LENDING CLUB CASE STUDY

SALIM BASHA AND SATYAM TRIPATHI

TABLE OF CONTENTS

Table Of Contents:

PROBLEM STATEMENT

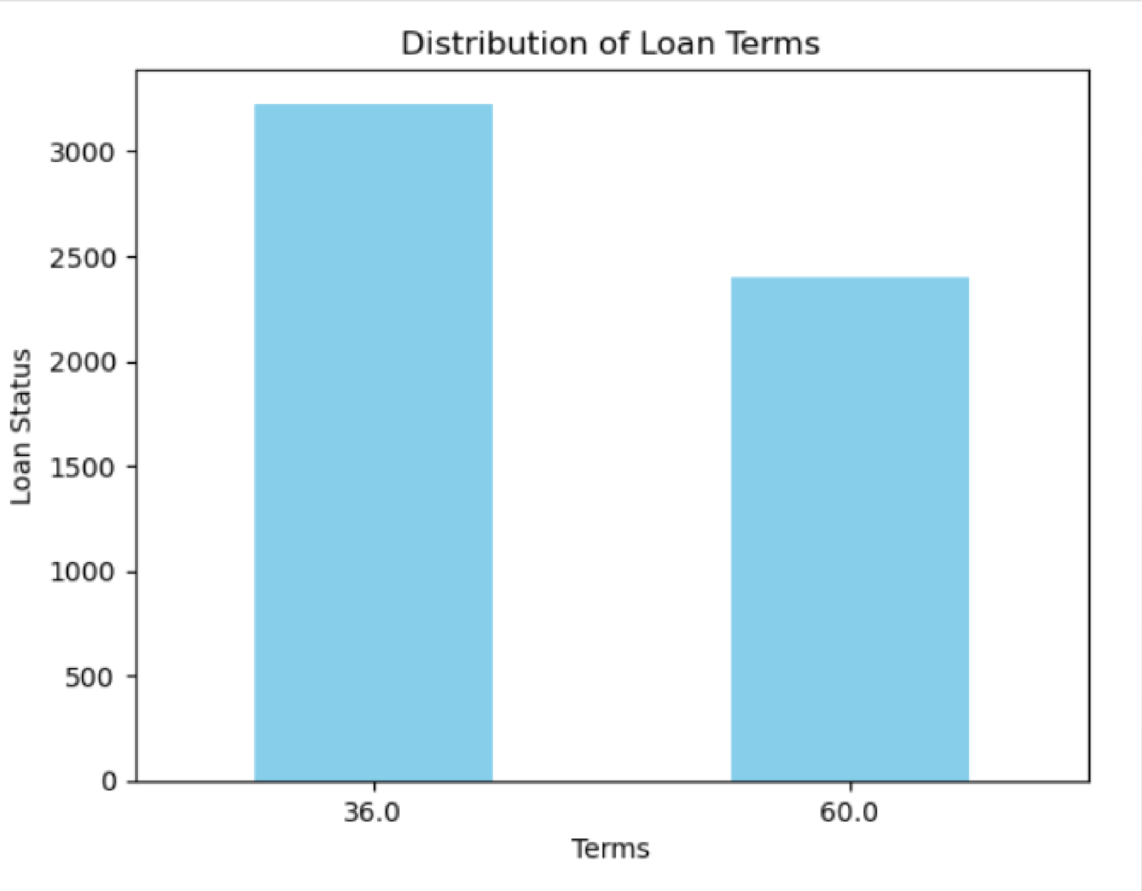
- The objective is to minimize financial risks associated with lending by identifying factors that predict loan defaults through Exploratory Data Analysis(EDA).
- The goal is to reduce financial losses by:
  - Rejecting risky applications.
  - Offering loans with higher interest rates to risky applicants.
  - Minimizing credit loss caused by defaulters.

DATA CLEANUP & TRANSFORMATION

- Checking the missing values in row dataset
- Checking the missing values in column dataset
- Dropping null columns from the dataset
- Dropping all the columns which has only single values in all the columns
- Fixing the null value with mean or median
- Removing % symbol from int_rate, revol_util columns
- Conversion of Datatypes

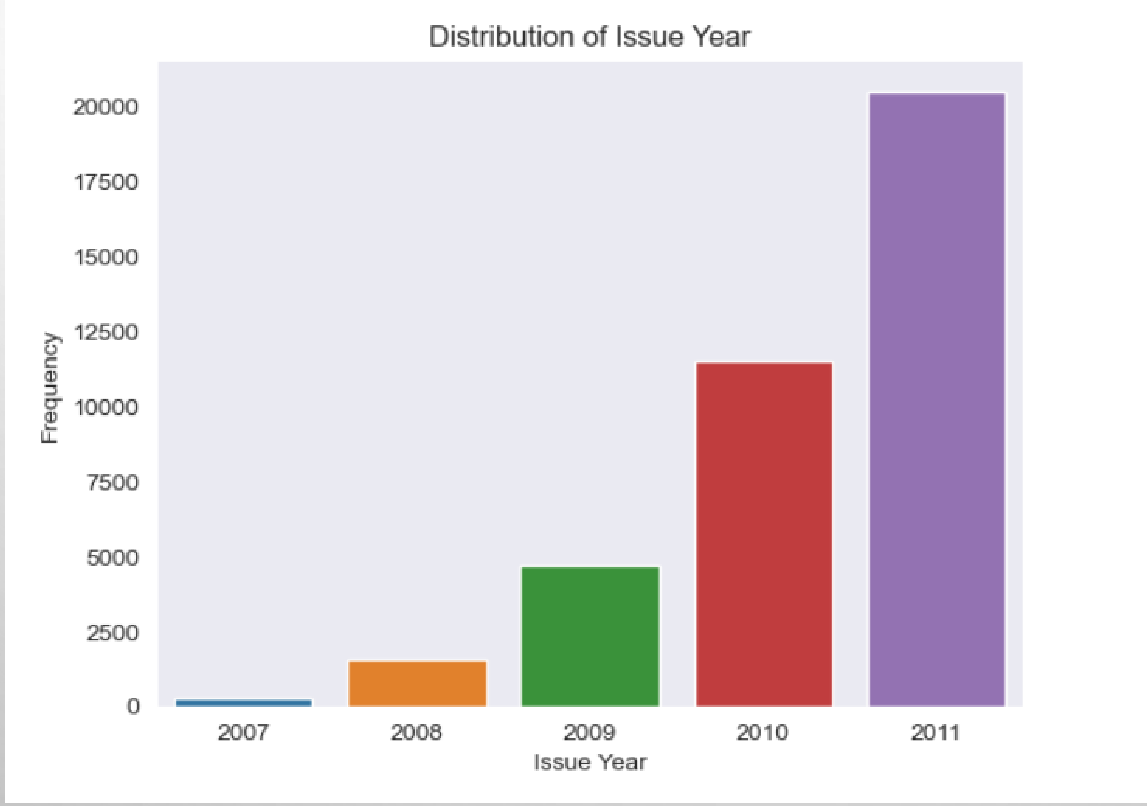
ANALYSIS - UNIVARIATE, BIVARIATE, MULTIVARIATE

- 36 months loan defaults more than 60 months loans



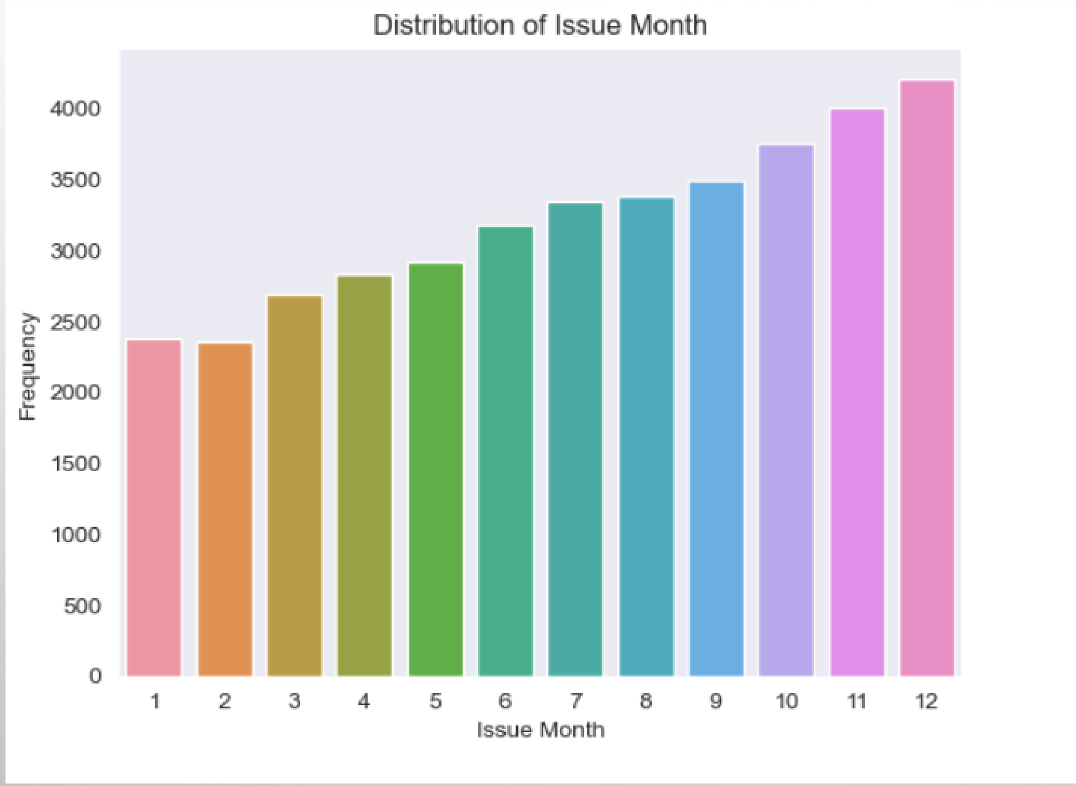
ANALYSIS - UNIVARIATE, BIVARIATE, MULTIVARIATE

- The number of loans issued has consistently increased year over year, reflecting a growing demand for loans over time



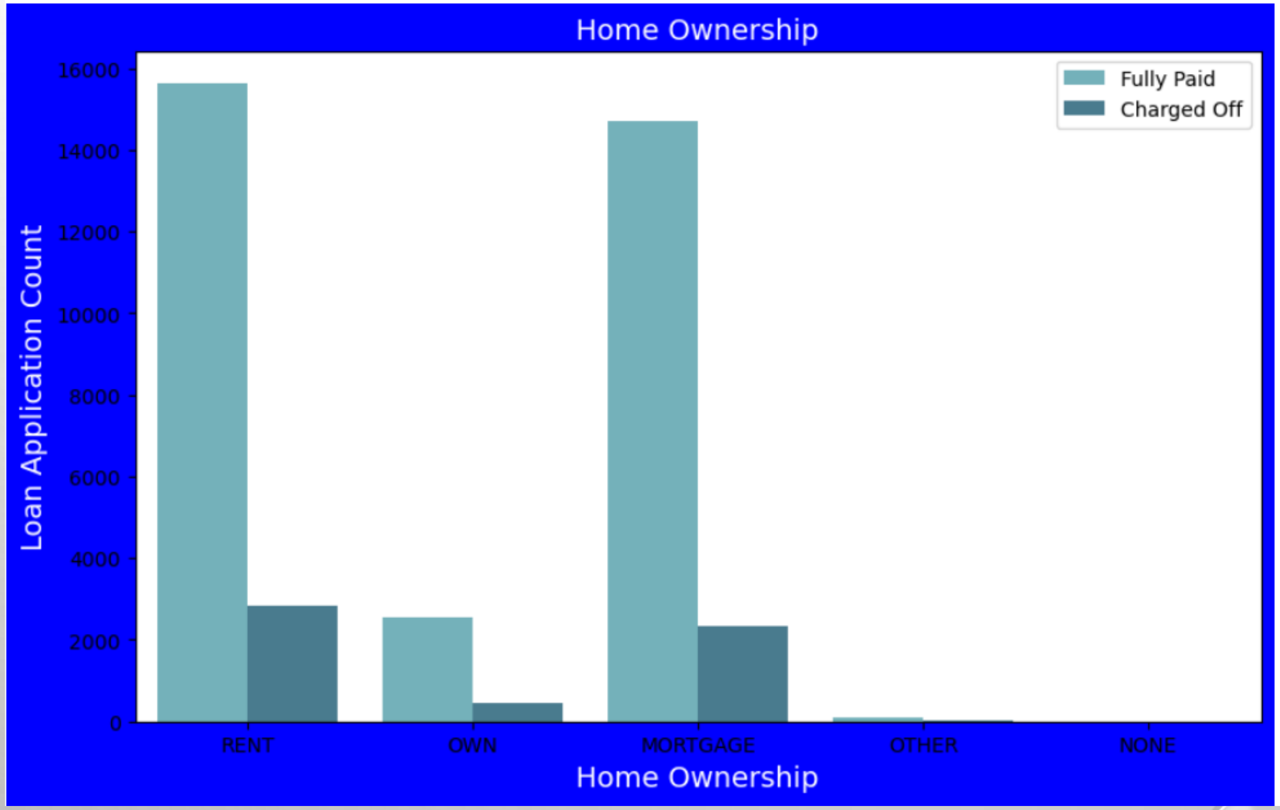
ANALYSIS - UNIVARIATE, BIVARIATE, MULTIVARIATE

- Loan issuance shows an upward trend within each year, with the loan count increasing steadily from January to December



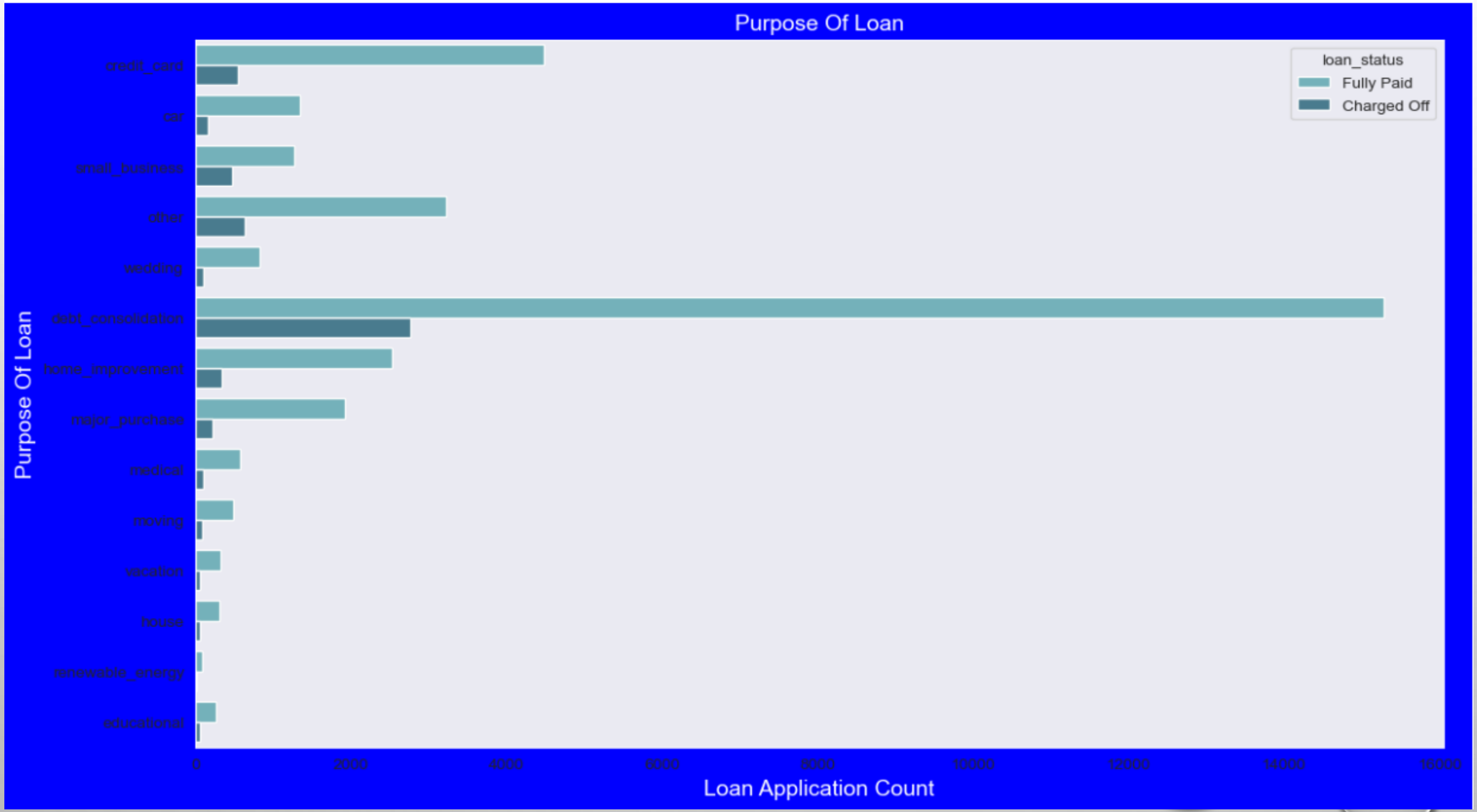
ANALYSIS - UNIVARIATE, BIVARIATE, MULTIVARIATE

- Below plot shows that most of them living in rented home or mortgaged their home
- Applicant numbers are high from these categories so charged off is high too.



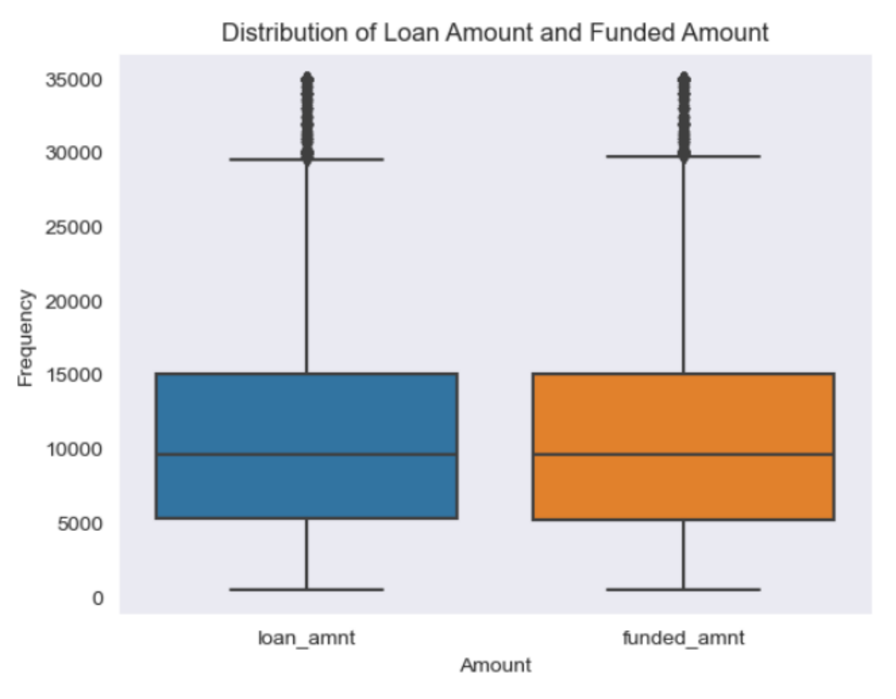
ANALYSIS - UNIVARIATE, BIVARIATE, MULTIVARIATE

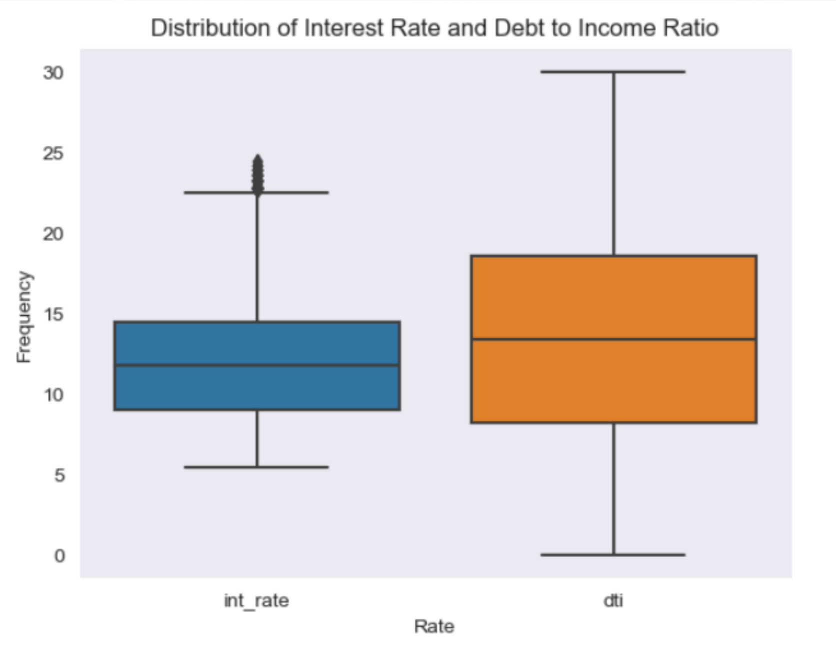
- Below plot shows that most of the loans were taken for the purpose of debt consolidation & paying credit card bill



ANALYSIS - UNIVARIATE, BIVARIATE, MULTIVARIATE

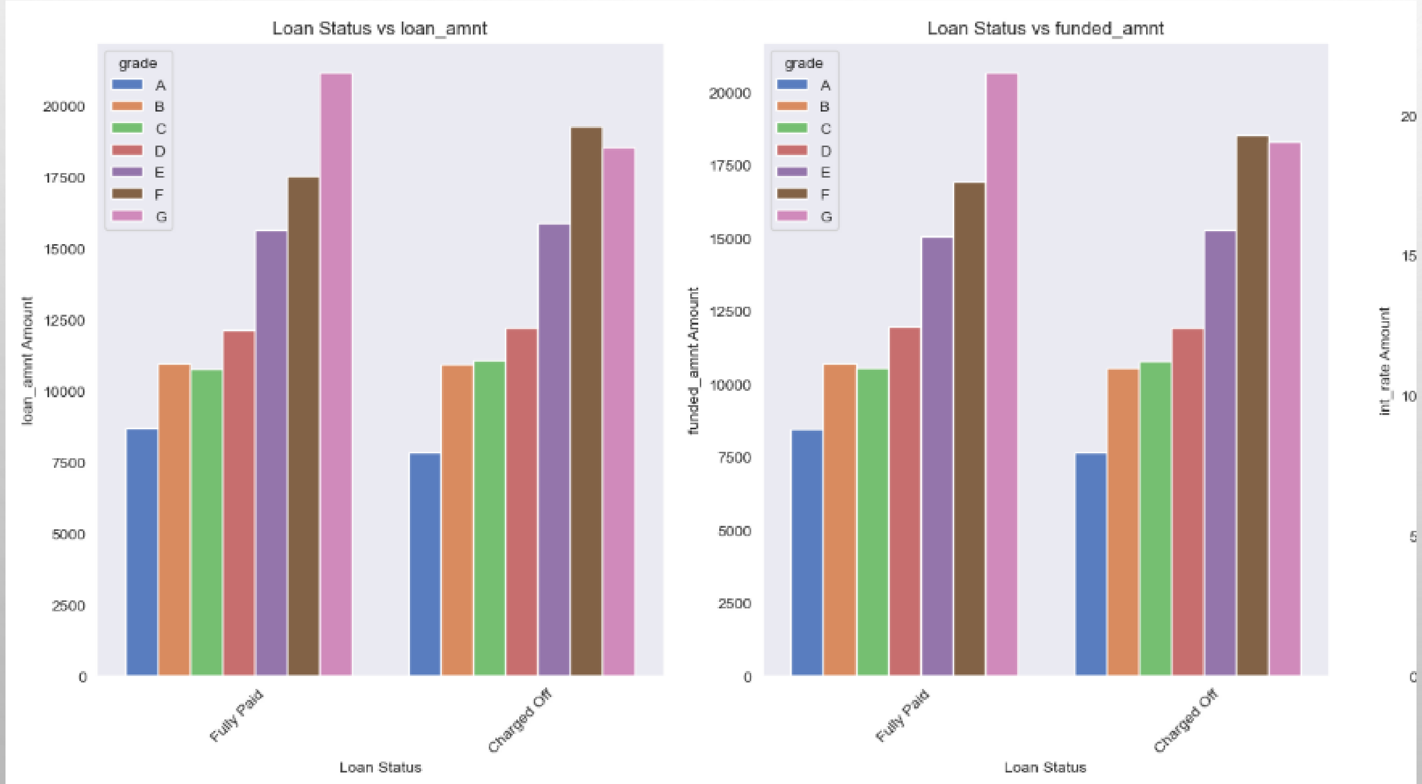
- The loan amount is mostly between 5500 to 15000 where outliers are mostly above 30000
- The funded amount is mostly between 5400 to 15000 where outliers are mostly above 30000
- The interest rate is mostly bwtween 9.25% to 14.59% where outliers are above 22.5%
- The debt to income ration is mostly between 8.17% to 18.6%





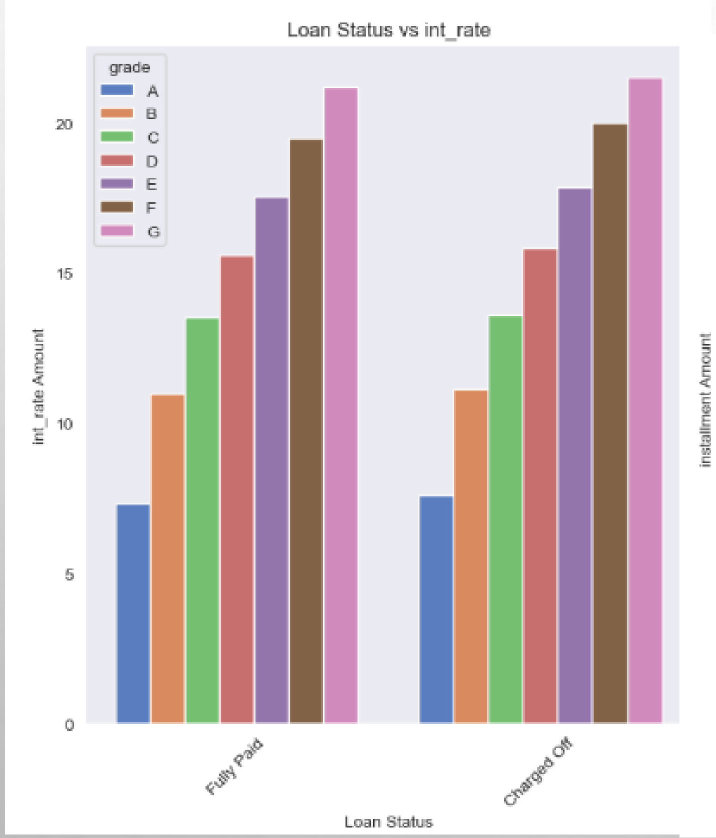
ANALYSIS - UNIVARIATE, BIVARIATE, MULTIVARIATE

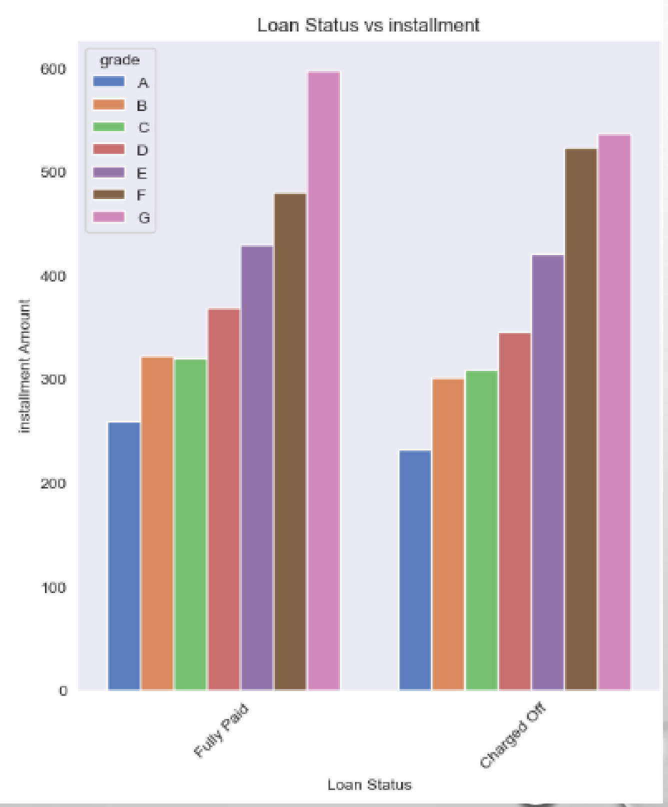
- **Loan Amount vs Loan Status:** Borrowers with higher grades (like A and B) generally take smaller loan amounts, possibly due to their better creditworthiness
- **Funded Amount vs Loan Status:** Borrowers with lower grades tend to receive larger funded amounts, which might increase their chances of defaulting



ANALYSIS - UNIVARIATE, BIVARIATE, MULTIVARIATE

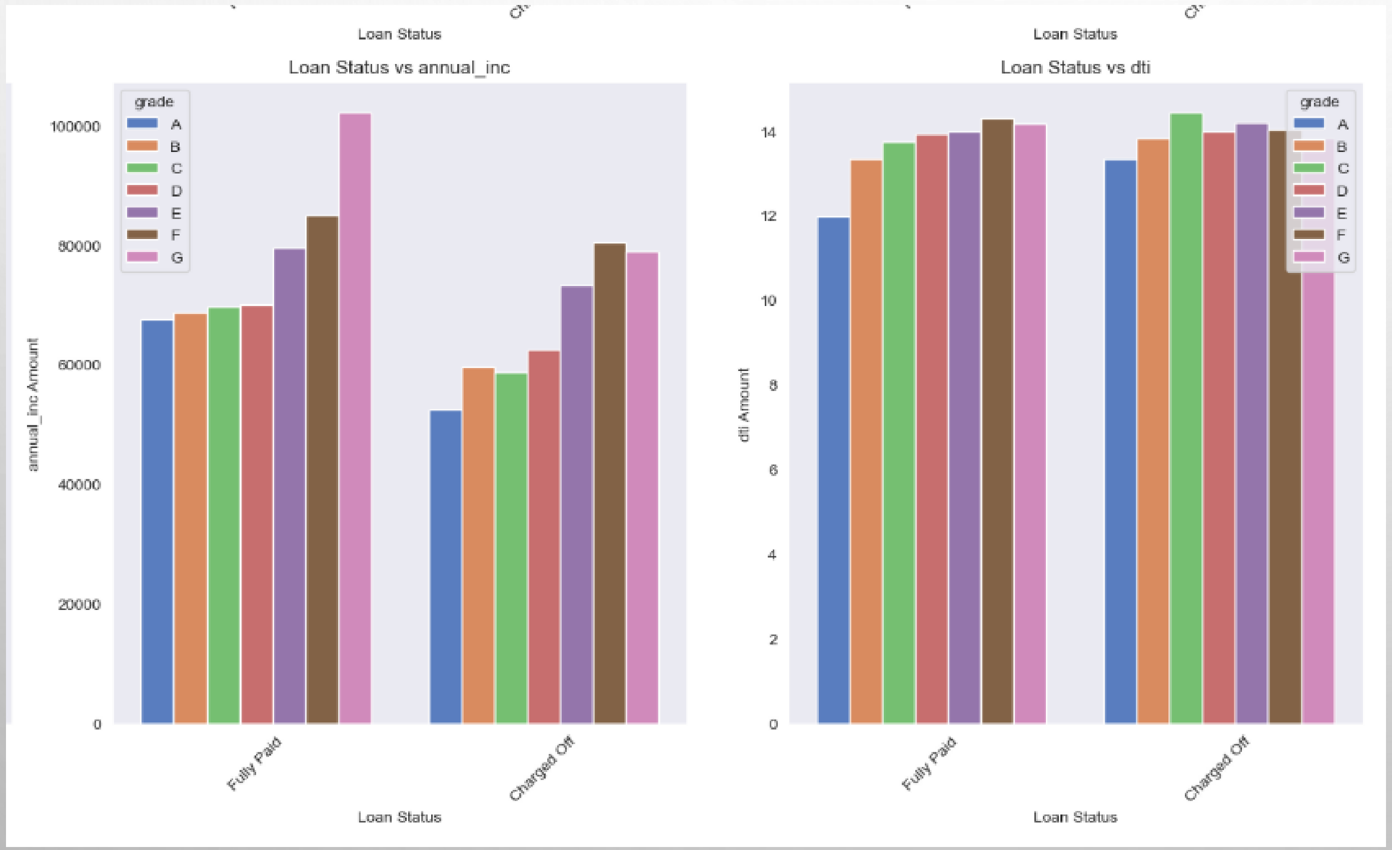
- **Interest Rate vs Loan Status:** Higher interest rates in lower grades may contribute to higher Charged Off.
- **Installment vs Loan Status:** Since installments increases with grade, higher installments, especially in lower-grade loans, may increase likelihood of Charged Off.





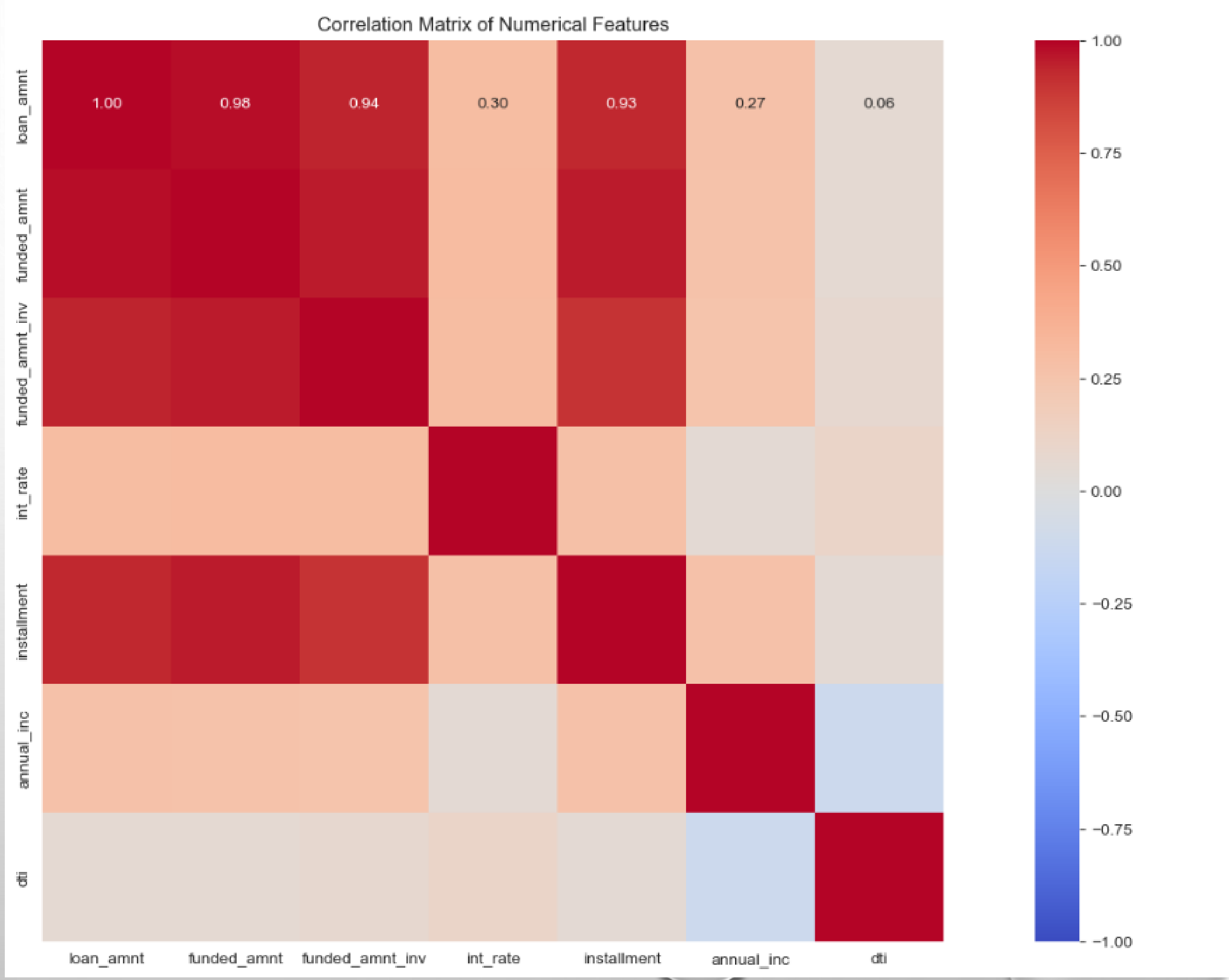
ANALYSIS - UNIVARIATE, BIVARIATE, MULTIVARIATE

- **Annual Income vs Loan Status:** Even borrowers with high incomes but lower grades are more likely to default or keep loans in the Current status
- **Debt-to-Income Ratio (DTI) vs Loan Status:** Lower-grade loans carry higher risk because of high DTI ratios, indicating borrowers might be over-leveraged



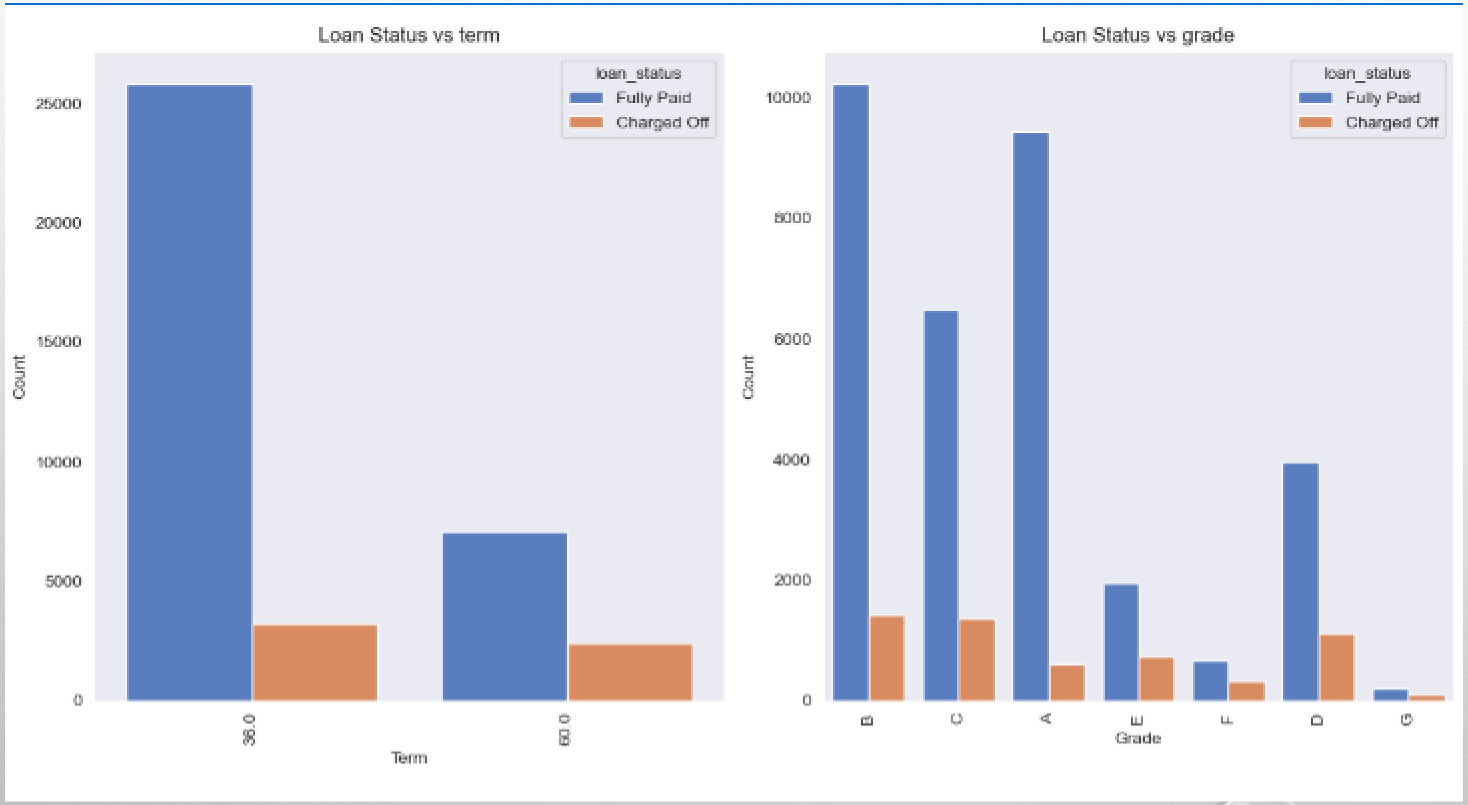
ANALYSIS - UNIVARIATE, BIVARIATE, MULTIVARIATE

- **Loan Amount and Funded Amount (0.98):** These two are almost identical, as the funded amount is closely linked to the approved loan amount
- **Loan Amount and Installment (0.93):** Bigger loan amounts lead to higher installment payments
- **Funded Amount and Installment (0.96):** The funded amount strongly impacts the size of the installments
- **Funded Amount and Funded Amount by Investors (0.96):** These two are highly related and increase proportionally with each other



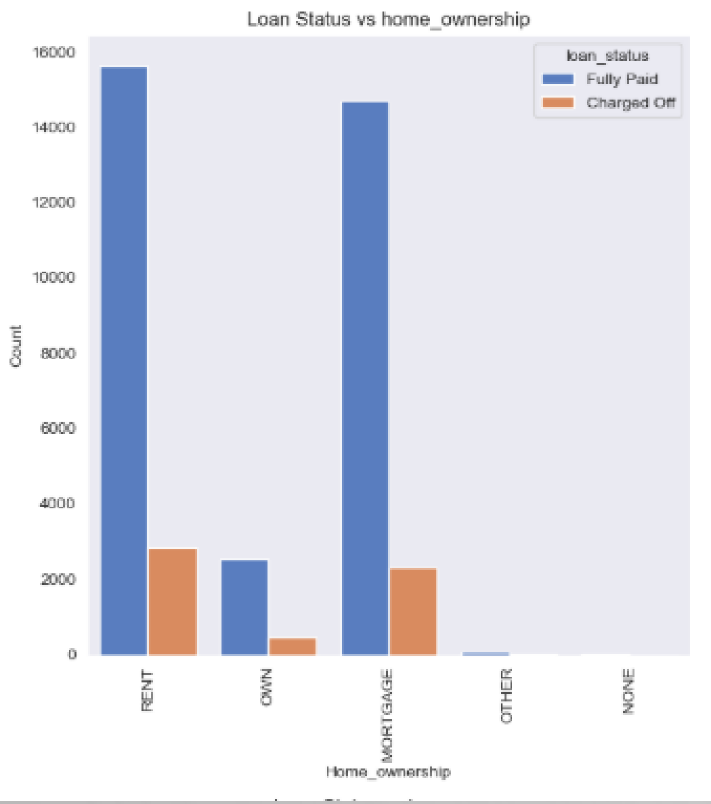
ANALYSIS - UNIVARIATE, BIVARIATE, MULTIVARIATE

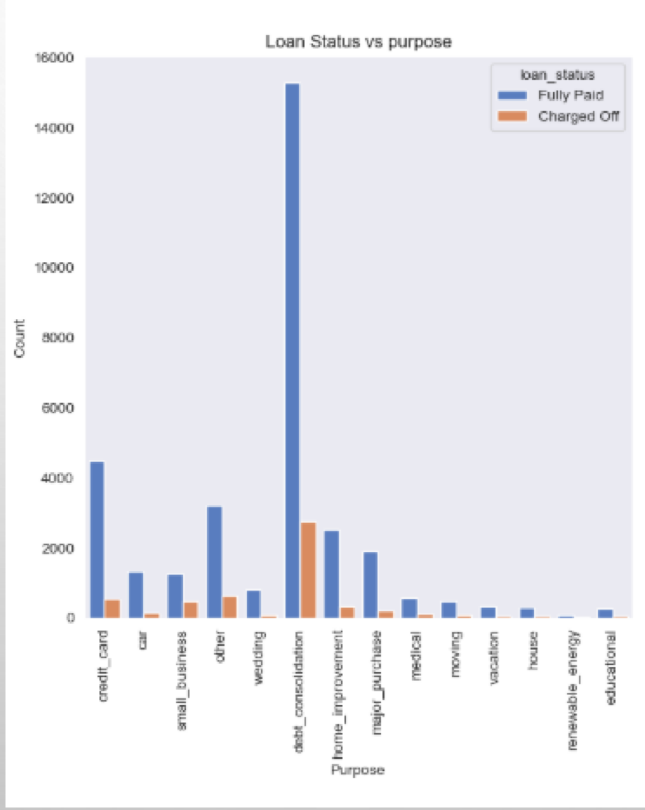
- **Term vs Loan Status:** 36-month loans have a lower chance of default compared to 60-month loans
- **Grade vs Loan Status:** A-grade loans are the safest, with very low default risks compared to other grades



ANALYSIS - UNIVARIATE, BIVARIATE, MULTIVARIATE

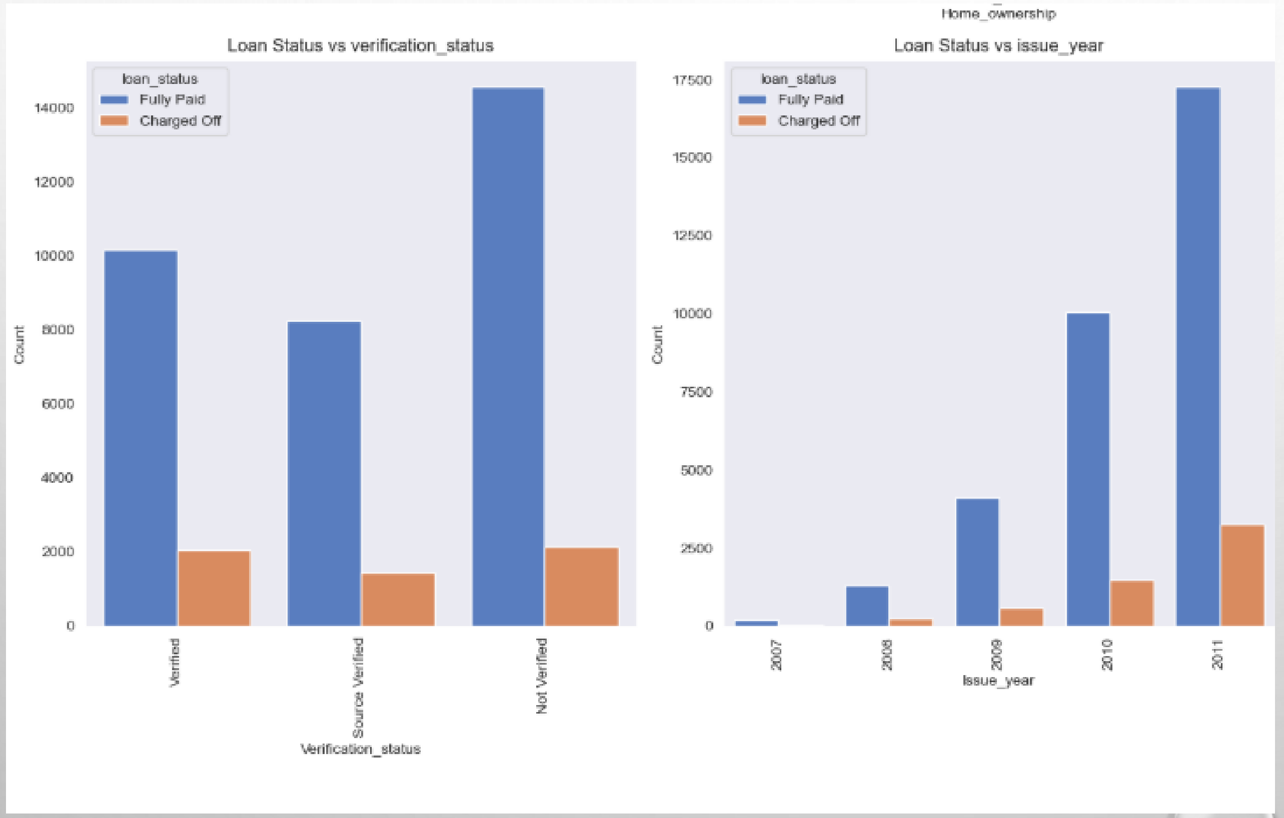
- **Home Ownership vs Loan Status:** Homeowners are more likely to repay their loans fully compared to renters or those with mortgages
- **Purpose vs Loan Status:** While debt consolidation loans are most commonly repaid, they also have the highest number of defaulters





ANALYSIS - UNIVARIATE, BIVARIATE, MULTIVARIATE

- **Verification Status vs Loan Status:** Interestingly, loans given without verified income have a lower likelihood of default
- **Issue Year vs Loan Status:** Both fully paid loans and defaulters have increased year by year



CONCLUSION

- **Risk Factors:** Loans with lower grades (E, F, G) tend to have higher loan amounts, interest rates, and DTI ratios, increasing the likelihood of default
- **Income Influence:** Borrowers with higher incomes are more likely to fully repay loans
- **Current Loans:** These loans often have higher loan amounts, funded amounts, and installments, particularly in lower grades, indicating they are still at risk of default
- **Key Findings:** Loans with shorter terms, higher grades, home ownership, unverified income, and a debt consolidation purpose are more likely to be fully paid