

Lending Club Case Study

Group Members :

Jeeva

Karan

1. **Problem Statement And Approach**
2. **Univariate Analysis**
3. **Segmented Univariate Analysis**
4. **Bivariate Analysis**
5. **Conclusion**

Problem Statement

Lending Club wants to identify the risky loan applicant, so that such loans can be reduced thereby cutting down the amount of credit loss.

Our aim is to find the driving factors behind loan getting default



Problem Solving Methodology

Data Understanding

- Loan Data set contains the complete data for all loans issued.
- Data dictionary describes the meaning of each variable

Data Cleaning

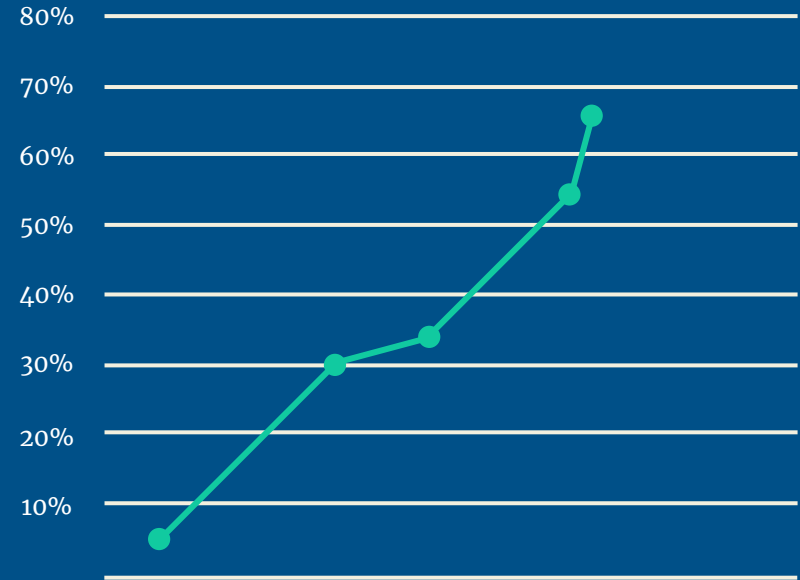
- We dropped the columns having more than 50% of missing values.
- Imputed the default values where all the values in the columns are unique.

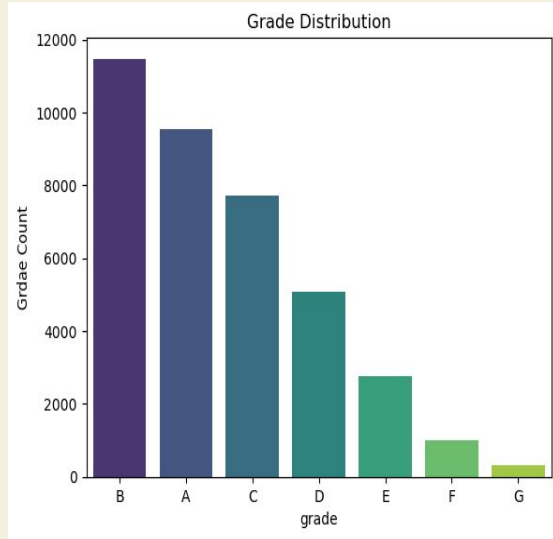
Categorized the Columns

- Unordered Categorical
- Ordered Categorical
- Numeric/Quantitative

Univariate Analysis

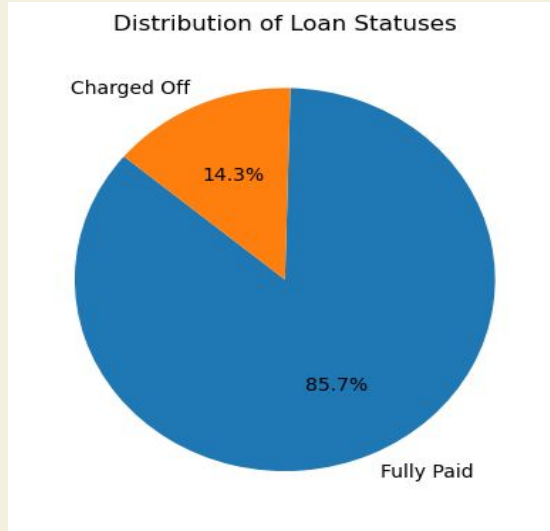
Analysing each column plotting the distribution of each columns





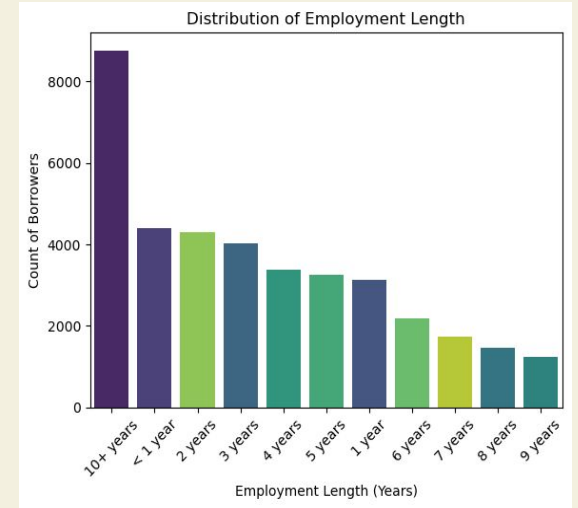
Grade Distribution

Loans are mostly allocated to the borrowers with grade B



Loan Status

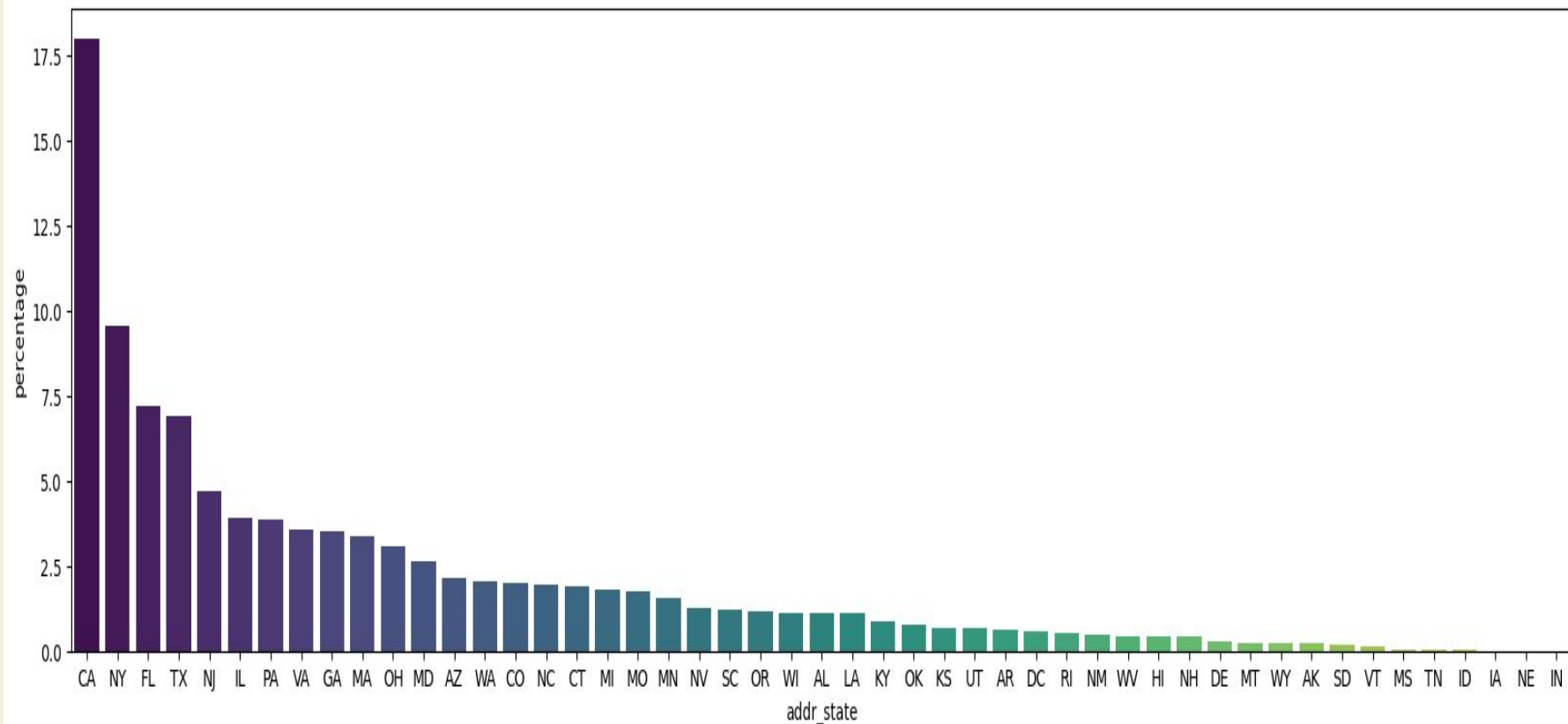
Charged of borrowers are around 14.3% and fully paid is around 85.7%



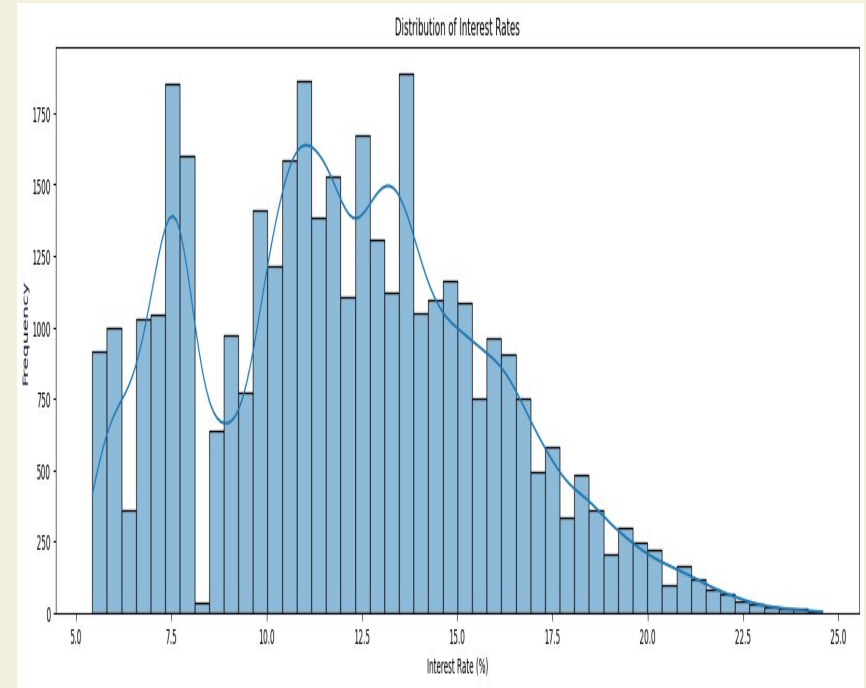
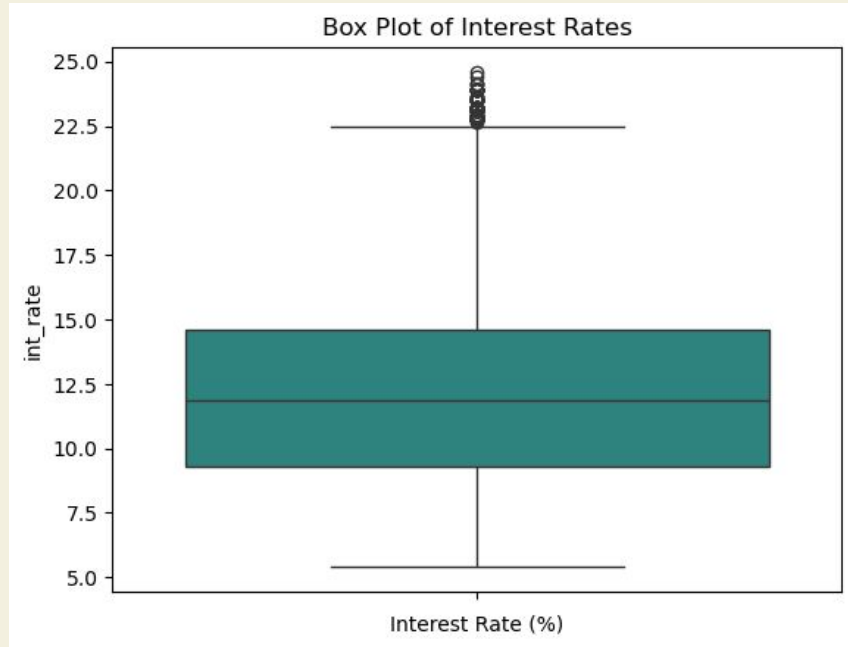
Employment Length

Borrowers with employment lengths exceeding 10 years tend to borrow more loans.

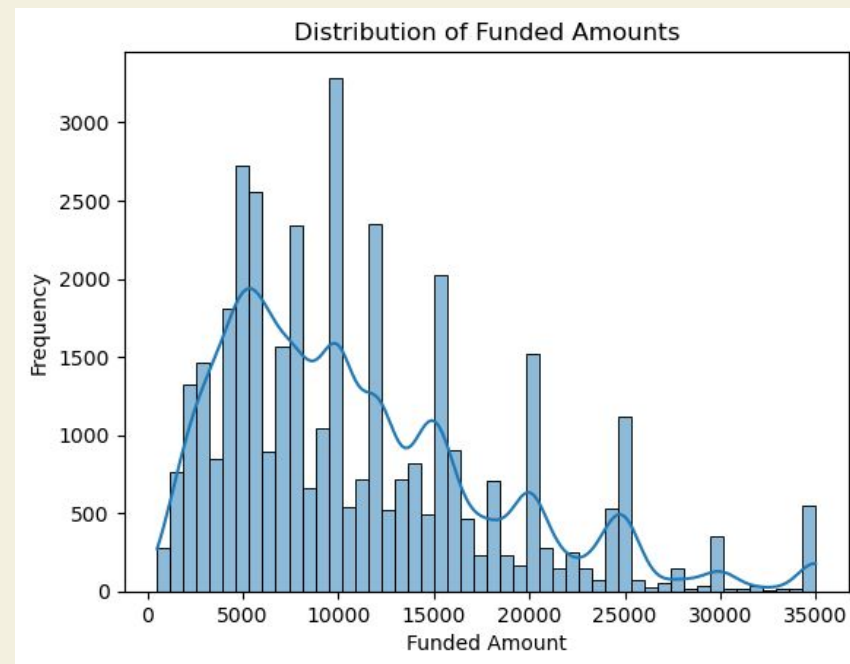
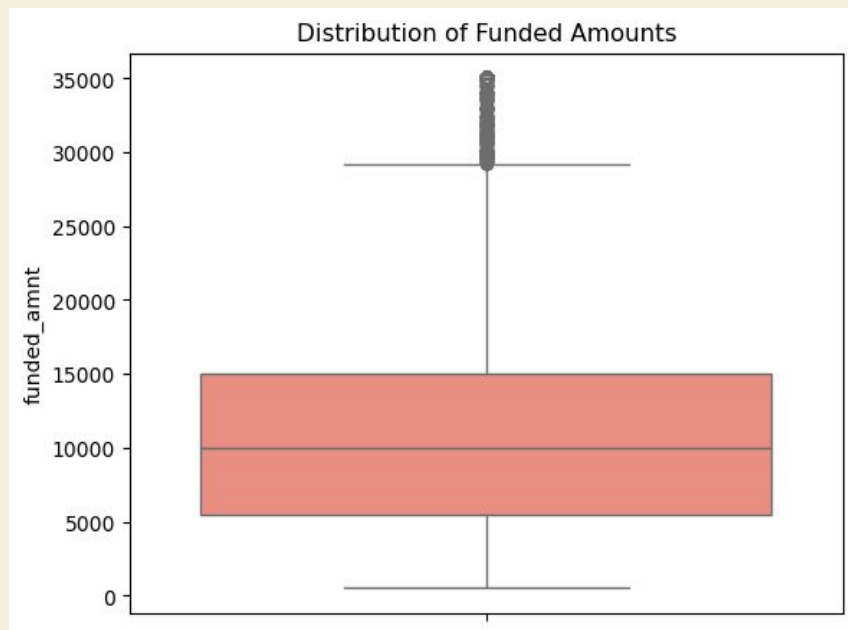
Address State Distribution



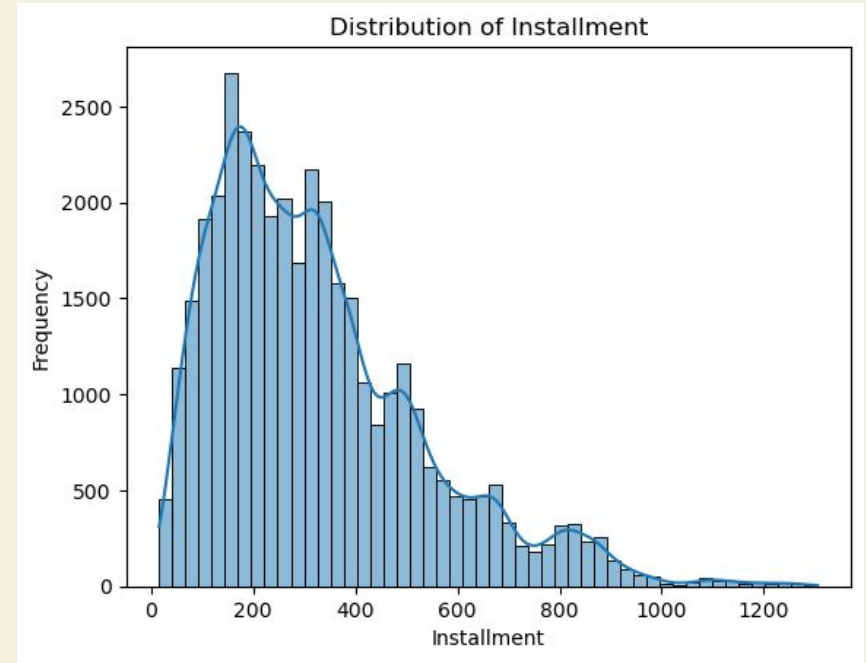
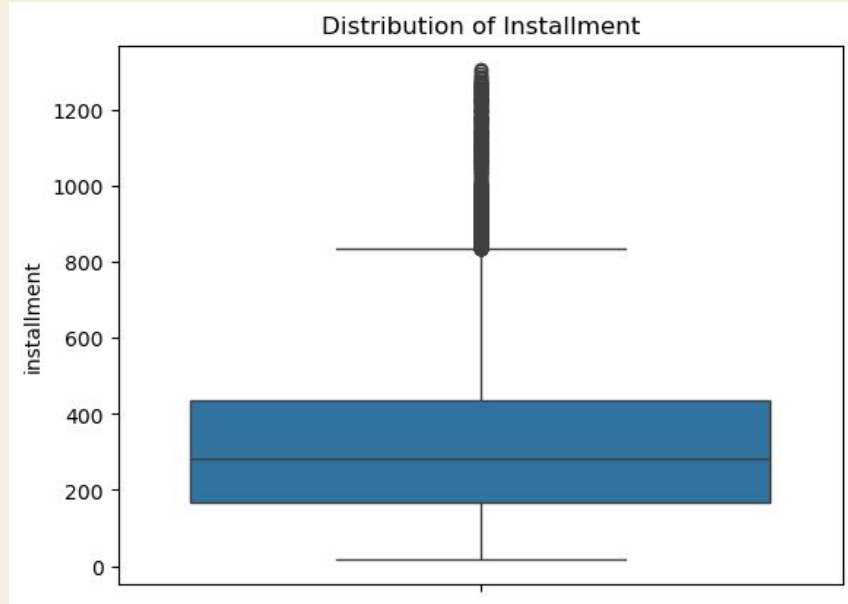
Most of the lending individuals are from California and New York.



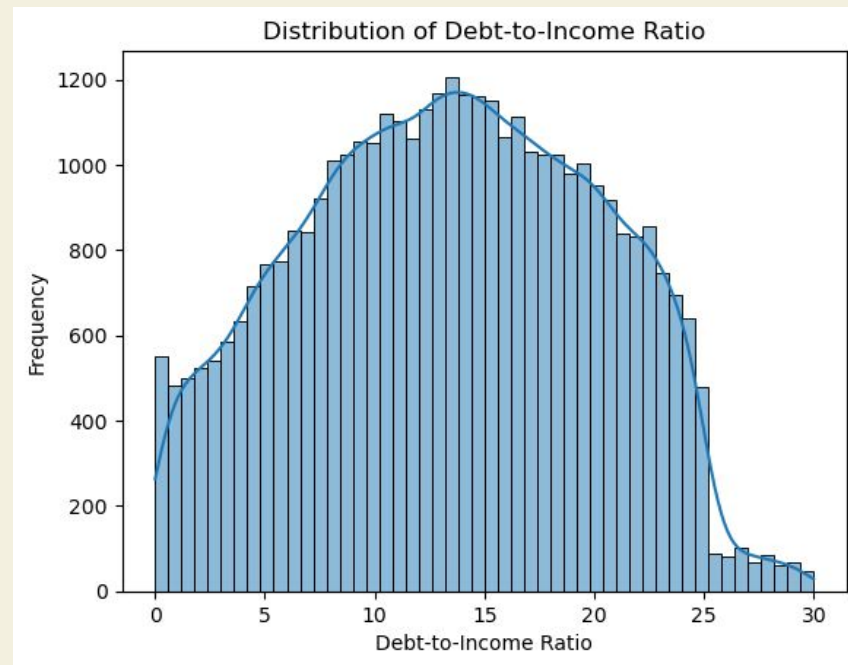
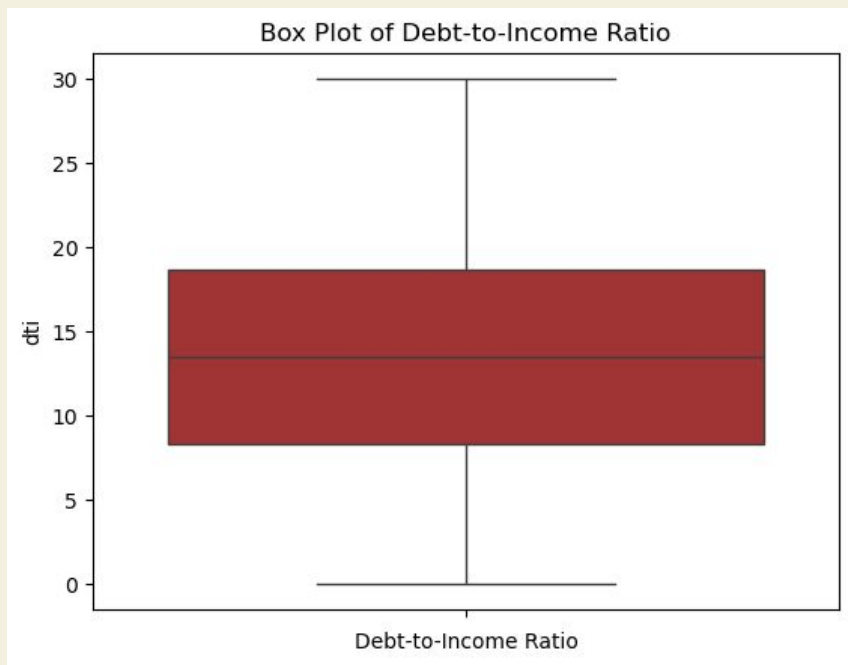
The average interest rate is 12.07%, with most loans having rates between 5.42% and 14.61%. The lowest rate is 5.42%, and the highest is 24.59%.



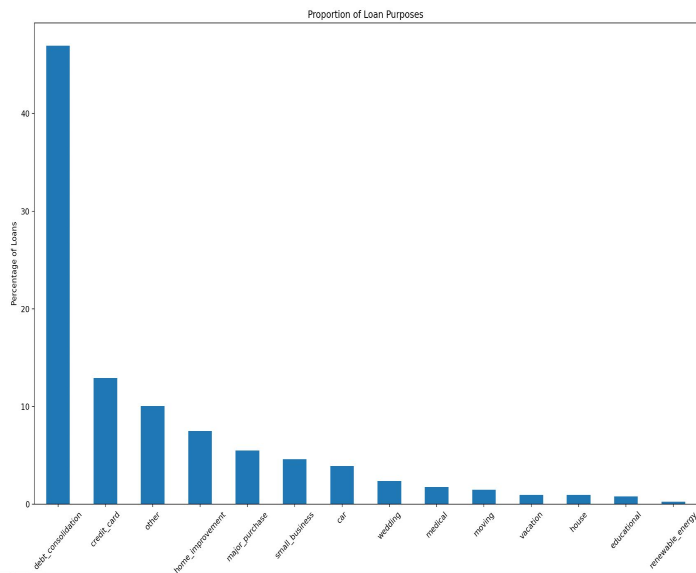
From the summary, most borrowers tend to take loan amounts around \$10,000, as indicated by the median and the interquartile range between \$5,500 and \$15,000. Very few borrowers have loan amounts exceeding \$30,000, with the maximum loan amount reaching \$35,000.



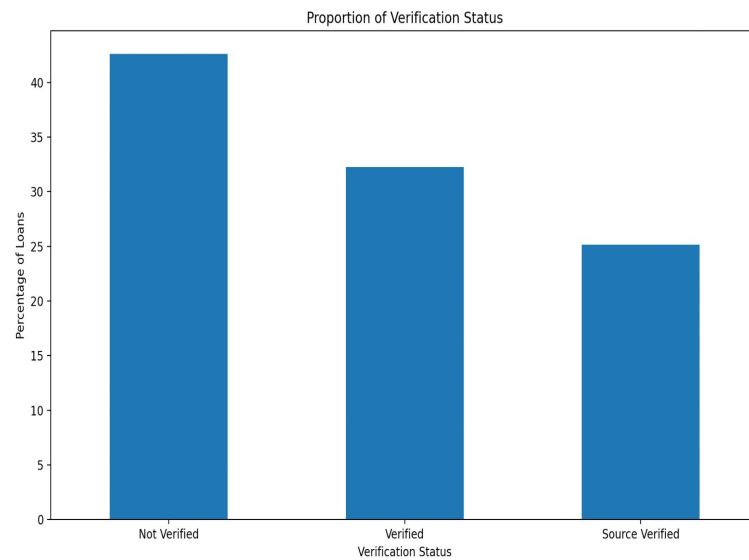
The installment column shows that the average monthly payment is \$327.05, with most borrowers paying between \$168.67 and \$434.75. While the minimum payment is as low as \$16.08, some borrowers have high payments up to \$1,305.19, indicating significant variability in monthly installments.



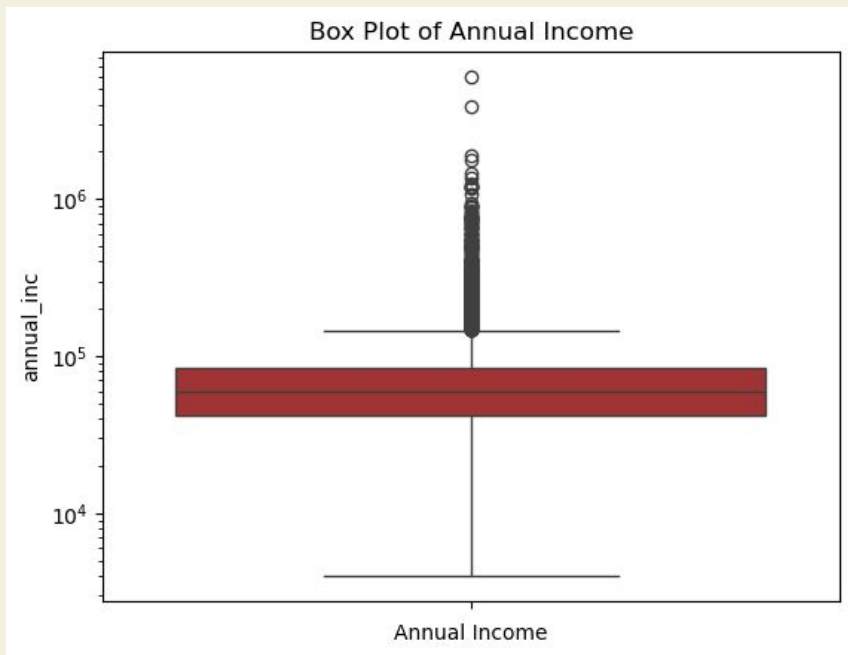
The average DTI (Debt-to-Income Ratio) is 13.37%, with most borrowers having a DTI between 0% and 18.63%. The lowest DTI is 0%, and the highest is 29.99%



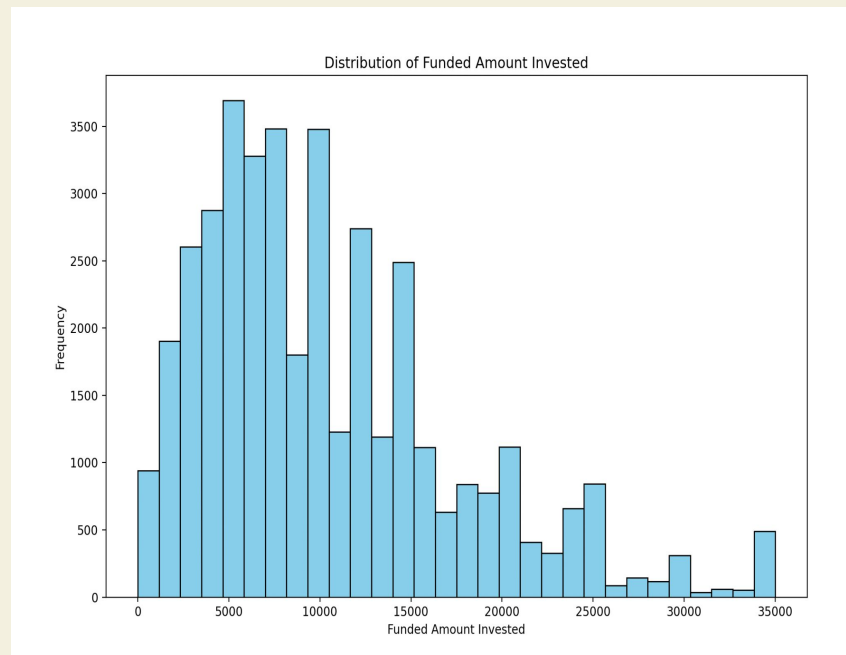
The main reason for taking a loan was to pay off debt, followed by covering credit card bills, with the least common reason being for renewable energy.



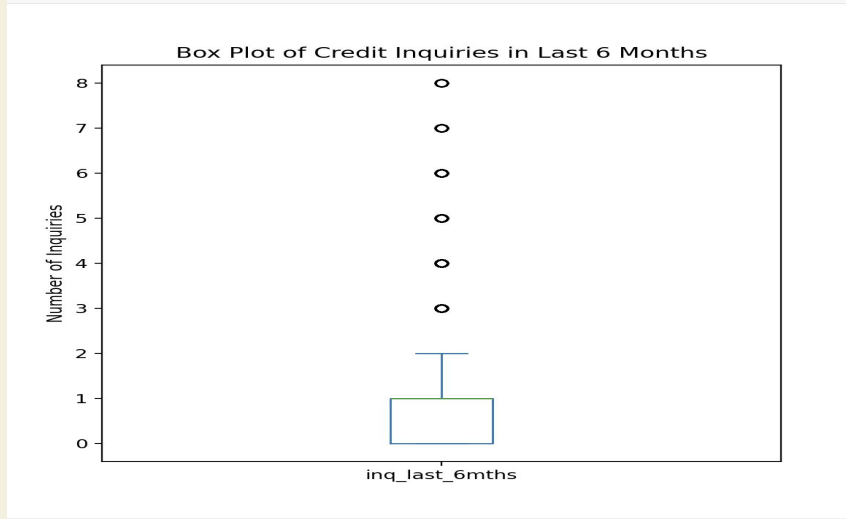
Surprisingly, the largest loans were granted to non-verified borrowers, followed by verified borrowers.



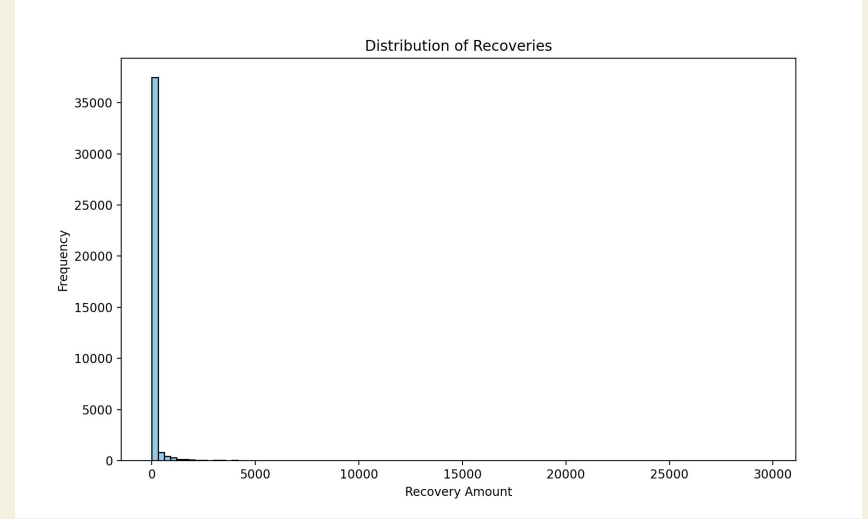
The majority of borrowers had an annual income ranging from 0 to 15,000.



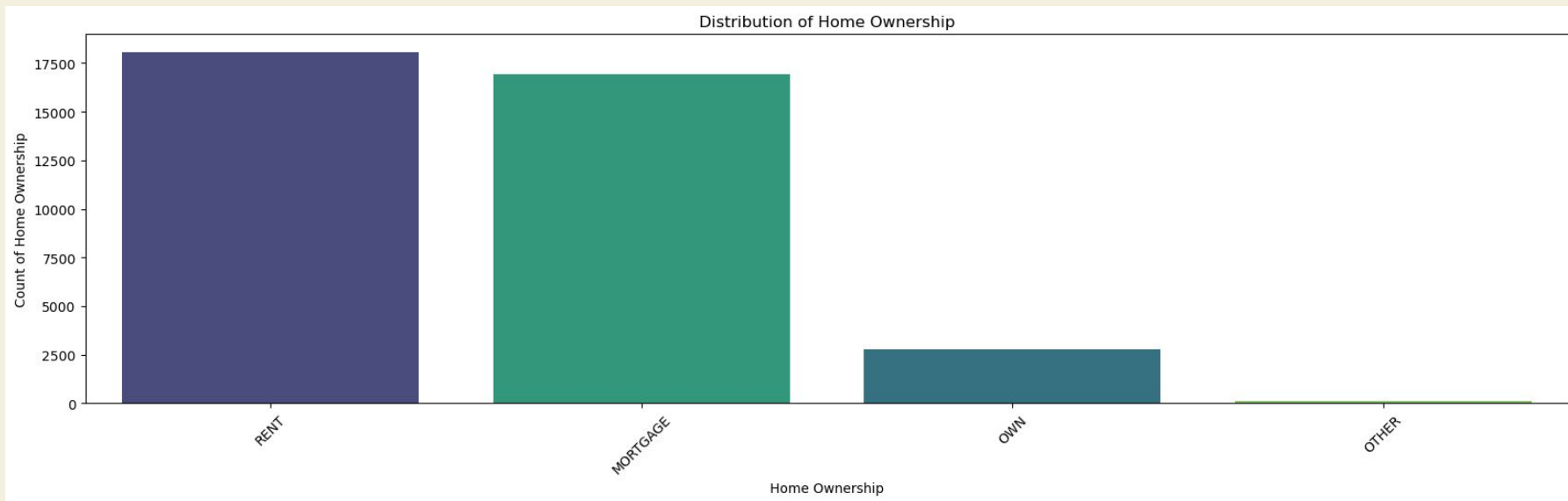
The majority of loans were approved in the range of 5,000 to 10,000.



The graph shows the number of inquiries made to secure a loan, with the maximum reaching up to 8 inquiries.



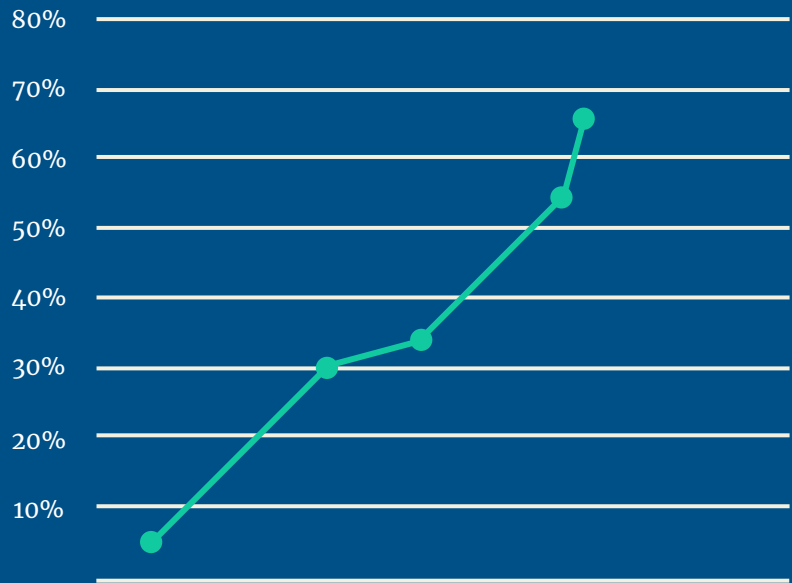
The maximum loan amounts could not be recovered.



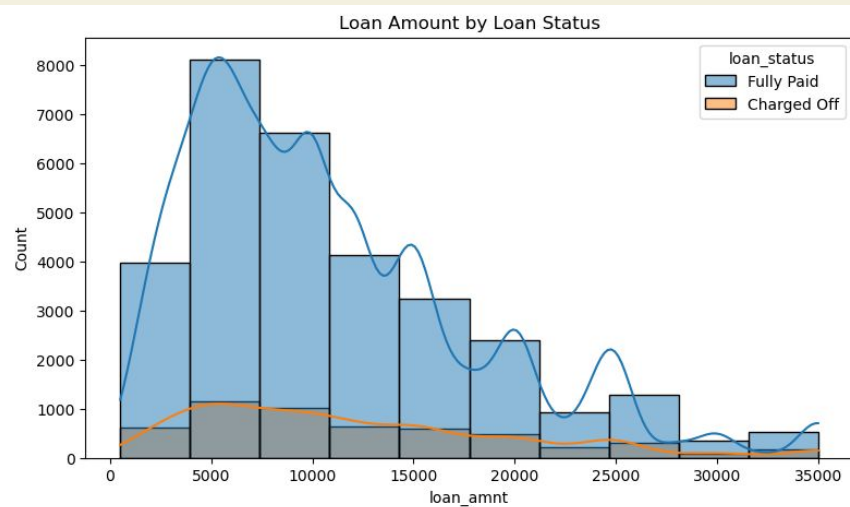
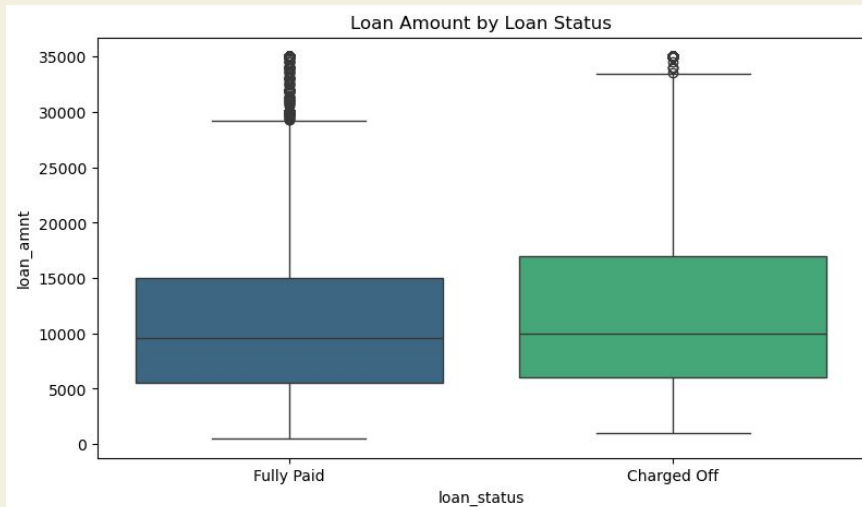
The highest loans were taken by people who rented, followed by those with mortgages.

Segmented Univariate Analysis

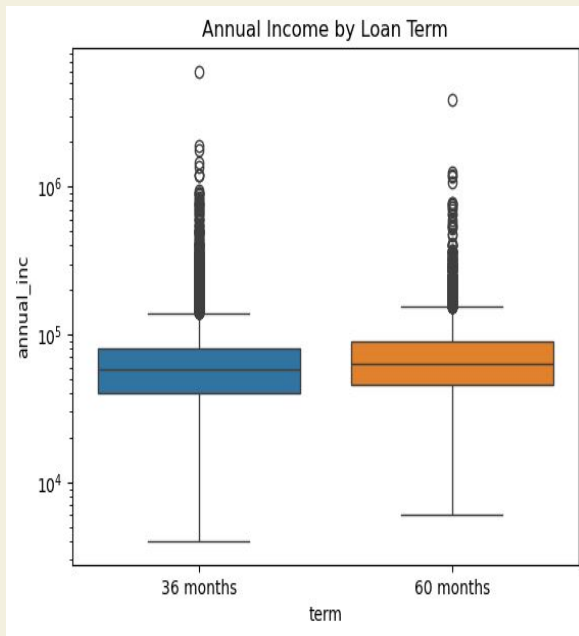
Analysis the continuous data column with respect
to the categorical column



Loan Amount by Loan Status

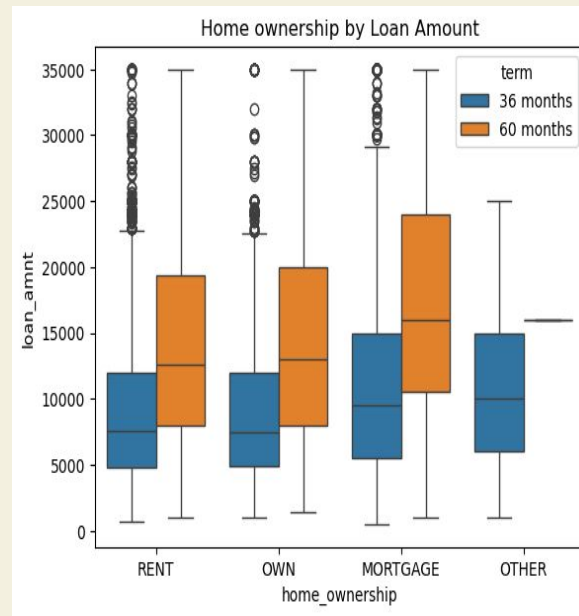


Charged Off loans have a higher average loan amount (\$12,256) compared to Fully Paid loans (\$10,965).



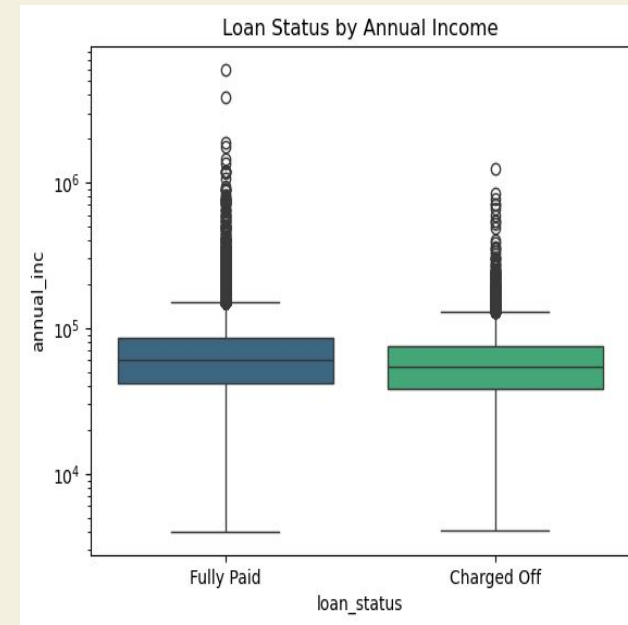
Annual Income by Loan Term

No much differences in the annual income by term



Loan Amount by Home Ownership

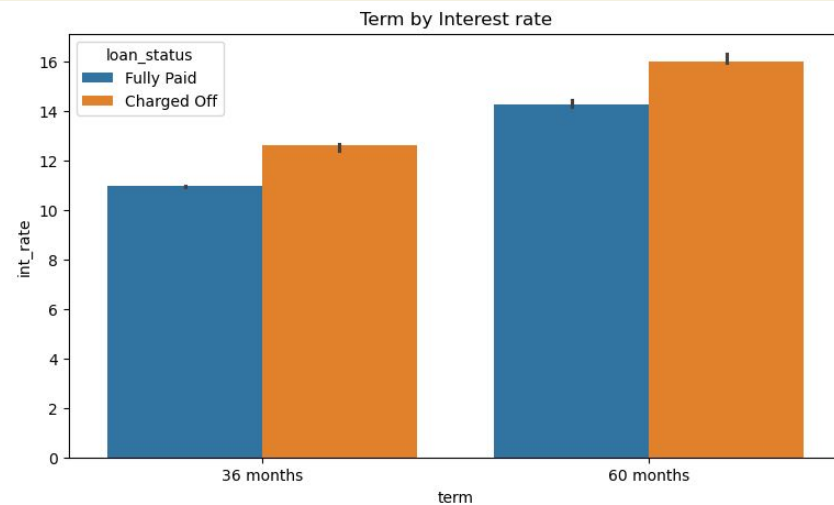
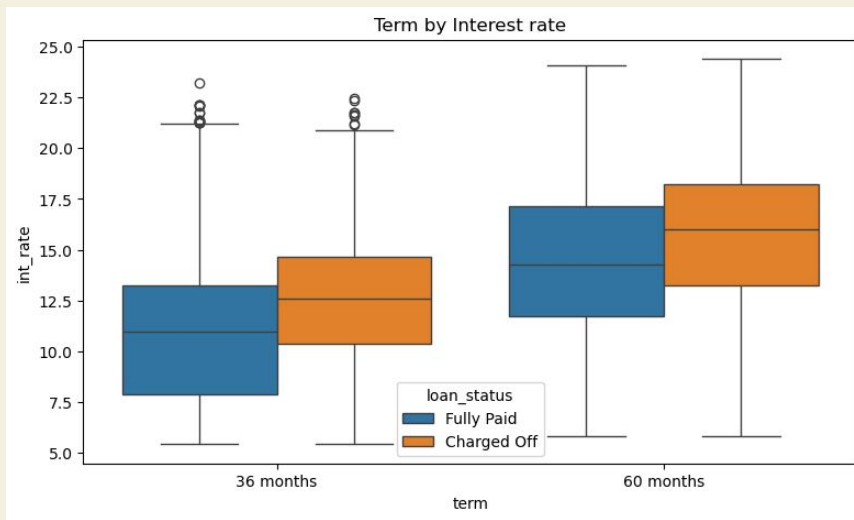
Higher the loan amount in the mortgage homeownership



Annual Income by Loan Status

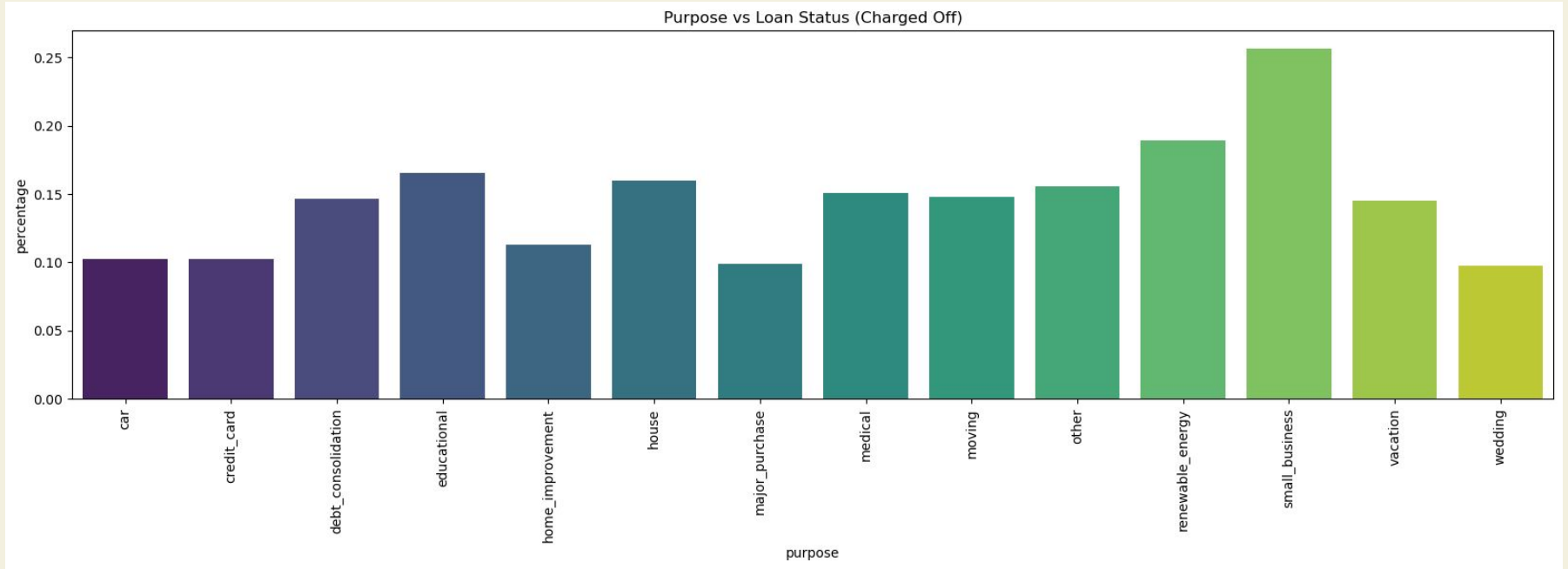
People with higher annual income have fully paid their loan

Term by Interest Rate



In both 36-month and 60-month loan terms, higher interest rates lead to higher default rates.

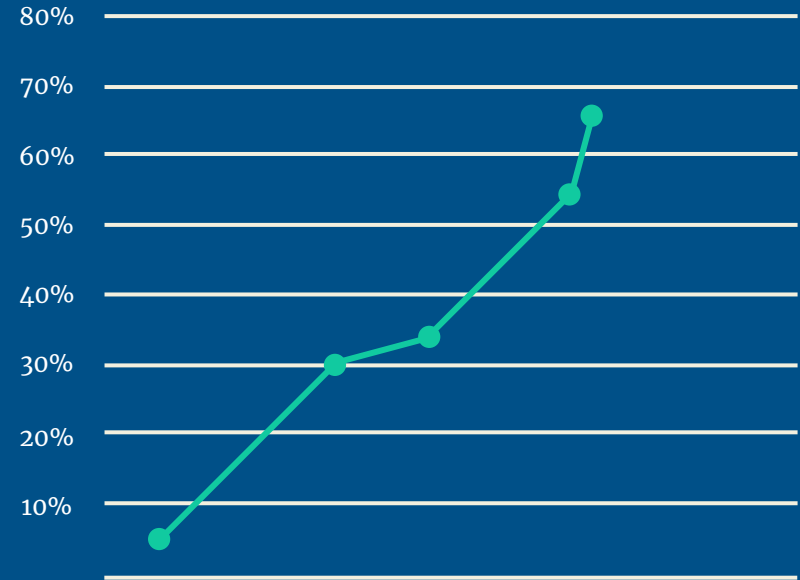
Purpose by Loan Status

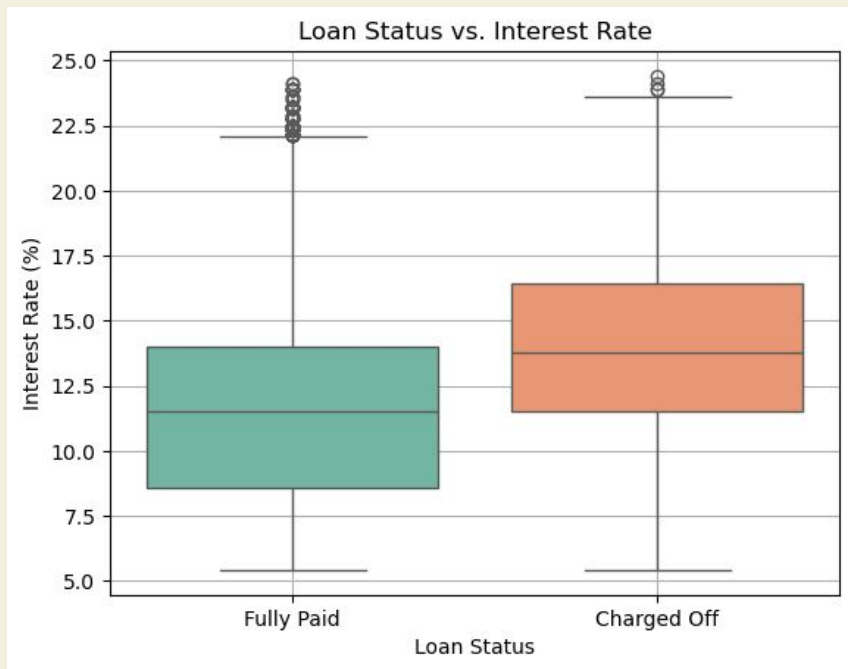


People who bought loan for the small business have defaulted more

Bivariate Analysis

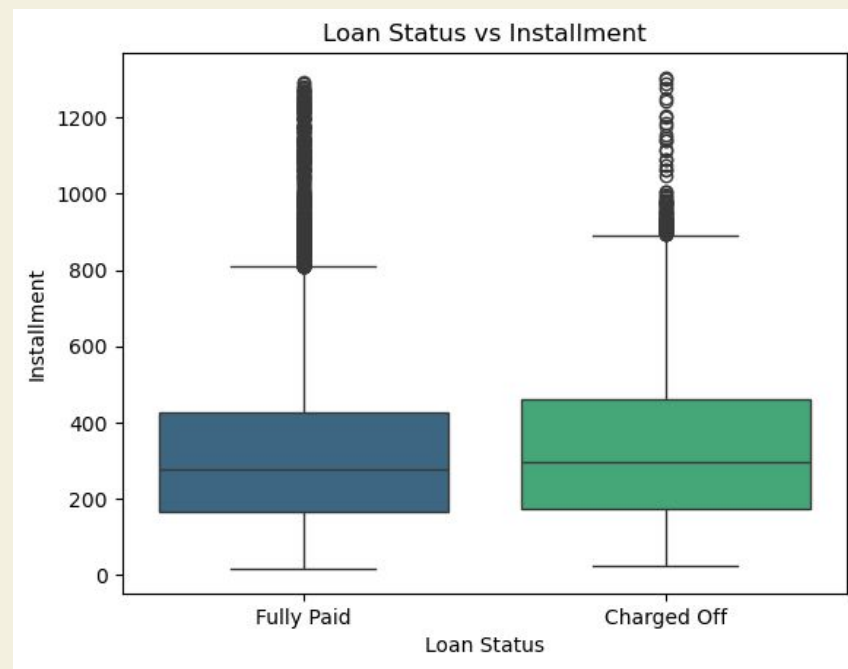
Analysing two variable behaviour





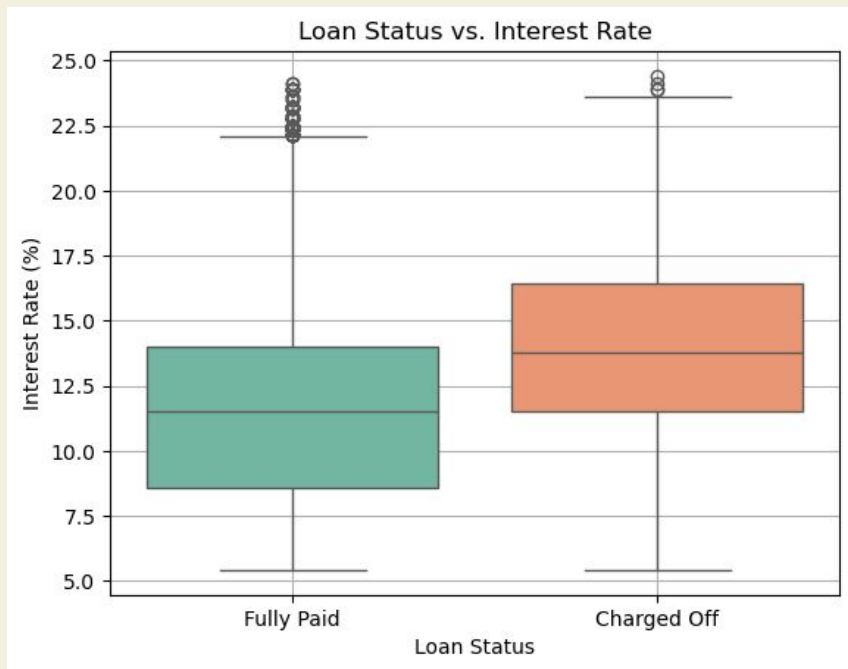
Loan Status vs Interest Rate

Charged off loans have a higher average interest rate (13.93%) than fully paid loans (11.66%), with slightly more variation in rates. The median interest rate for charged off loans (13.79%) is also higher, indicating that loans with higher interest rates tend to default more



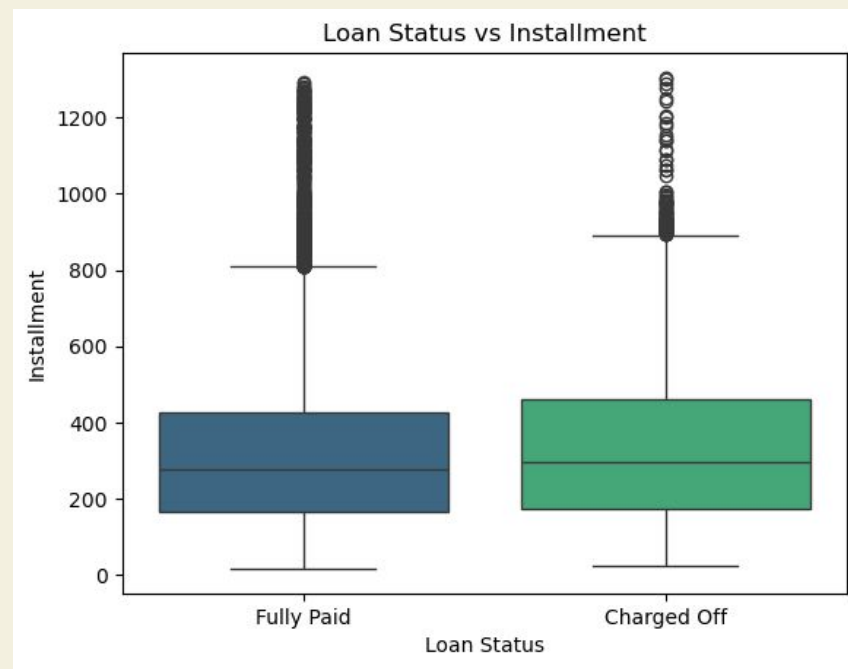
Loan Status vs Installment

Loan Status vs. Installment, charged off loans have slightly higher average installments (\$338.77) compared to fully paid loans (\$322.59). Charged off loans also show more variation in installment amounts (higher standard deviation). The median installment is higher for charged off loans (\$296.75 vs. \$278.15), suggesting that larger loans may have a higher risk of default.



Loan Status vs Interest Rate

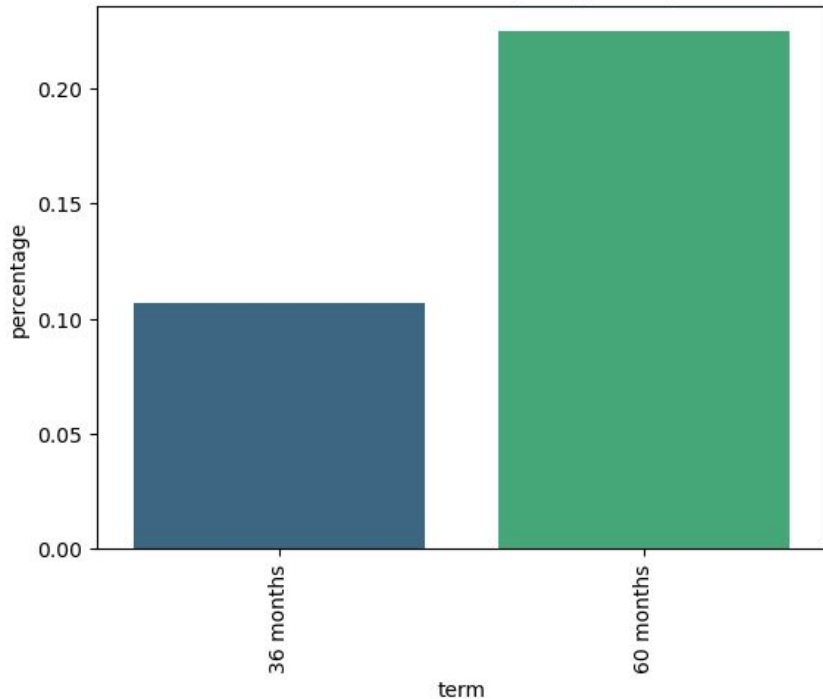
Charged off loans have a higher average interest rate (13.93%) than fully paid loans (11.66%), with slightly more variation in rates. The median interest rate for charged off loans (13.79%) is also higher, indicating that loans with higher interest rates tend to default more



Loan Status vs Installment

Loan Status vs. Installment, charged off loans have slightly higher average installments (\$338.77) compared to fully paid loans (\$322.59). Charged off loans also show more variation in installment amounts (higher standard deviation). The median installment is higher for charged off loans (\$296.75 vs. \$278.15), suggesting that larger loans may have a higher risk of default.

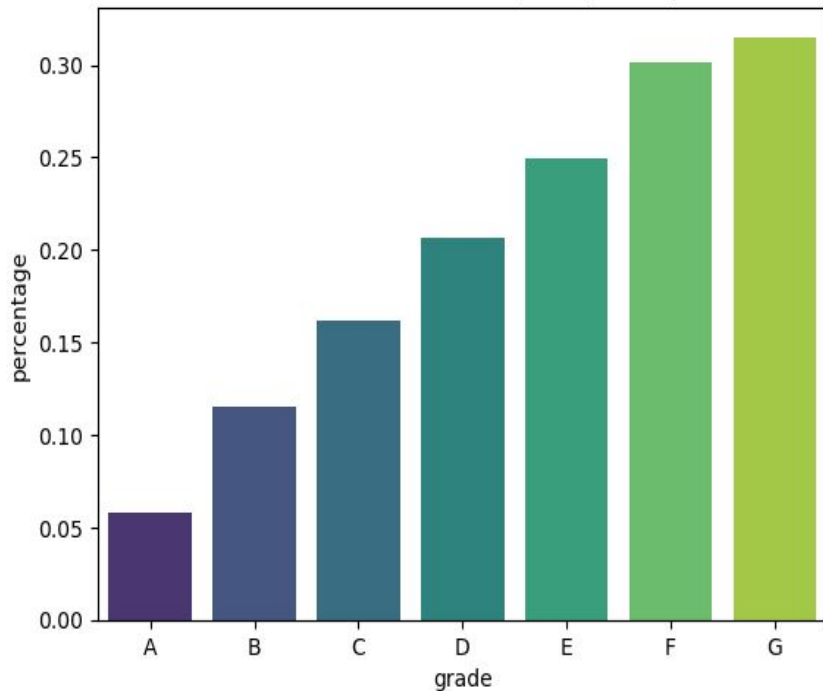
Term vs Loan Status (Charged Off)



Loan Status vs Term

Charged off loans are more when the term is 60 months

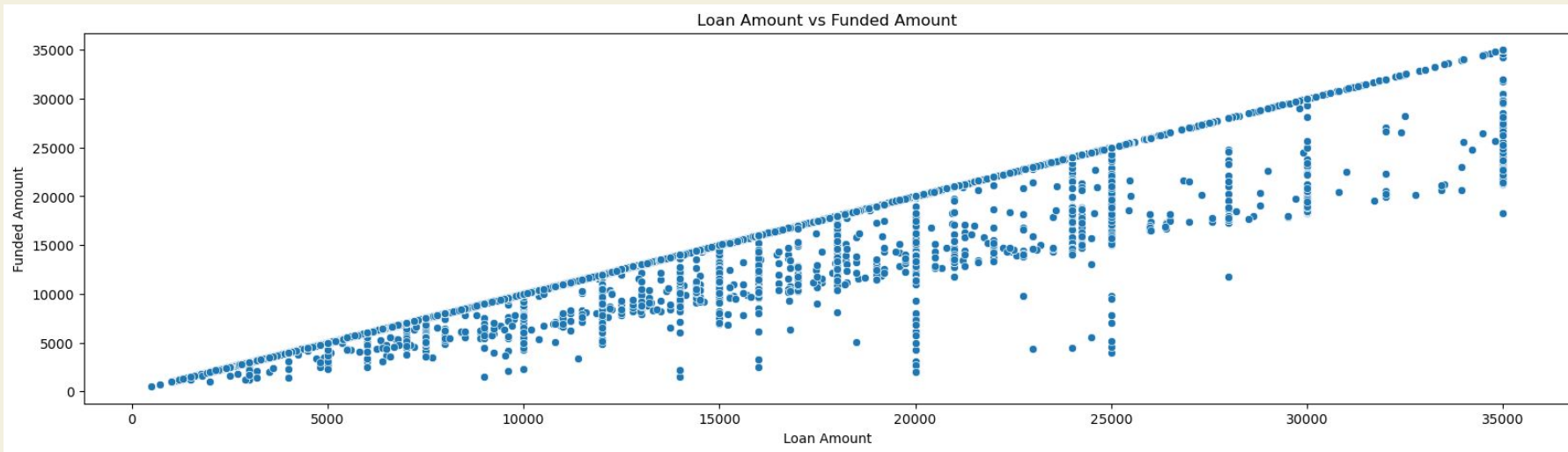
Grade vs Loan Status (Charged Off)



Loan Status vs Grader

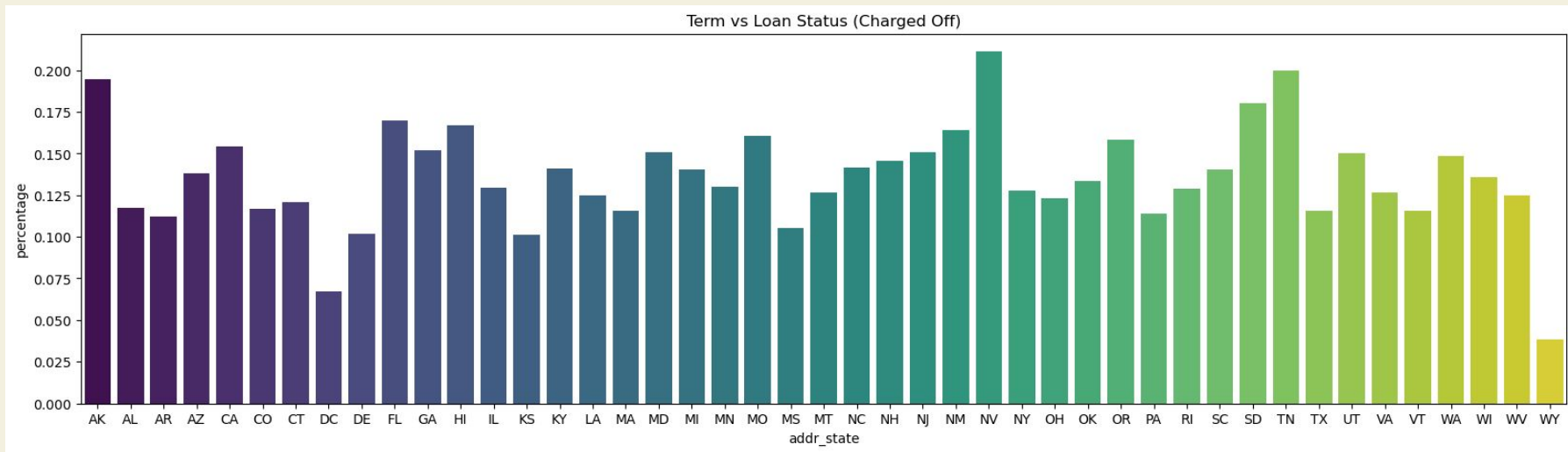
Charged Off loans are more when the grade is F and G.

Loan Amount vs Funded Amount



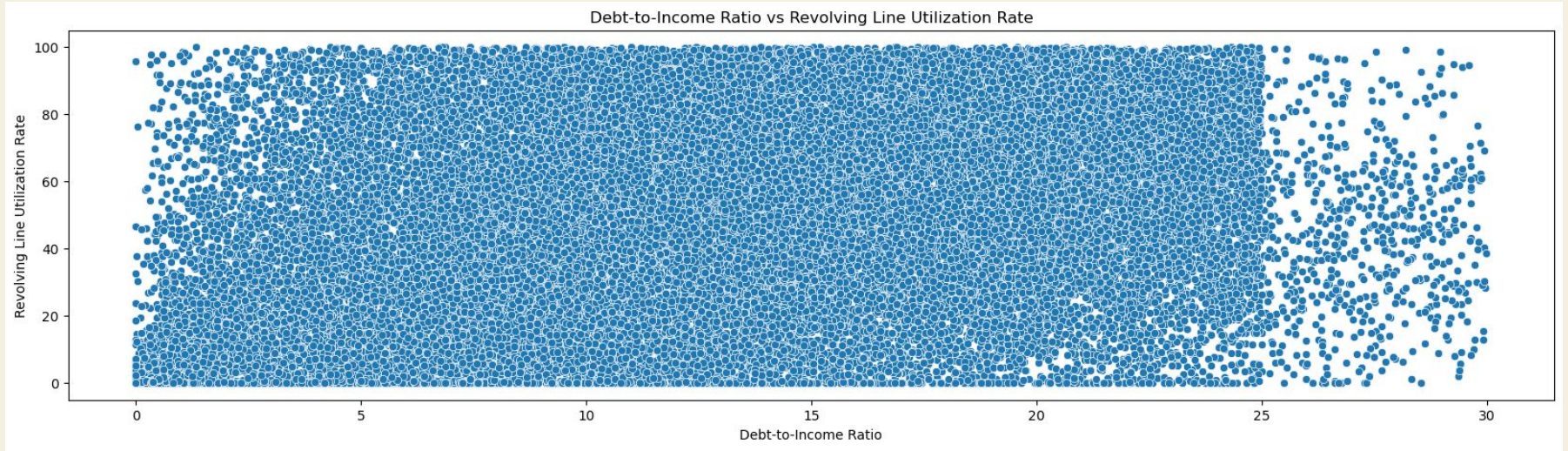
The correlation between Loan Amount and Funded Amount is 0.98, showing a very strong relationship. This means that the funded amount is almost always close to the loan amount. When the loan amount goes up, the funded amount increases similarly.

Address State vs Loan Status



People living in the state NV tend to have more charged off loans

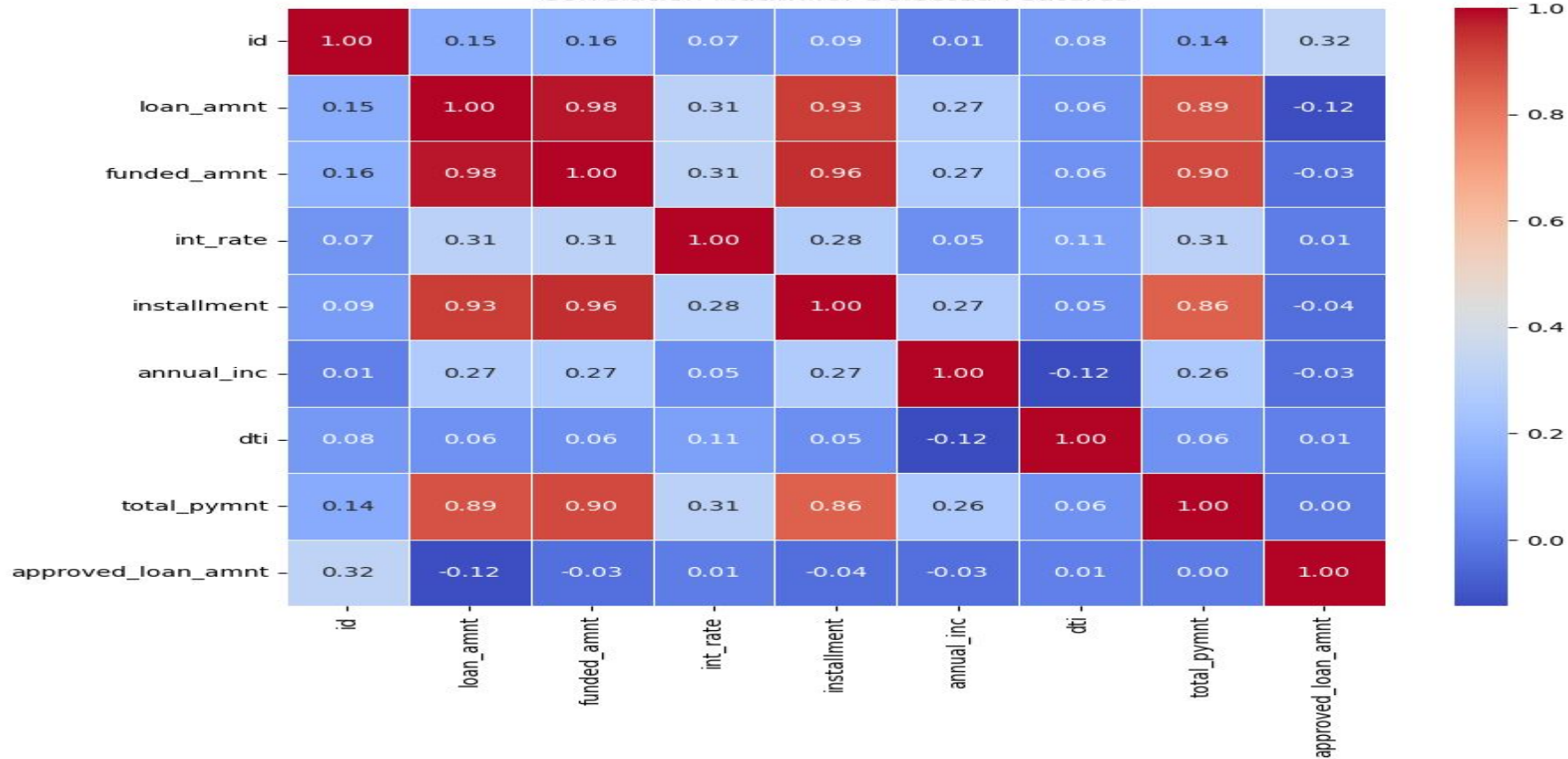
Debit to Income Ratio vs Revolving Credit Utilization



The correlation between DTI (Debt-to-Income Ratio) and Revolving Utilization (Revol_Util) is 0.28, indicating a weak positive correlation. This suggests that as the debt-to-income ratio increases, revolving utilization tends to increase slightly as well, but the relationship is not very strong.

Overall Correlation

Correlation Matrix for Selected Features



Conclusion

Loan Acceptance vs. Loan Rejection:

- Loans with amounts ranging between \$5,000 and \$15,000 tend to be approved.
- Loans exceeding \$30,000 are more likely to be rejected.

Loan Purpose:

- Loans for purposes such as car purchases, credit card debt consolidation, and vacations are more likely to be accepted.
- Loans for small business purposes tend to face higher rejection rates.

Loan Term:

- Loans with a 36-month term are more likely to be accepted.
- Loans with a 60-month term are more frequently rejected.

Credit Grade:

- Borrowers with grades A and B are more likely to have their loans approved.
- Borrowers with grades E, F, and G tend to face higher rejection rates.

Annual Income:

- Borrowers with annual incomes above \$100,000 are more likely to be approved for loans.
- Borrowers with annual incomes below \$25,000 are more likely to be rejected.

Interest Rates:

- Loans with lower interest rates tend to have higher approval rates.
- Loans with high interest rates are more likely to be rejected.

Geographic Trends:

- Borrowers living in New York tend to have higher rates of loan charge-offs.