LENDING CLUB CASE STUDY

Group Members:
Mahathi Mahalakshmi Satish
Pedaballi Madan Mohan Reddy

Problem Statement:

As a consumer finance company specializing in lending various types of loans to urban customers, our primary objective is to optimize our loan approval process to maximize profitability and minimize financial risks. The decision to approve or reject a loan application must be made based on the applicant's profile, considering two critical types of risks:

1.**Opportunity Loss Risk**: If we do not approve a loan for an applicant who is likely to repay it, we risk losing potential business and revenue.

2.**Default Risk**: If we approve a loan for an applicant who is likely to default, we risk incurring financial losses.

To address these risks effectively, we need to develop a robust and data-driven decision-making framework that accurately assesses the creditworthiness of applicants. This framework should help us:

•Identify applicants with a high probability of repayment to ensure we capture profitable lending opportunities.

•Detect applicants with a high likelihood of default to prevent financial losses.

By achieving a balanced approach between these two risks, we aim to enhance our loan approval process, increase profitability, and maintain financial stability.

ABSTRACT:

• Lending Club is the largest online loan marketplace.

• Offers personal, business, and medical financing.

• Provides borrowers with lower interest rates through a fast online interface.

• Analysis aims to predict loan defaults using past applicant data.

PROBLEM SOLVING METHODS IMPLEMENTATION:

1. **Data Cleaning**: Removing columns with null values and unnecessary variables, checking the percentage of null values, and removing the corresponding rows.
2. **Data Analysis**: Using the Data Dictionary to understand all columns and their domain-specific uses.
3. **Univariate Analysis**: Analyzing each column and plotting their distributions.
4. **Segmented Univariate Analysis**: Examining continuous data columns in relation to categorical columns.
5. **Bivariate Analysis**: Investigating the relationship between variables, such as term and loan status, with respect to loan amount.
6. **Recommendations**: Reviewing all plots and providing recommendations to minimize business losses by identifying columns that best predict loan defaults.
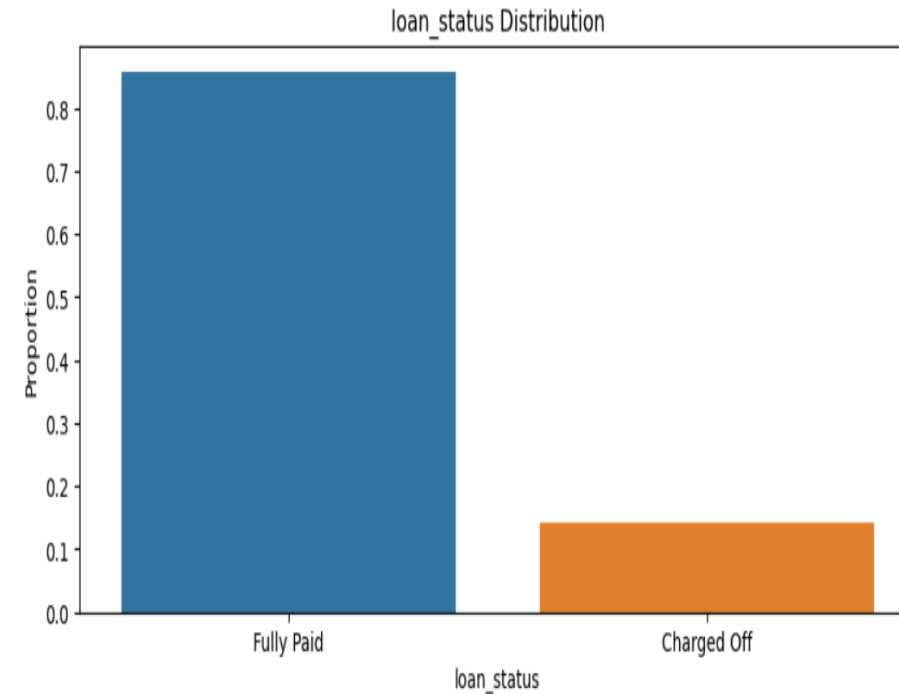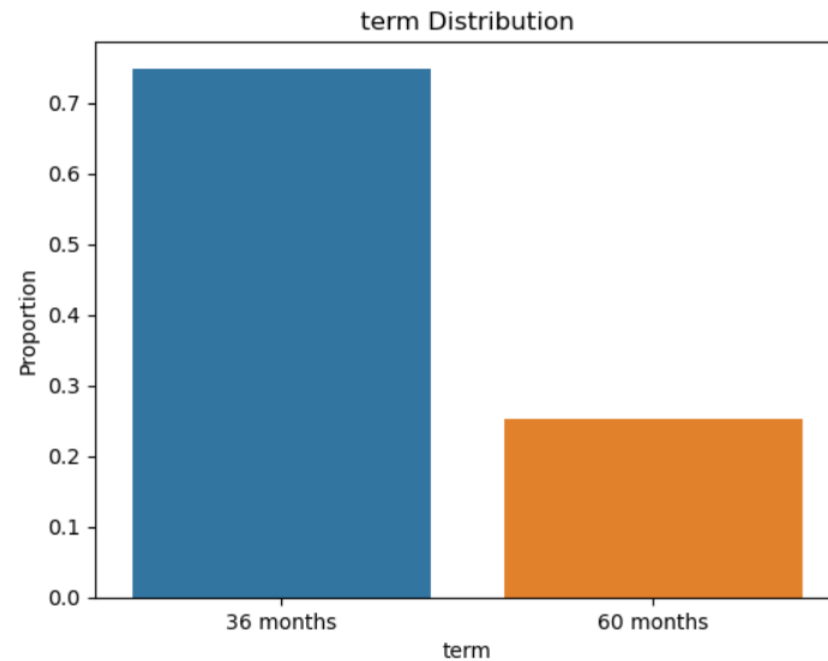
Analysis:
Issue dates Analysed by Year, Month and Days.
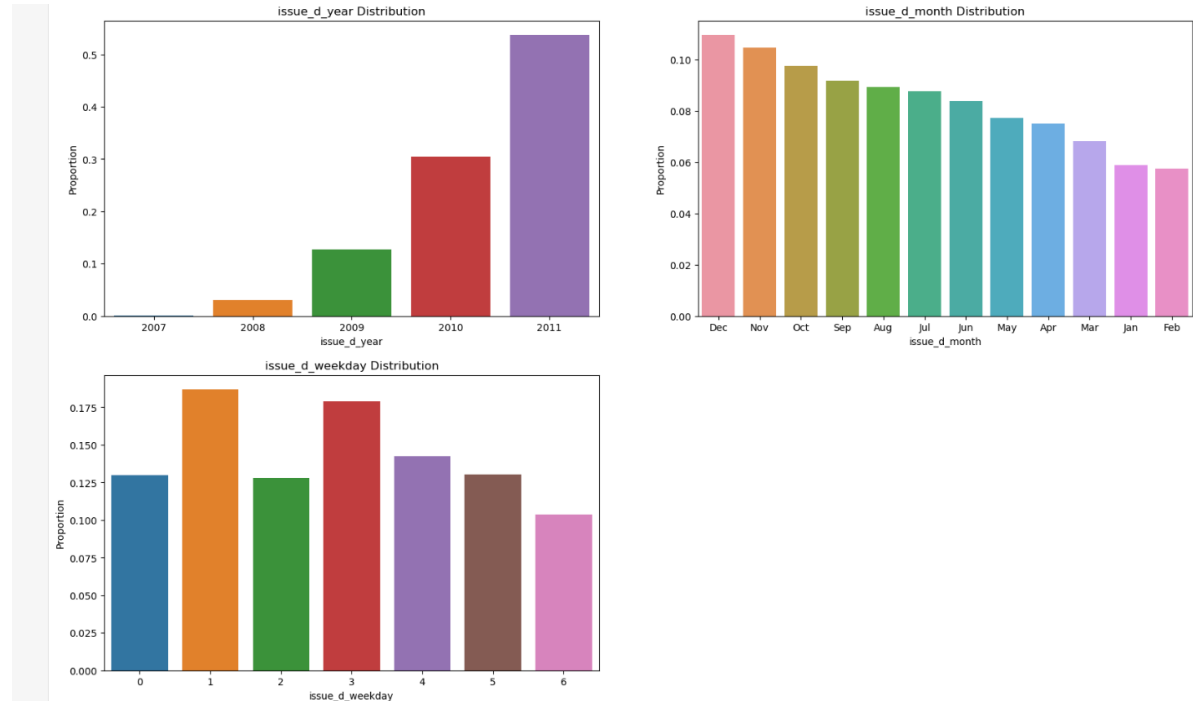
Lending Club has significantly expanded year by year, with the number of loans issued doubling annually. Additionally, the issuance of loans increases from January to December, peaking in the final quarter, likely due to vacations and Christmas.
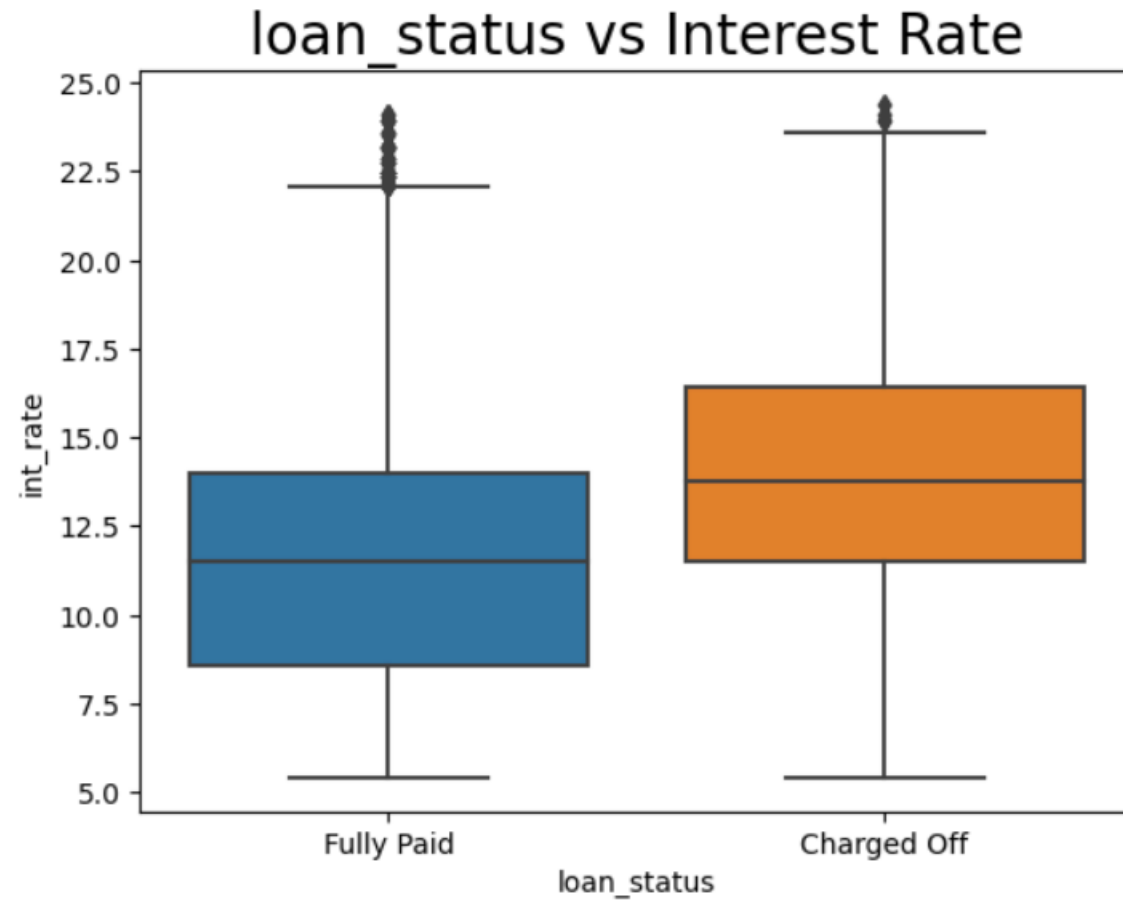
Analysis:
Analysing the Term distributions and loan status.

There are only two loan terms: 36 months and 60 months. Approximately 75% of borrowers chose the 36-month term. In the given dataset, about 15% of borrowers defaulted, while 85% fully repaid their loans.
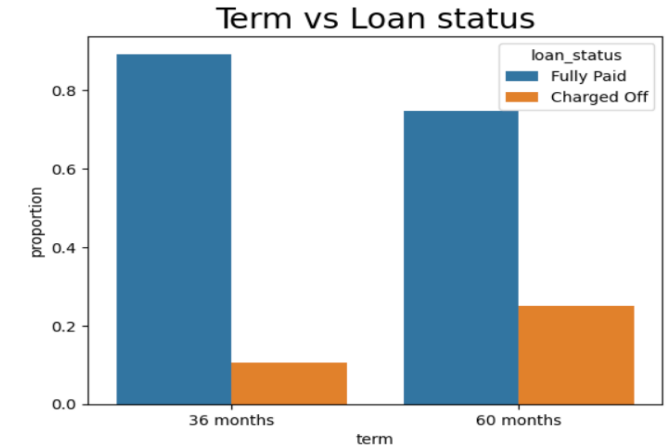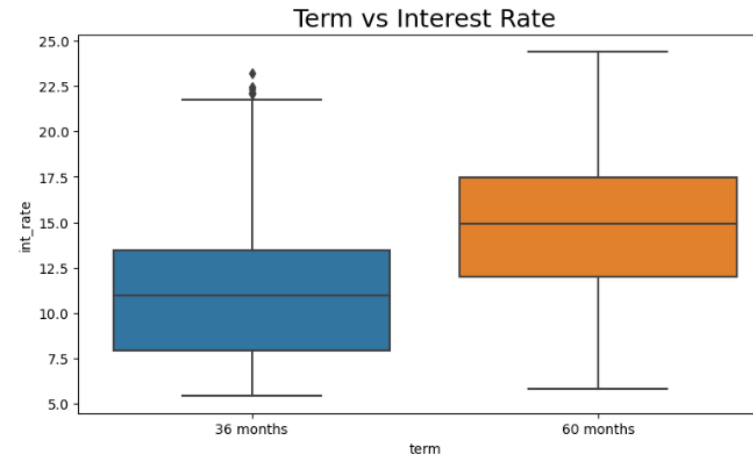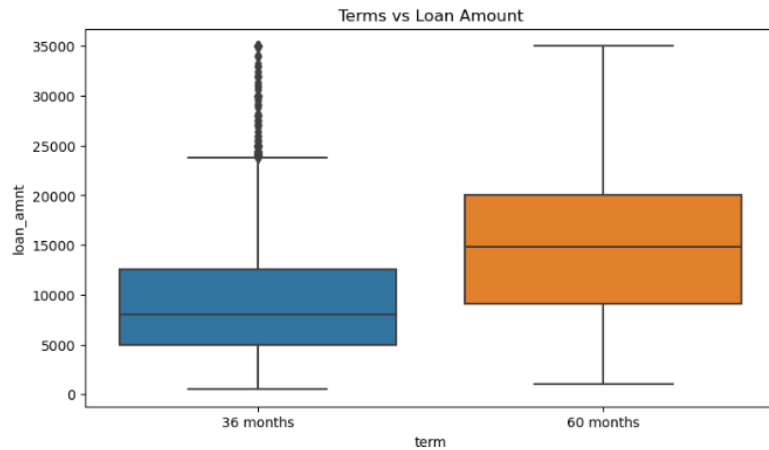
Analysis:
Analysing the loan status v/s Interest rates.

loan_status vs Interest Rate

Higher loan interest rates are associated with a greater likelihood of loan defaults.
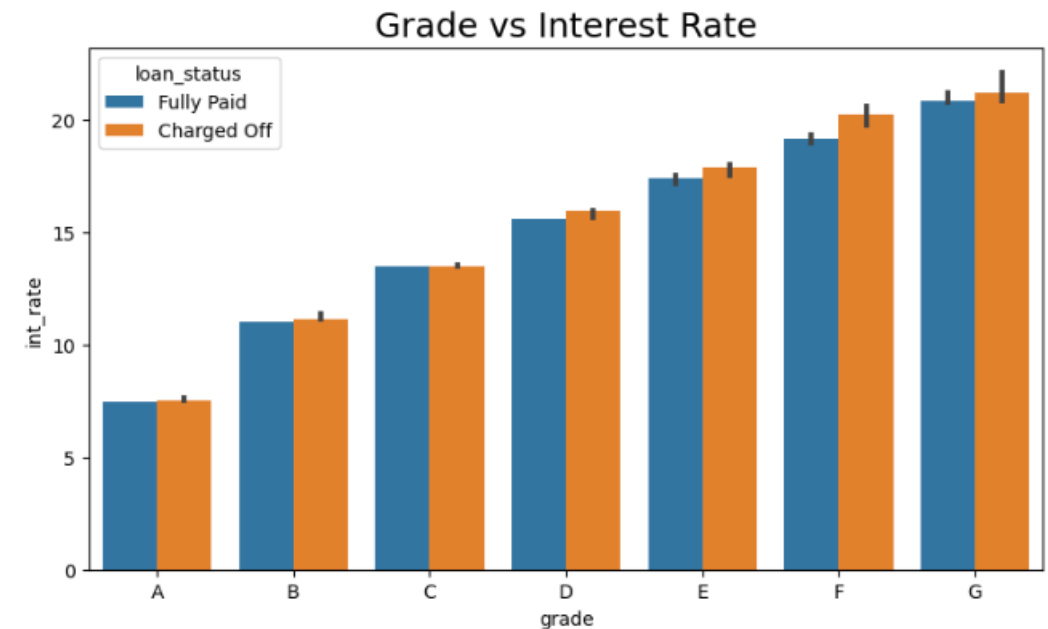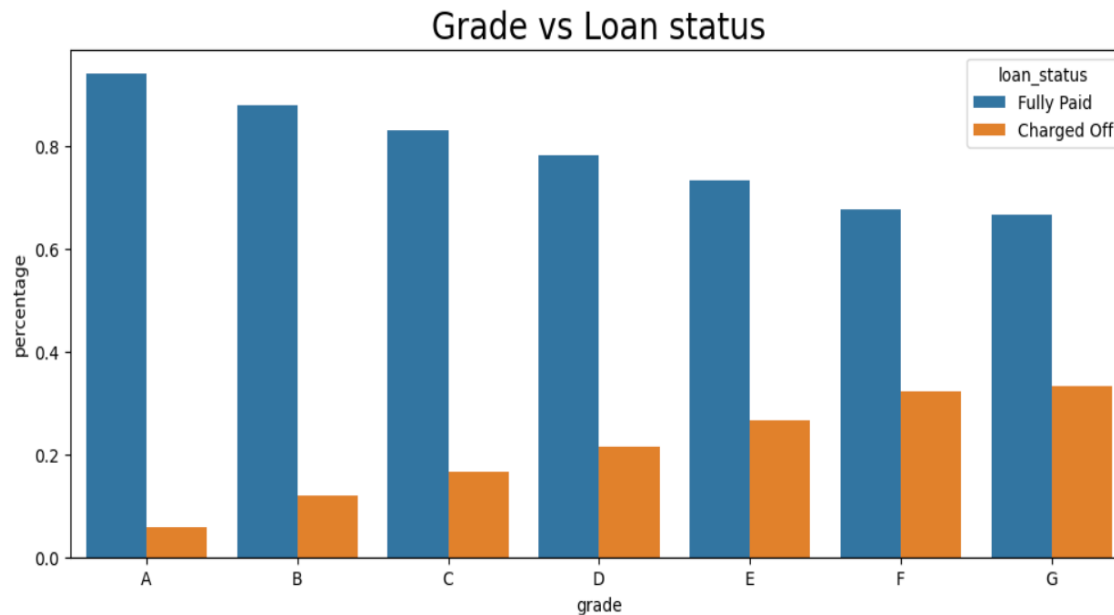
Analysis:

Analysing the Term v/s Loan amount, Interest rates and Loan status.

The default rate is higher for 60-month loans because borrowers typically took out larger loan amounts with higher interest rates, making it more difficult for them to repay the loans.
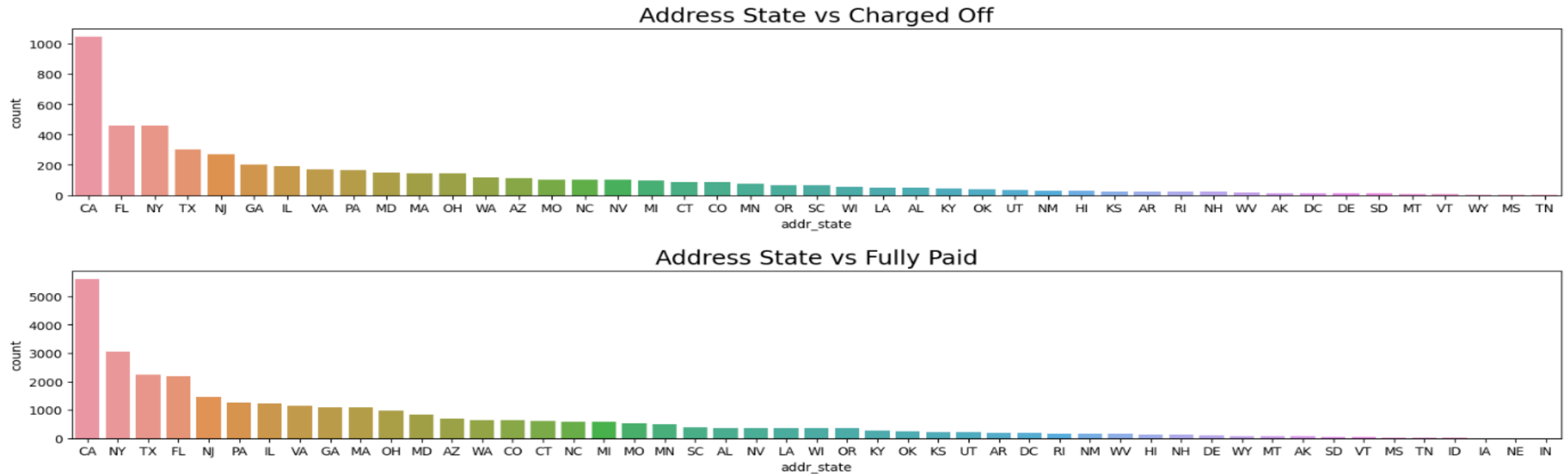
Analysis:

Analysing the Grade v/s Loan status and Interest rate.

Grades are a reliable indicator of a borrower's likelihood of defaulting on a loan. Lower grades (E, F, G) have a higher probability of default compared to higher grades (A, B). Additionally, borrowers with lower grades receive loans with higher interest rates, which may contribute to the increased default risk.
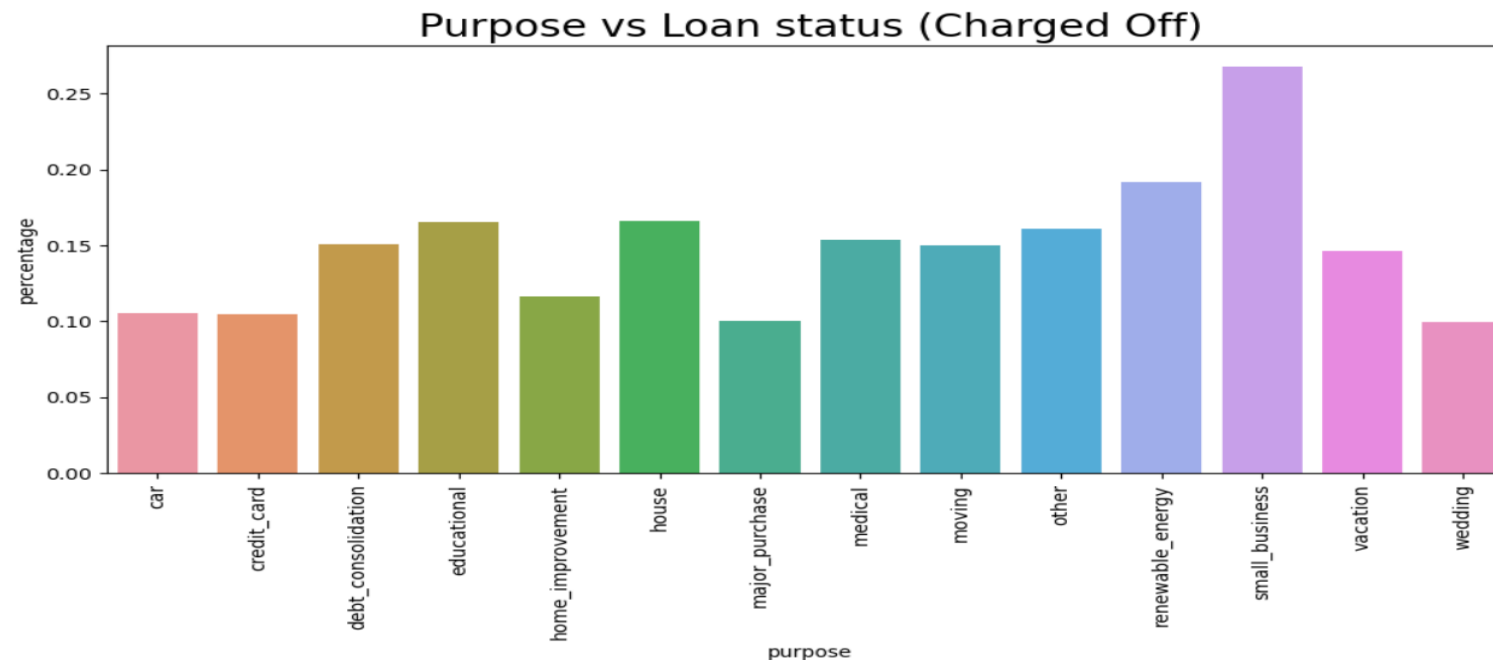
Analysis:
Analysing the Address states v/s Charged off and fully paid loans.

Loans given to borrowers in California, Florida, and Texas have higher default rates compared to other states.
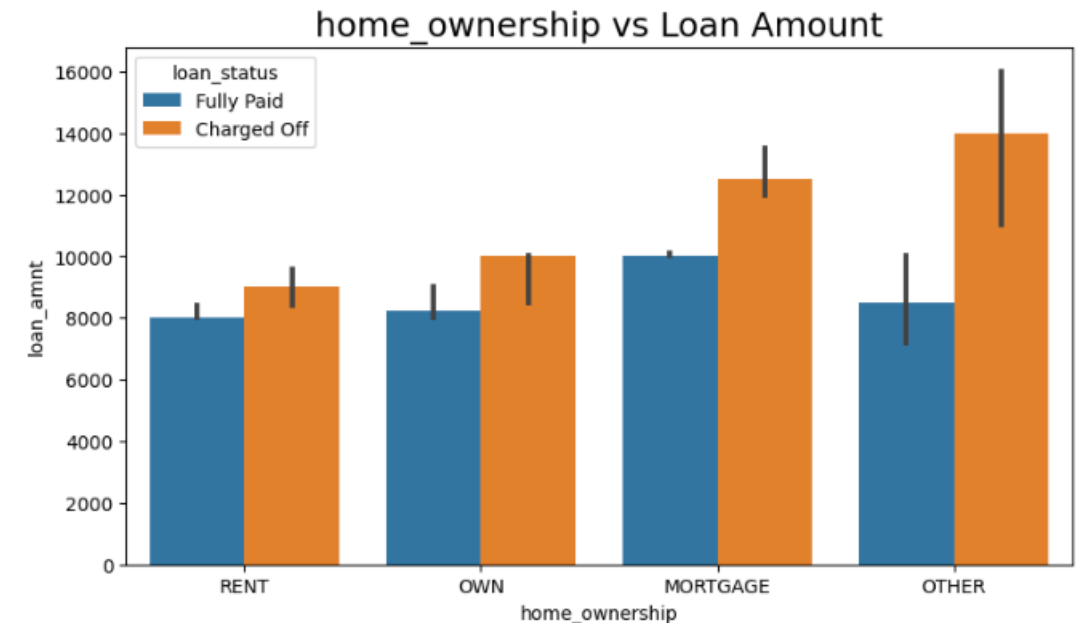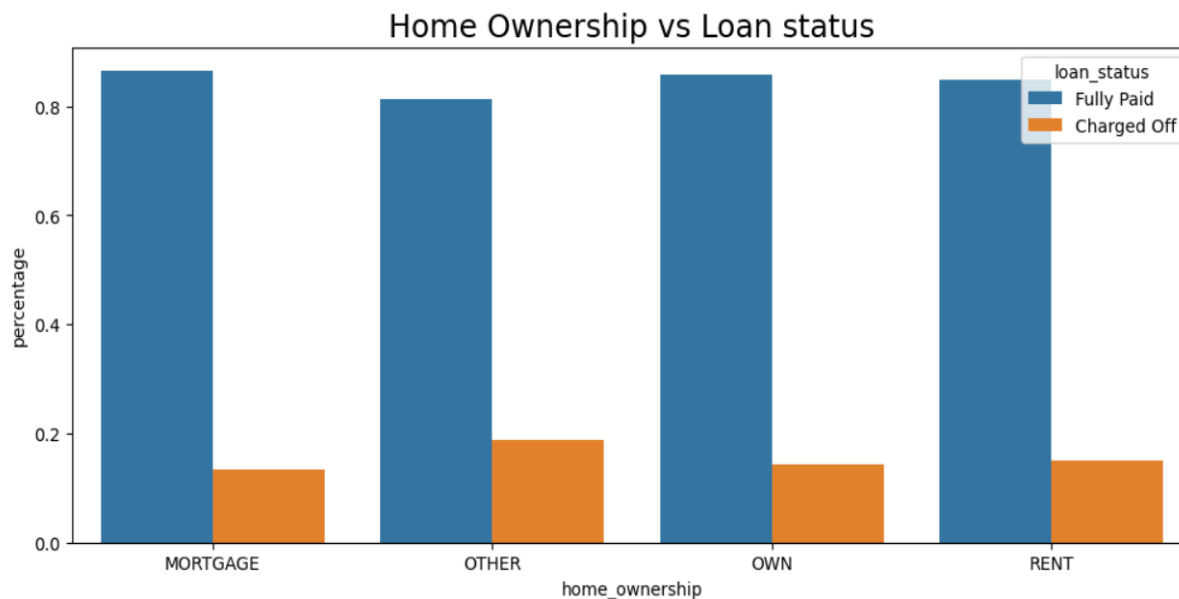
Analysis:

Analysing the Purpose of loan v/s Charged off loans.

Purpose vs Loan status (Charged Off)

Borrowers who took out loans for small business purposes have a higher rate of default.

Analysis:
**Analysing Home ownerships v/s Loan status and Loan amount.**

There is approximately a 20% chance of loan default in each homeownership category.
The second plot reveals that borrowers with higher loan amounts in the mortgage homeownership category have a higher default rate compared to others.

Analysis:
Analysing the Home ownership v/s Approval loan amount ratio.

Home Ownership vs Approval Loan Amount Ratio
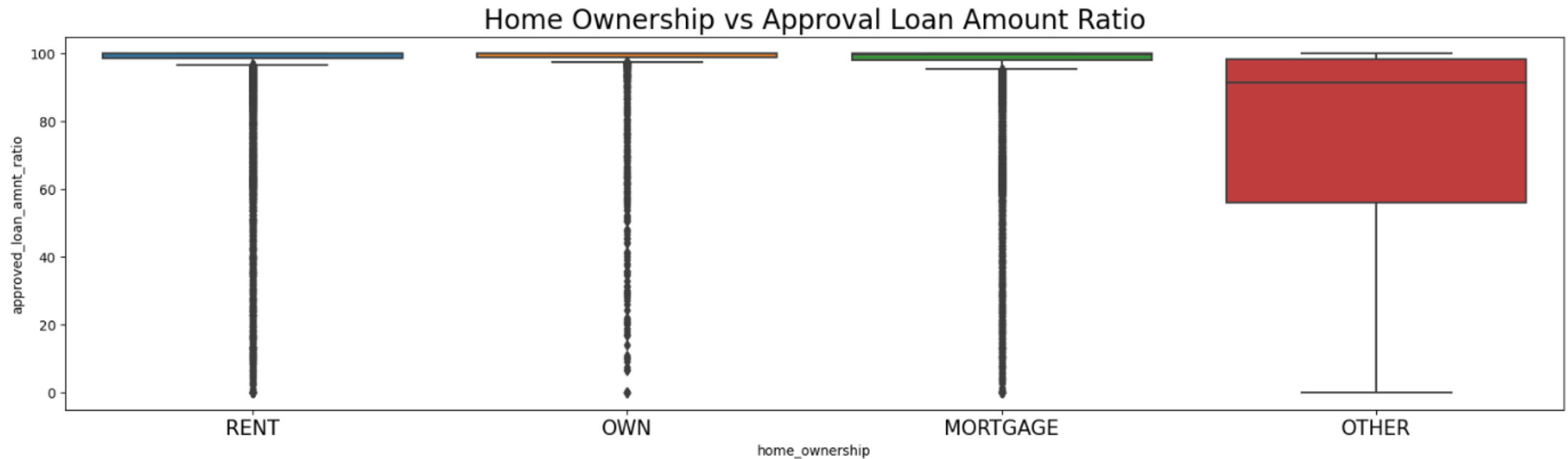
For the "Other" homeownership category, the approved loan amount (funded by investors) is less than the requested loan amount by borrowers.

CONCLUSIONS:

1. Lending Club should reduce the interest rates for 60-month loans, as they are more prone to default.

2. Grades are a good metric for detecting defaulters. Lending Club should examine more information from borrowers before issuing loans to lower grades (G to A).

3. Lending Club should limit the number of loans issued to borrowers from California, Florida, and New York to increase profitability.

4. Small business loans have a higher default rate. Lending Club should reduce or stop issuing loans for small businesses.

5. Borrowers with mortgage homeownership are taking higher loans and defaulting more frequently. Lending Club should stop approving loans for this category when the requested loan amount exceeds $12,000.

6. People with a higher number of public derogatory records are more likely to file for bankruptcy. Lending Club should ensure borrowers have no public derogatory records before issuing loans.