Loan Default Analysis

Understanding the Driving Factors Behind Loan Defaults

Agenda

- Problem Statement
- Analysis Approach
- Univariate Analysis
- Bivariate Analysis
- Correlation Analysis
- Key Insights
- Conclusion

Problem Statement

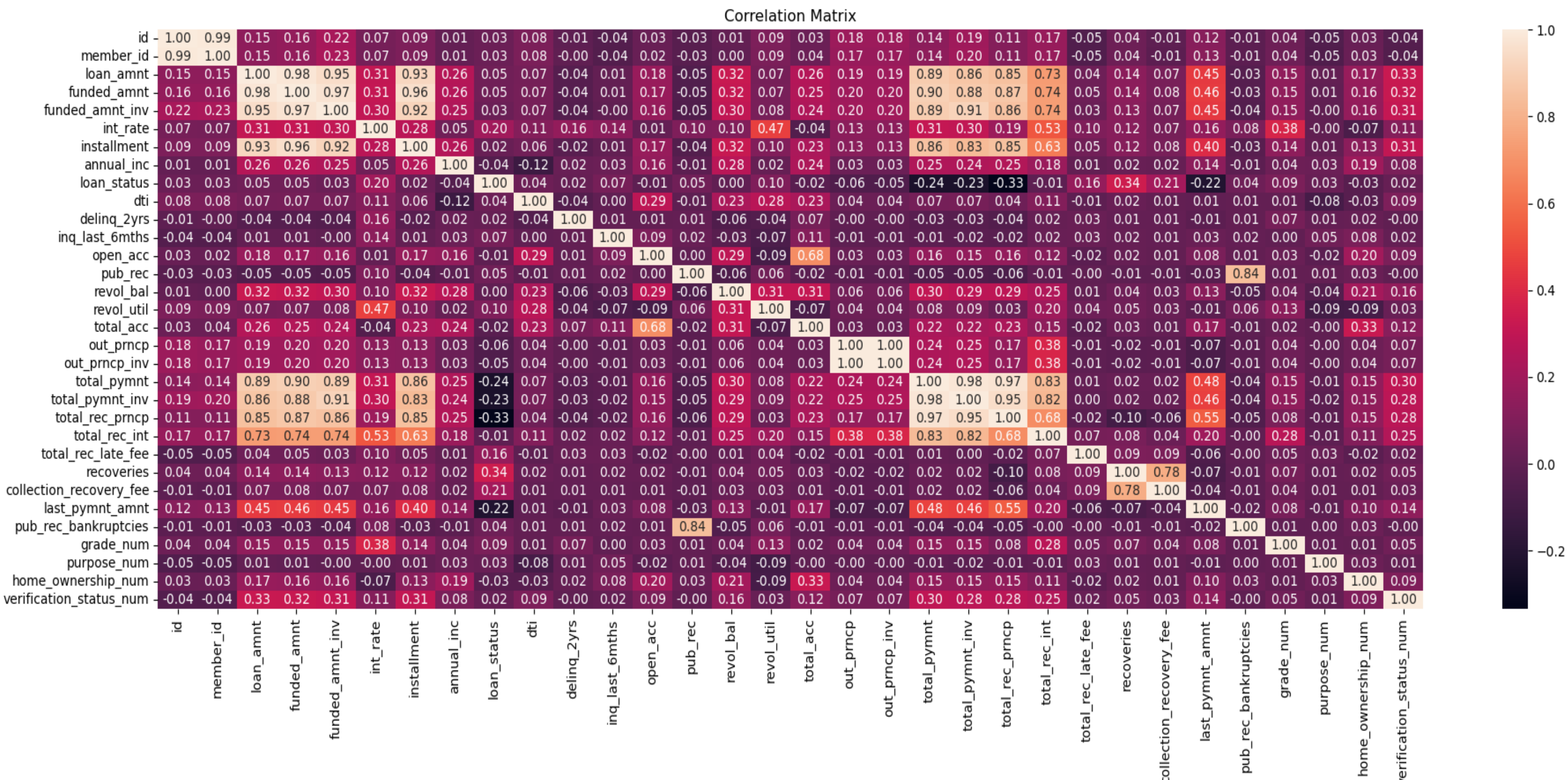
The objective of this analysis is to identify patterns that indicate if a person is likely to default on a loan. By understanding these patterns, the lending company can take actions such as denying the loan, reducing the loan amount, or lending to risky applicants at a higher interest rate, thereby reducing credit loss.

Analysis Approach

- Data Cleaning: Handle missing values by taking median or max values.
- Univariate Analysis: Analyze the distribution of individual features.
- Bivariate Analysis: Examine the relationship between features and the target variable (loan default).
- Correlation Analysis: Identify significant correlations between features.
- Visualization: Use plots to illustrate key findings and insights.

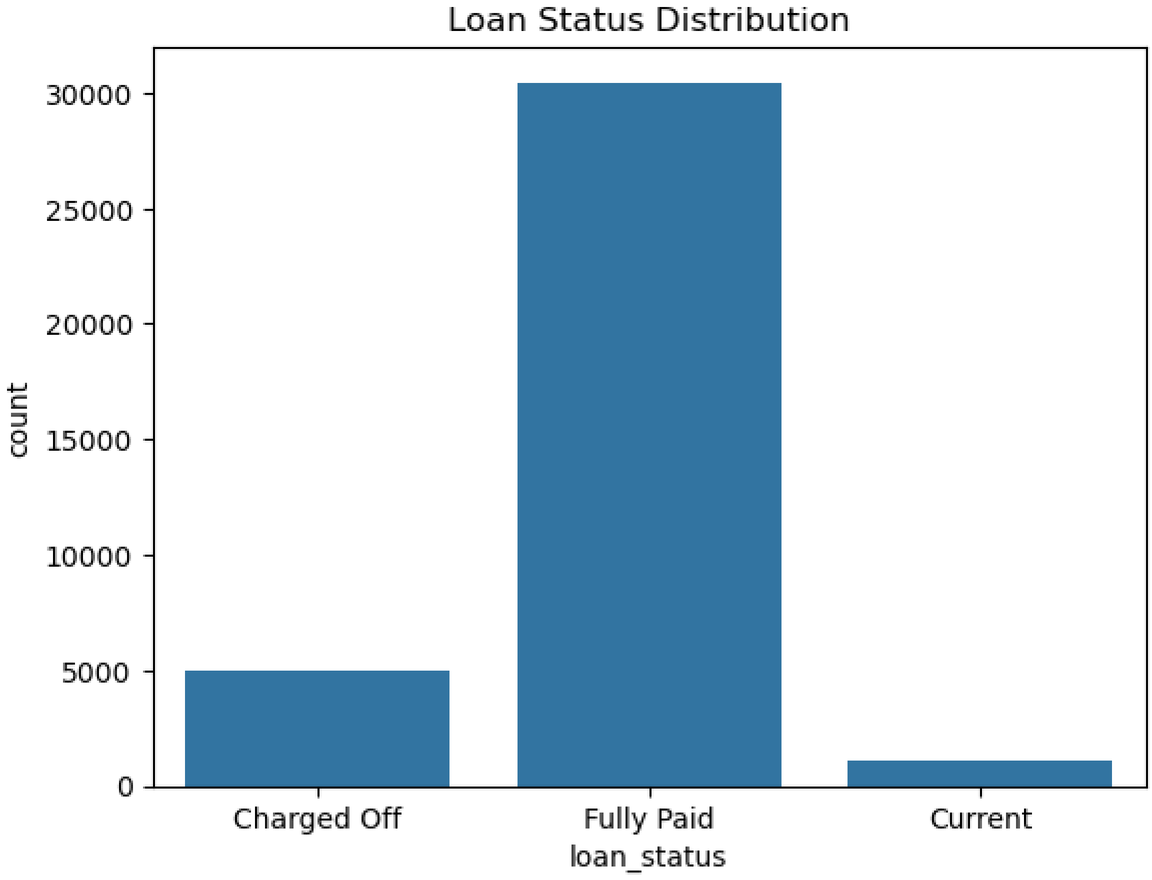
Multivariate Analysis - Correlation Matrix

- This chart provide with the view of overall data to select the area for analysis.
- Here columns like loan-amt, dti, installment, annual_inc, Grade, Verification status ,owners have correlation ratio that attracted to analyse the pattern for defaulting loans.



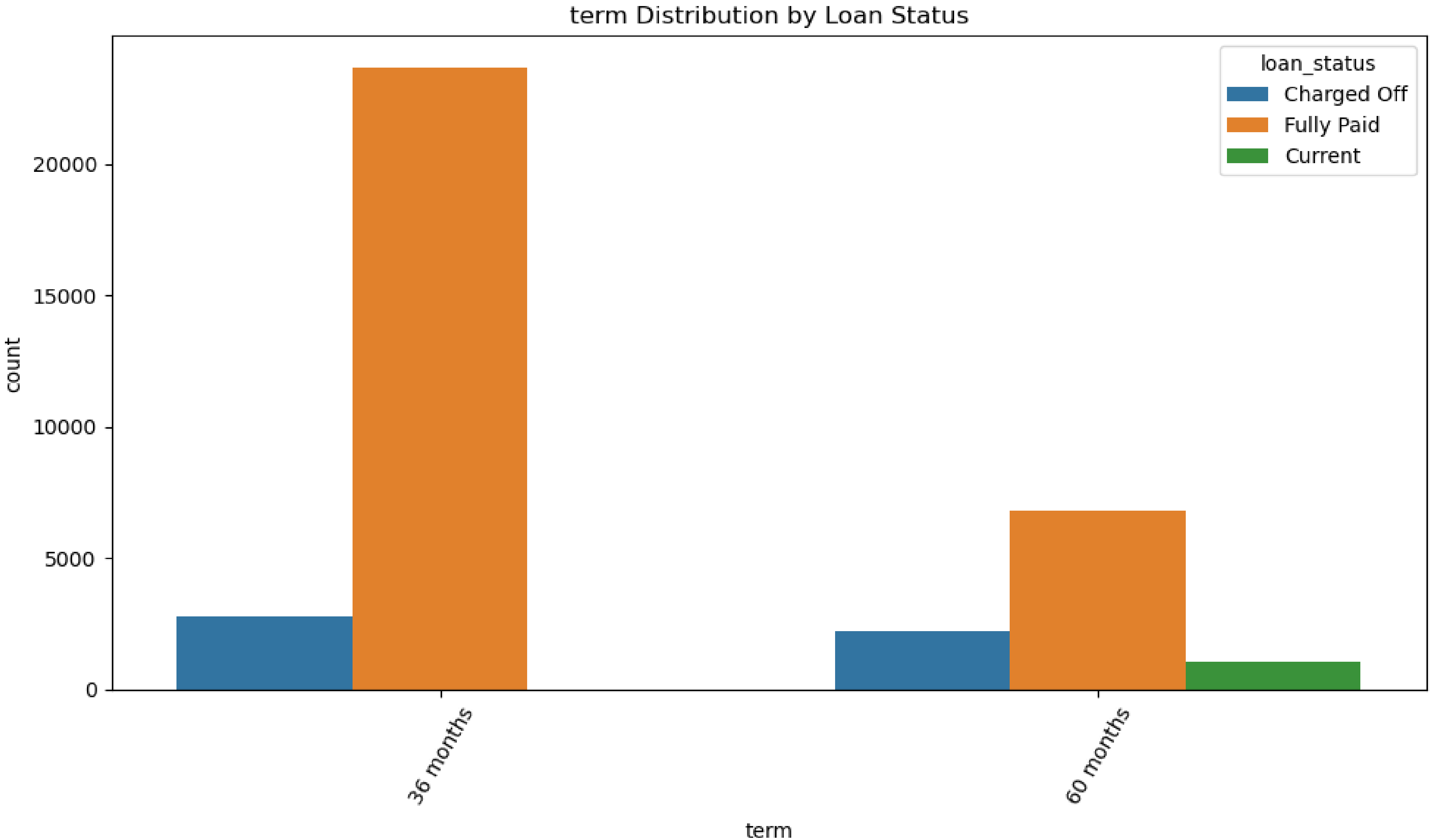
Loan Distribution Pattern

- Maximum loan is fully paid
- 20% of loan is defaulted
- 3% of loan is in progress



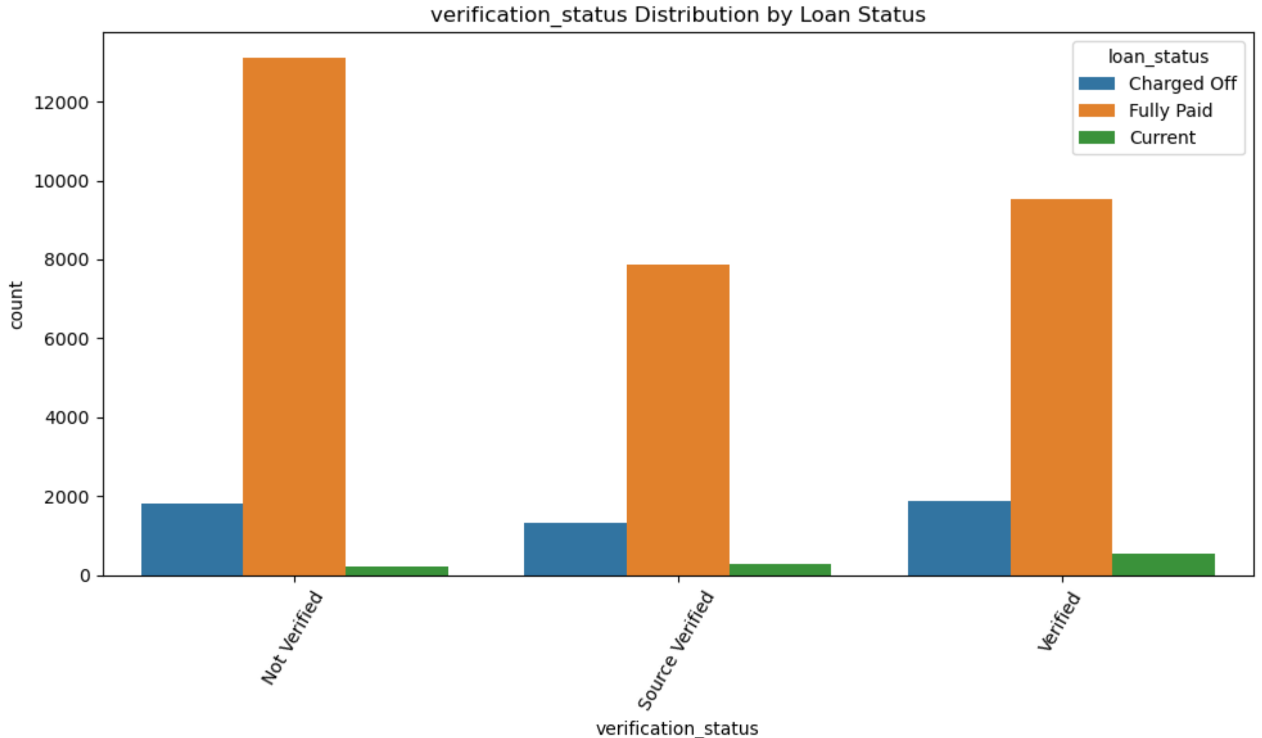
Distribution of Loan for Term

- Loan provided for 60 months tenures has defaulted more
- Customers credit limit should be keep check periodically for current running loans.



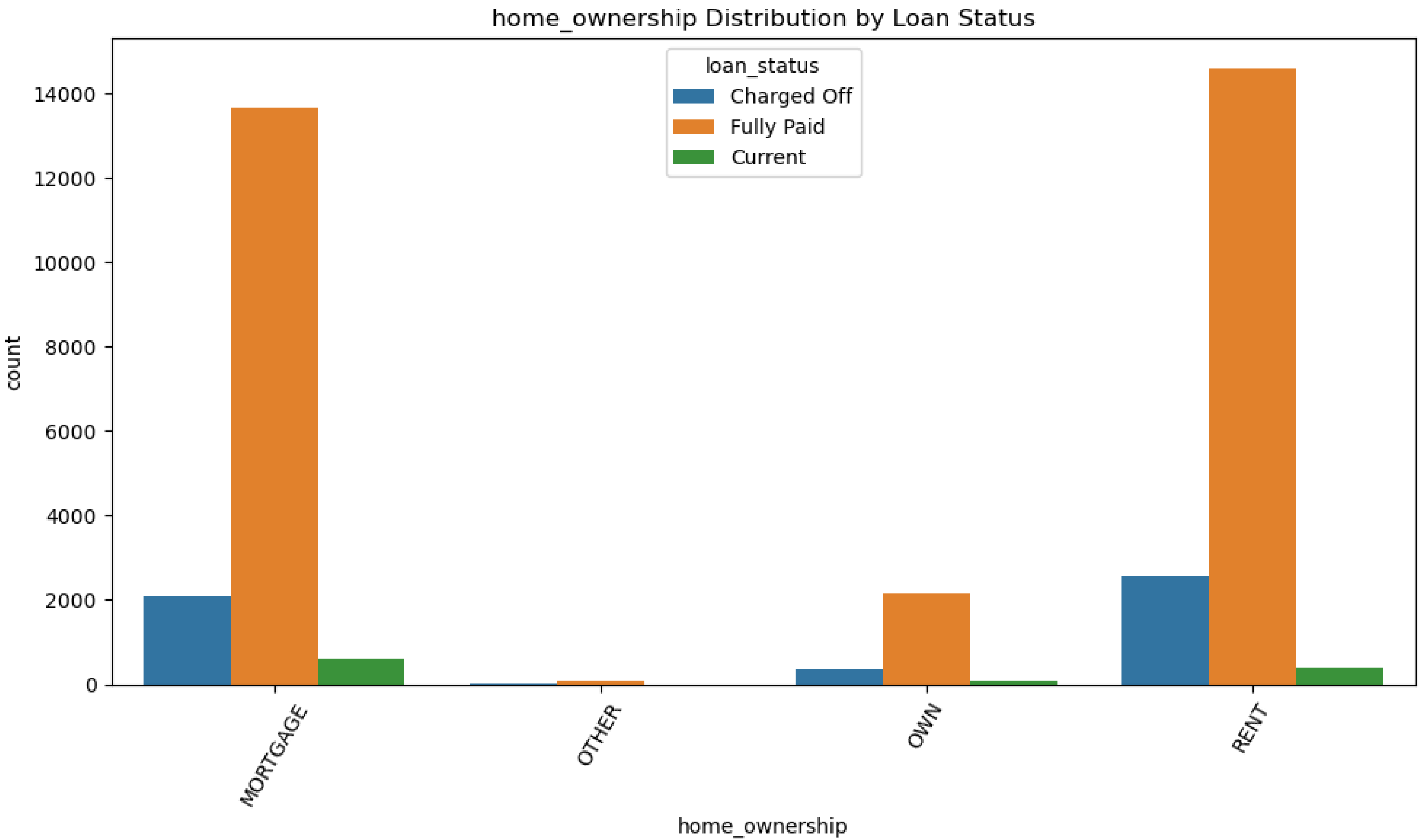
Verification status Distribution by Loan Status

- People that are not verified has taken more loans.
- People verified has defaulted more.
- So risk can be cover by increasing processing fee/charges for lending loans.
- Source verified and Verified is showing more defaulters that should be controlled against strict validation and backlog check



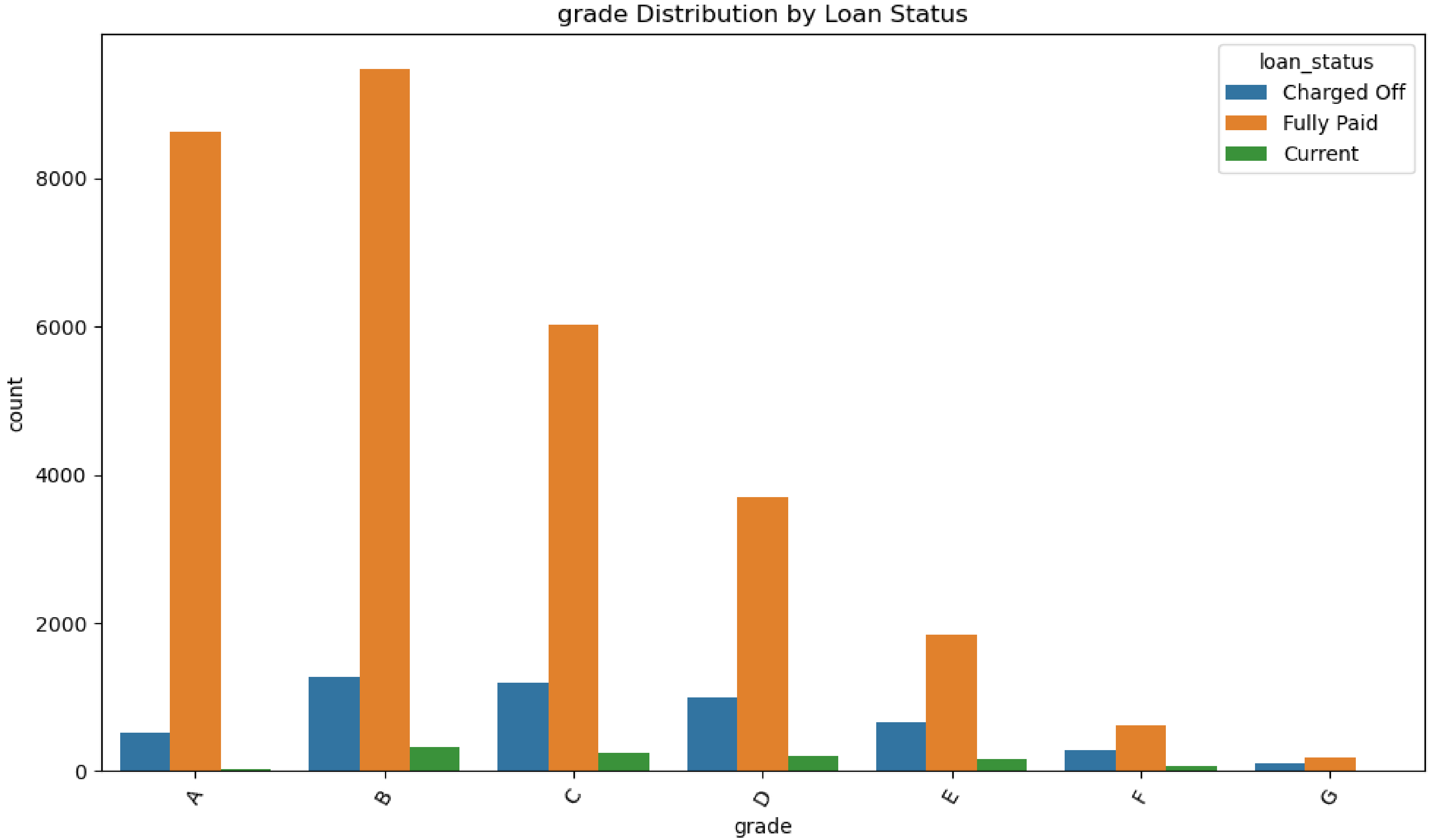
Distribution of Home Ownership

- Own ,Other customers has maximum defaulted in percentage compared to Mortgage and Rent
- Rented customers can be process with higher processing fee and rate of interest.



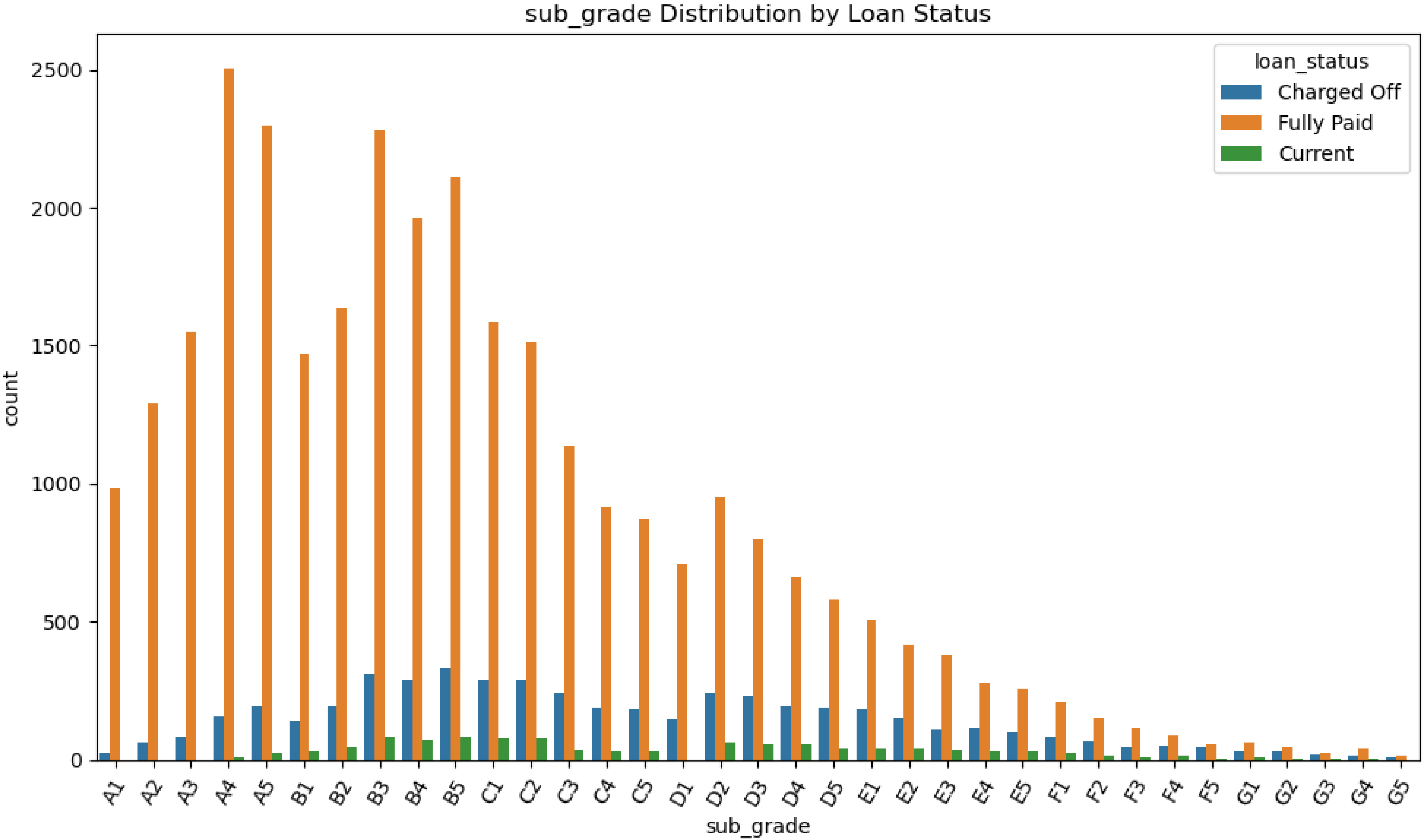
Distribution as per Grade

- Maximum ratio of loan distributed among D,E,F,G Grade is defaulted



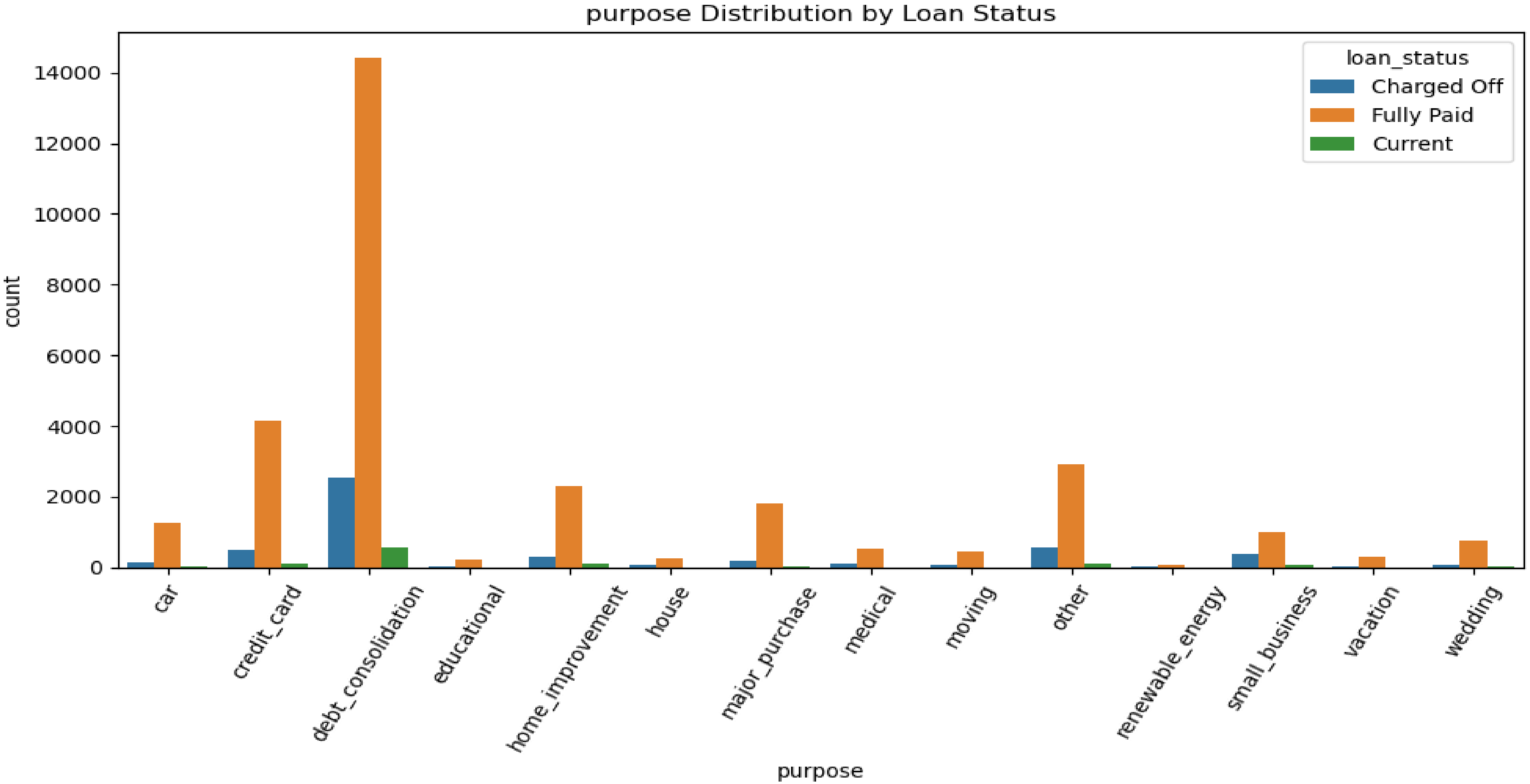
Distribution of loan among Sub Grade

- Sub grade from D1 to G5 has more defaulter
- While lending loan to this group can process with mortgage or higher rate of interest.



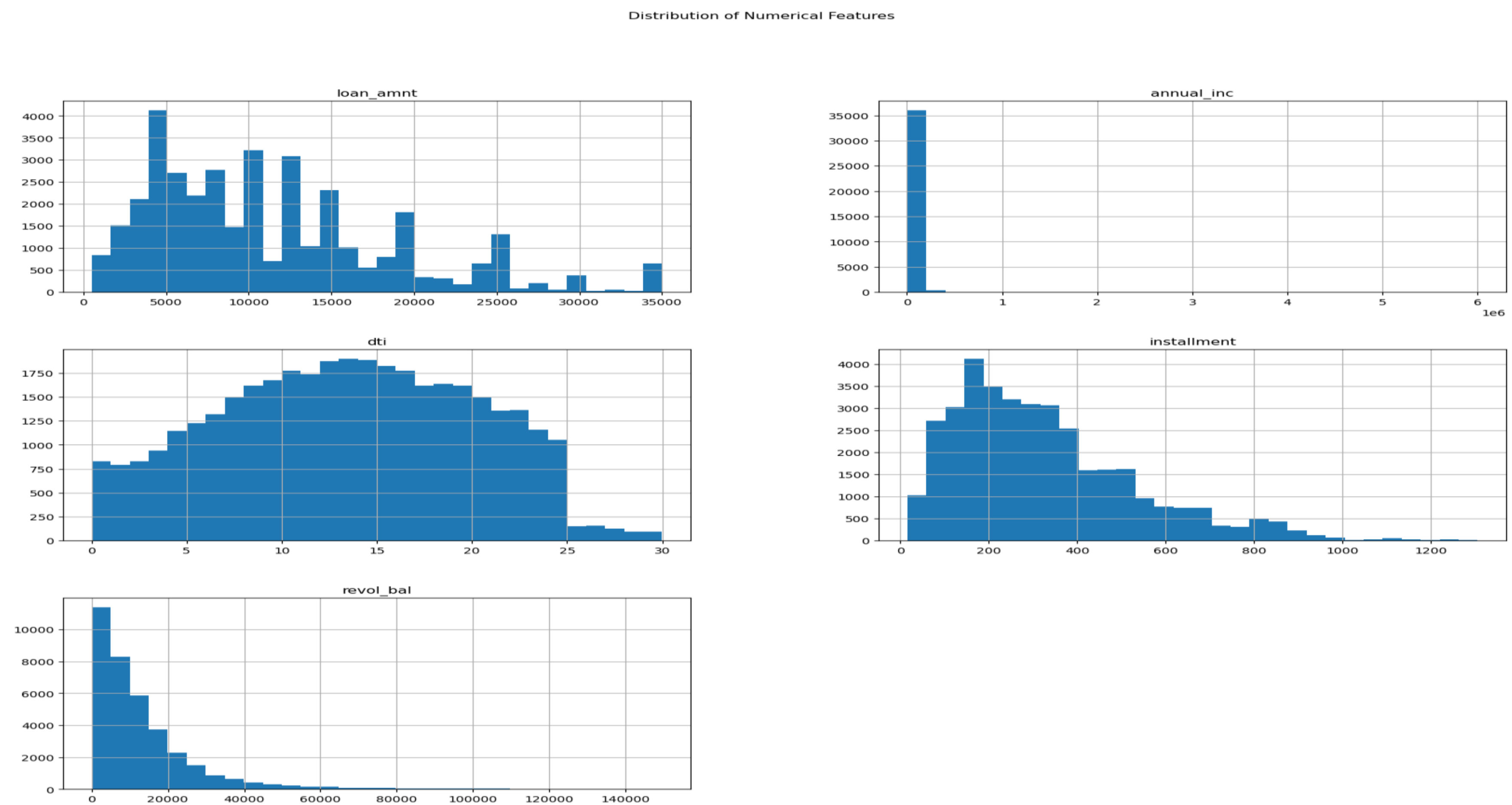
Distribution of Loan for loan purpose

- Debt consolidation has given business but has noticeable defaulters as well
- Rest customers are giving less business but has defaulted more in proportion of business done.



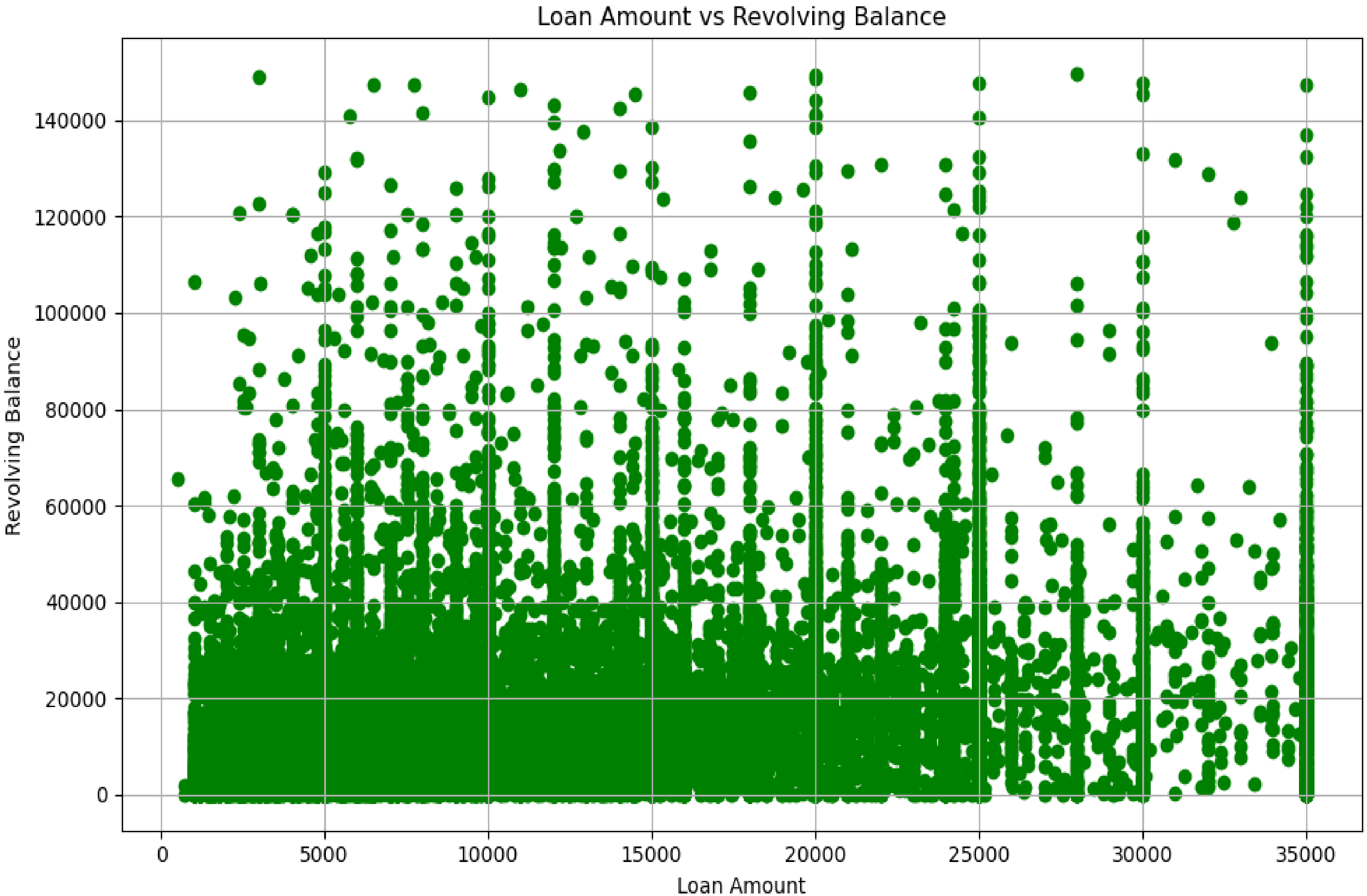
Univariate Analysis - Numerical Features

- Loan amount is positively skewed
- Loan is taken by maximum people having income less than 50k
- Loan is taken by the people having DTI ration between 5 to 25
- Instalments preferred is in between 100 to 500
- Revol_bal is positively skewed



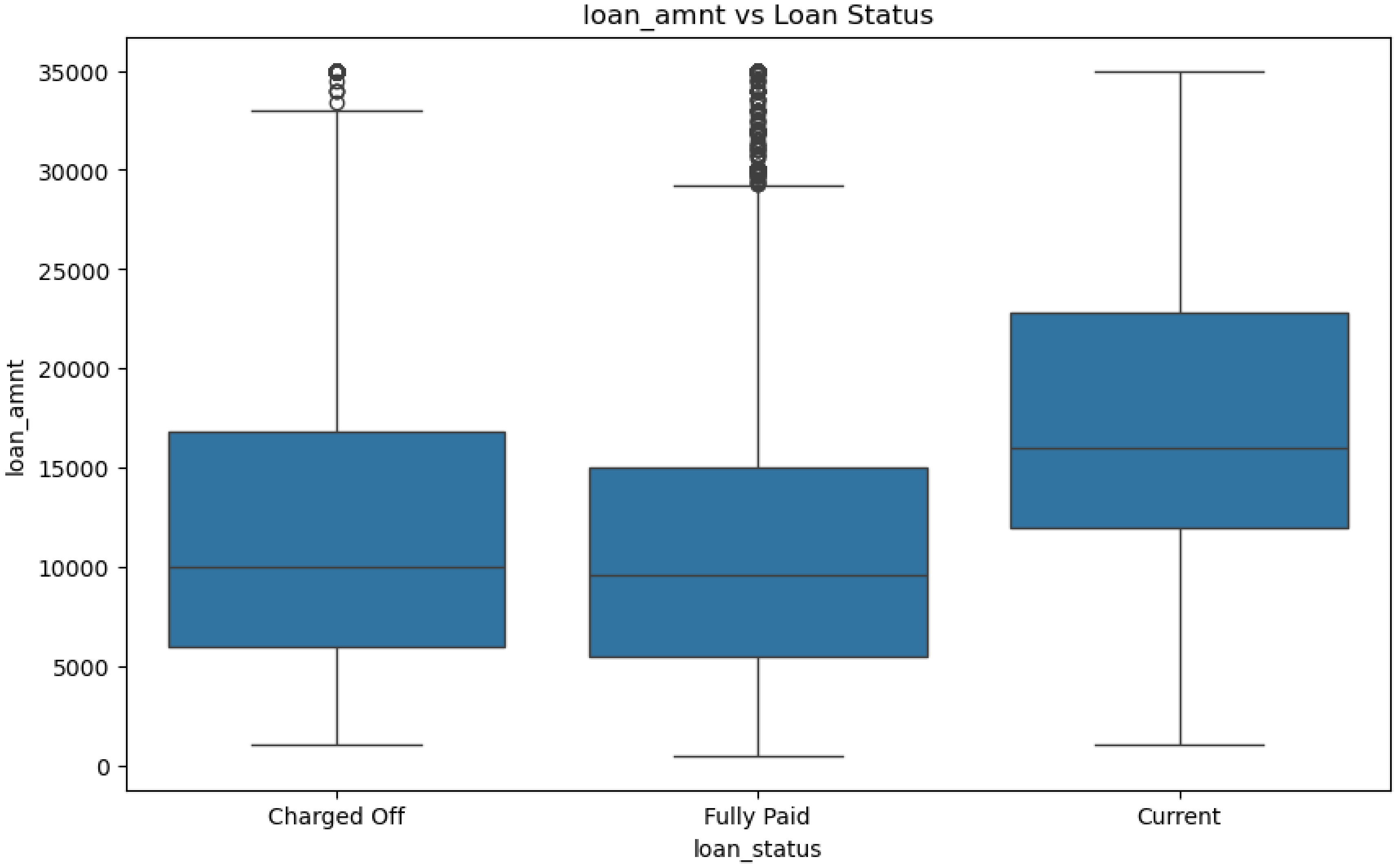
Bivariate Analysis - Loan Amount Vs Revolving Balance

- Loan demanded seems in lower range of Revolving balance customer.
- Max demand of loan is with Revolving balance 40K and Loan Amount in range of 25K



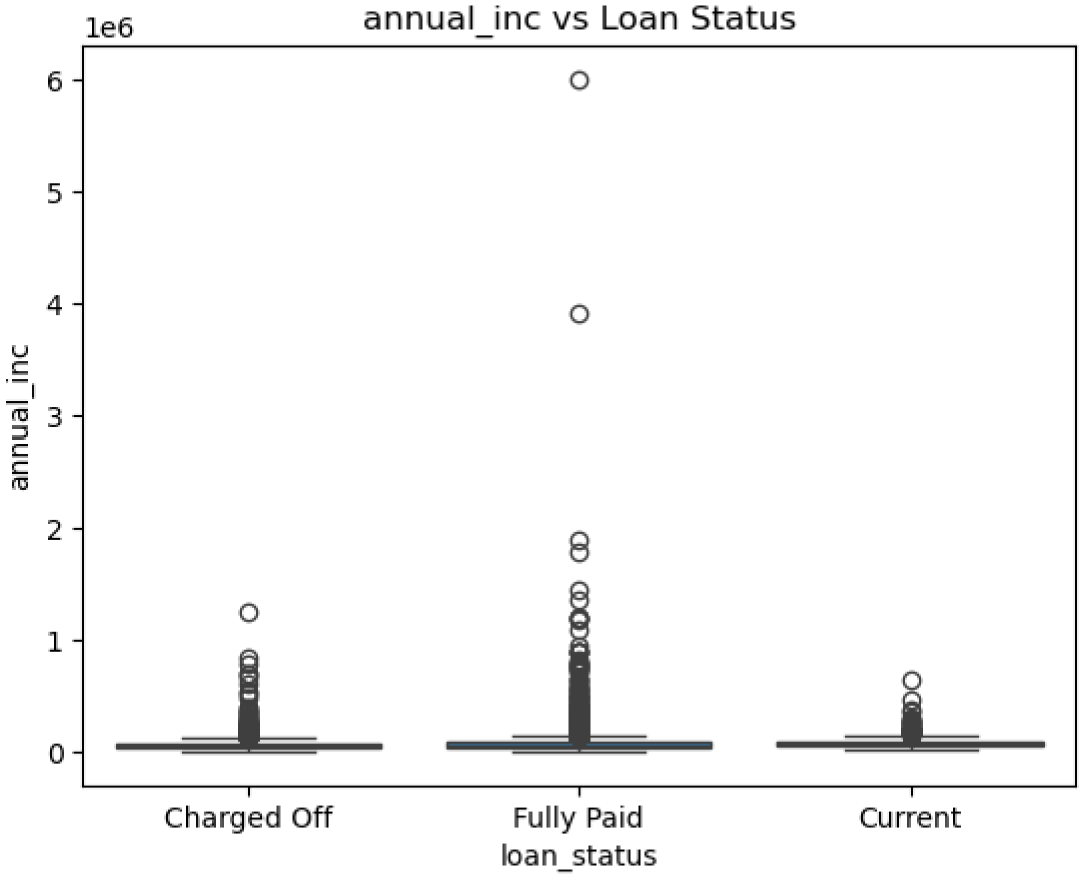
Loan Amount disbursed vs Loan status

- Charged off boxplot is positive skewed hence there is possibility that more loans in the range of 5K to 18K get defaulted for current running loans



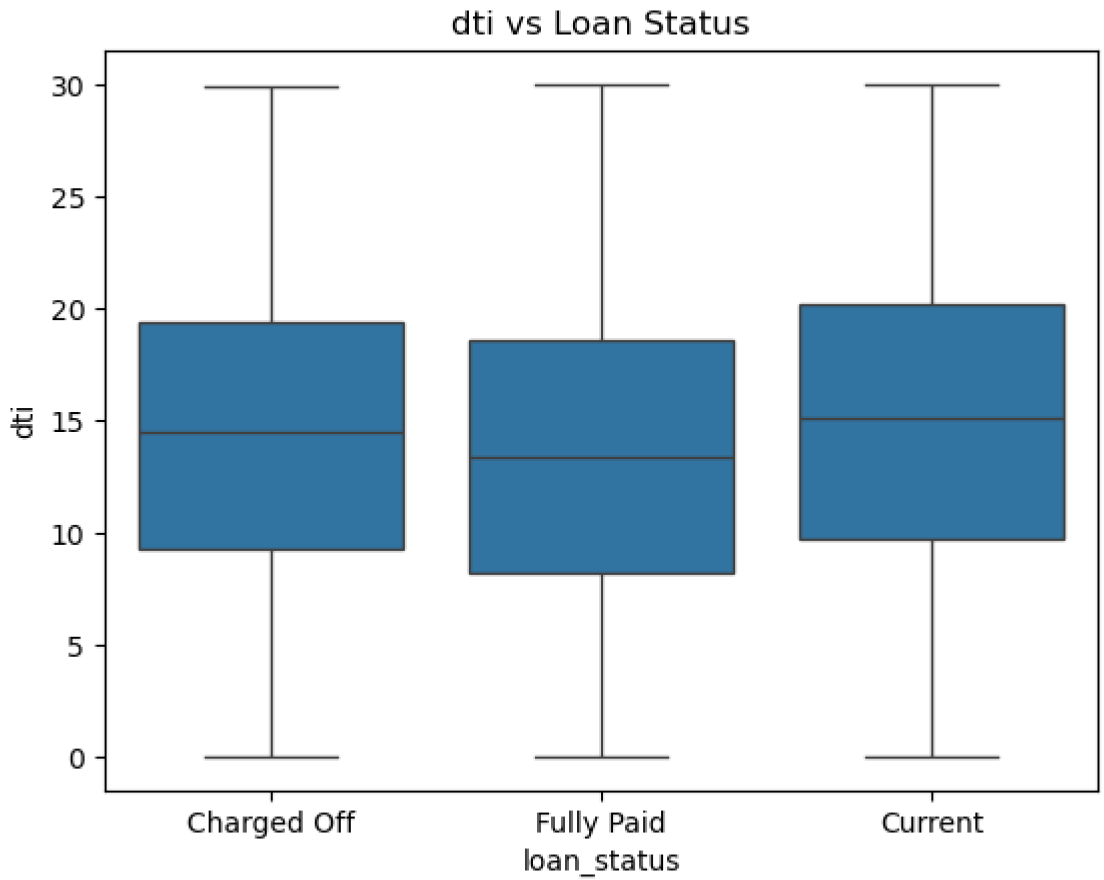
Annual Income vs Loan Status

- Lower income group is more to default loans



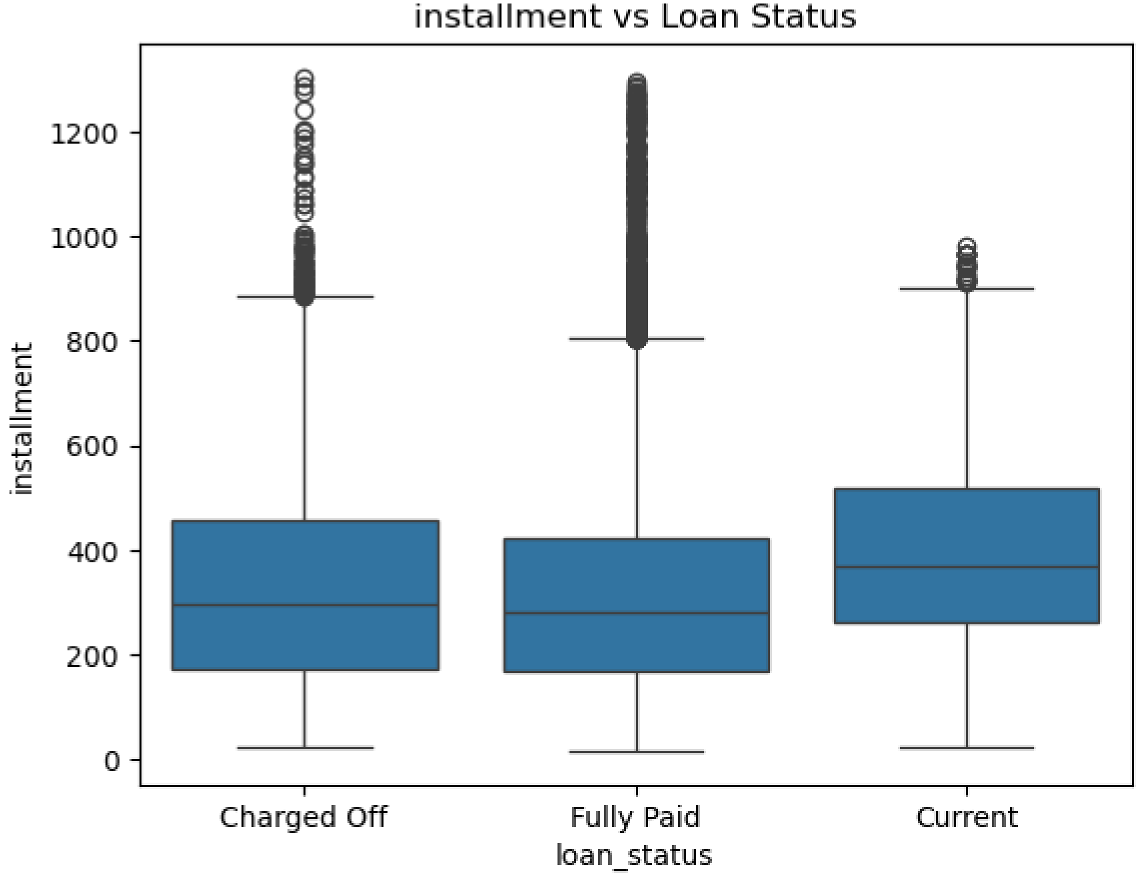
DTI Ration Vs Loan status

- Current loan going customers is tending more to default in the range of DTI range of 10-15

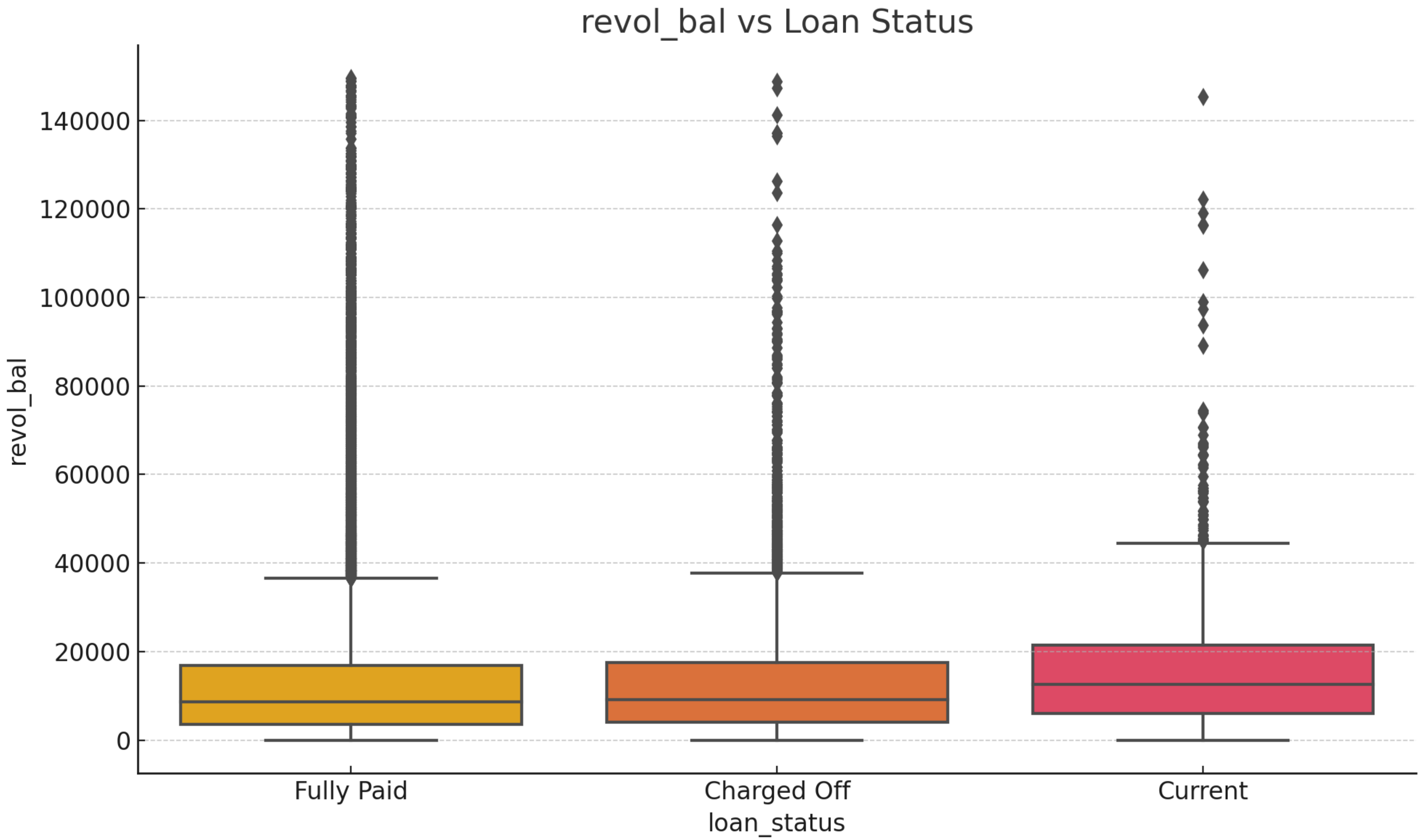


Installment vs Loan Status

- Customers in the range of EMI 200 and 400 is more to tend as defaulter due to any financial situation.



revol_bal vs Loan Status



Key Insights

- Higher loan amounts and higher debt-to-income ratios are associated with higher default rates.
- Borrowers with lower annual incomes tend to default more often.
- Certain loan purposes, such as small business and debt consolidation, have higher default rates.
- Higher grade and sub-grade loans tend to default less.
- Home ownership status and verification status also influence default rates.
- Correlation analysis shows significant relationships between certain features and loan defaults.

Thank You