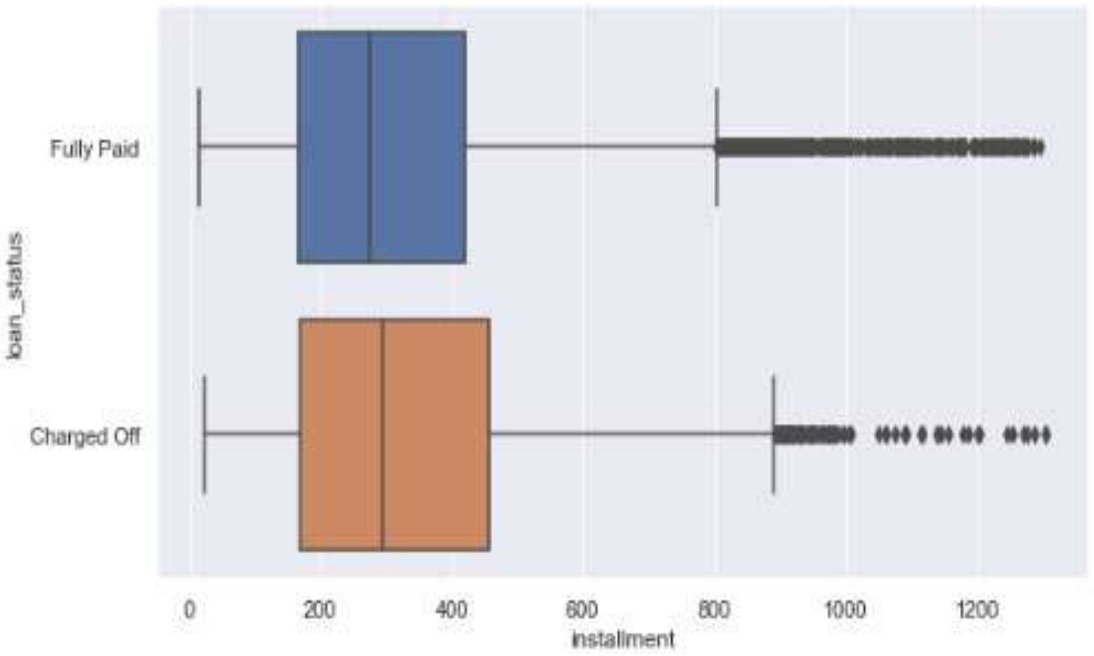
Analysis – Employee Experience



- Employees having more than 10+ years of experience have higher frequency of getting loans, followed by lower experience group like 3 years and less experience
- Number of defaulters is the maximum for the 10+ years experience borrowers

Analysis - Installment





- Charged off loans have higher installments

Recommendation

- Loans with tenure of 60 months have higher interest rate and such loans tend to default more. So a point to consider would be to reduce the interest rate on such loans
- States like California, Florida, New York, Texas have maximum number of defaulters. The scrutiny process for customers in these states should be more stringent to verify eligibility of loan customers
- Current grading system is helpful in identifying probable defaulters
- Loans taken with purpose of debt reconciliation should be verified properly before the loan is approved
- Borrowers having 10+ years of experience tend to default more. So additional checks needs to be added for such cases to verify if the borrower will be able to pay off the loan

Conclusion

Variables to be considered for loan application:

- Term and interest rate
- Purpose of loan
- Borrower’s experience
- Grade
- Installment
- Annual income