LENDING CASE STUDY

CONTEXT

You work for a consumer finance company specializing in lending various types of loans to urban customers through the "Lending Club." When a loan application is received, the company must decide whether to approve the loan based on the applicant's profile. This decision hinges on two key risks:

1. **Risk of Losing Business (False Negative):** If an applicant who is likely to repay the loan is denied, the company misses out on potential revenue and profit.
2. **Risk of Financial Loss (False Positive):** If an applicant likely to default is approved, the company risks financial loss due to non-repayment.

The objective is to analyze historical data on loan applicants—whether they defaulted or not—and identify patterns that can help predict whether new applicants will default. This will assist in making informed loan approval decisions and may lead to the following actions:

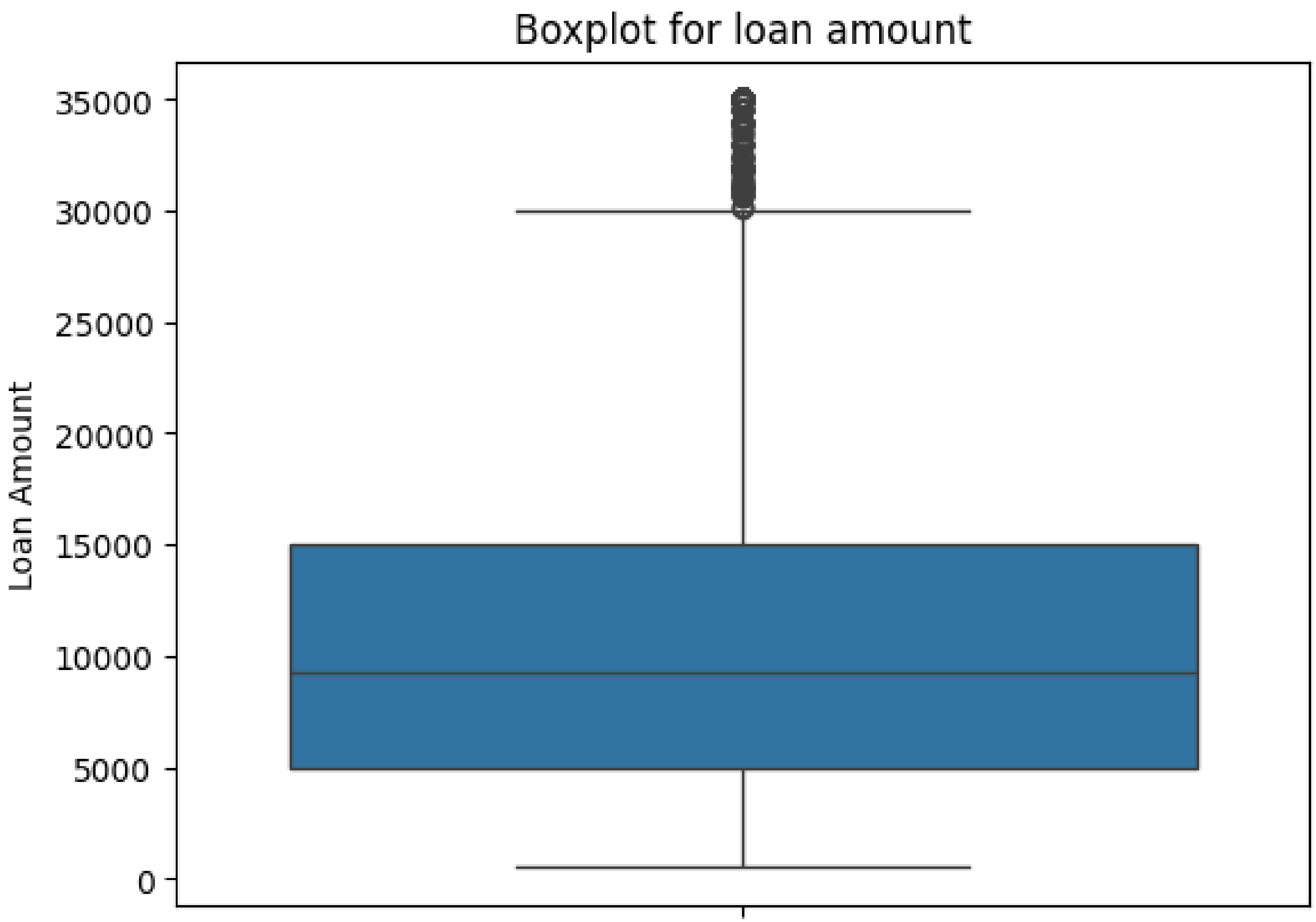
- Deny loans to high-risk applicants.
- Reduce loan amounts for risky individuals.
- Offer loans to risky applicants but at higher interest rates to offset potential risks.

PROBLEM STATEMENT:

The goal is to leverage exploratory data analysis (EDA) to understand the relationship between consumer attributes (e.g., income, credit history, employment status) and loan attributes (e.g., loan amount, term) with the tendency to default. This will improve decision-making and risk management in lending.

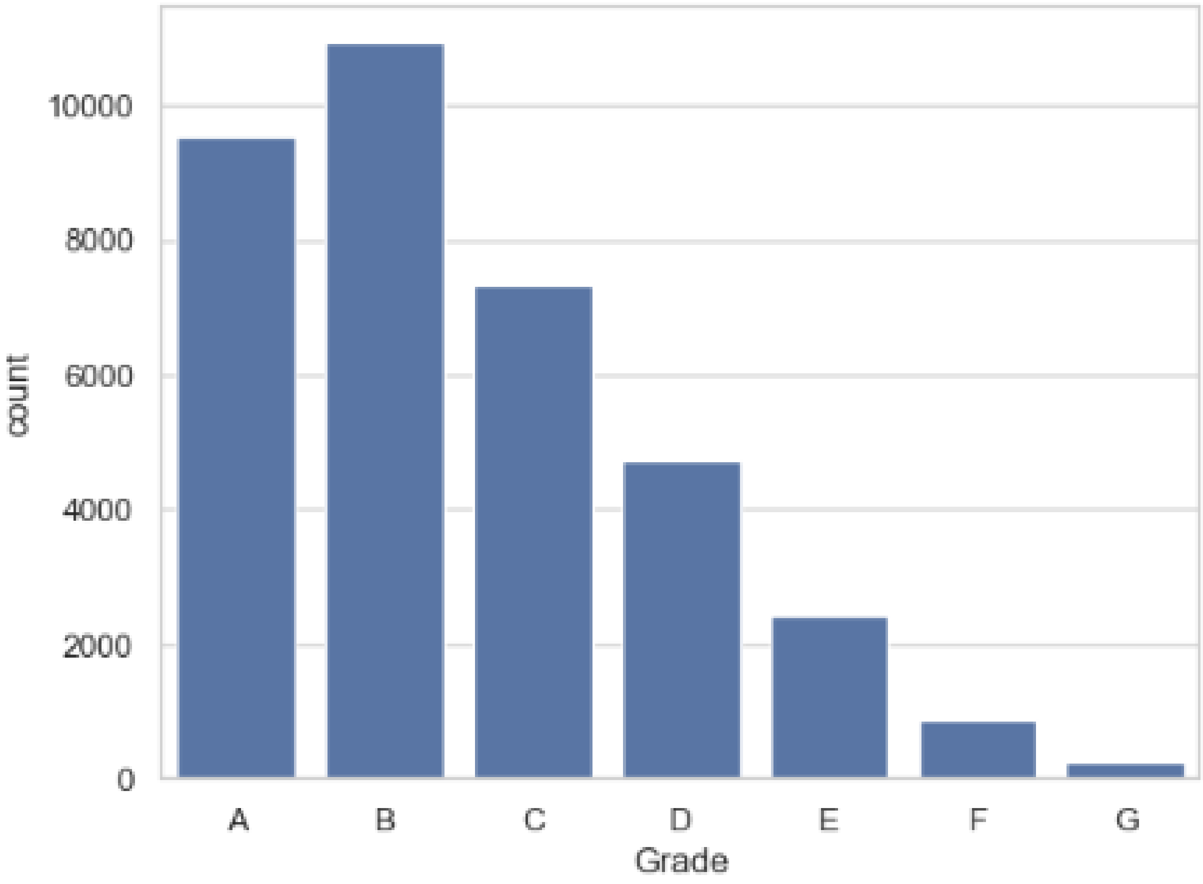
KEY INSIGHTS:

1. **Loan Amount Distribution:**
    1. Loan amounts range from \$0 to \$35,000, with an average of \$10,000.
    2. Outliers are observed near the upper limit, but they are not extreme enough to be removed.



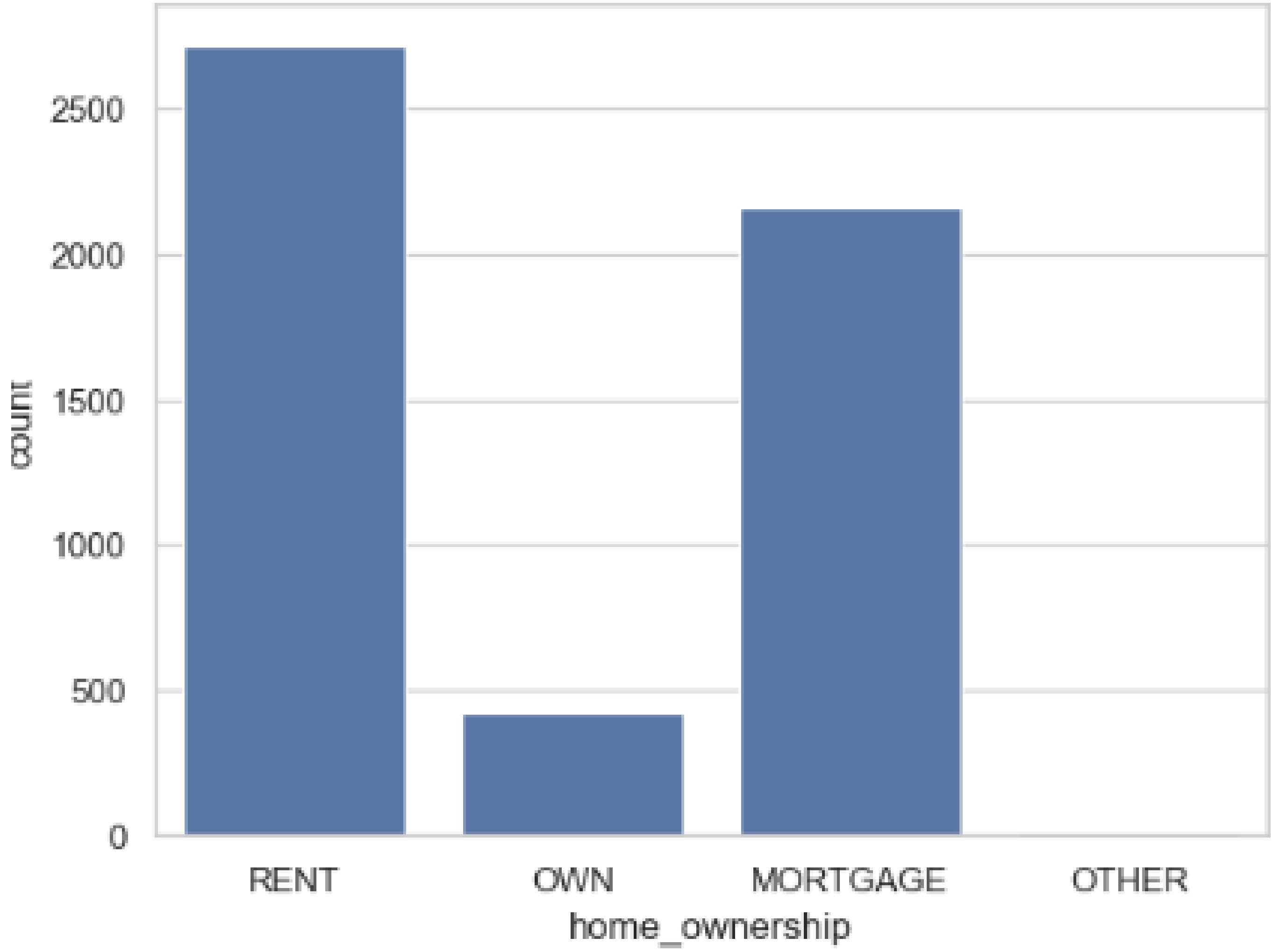
LOAN GRADE PREFERENCE

- Borrowers mostly invest in loans graded A to C, with fewer opting for D to G. The trend is generally downward from A to G, except for B.
- Lower interest rates correspond to higher loan grades, which indicates borrowers prefer loans with lower interest rates.

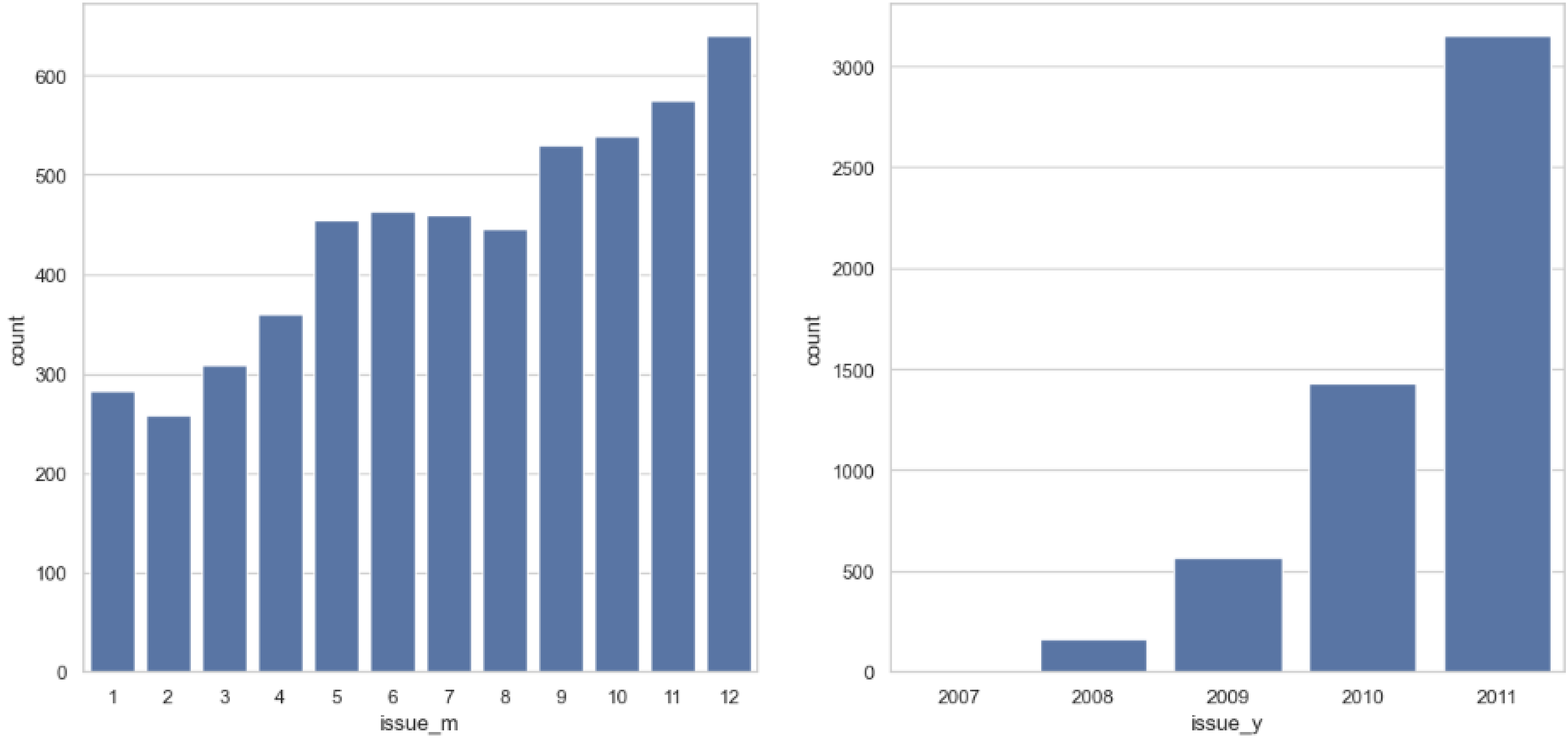


HOMEOWNERSHIP STATUS:

- Homeowners are less likely to default on their loans, while those who rent have a higher likelihood of defaulting.



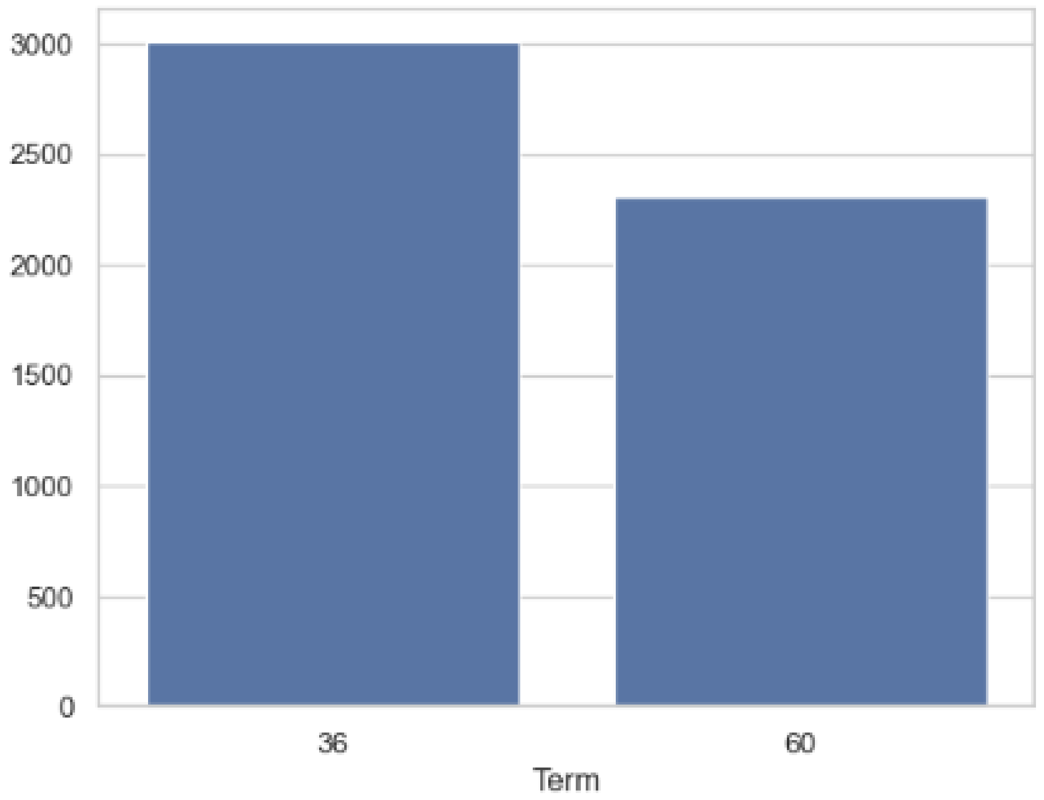
LOAN ISSUE DATE:



- The highest number of defaults occurred in loans issued in December. The year 2011 saw a higher number of defaults compared to other years.

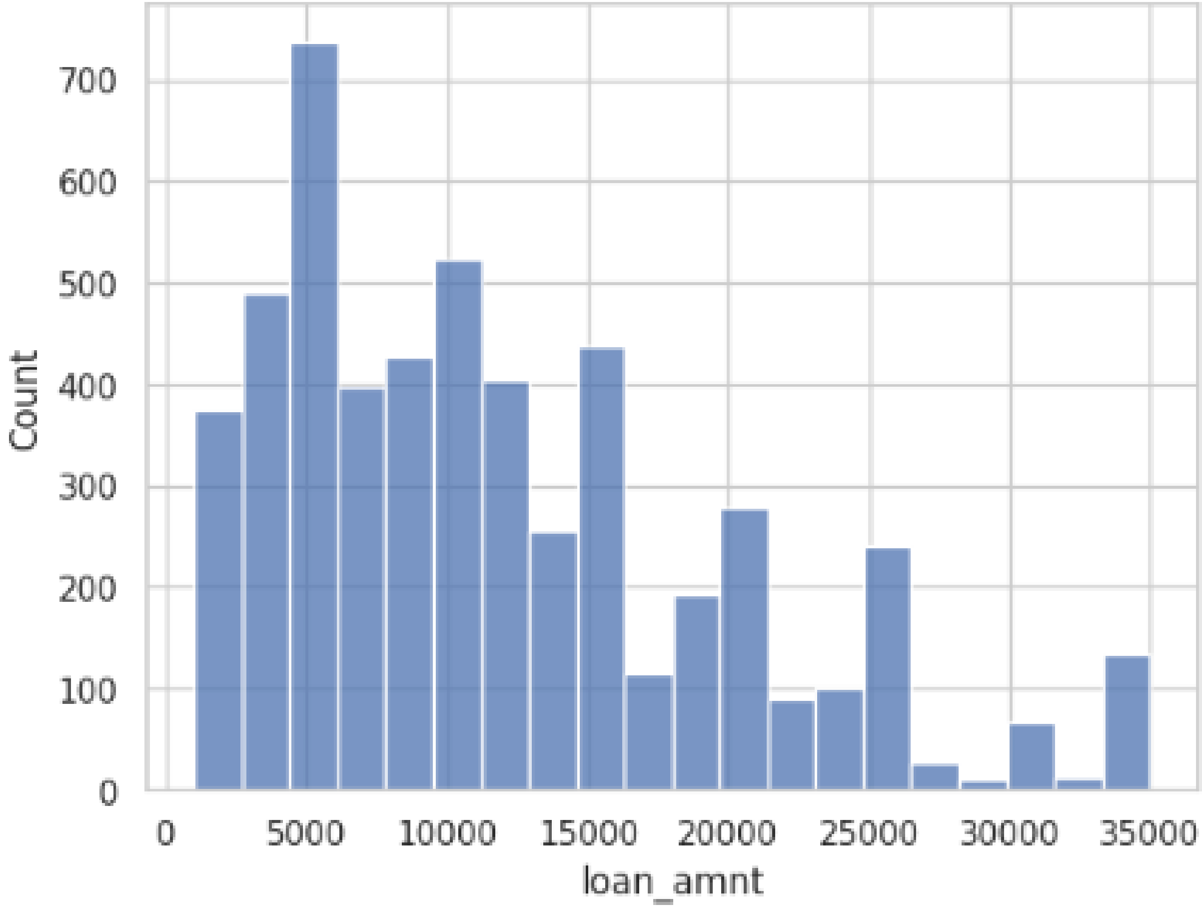
LOAN TERM ANALYSIS:

- Loans with a 36-month term tend to have a higher default frequency than those with other terms.

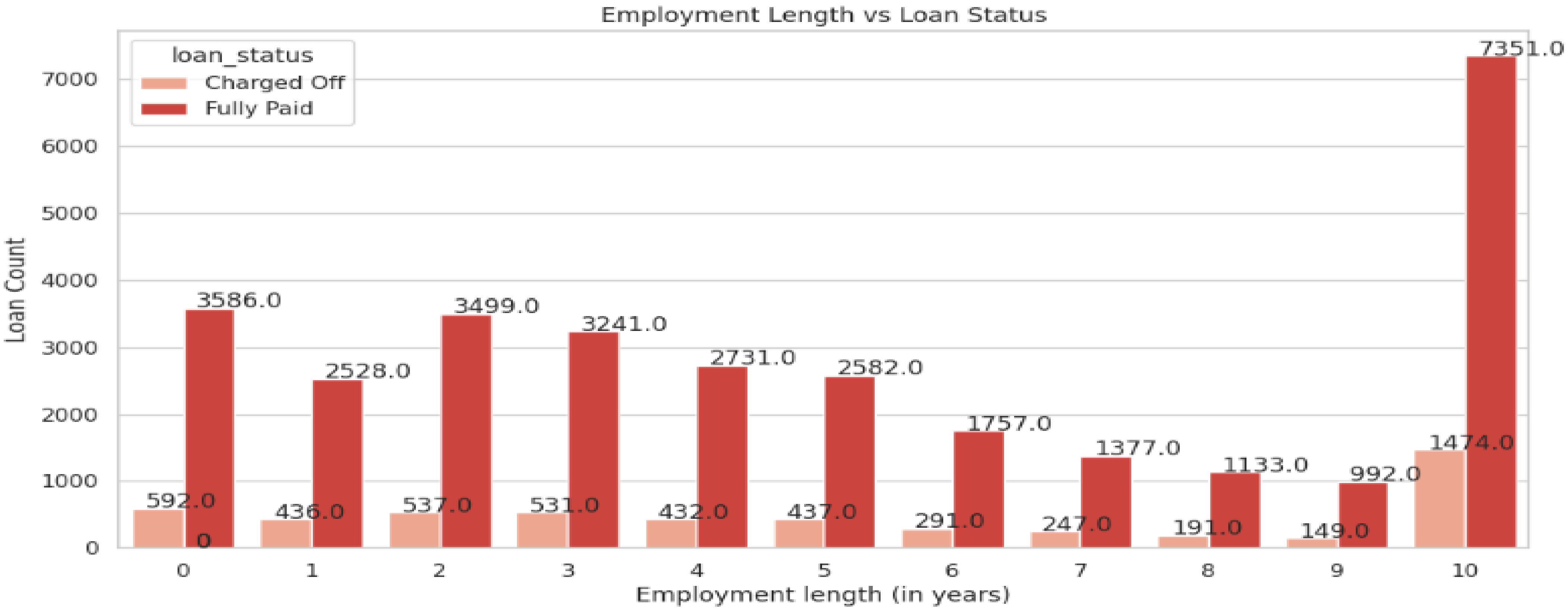


LOAN AMOUNT AND DEFAULT RISK:

- Loans under $15,000 have a higher chance of being charged off.



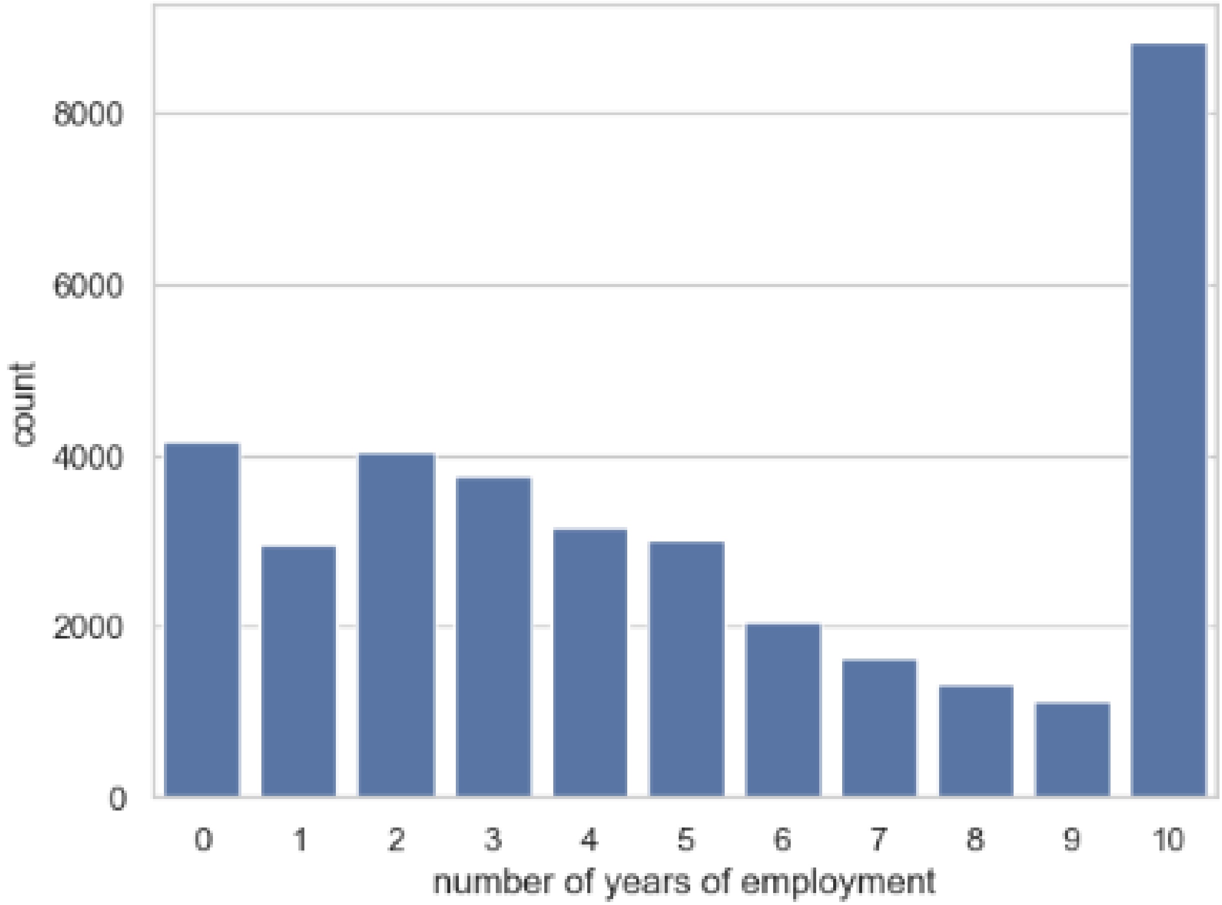
EMPLOYMENT EXPERIENCE:



- Most loan applicants have over 10 years of experience, but this group also shows a higher default risk. This may be due to their age (32+), where more financial responsibilities drive the need for loans.

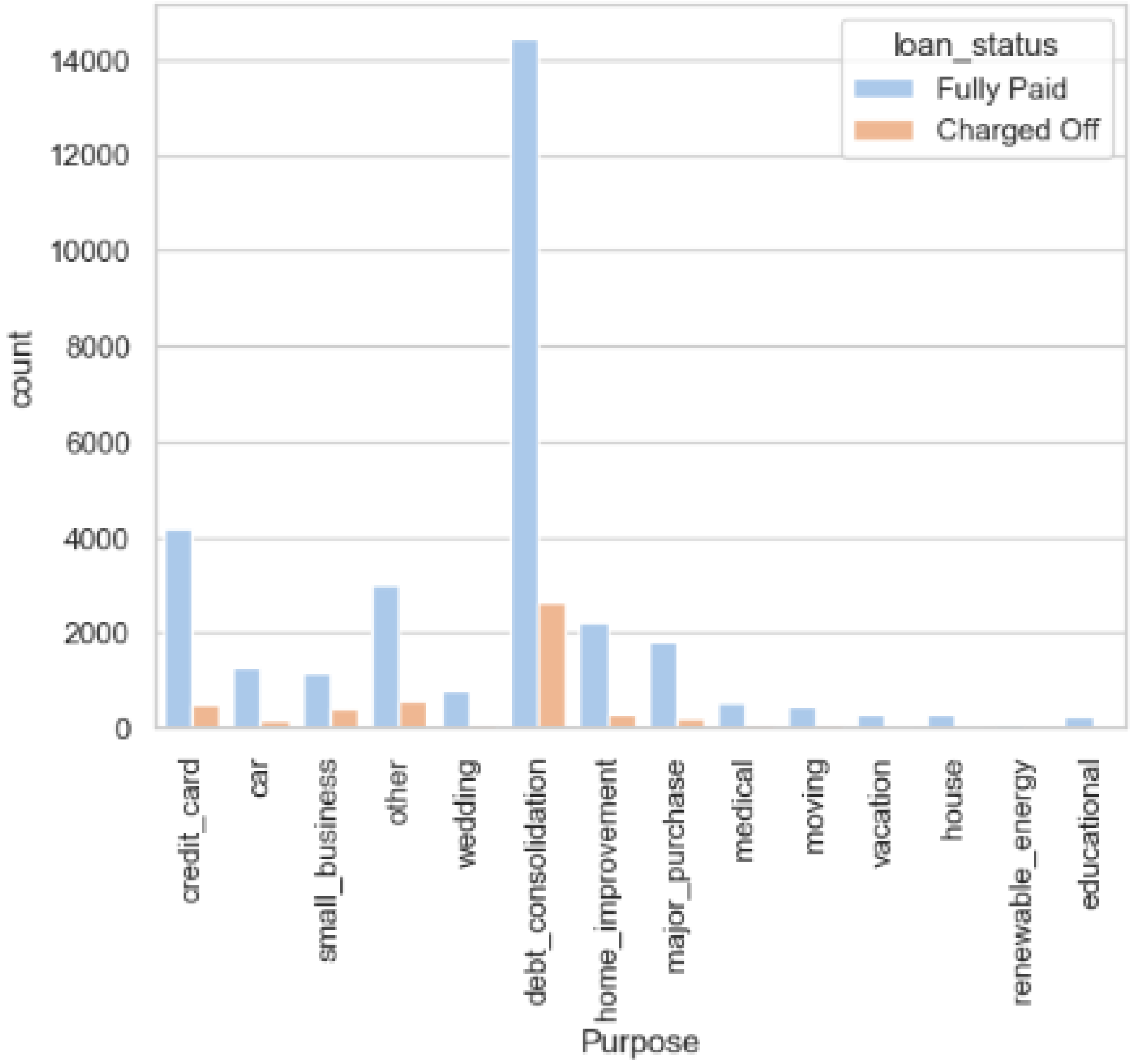
LOAN VERIFICATION

- Verified loan applications tend to have higher mean loan amounts, suggesting the company prioritizes verifying higher-value loans.



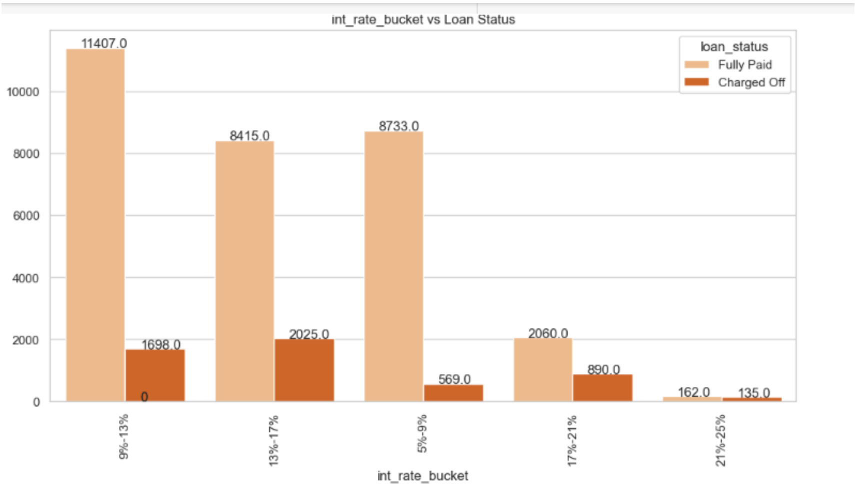
LOAN PURPOSE:

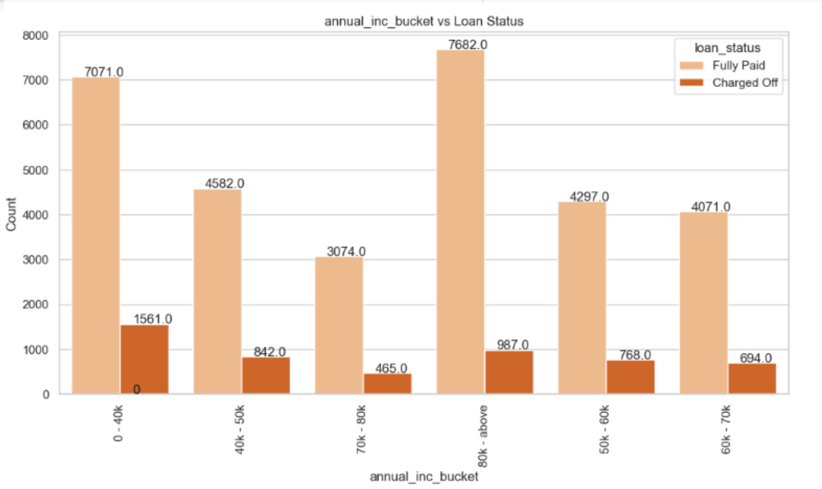
- Debt consolidation is the most common reason for borrowing, and most loans for this purpose are fully paid off. Borrowers with significant debt are more likely to take out loans.

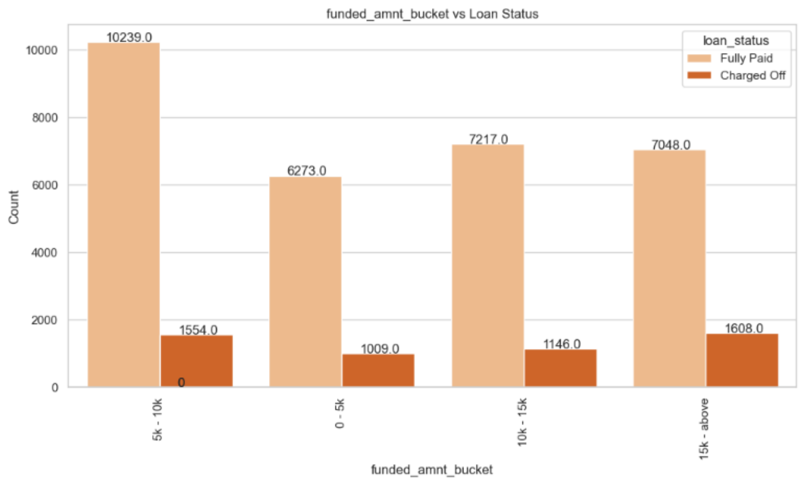


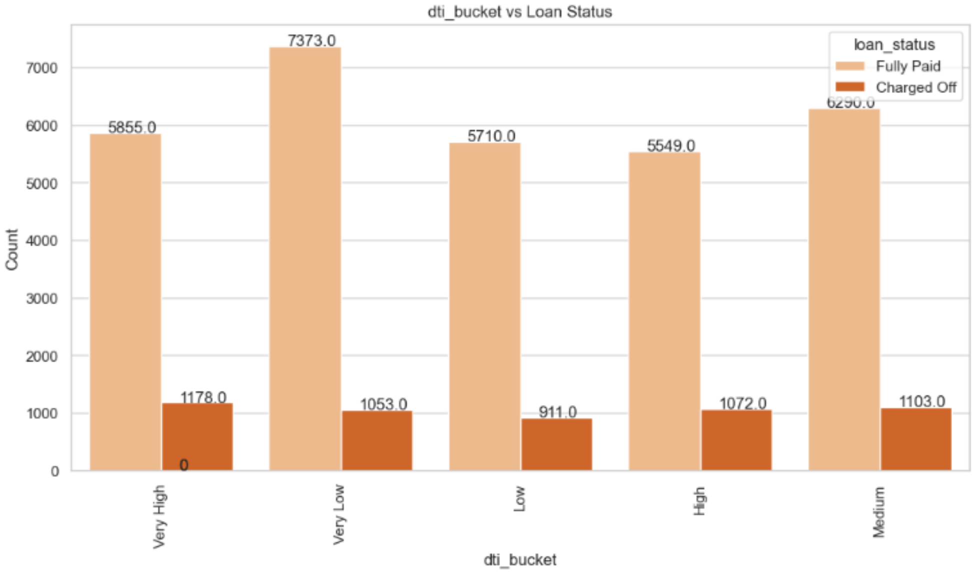
DEFAULT RISK FACTORS:

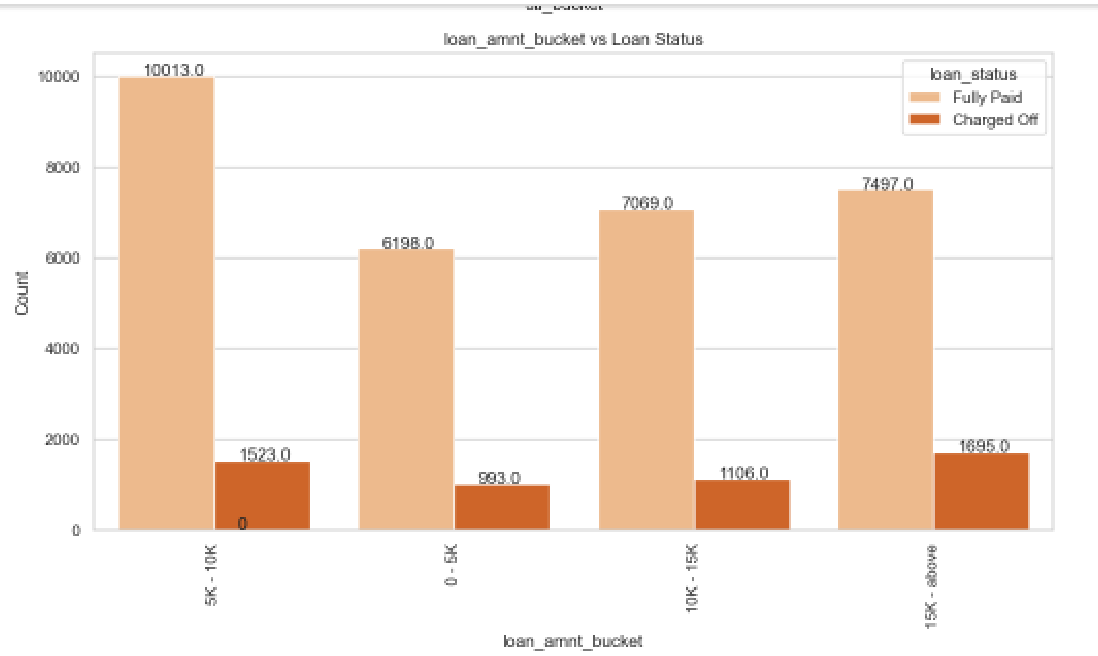
1. Applicants who defaulted generally received loans of $15,000 or more.
2. Those who charged off often had high Debt-to-Income (DTI) ratios and reported annual incomes below $40,000.
3. Many defaulted on loans with interest rates ranging from 13% to 17%.


int_rate_bucket vs Loan Status
loan_status
Fully Paid
Charged Off
11407.0
1698.0
8415.0
2025.0
8733.0
569.0
2060.0
890.0
162.0
135.0
0
9%-13%
13%-17%
5%-9%
17%-21%
21%-25%
int_rate_bucket


annual_inc_bucket vs Loan Status
loan_status
Fully Paid
Charged Off
Count
7071.0
1561.0
4582.0
842.0
3074.0
465.0
7682.0
987.0
4297.0
768.0
4071.0
694.0
0
0 - 40k
40k - 50k
70k - 80k
80k - above
50k - 60k
60k - 70k
annual_inc_bucket


funded_amnt_bucket vs Loan Status
loan_status
Fully Paid
Charged Off
Count
10239.0
1554.0
6273.0
1009.0
7217.0
1146.0
7048.0
1608.0
0
5k - 10k
0 - 5k
10k - 15k
15k - above
funded_amnt_bucket


dti_bucket vs Loan Status
loan_status
Fully Paid
Charged Off
Count
5855.0
1178.0
7373.0
1053.0
5710.0
911.0
5549.0
1072.0
6290.0
1103.0
0
Very High
Very Low
Low
High
Medium
dti_bucket


loan_amnt_bucket vs Loan Status
loan_status
Fully Paid
Charged Off
Count
10013.0
1523.0
6198.0
993.0
7069.0
1106.0
7497.0
1695.0
0
5K - 10K
0 - 5K
10K - 15K
15K - above
loan_amnt_bucket