



International
Institute of Information
Technology Bangalore



GRAMENER CASE STUDY

PRATEEK GHARAT

KARTHIK KRISHNAN

MANISH CHANDRA

ANAND JOSHI

ABSTRACT

The aim of Gramener Case Study is to develop a basic understanding of risk analytics in banking and financial services and understand how data is used to minimize the risk of losing money while lending to customers using EDA.

The data given contains the information about past loan applicants and whether they 'defaulted' or not. 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.

PROBLEM SOLVING METHODOLOGY



BUSINESS UNDERSTANDING

A Consumer finance company 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.

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.

The data given contains the information about past loan applicants and whether they 'defaulted' or not. 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.

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

Loan Status	Description
Loan Accepted (Fully Paid)	Applicant has fully paid the loan (the principal and the interest rate)
Loan Accepted (Current)	Applicant is in the process of paying the instalments.
Loan Accepted (Charged-Off)	Applicant has not paid the instalments in due time for a long period of time
Loan Rejected	The company rejected loan (as candidate didn't meet their requirements)

DATA UNDERSTANDING

Pre Loan Attributes

loan_amnt
funded_amnt
funded_amnt_inv
term
int_rate
installment
grade
sub_grade
purpose
earliest_cr_line
inq_last_6mths

Post Loan Attributes

out_prncp
out_prncp_inv
total_pymnt
total_pymnt_inv
total_rec_prncp
total_rec_int
total_rec_late_fee

Applicant Attributes

emp_title
emp_length
home_ownership
annual_inc
zip_code
addr_state

Background Check Attributes

dti
delinq_2yrs
open_acc
pub_rec
revol_bal
revol_util
total_acc
recoveries
collection_recovery_fee
last_pymnt_d
last_pymnt_amnt
last_credit_pull_d
pub_rec_bankruptcies

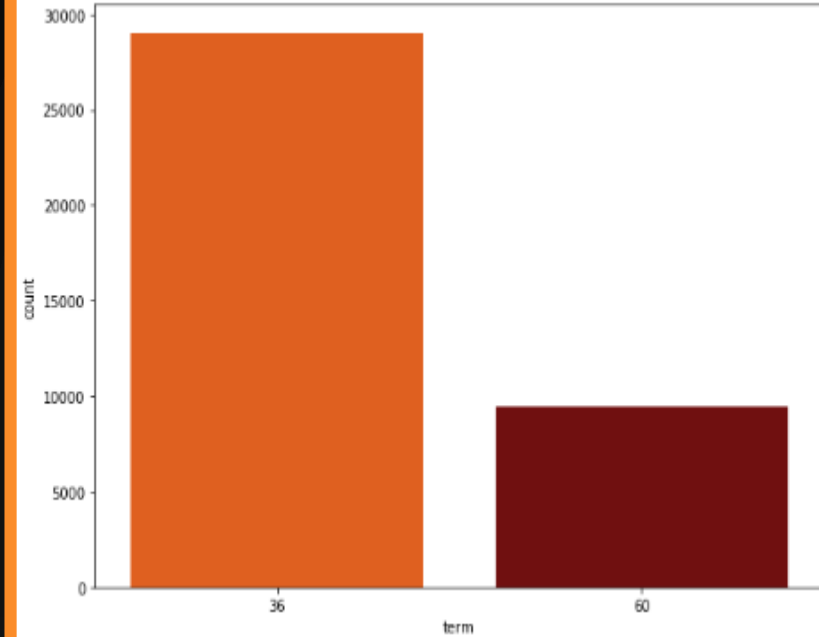
We have categorised all the attributes as Pre Post Loan, Applicant and background check to better understand the data.

DATA CLEANING & PREPARATION

- Identification of missing values and duplicate in data set.
- Treatment of missing values by dropping columns having missing values more than 90%.
- Dropping off few irrelevant columns (which contain extra information related to loan) for analysis.
- Cleaning of special characters in rows of specific columns and updating data type as per analysis.
- Converting the date to date time format and extracting months and year.
- As aim is to identify patterns which indicate if a person is likely to default so we would drop the values for loans which are current as the applicant is in the process of paying the instalments.
- Creating the new columns and bucketing the values (creating slots) for easy understanding of data for analysis.

UNIVARIATE ANALYSIS

Term of loan



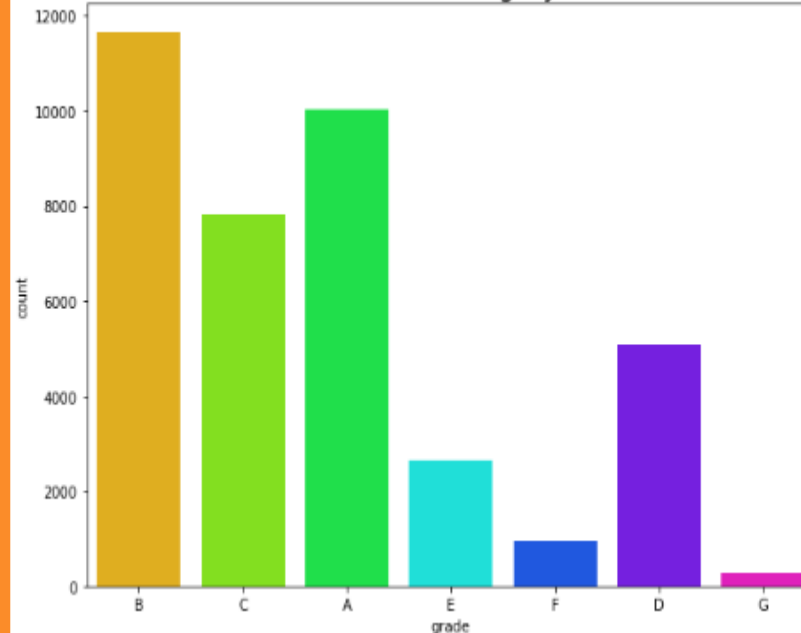
It can be seen that major loans are applied grade B followed by grade A.



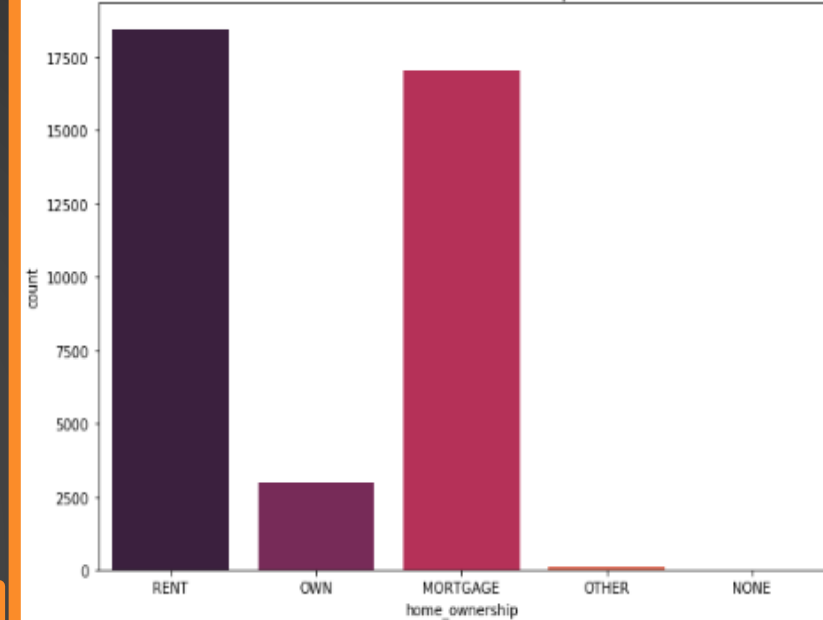
It can be seen that applicant majorly apply for short term i.e 36 months loans.



Grade Category



Home Ownership

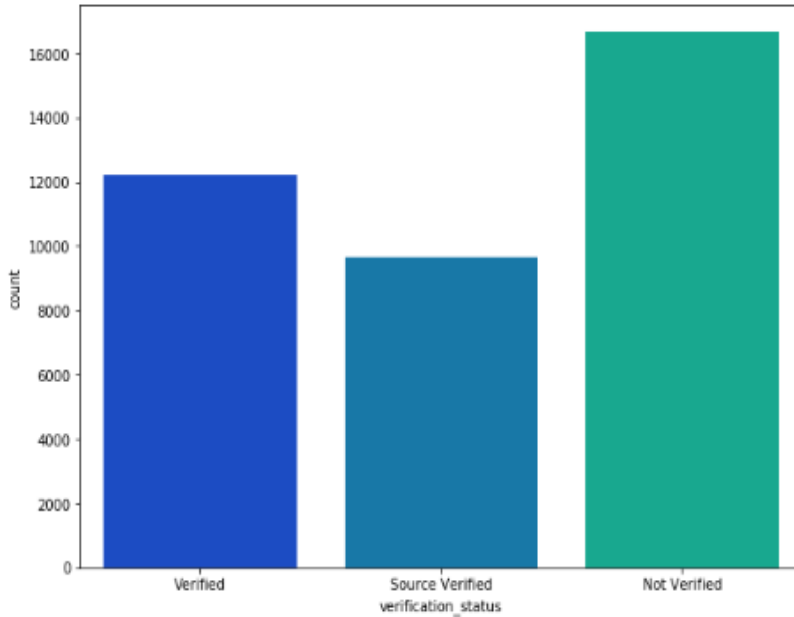


It can be seen that major loans applicant stay in rented homes.



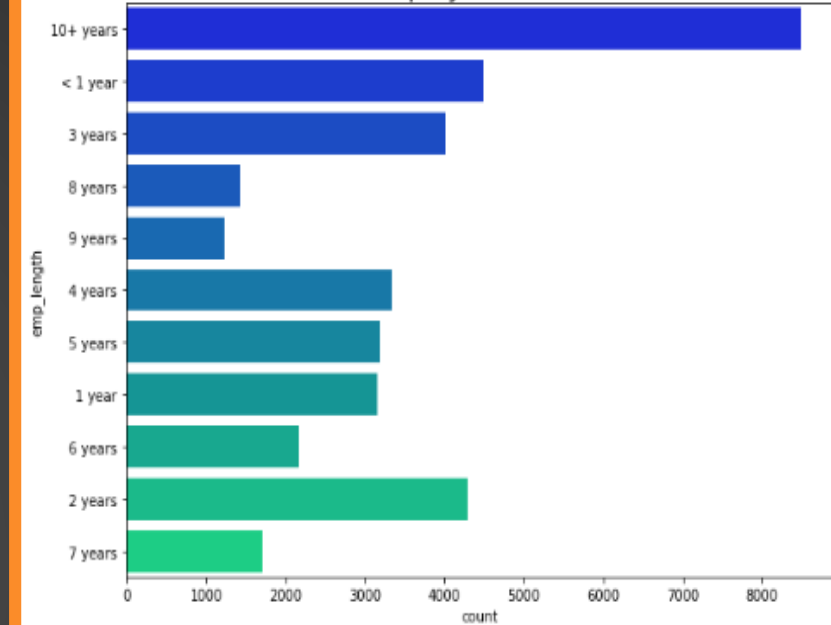
UNIVARIATE ANALYSIS

Verification Status



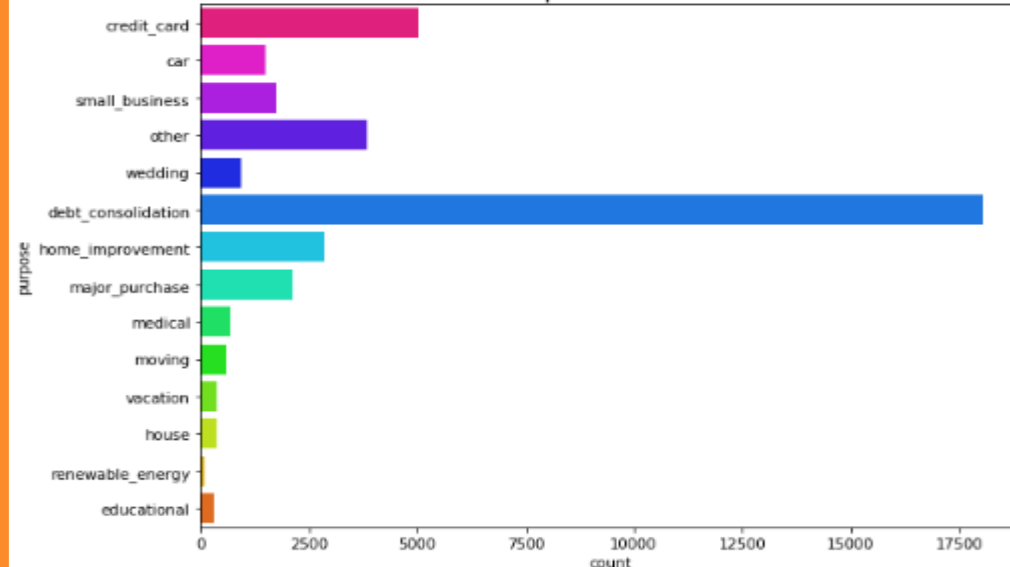
It can be seen that major loans are applied for debt consolidation.

Employee Tenure



It can be seen that majority of loans are not verified.

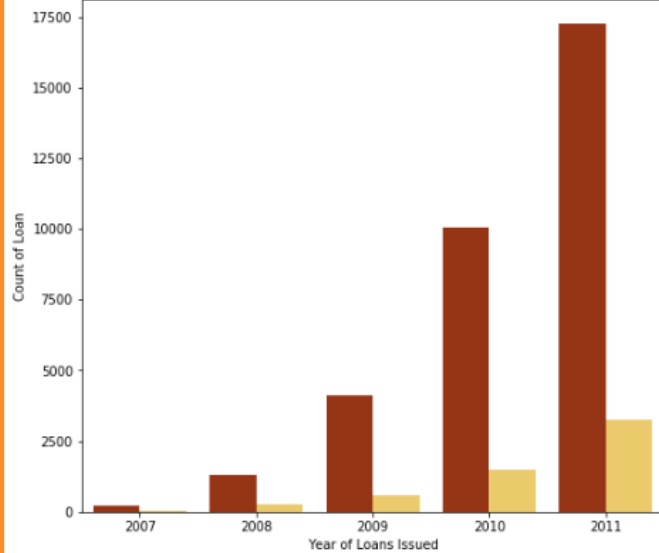
Purpose of loan



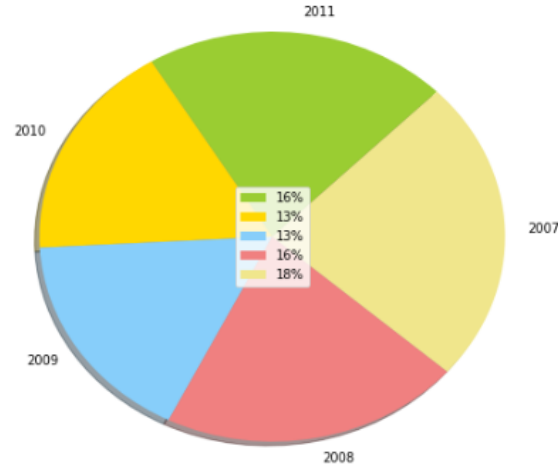
It can be seen that applicant with 10+ year of experience apply mostly for loans.

SEGMENTED UNIVARIATE ANALYSIS

Yearly loan status

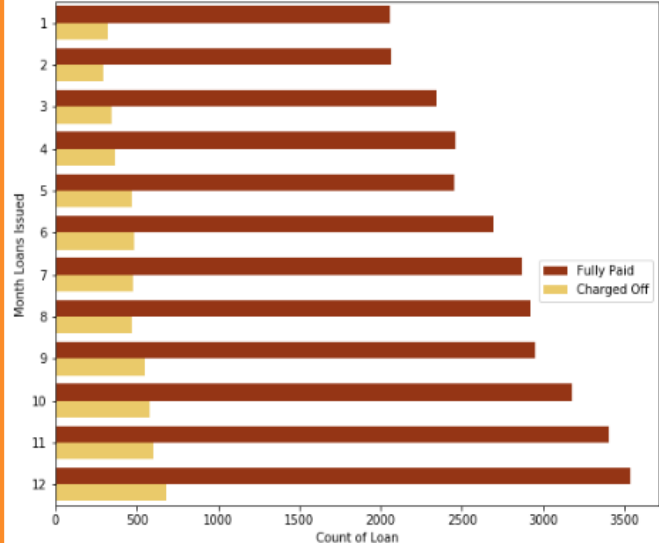


Percentage of Charged Off Loans on yearly basis

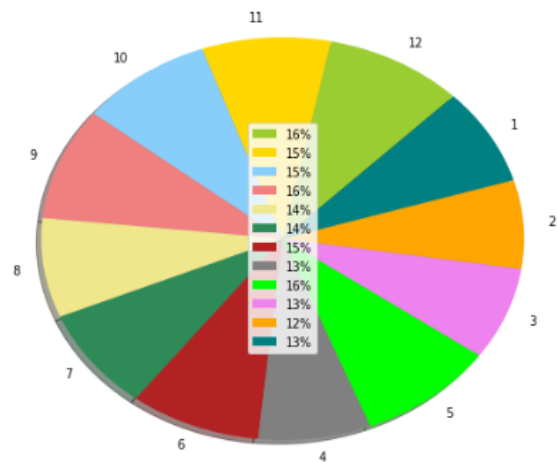


In the year 2007, 17.92% of loans were charged off as compared to other years.

Monthly loan status



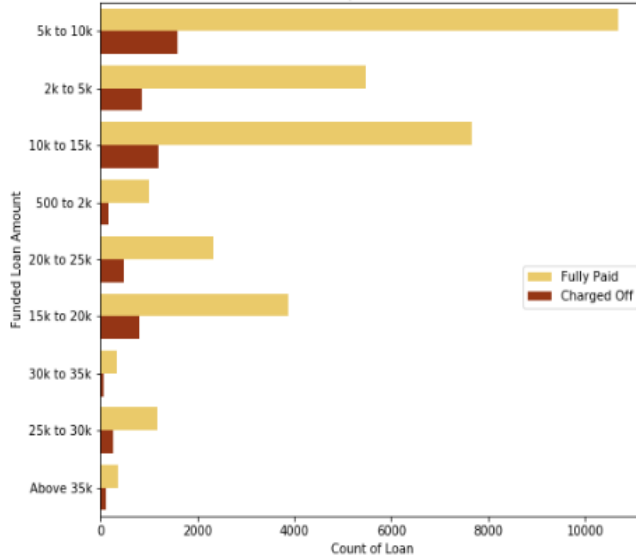
Percentage of Charged Off Loans on monthly basis



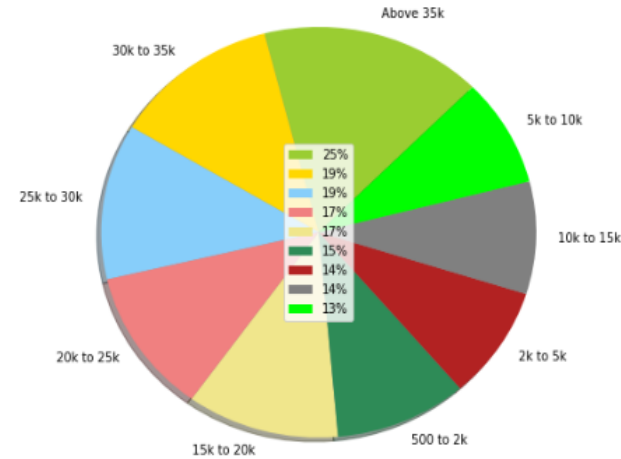
In the month of December 16.07% of loans were charged off.

SEGMENTED UNIVARIATE ANALYSIS

Loan status as per amount funded

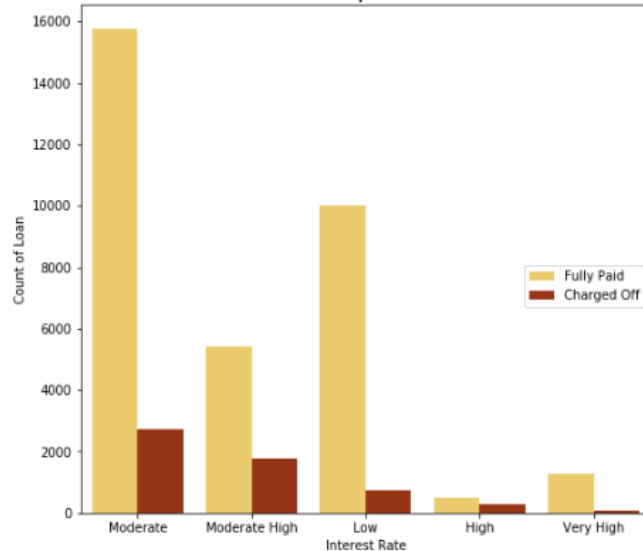


Percentage of Charged Off Loans as per amount funded

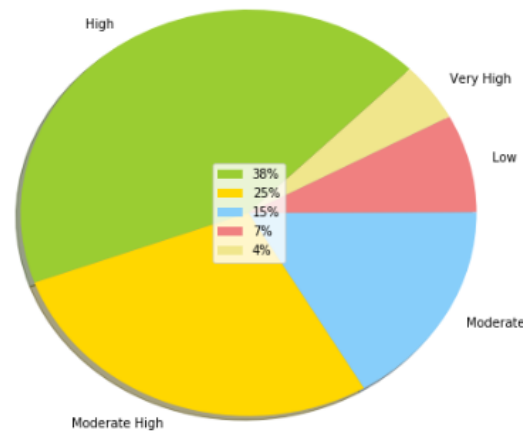


Loan amount which are funded above 35000 had a major charged off proportion of 25.25%.

Loan status as per Interest rate



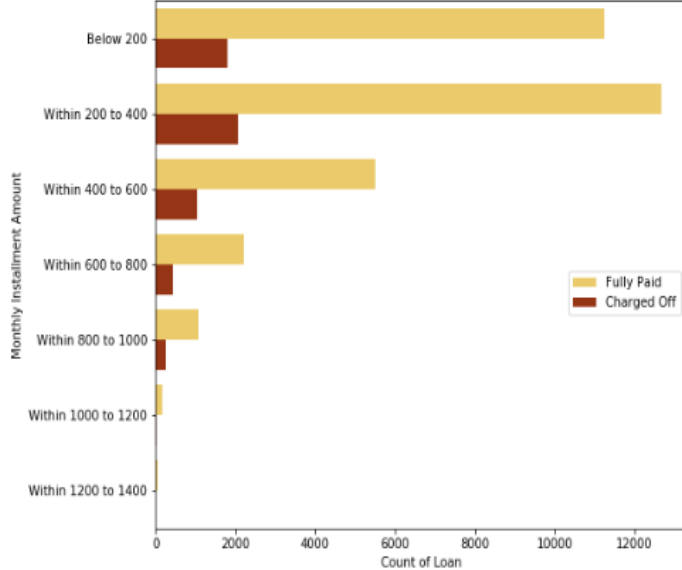
Percentage of Charged Off Loans as per Interest rate



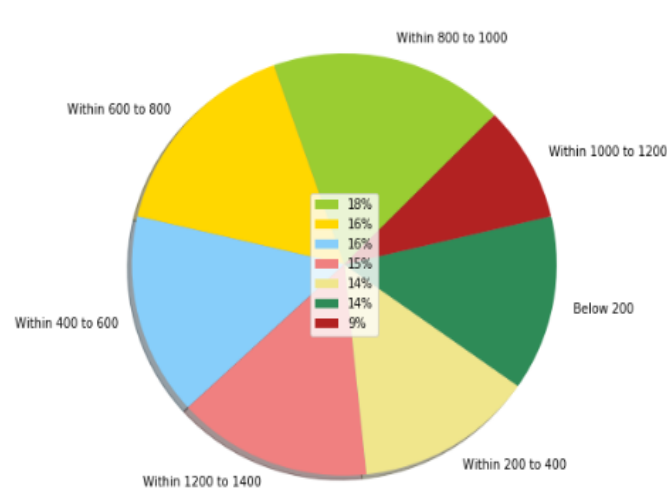
Loans with Interest rate between 20% to 25% had a major charged off proportion of 38.33%.

SEGMENTED UNIVARIATE ANALYSIS

Loan status as per Monthly Installment Amount

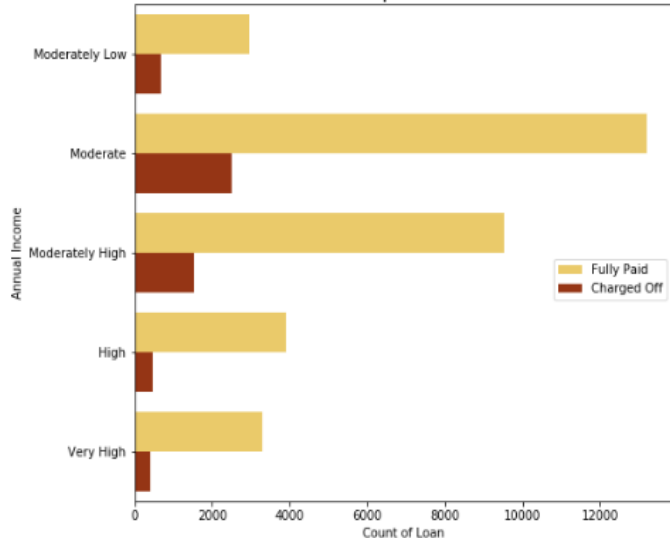


Percentage of Charged Off Loans as per Monthly Installment Amount

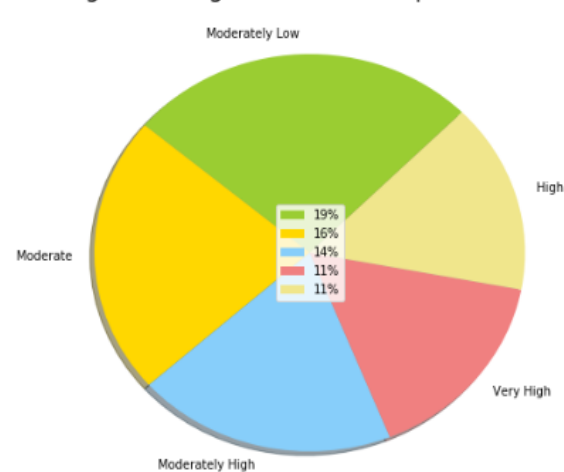


Loans with monthly installment amount between 800 to 1000 had a major charged off proportion of 18.41%.

Loan status as per Annual Income



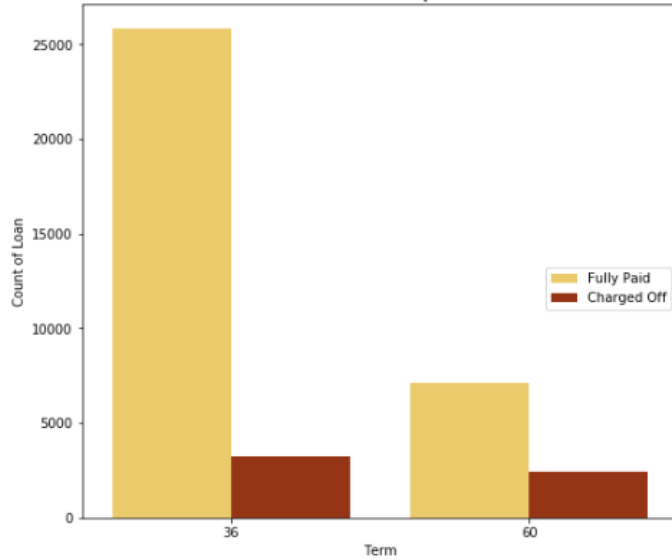
Percentage of Charged Off Loans as per Annual Income



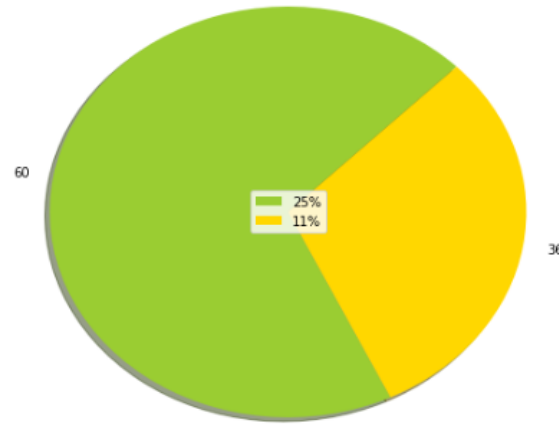
Loans of applicant with annual income 10000 to 30000 had a major charged off proportion of 18.61%.

SEGMENTED UNIVARIATE ANALYSIS

Loan status as per term

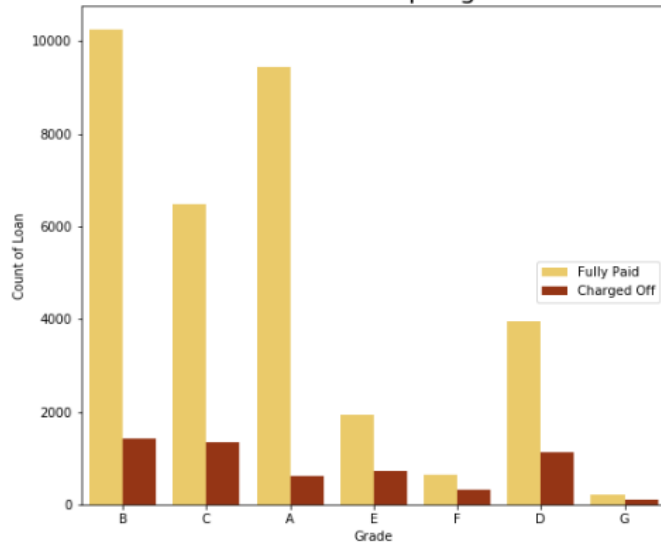


% of Charged Off Loans as per term

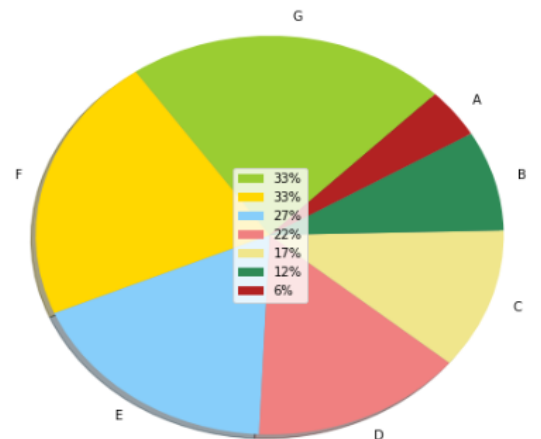


60 month term had a major charged off proportion of 25.29%.

Loan status as per grade



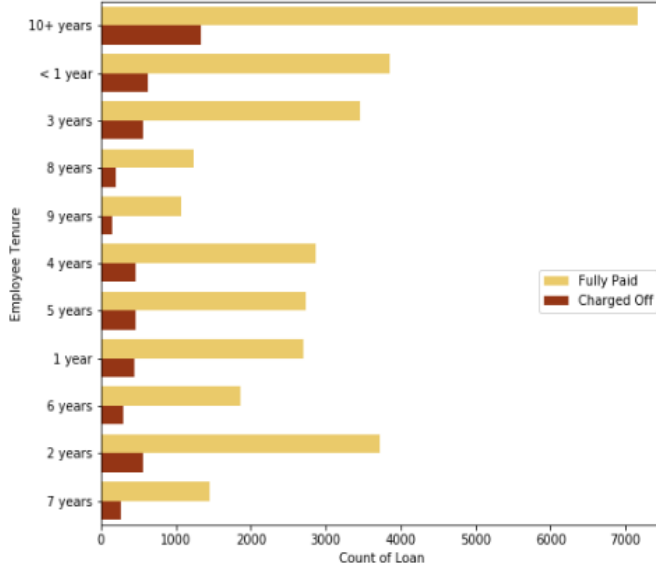
% of Charged Off Loans as per grade



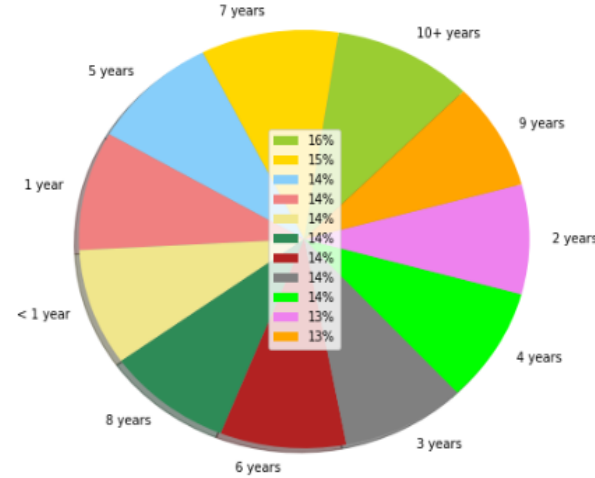
Grade G had a major charged off proportion of 33.33%.

SEGMENTED UNIVARIATE ANALYSIS

Loan status as per employee tenure

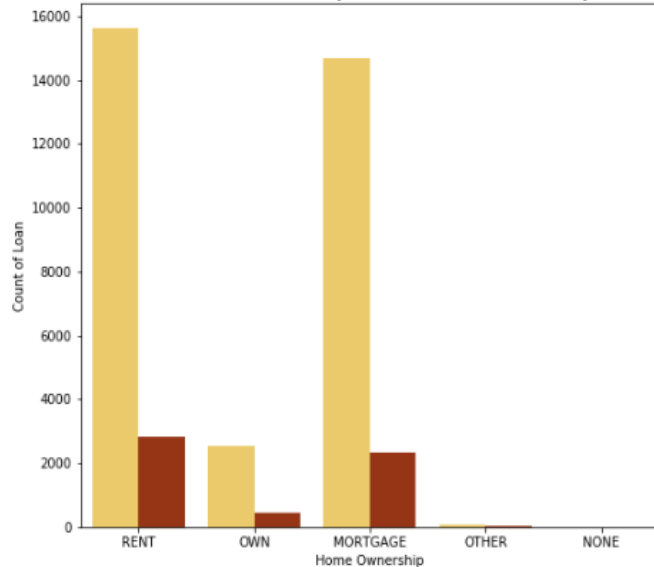


% of Charged Off Loans as per employee tenure

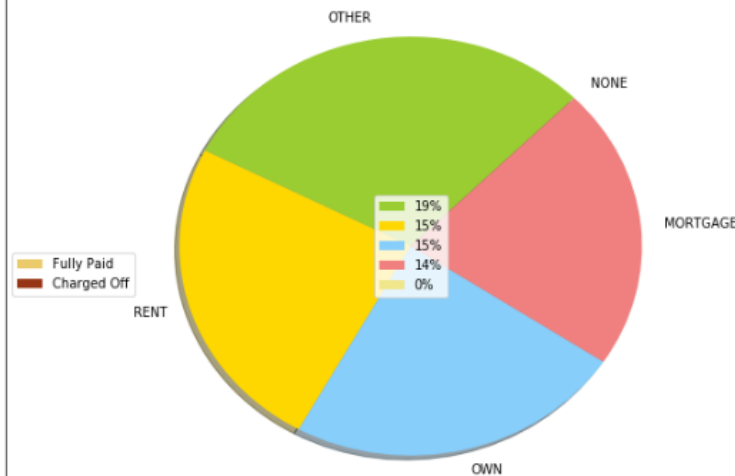


Applicant having work experience of 10+ years had a major charged off proportion of 15.67%.

Loan status as per home ownership

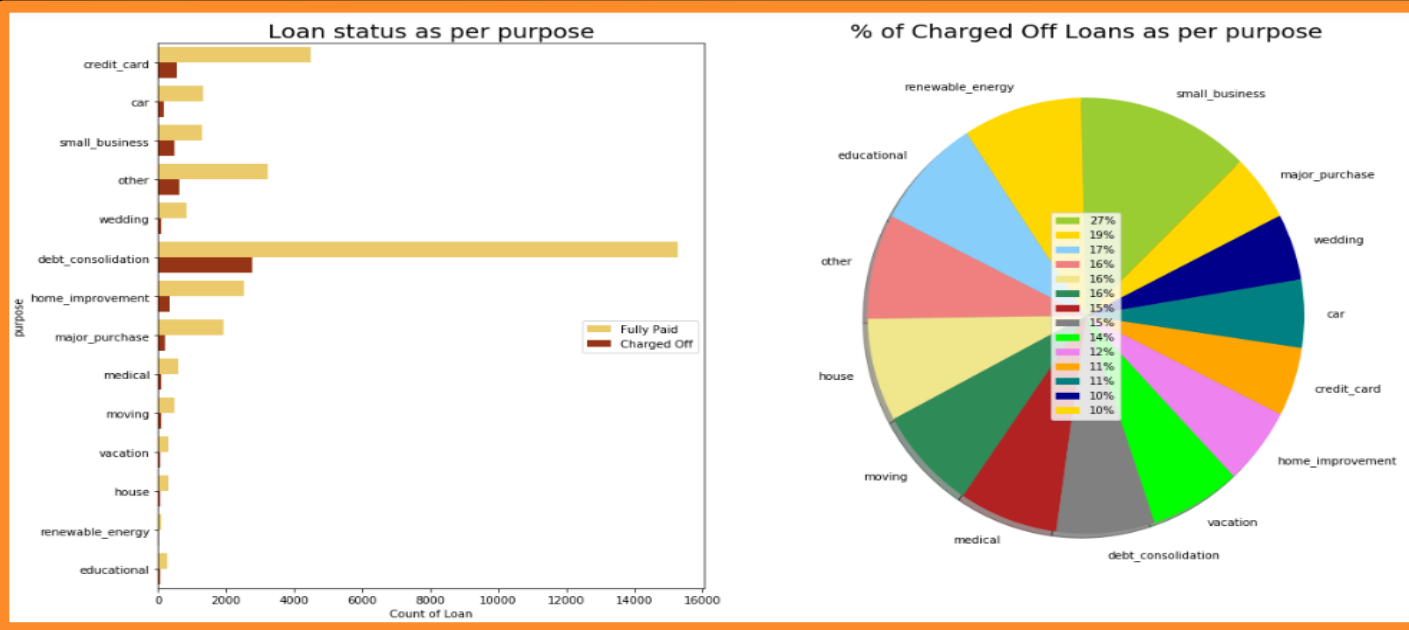


% of Charged Off Loans as per home ownership

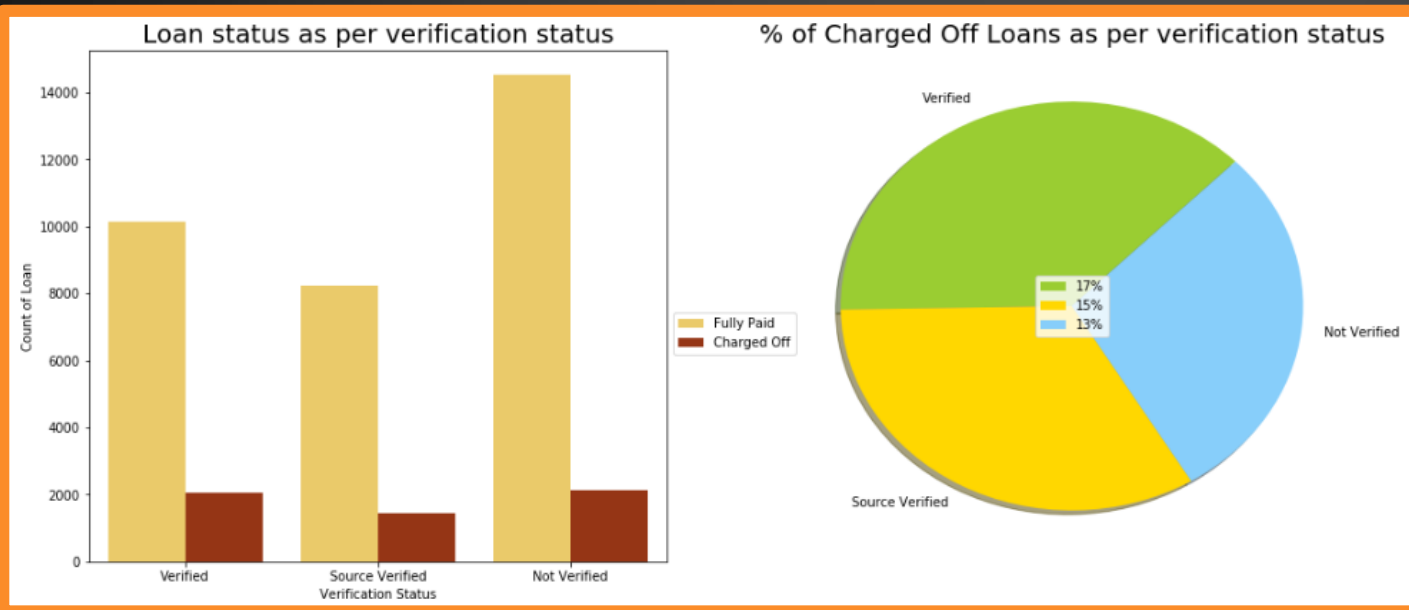


Applicant with home ownership as other had a major charged off proportion of 18.75%.

SEGMENTED UNIVARIATE ANALYSIS



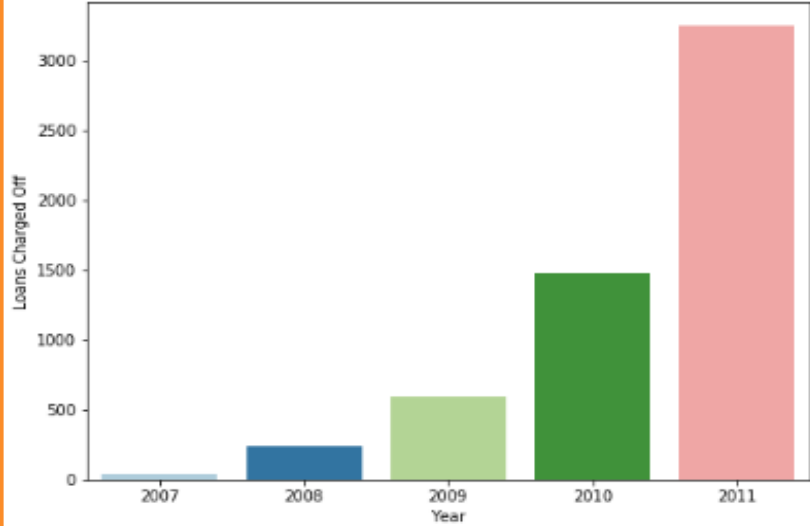
Loan for small business had a major charged off proportion of 27.09%.



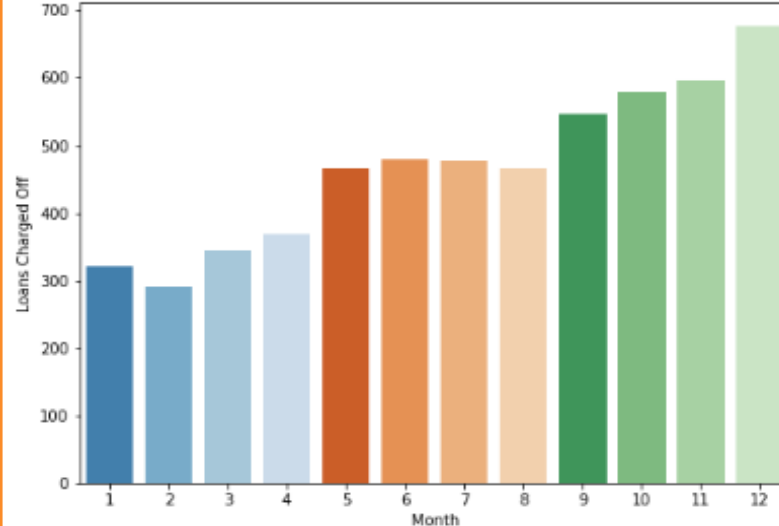
The verified loans had a major charged off proportion of 16.76%.

SEGMENTED UNIVARIATE ANALYSIS

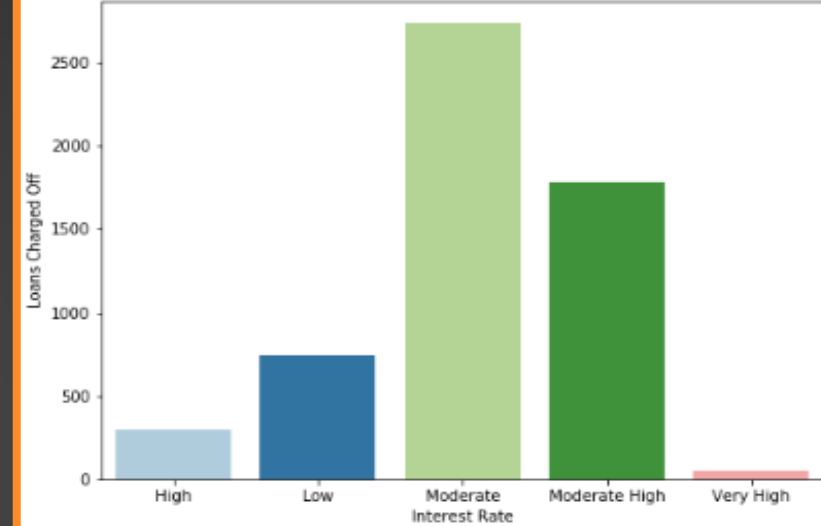
Loans Charged Off as per year



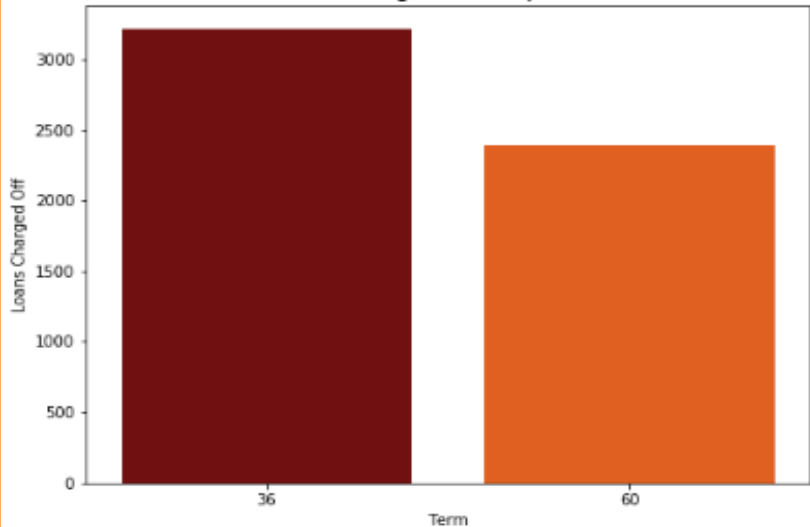
Loans Charged Off as per months



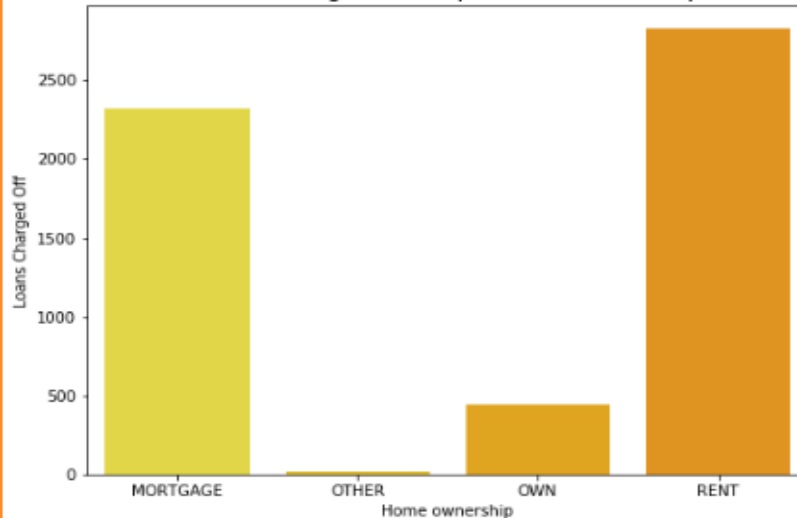
Loans Charged Off as per Interest Rate



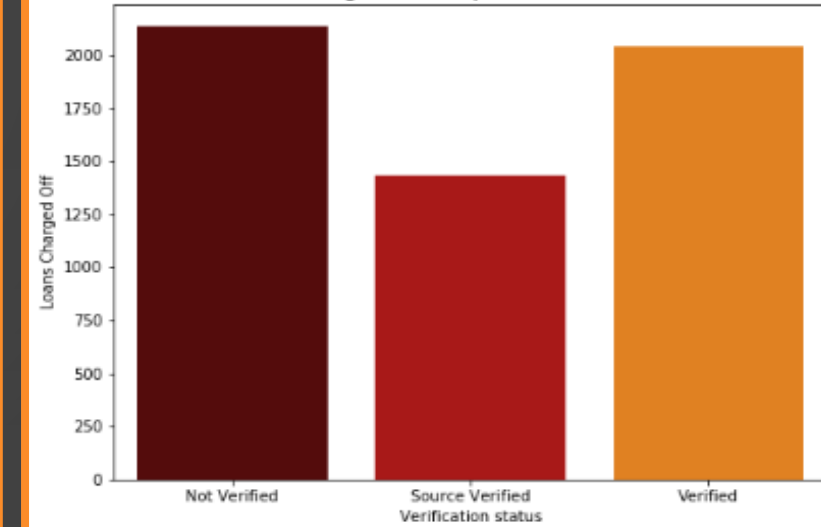
Loans Charged Off as per Term



Loans Charged Off as per home ownership



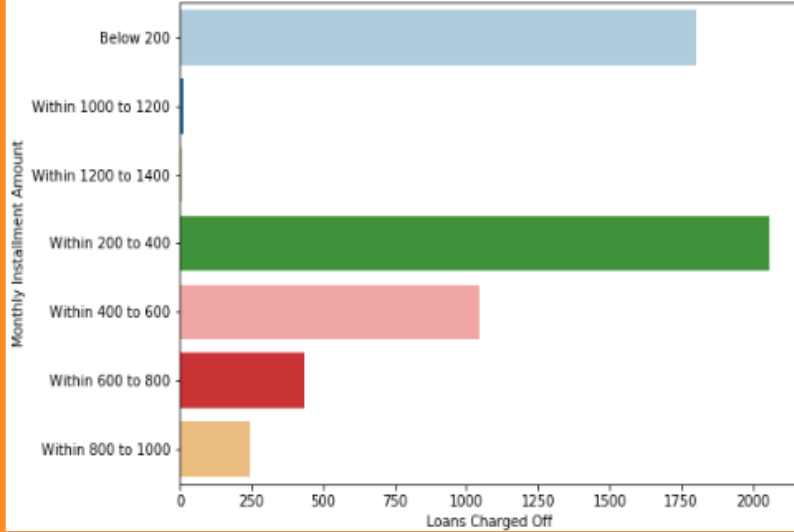
Loans Charged Off as per verification status



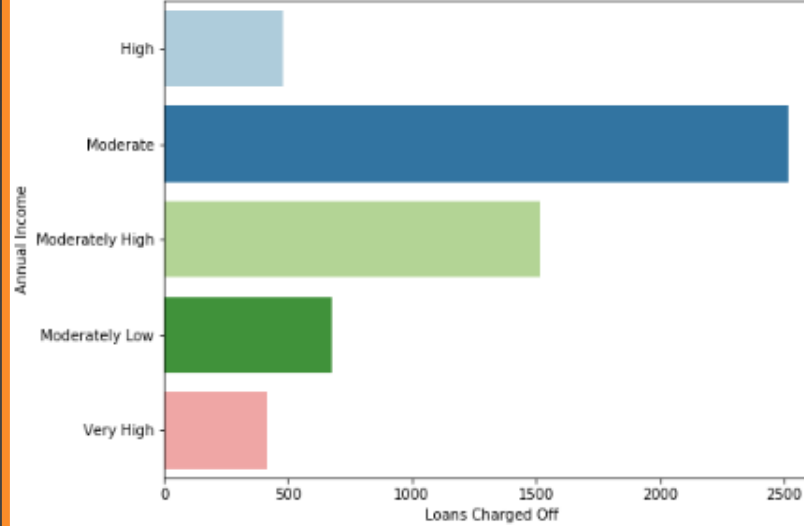
Count status of charged off loans as per the various parameters.

SEGMENTED UNIVARIATE ANALYSIS

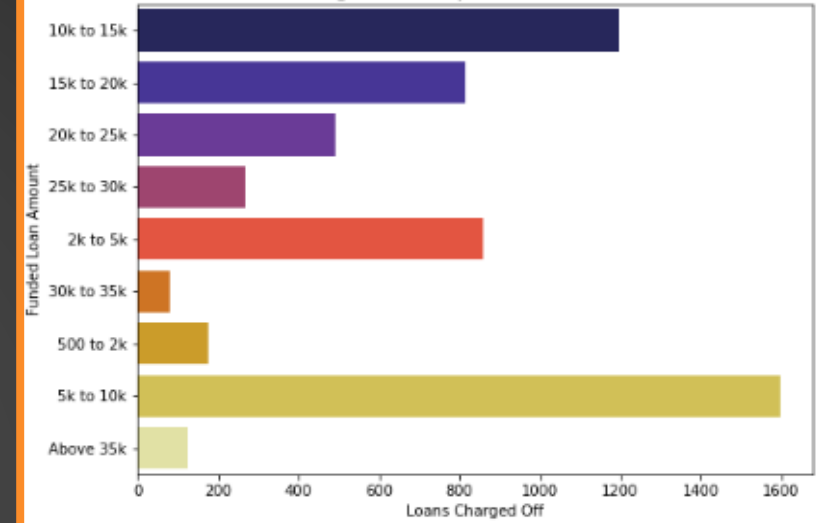
Loans Charged Off as per Monthly Installment Amount



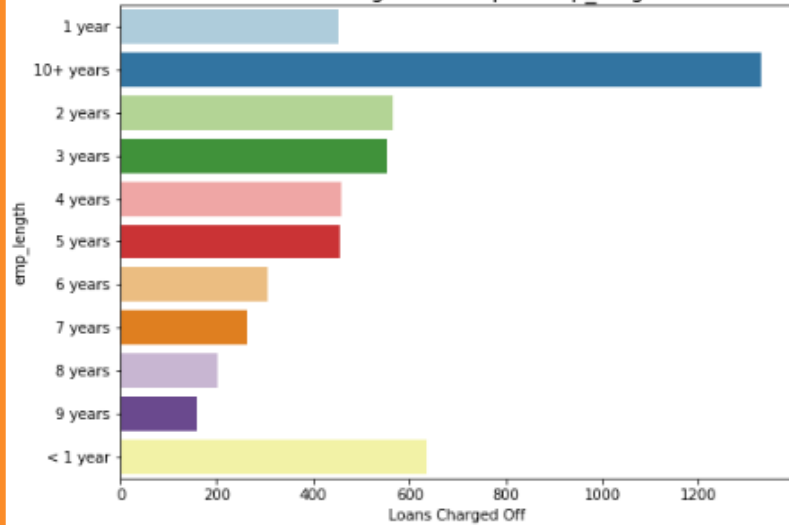
Loans Charged Off as per Annual Income



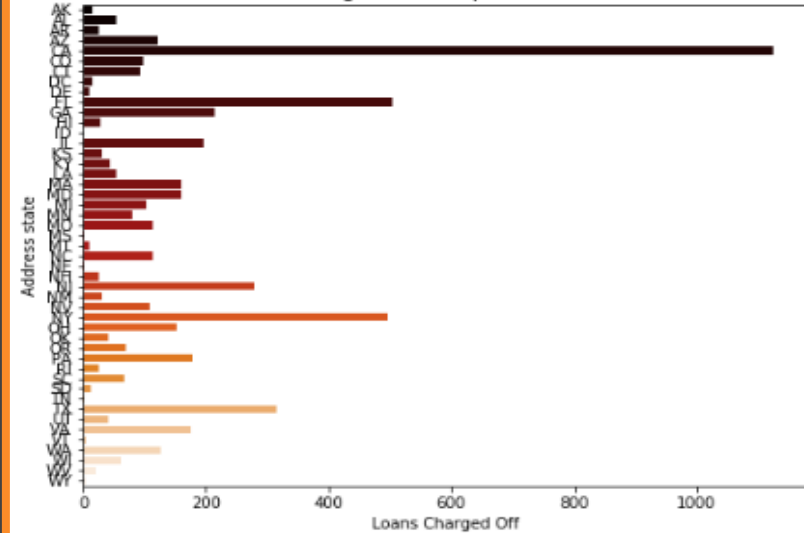
Loans Charged Off as per Funded Loan Amount



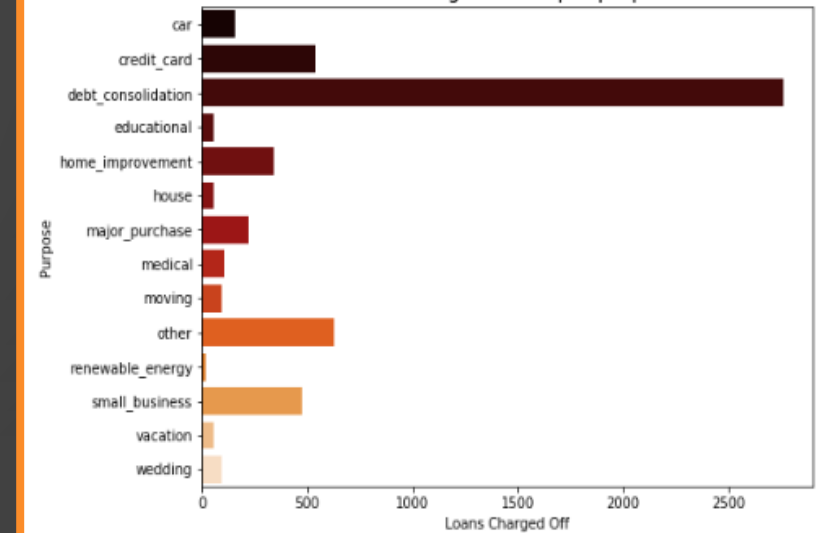
Loans Charged Off as per emp_length



Loans Charged Off as per address state



Loans Charged Off as per purpose



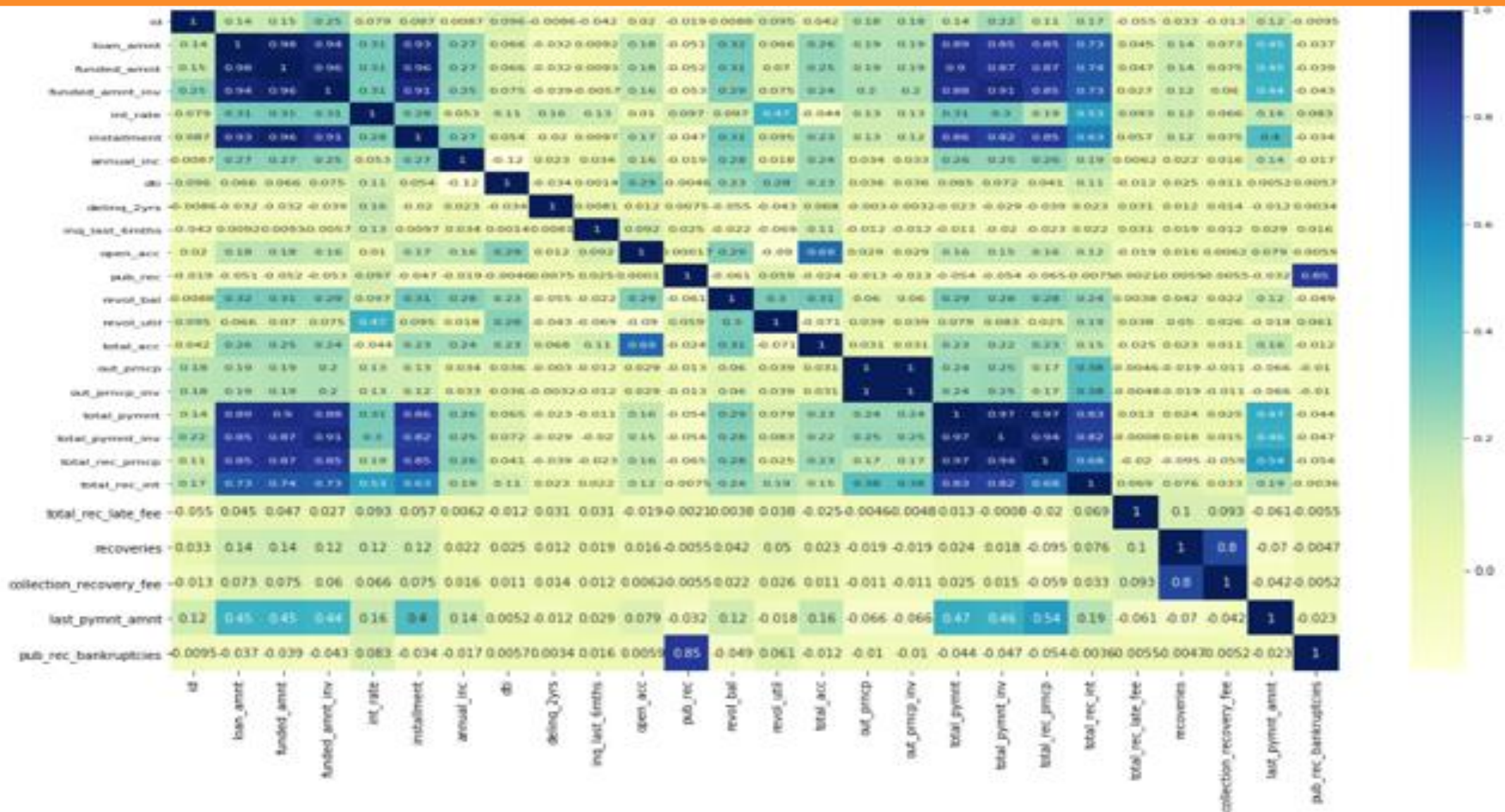
Count status of charged off loans as per the various parameters.

BIVARIATE ANALYSIS

Below is the chart for relationship between two attributes and loan charged off proportion

Attribute 1	Attribute 2	% Charged off
Grade (F)	Sub Grade(F5)	47.78%
Interest Rate(High)	Loan Amount (Above 35k)	46.34%
Annual Income (Moderate)	Loan Amount (Above 35k)	80%
Purpose (Medical)	Loan Amount (Above 35k)	66.66%
Purpose (Renewable Energy)	Interest Rate (High)	100%
Purpose (Moving)	Home Ownership (Other)	100%
Purpose (Small Business)	Verification Status (Verified)	30.31%
Loan Amount (Above 35k)	Verification Status (Verified)	26.57%
Term (60 Months)	Interest Rate (High)	38.56%
Employee Tenure (1 Year)	Interest Rate (High)	48.97%
Employee Tenure (10+ Year)	Verification Status (Verified)	18.39%
Address State (NE)	Verification Status (Not Verified)	60%

CORRELATION



CONCLUSION

Important Variables

- Interest rate
- Monthly Installment
- Loan Term
- Annual Income
- Verification status
- Employee tenure
- Loan Purpose
- Home Ownership
- Address state

- ✓ Interest rate between 20% to 25% had highest % (38.33%) charged off and rates between 10% to 15% had highest count (18494) of charged off.
- ✓ Monthly Installment within 800 to 1000 had highest % (18.41%) charged off and Installments within 200 to 400 had highest count (14722) of charged off.
- ✓ Loan Term with 60 months had highest % (25.29%) charged off and 36 months had highest count (3214) of charged off.
- ✓ Annual Income within 10k to 30k had highest % (18.61%) charged off and Income within 30k to 60k had highest count (15729) of charged off.
- ✓ Verified loans had highest % (16.76%) charged off and not verified loans had highest count (2135) of charged off.
- ✓ Employee tenure with 10+ year of experience had highest % (15.67%) charged off and highest count (1330) of charged off.
- ✓ Loan purpose for small business had highest % (27.09%) charged off and debt consolidation had highest count (2762) of charged off.
- ✓ Home ownership as other had highest % (18.75%) charged off and rent had highest count (2829) of charged off.
- ✓ Address state NE had highest % (60%) charged off and CA had highest count (1124) of charged off.

THANK YOU