Lending risk analysis
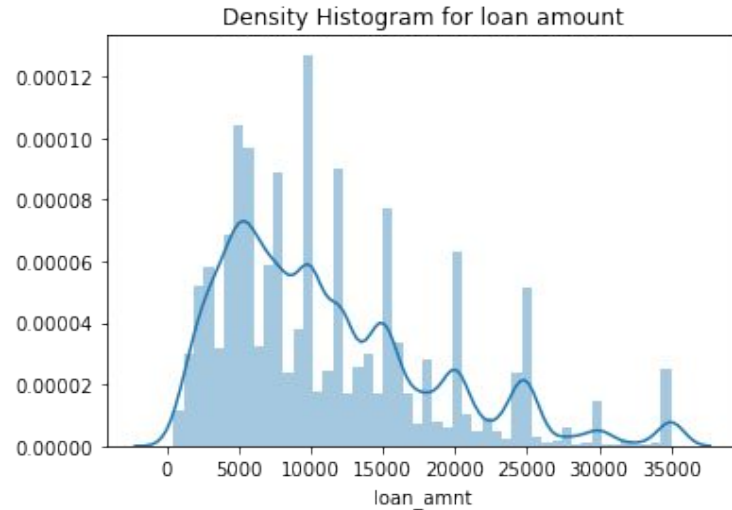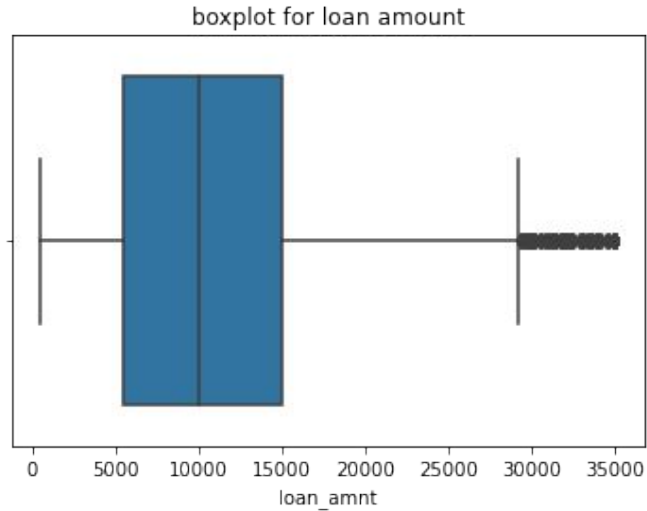
Group study done by :
Amit Choudhary
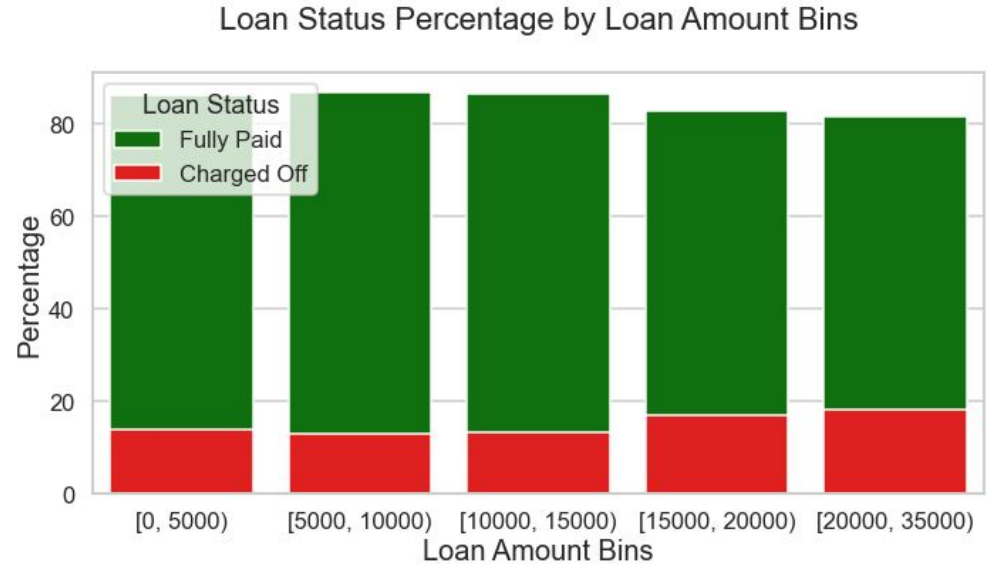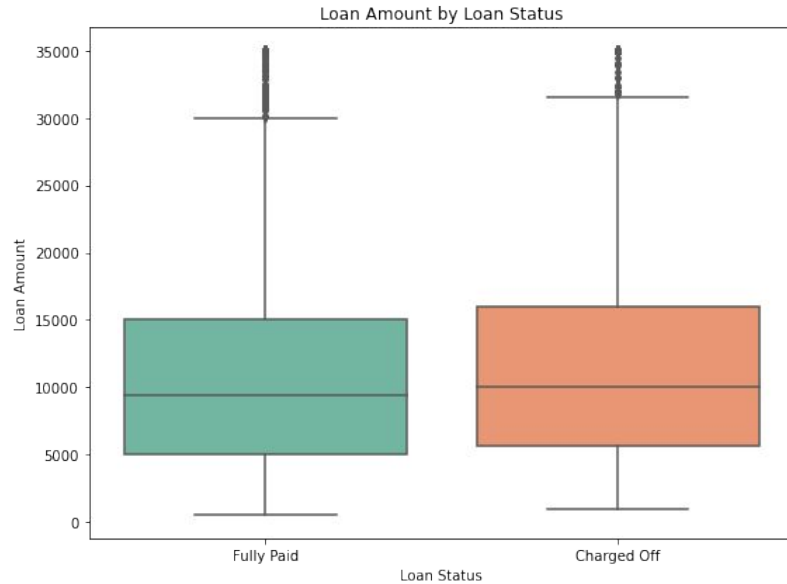**Nidhi Poojary**

Contents

- Problem statement
- Univariate Analysis
- Bivariate Analysis
- Insights

Loan Amount Distribution


1. The IQR for loan amount lies between 5000-15000
2. The histogram is elevated at 5000,10000, 15000, 20000, 250000

Analysis on Loan Amount

Loan Amount by Loan Status


Loan Status Percentage by Loan Amount Bins

1. In the box plot, we can see that the difference between 75th quartile is negligible for the borrowers who defaulted and who have paid the loan.
2. Where the loan amount bin is more than 20000, is when defaulting percentage significantly increases

Density Histogram for interest rate
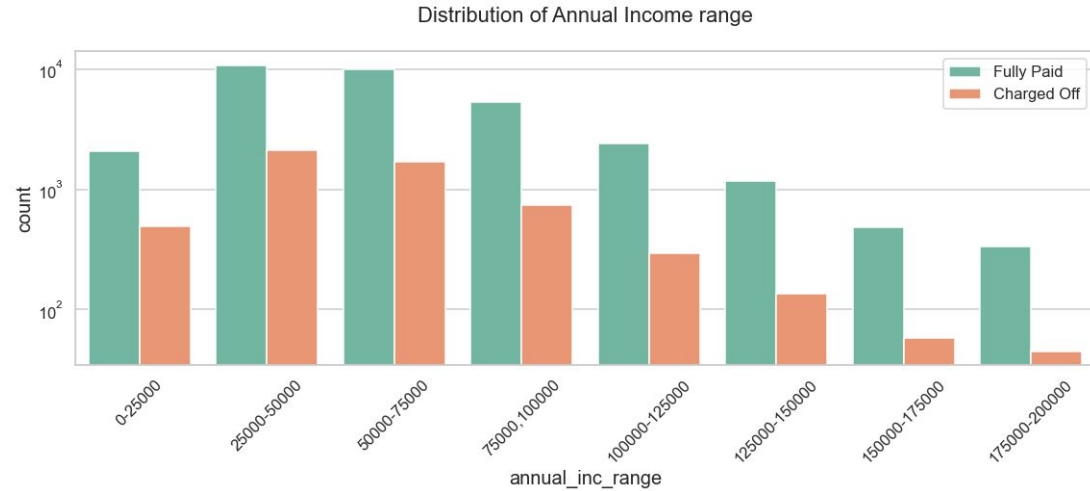

Interest rate by Loan Status
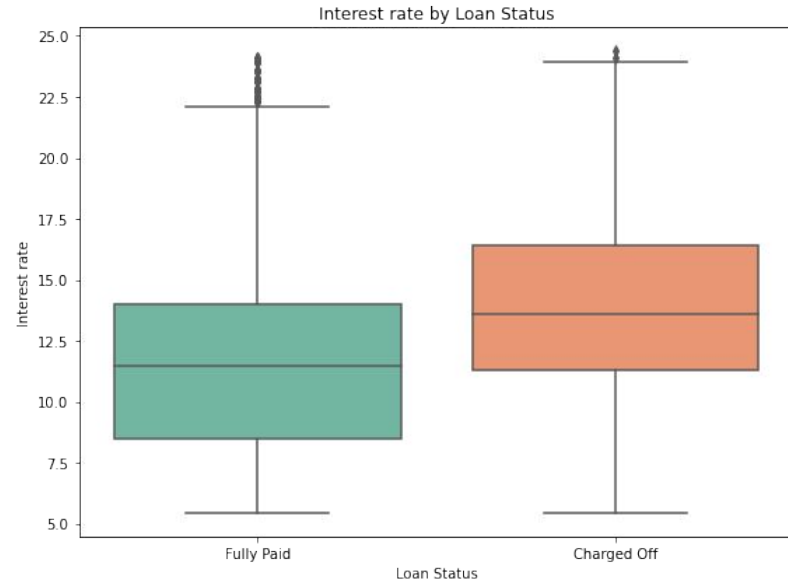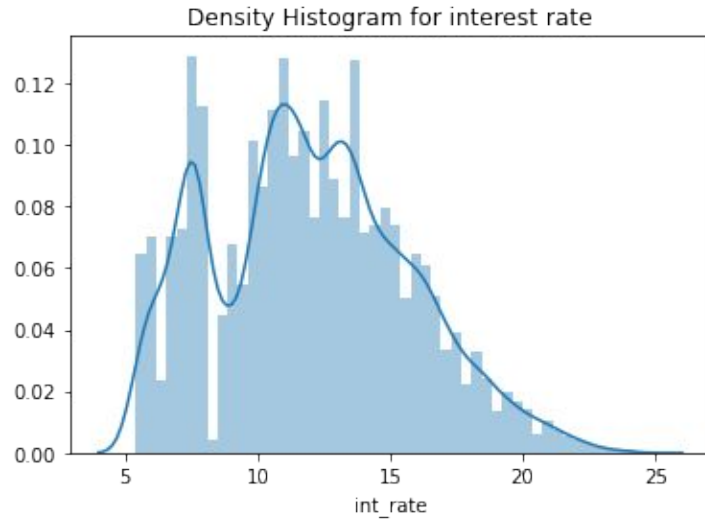
1.   It is noticed in the box plot that the defaulters has a greater interest rate.

Loan Status Percentage by Interest Rate Bins
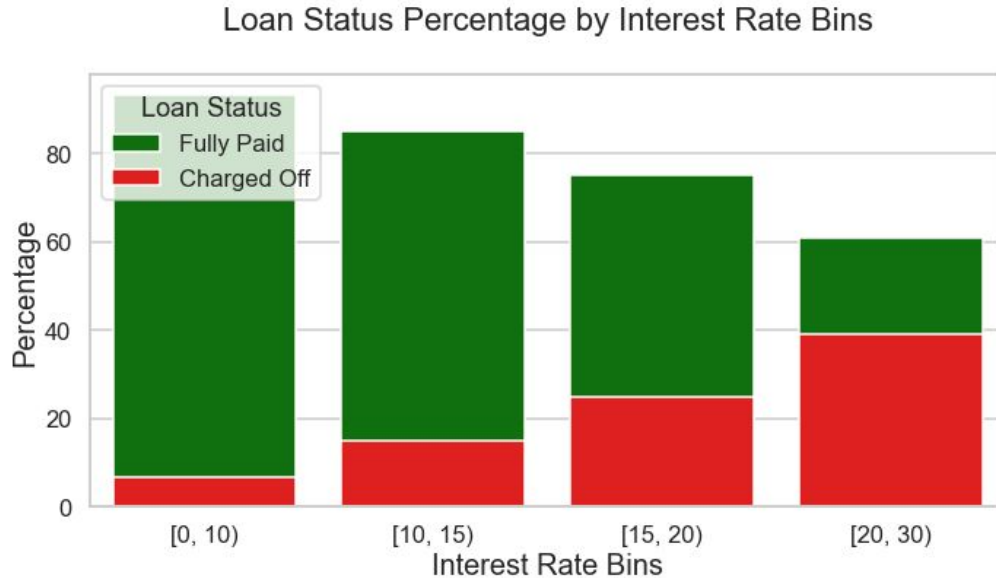
1. The defaulters percentage increases steeply as the interest rate increases and is maximum for interest rate in the [20,30) bin

boxplot for annual income



Distribution of Annual Income range
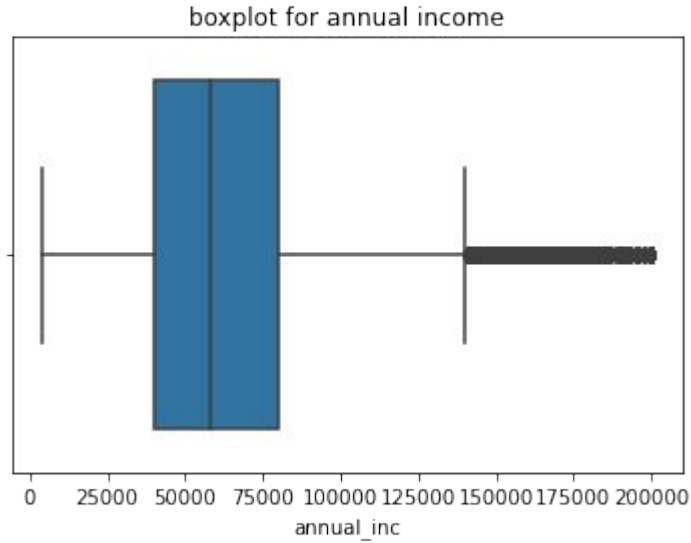
1. The interquartile range lies between the 40000 and 80000 for the annual income.
2. The dataset has been pruned to exclude outliers who had an annual income of more than 200000
3. The dataset in majority comprises of borrowers with annual income between 20000-100000

Annual Income by Loan Status
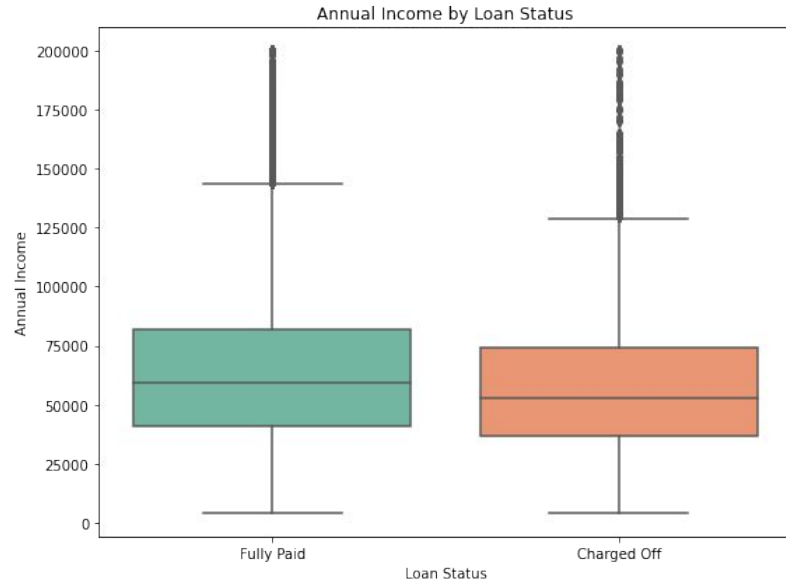
1. We can infer from the box plot that the defaulters have a slightly lower annual income range than the people who fully paid the loan.
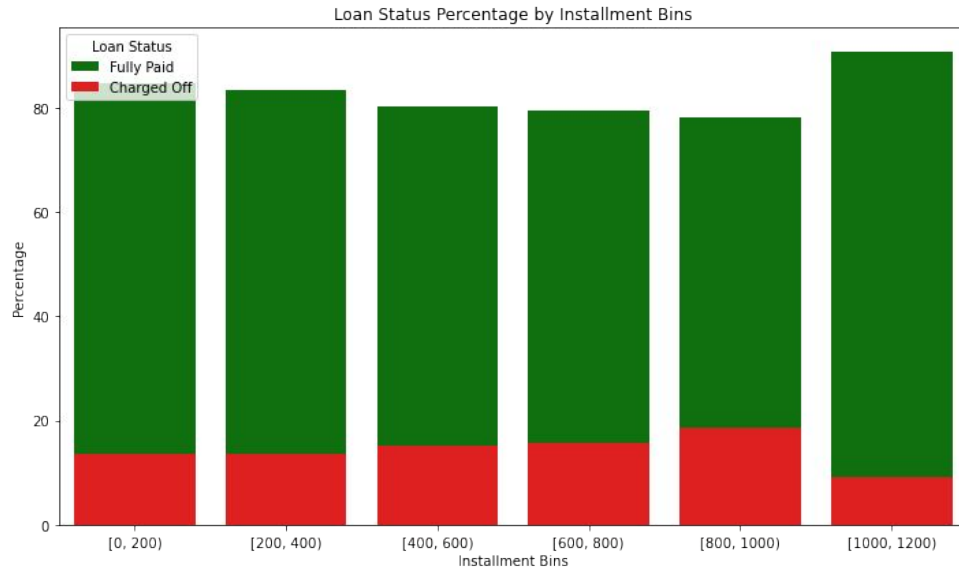
1. There is no clear relation between the installments and the defaulting percent. It stays gradual for the successive installments bins and suddenly decreases for installments greater than 1000

2. The term has a clear impact as we can see 60 months term has a higher percentage of defaulters

Loan Status Percentage by Months Since Last Delinquency Bins
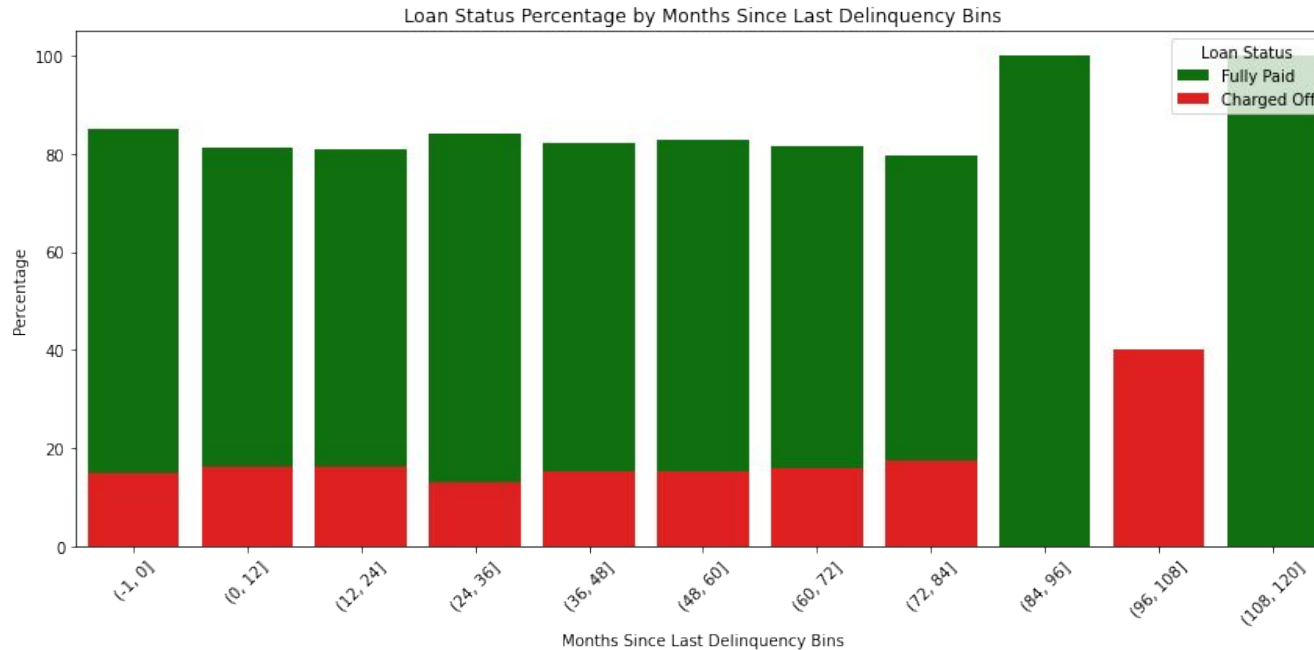
1. There is no clear relation between the months since last delinquent and the defaulting percent. It stays gradual for the successive bins and increases steeply for months (96,108]
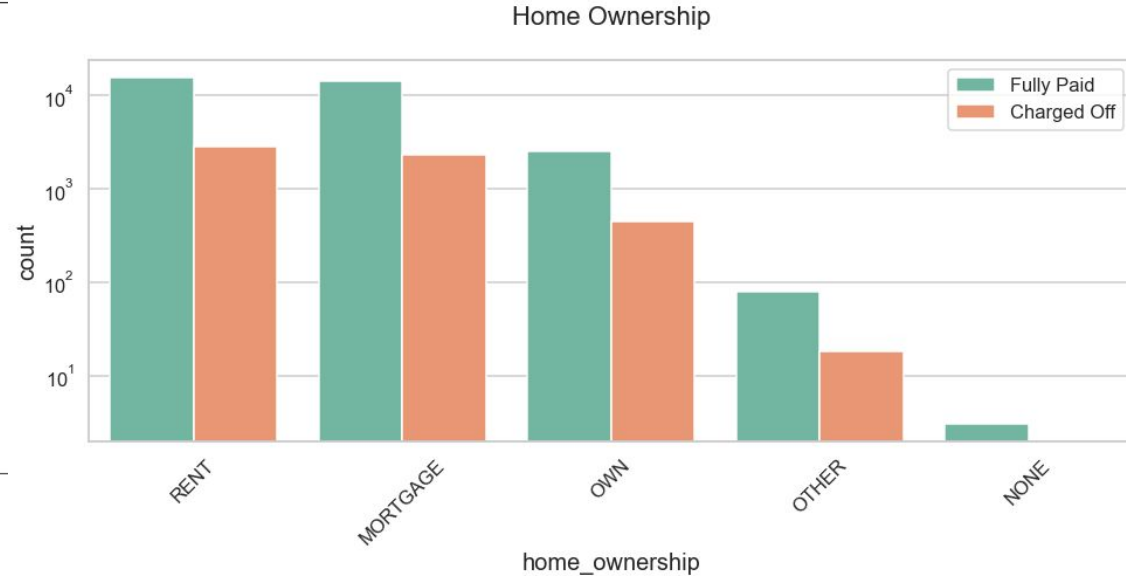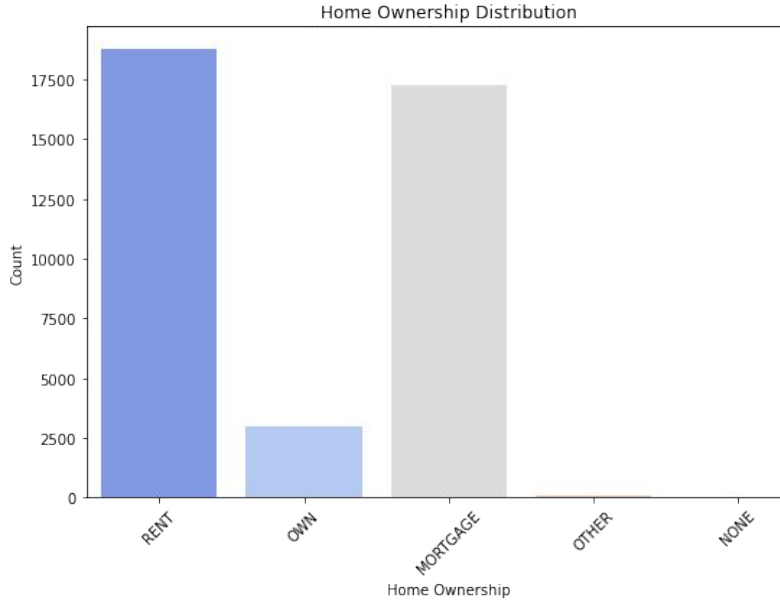
Employment Length Distribution
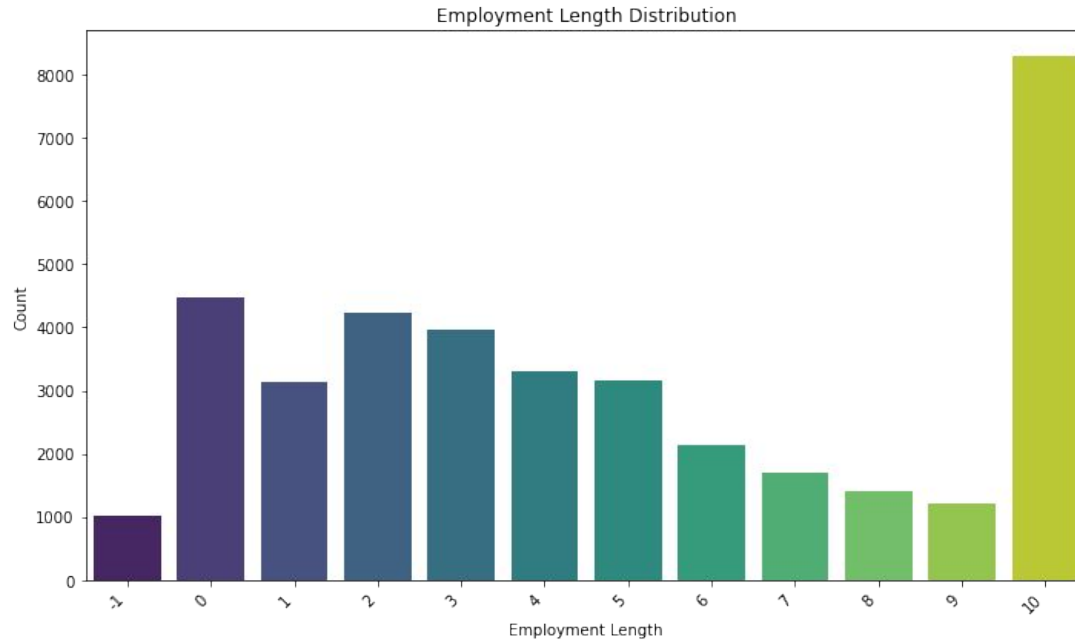
1. The 10+ years of employment have the most number of loans.
2. Note : -1 denotes the missing values

Employment length
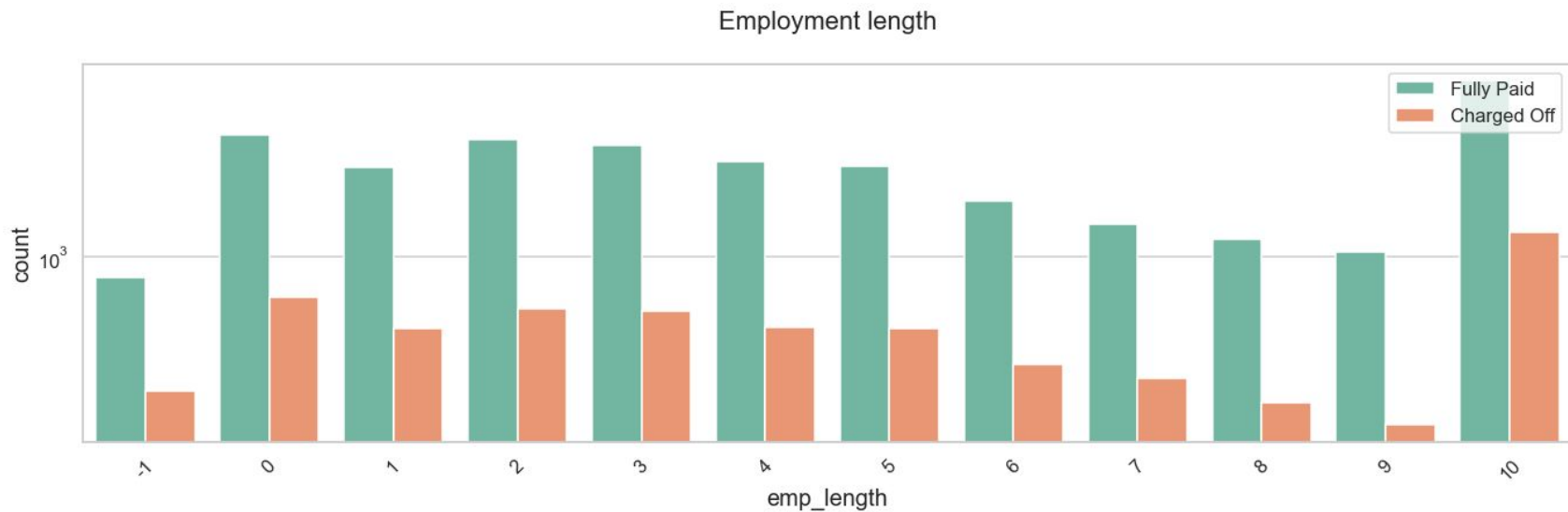
1. The percentage of defaulters remains same mostly throughout the years. It increases steeply for borrowers with more than 10 years of employement

Home Ownership Distribution

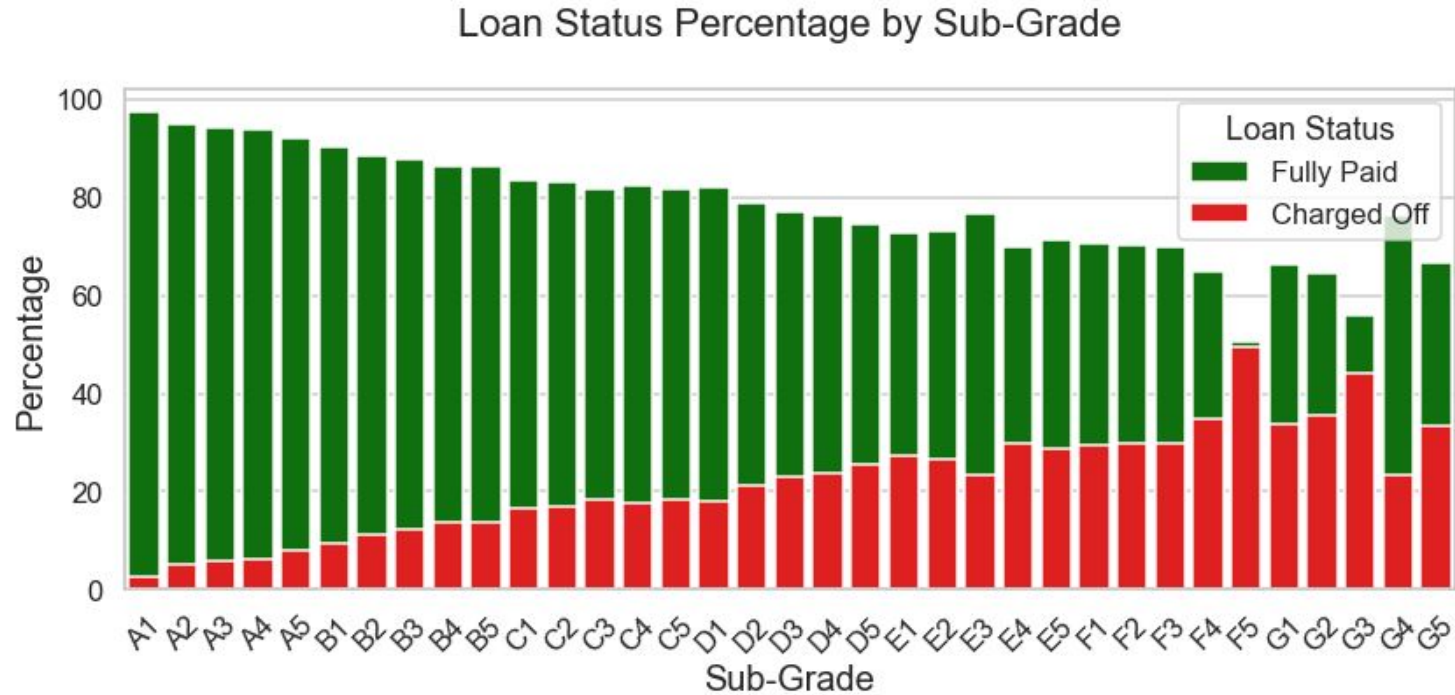Home Ownership

1. Borrowers paying rent and mortgage have greater loans
2. The loan defaulted percent remains mostly constant for mortgage, own and rent

Analysis on Grade
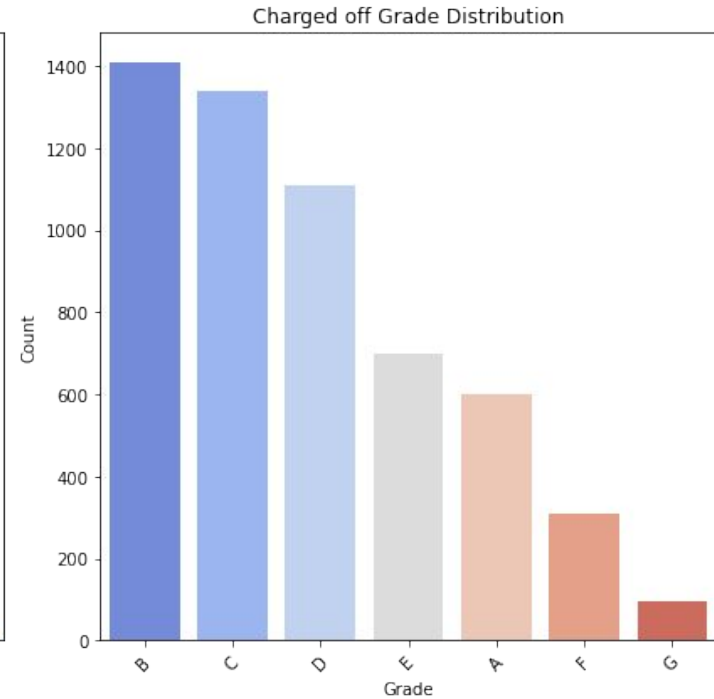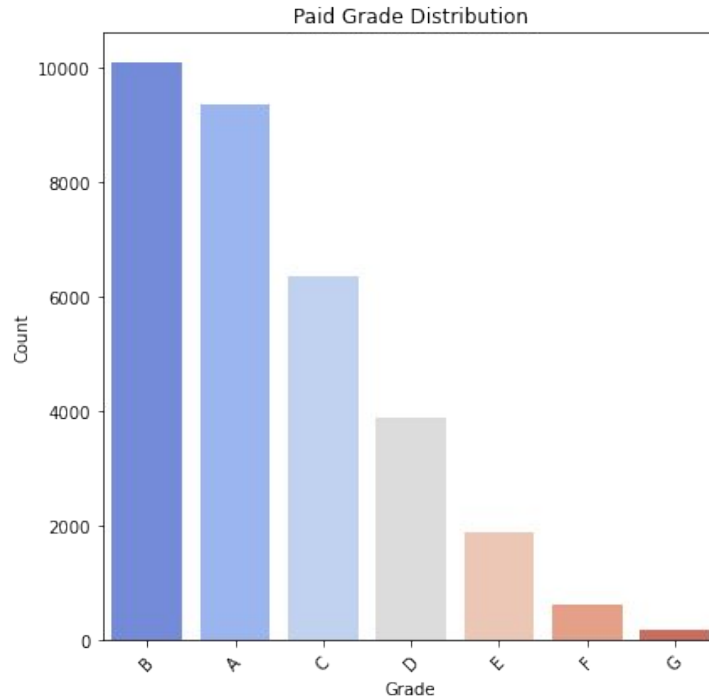
1. A- D grades have maximum number of loans.
2. Grade B, C and D borrowers have a higher defaulting rate

Loan Status Percentage by Sub-Grade

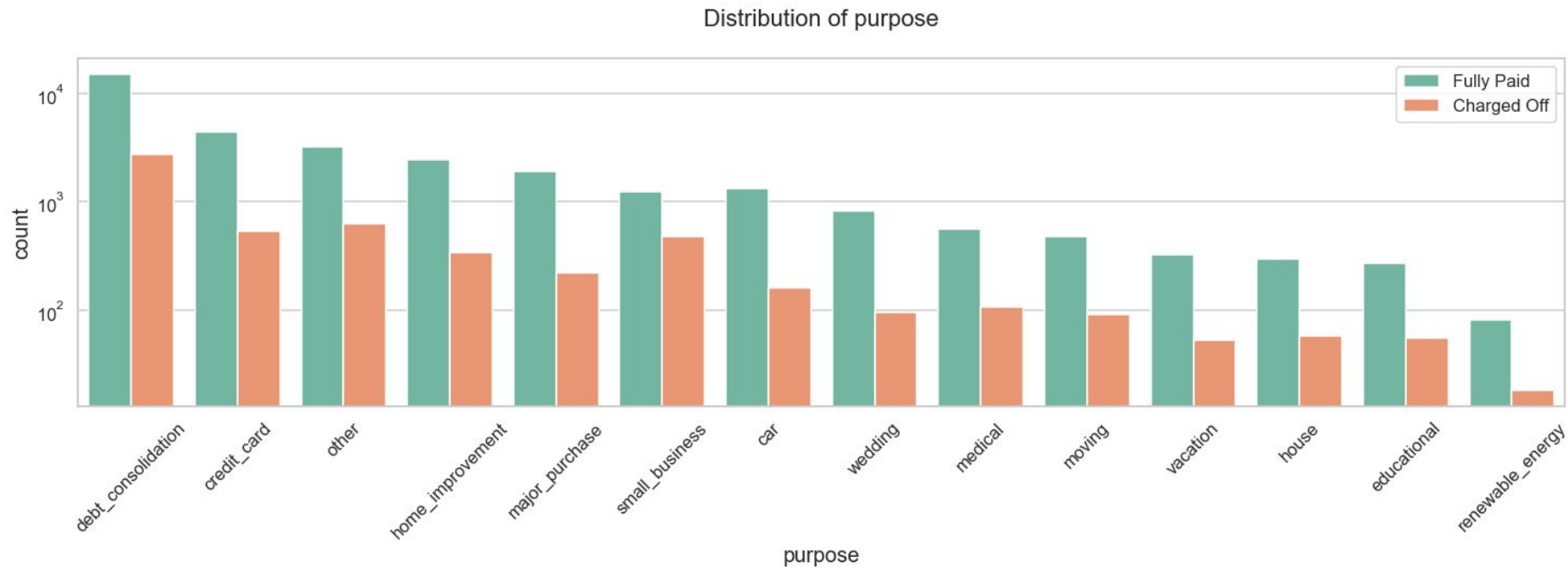1. The loan defaulted percent for sub grades between E and G are higher

Loan Purpose Distribution
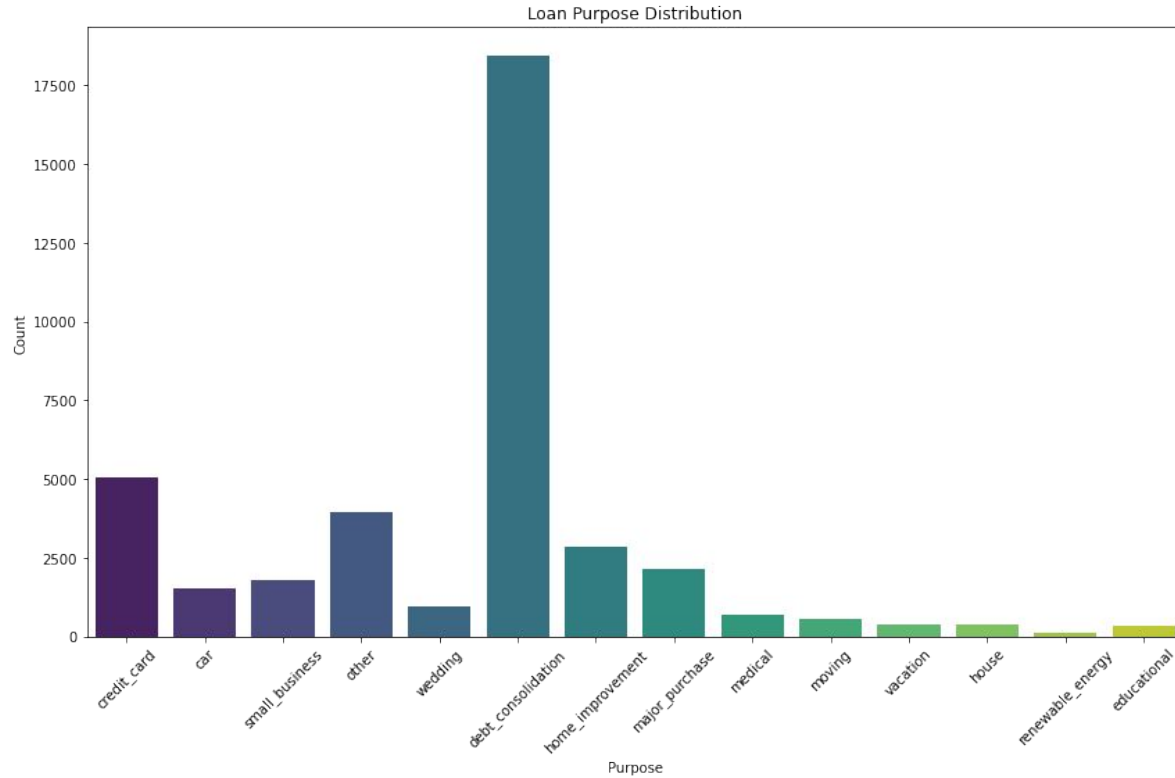
1. Maximum loans are taken for debt consolidation

Analysis on Loan Purpose

Distribution of purpose

1. The defaulters count for debt_consolidation and small business is the highest

Loan Status Percentage by State
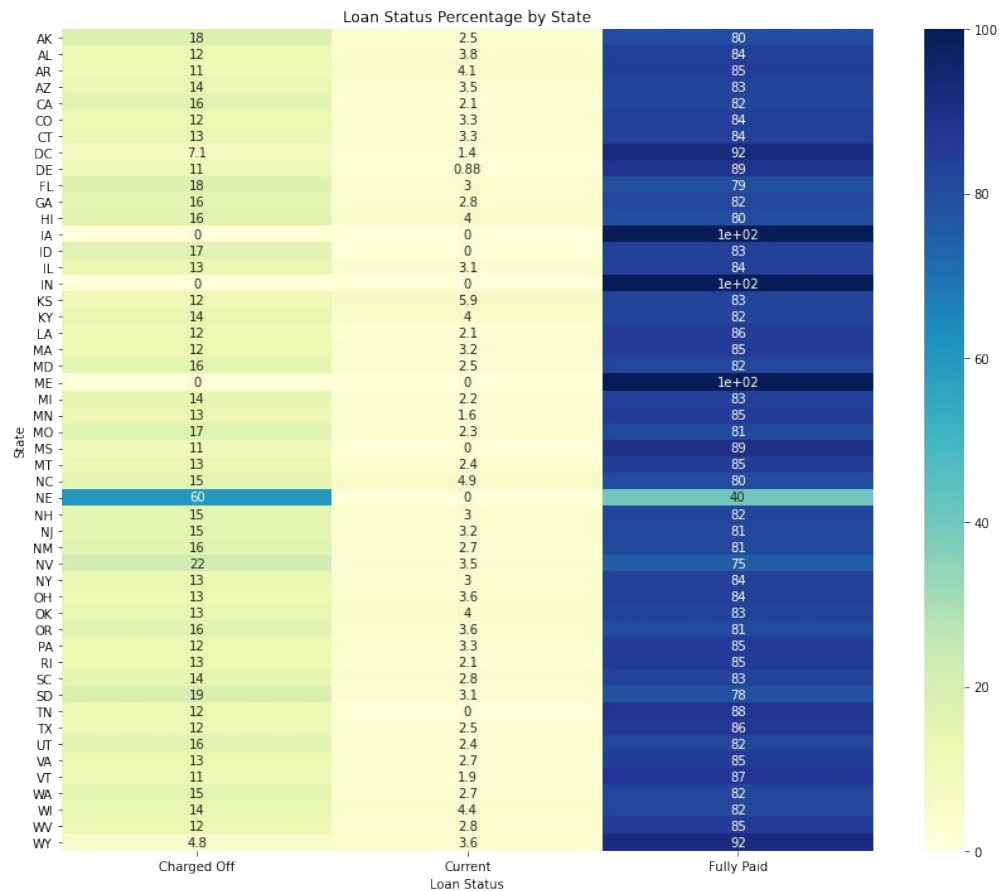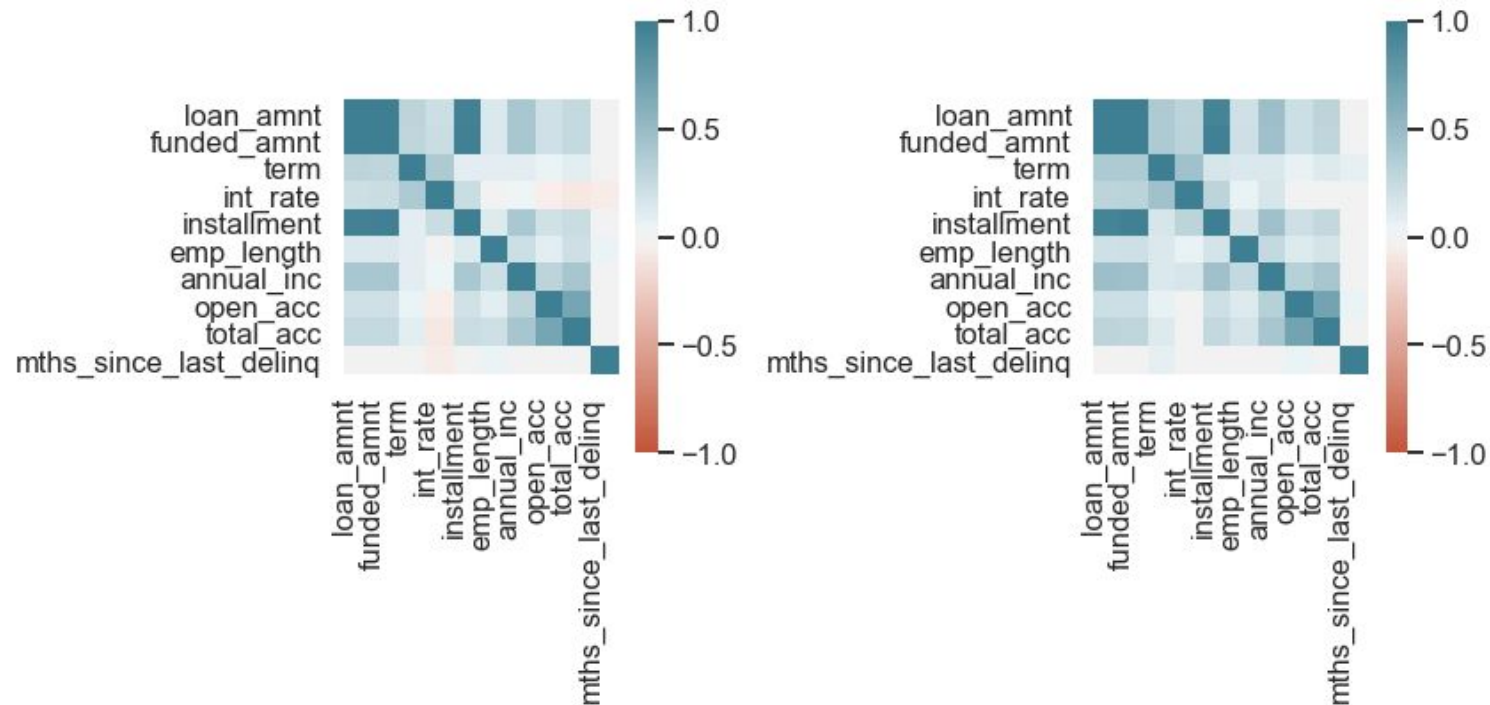
1. The greatest defaulting percentage of 60% is from NE.
2. The percent is also high for NV and AK

Verification Status

1.   The count of verified defaulters is slightly lower than the un verified defaulters

Correlation

1. loan amount and interest rates are more positively correlated for charged off loans
2. loan amount and term is positively correlated.

Insights - decisive factors for when a borrower may default

Continuous Variable:

- LOAN_AMOUNT : Loan amount greater than 15000 have higher default rate
- INTEREST_RATE : As Interest rate increases the default rate increases steeply
- ANNUAL_INCOME : As the annual income increase the default rate decreases
- MONTHS_SINCE_LAST_DELINQ : Crime committed between 90 to 110 days have higher default percent

Insights - decisive factors for when a borrower may default

Categorical Variable:

- TERM : 60 months term have a higher default rate than 36 months term
- SUB_GRADE : The lower subgrades(G5,G4..) have higher defaulting rate
- VERIFICATION STATUS : Non verified buyers have higher defaulting rate
- PURPOSE : Debt consolidation and small businesses have more chances of defaulting
- STATE : Percent of loan defaulted is very high for state NE