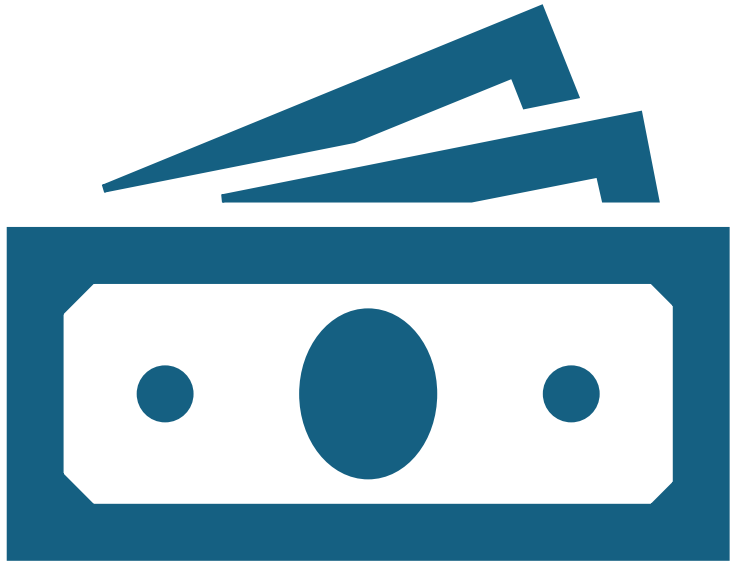
Lending Club Case Study

- Analysis of Risky Loan Applicants

- Komal Yadav
- Abhilash Kukawalkar



Objective

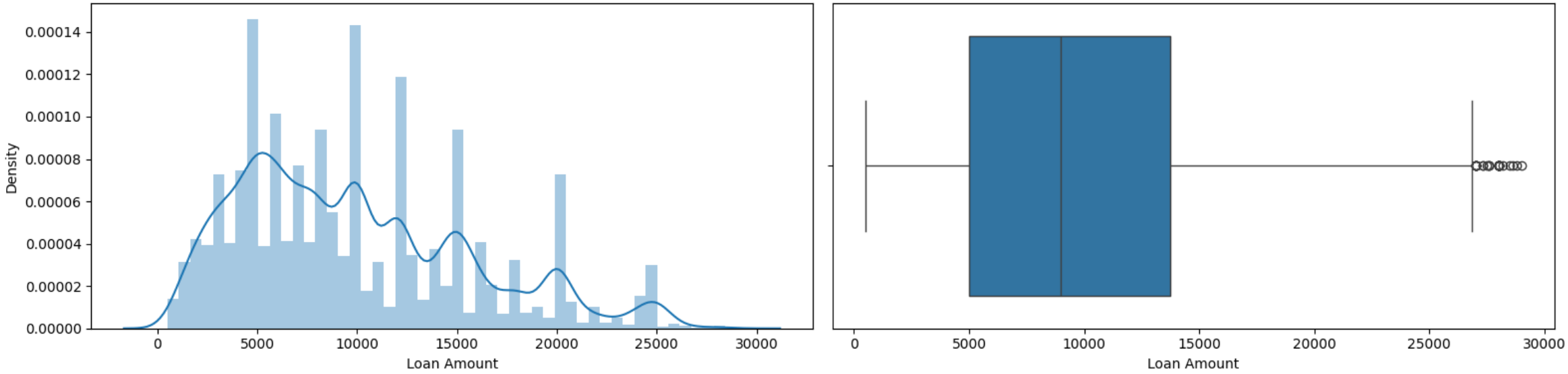
- This analysis aims to identify patterns and factors that indicate if a person is likely to default on a loan, using data from past loan applicants. The goal is to help the company make informed decisions to reduce credit loss.

Loan_status

loan_status	count
Fully Paid	28477
Charged Off	4714

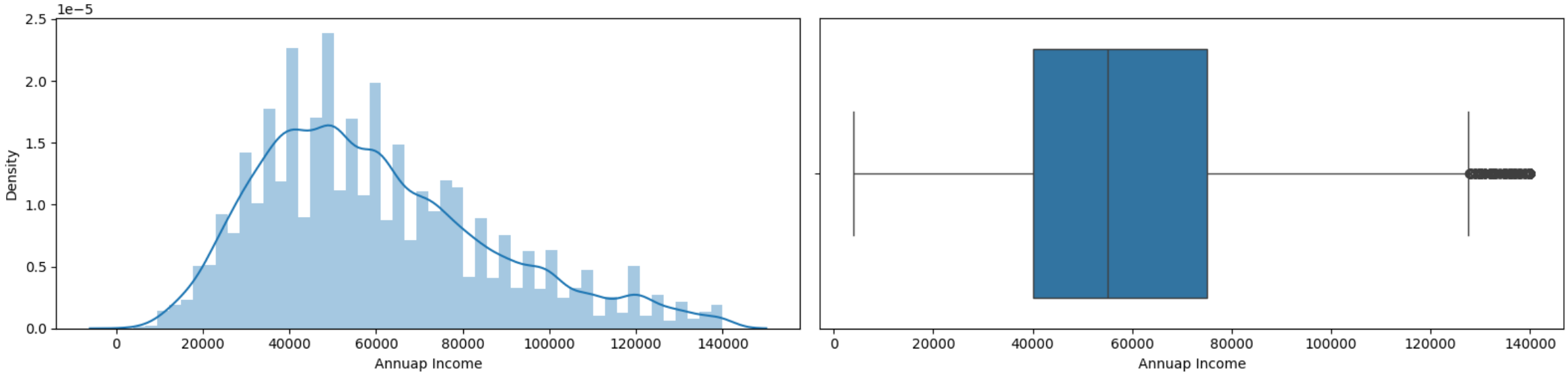
Observation

- The distribution of loan amounts is right-skewed, with a majority of loans falling between $5,000 and $15,000.
- # The mean loan amount is around $11,000.
- # There are a few outliers on the higher end, indicating some loans significantly larger than the average.



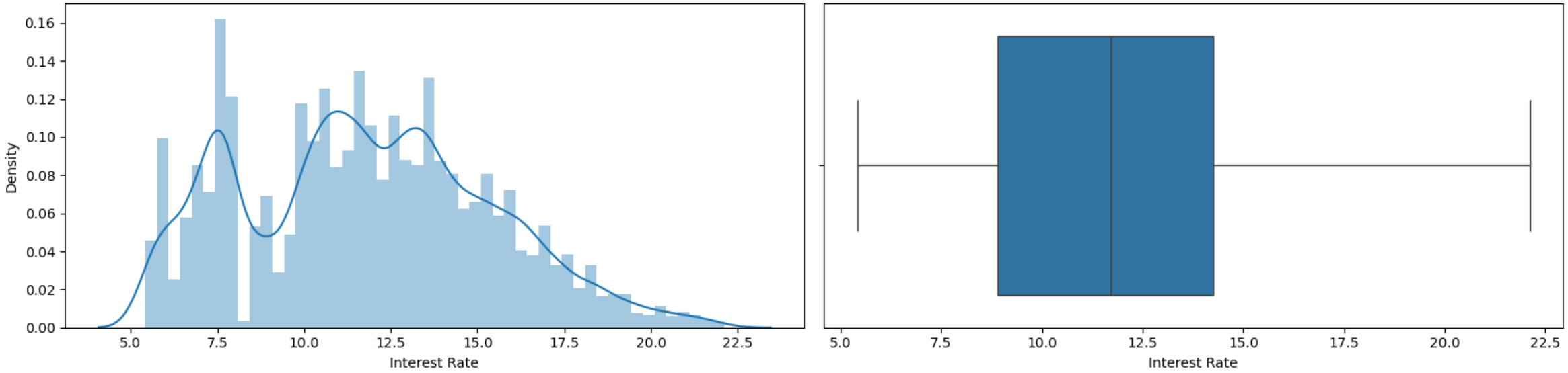
Observation

- The distribution of annual incomes is heavily right-skewed, indicating a concentration of borrowers with lower incomes and a long tail of high earners.
- The mean annual income is around $65,000, but the median is likely lower due to the skew.
- There are potential outliers on the higher end, suggesting a small number of borrowers with significantly higher incomes compared to the majority.



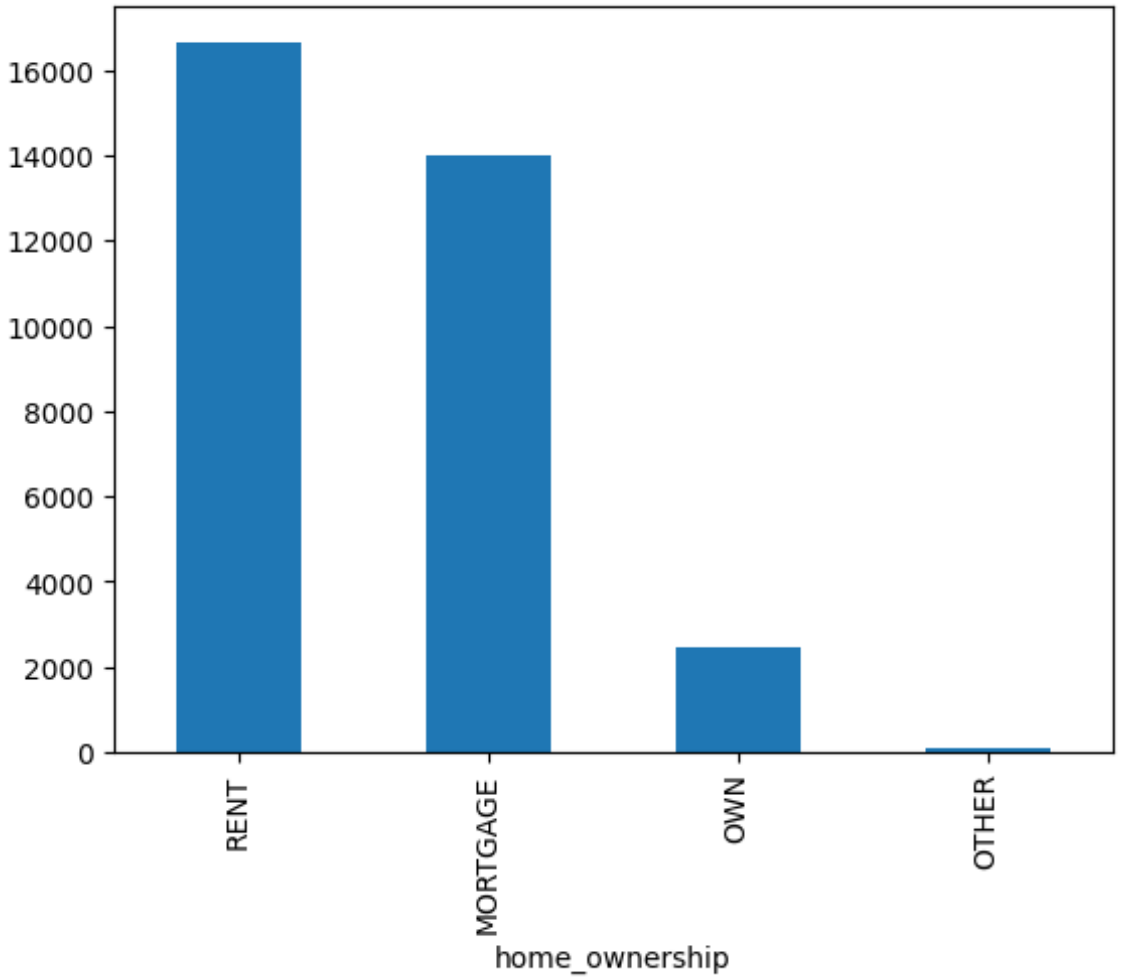
Observation

- The distribution of interest rates appears to be bimodal, with peaks around 10% and 13%.
- The mean interest rate is approximately 12%.
- There are no significant outliers in the interest rate distribution.
- The interest rate for most of the loans lies between 9%-14%.



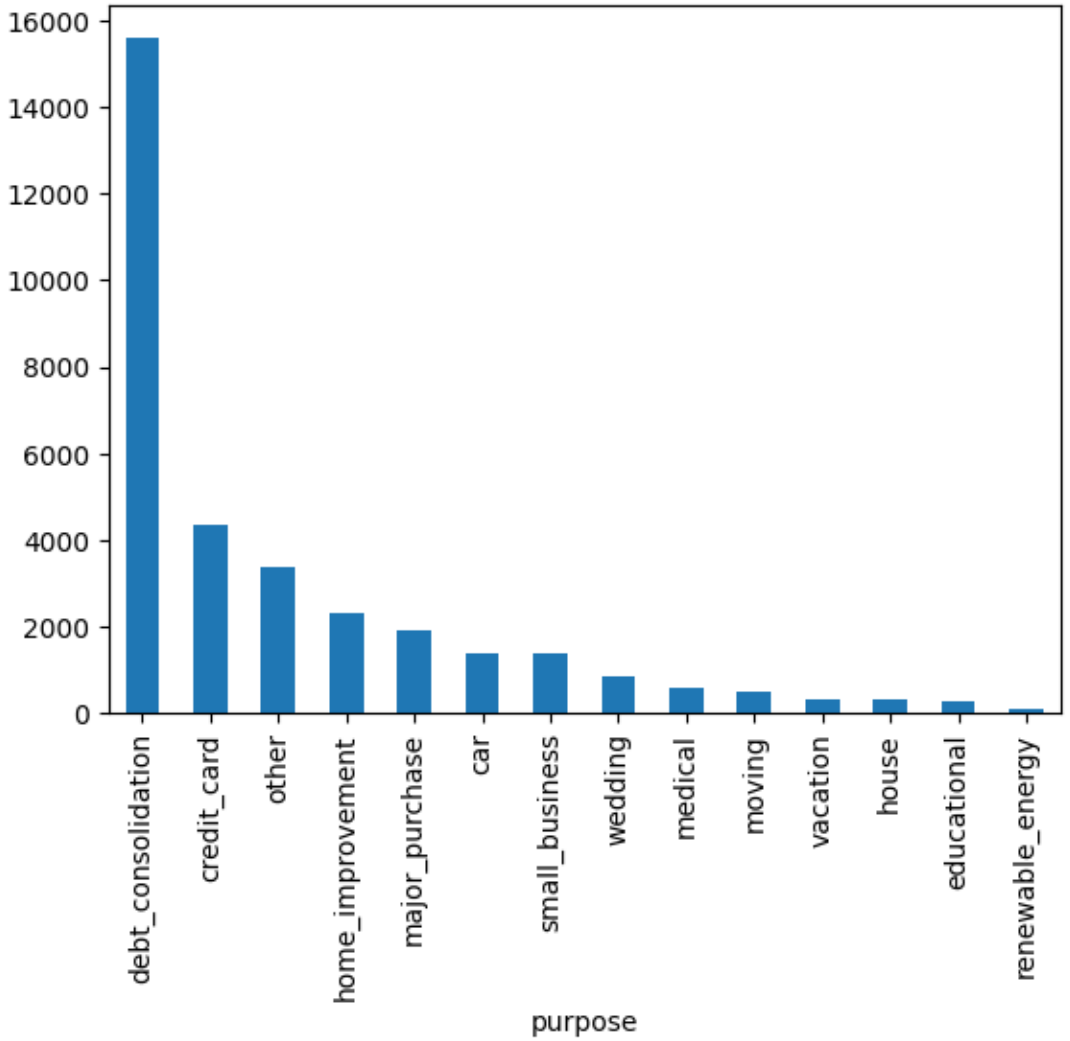
Observation

- Most of the loan applicants either living in rented house or mortgaged their house.



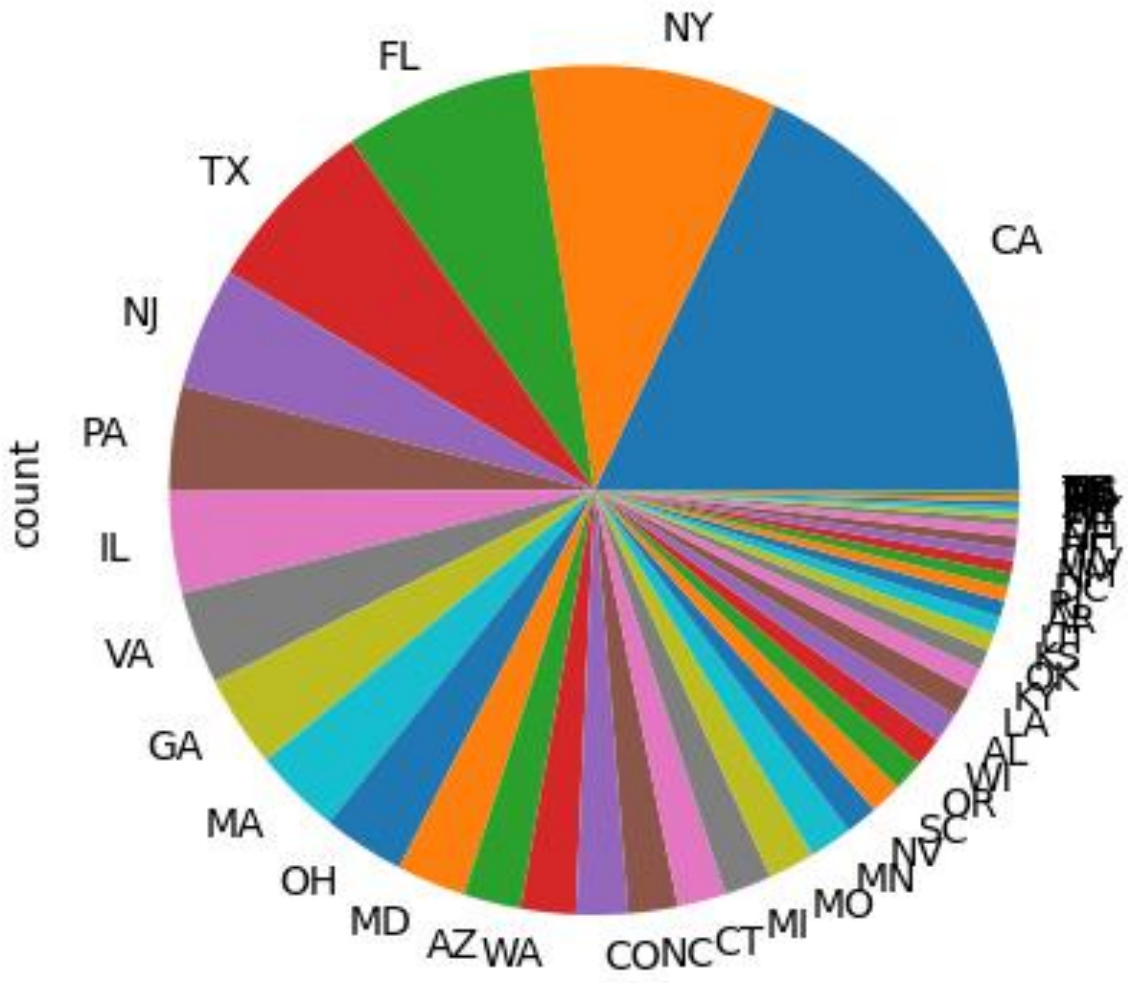
Observation

- Most of the loan applicants are for debt_consolidations.



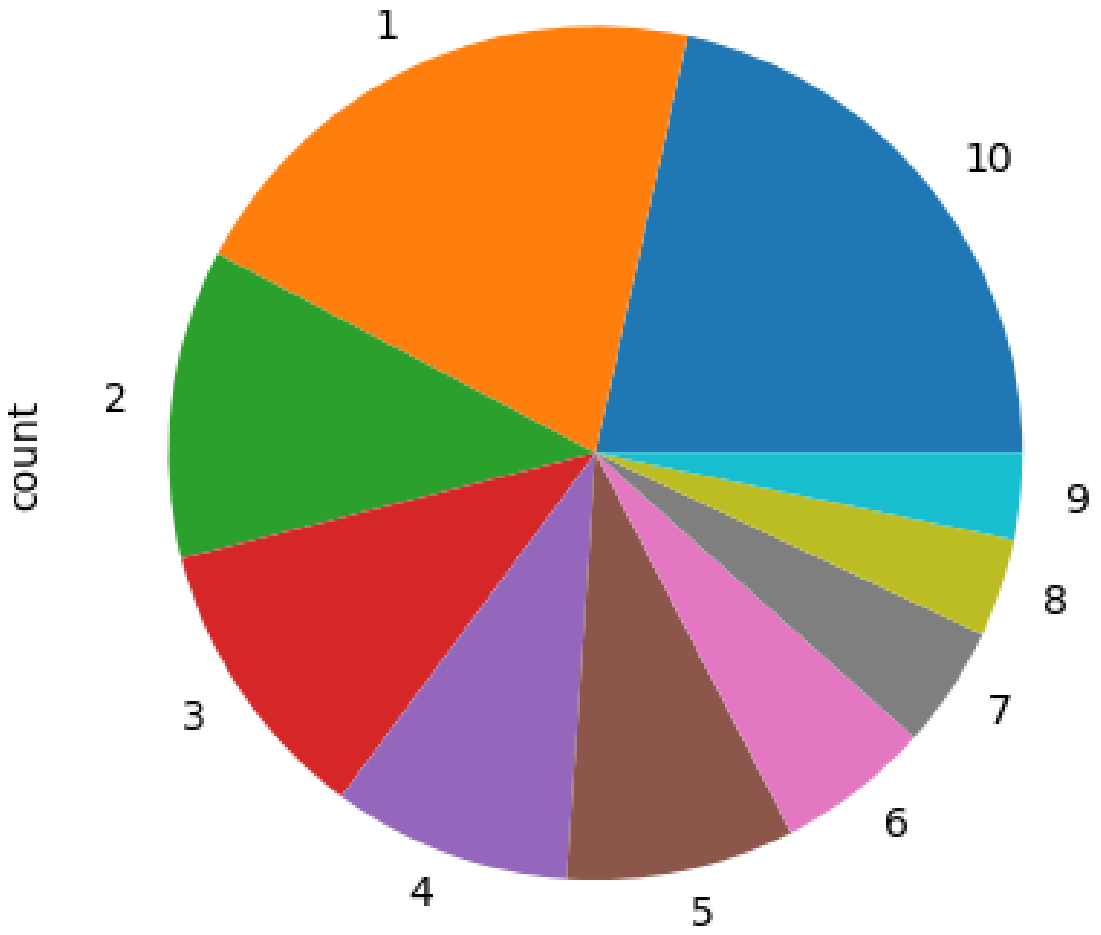
Observation

Most of the Loan applicants are from CA(State).



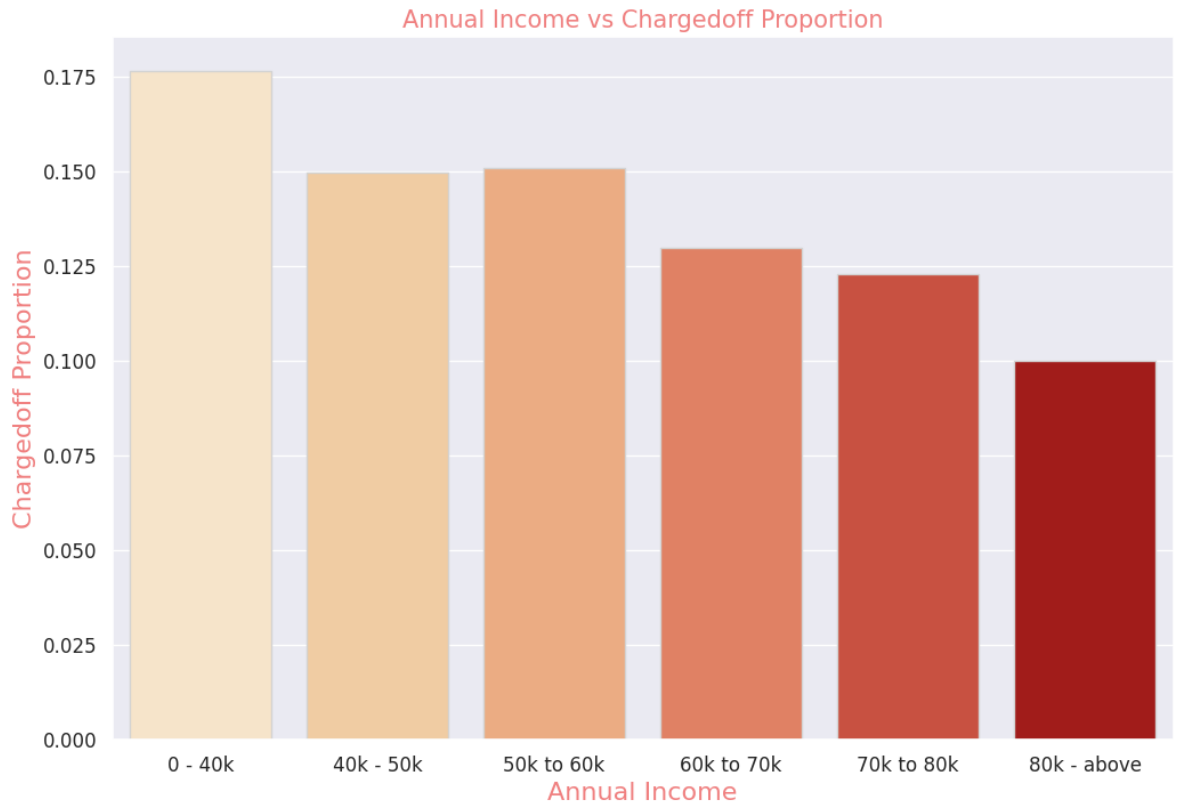
Observation

Most of the applications are having 10+ yrs of Exp.



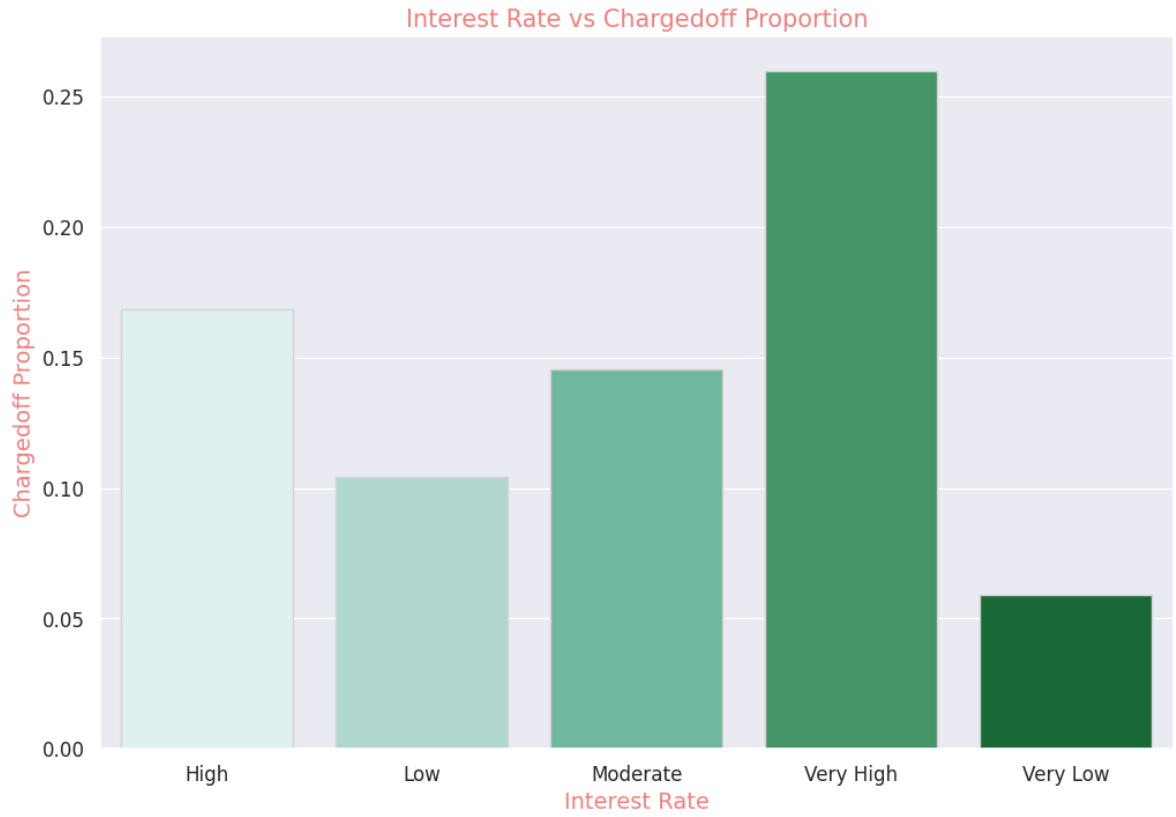
Bivariate Analysis

- Applicants with lower annual income are more likely to default on loans compared to those with higher annual incomes.
- This suggests that income level is a significant factor in assessing credit risk.



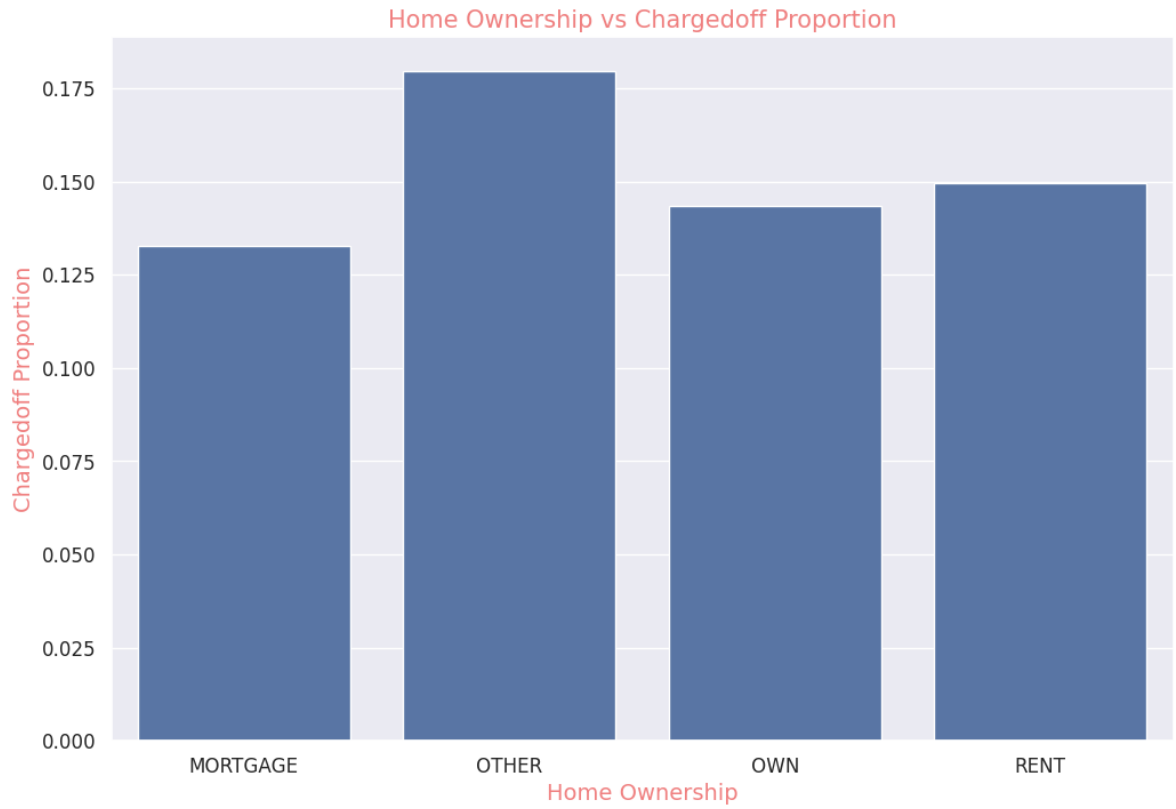
Observations:

- Loans with higher interest rates have a significantly higher proportion of charge-offs.
- This indicates that interest rate is a strong predictor of default risk, with higher rates associated with increased likelihood of loan default.



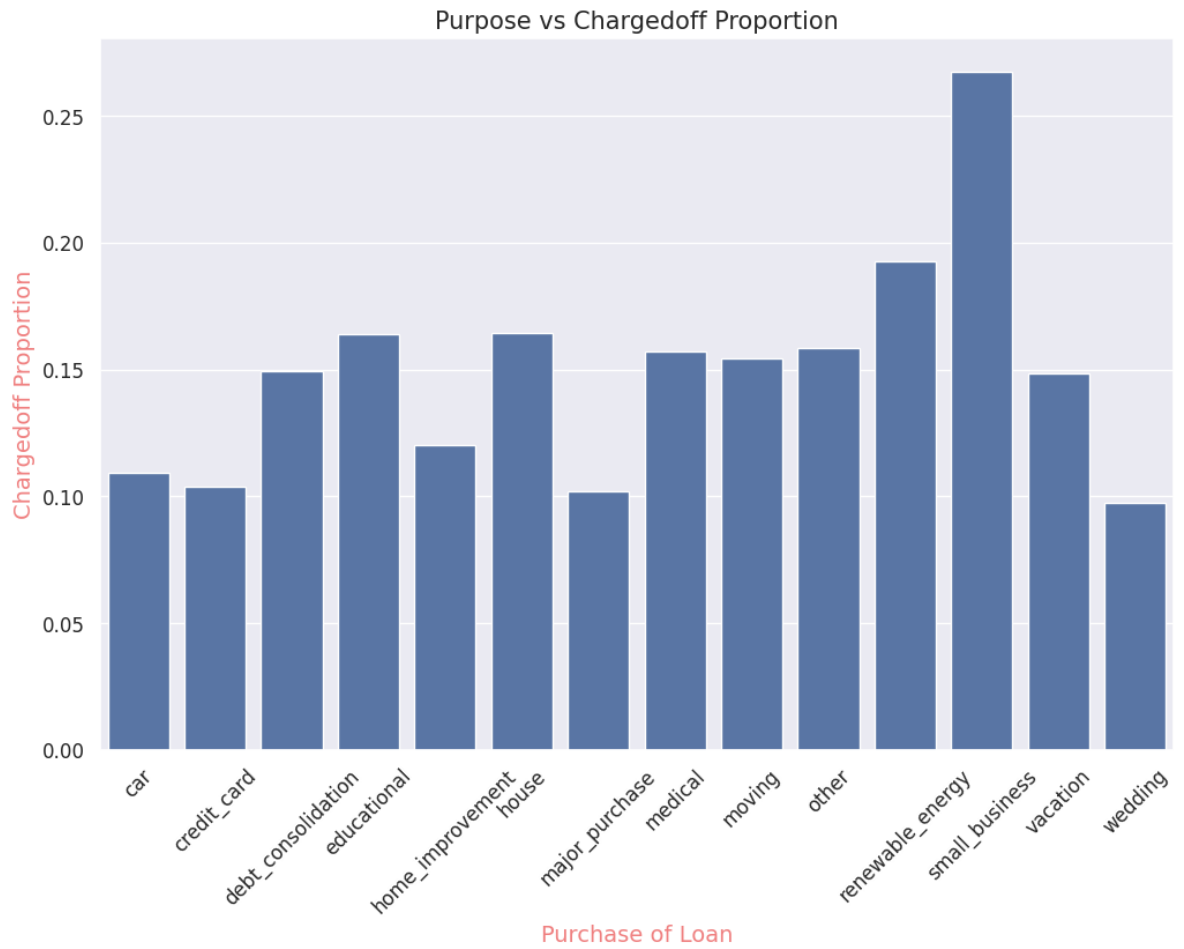
Observation

- Those who are not owning the home is having high chances of loan defaults.



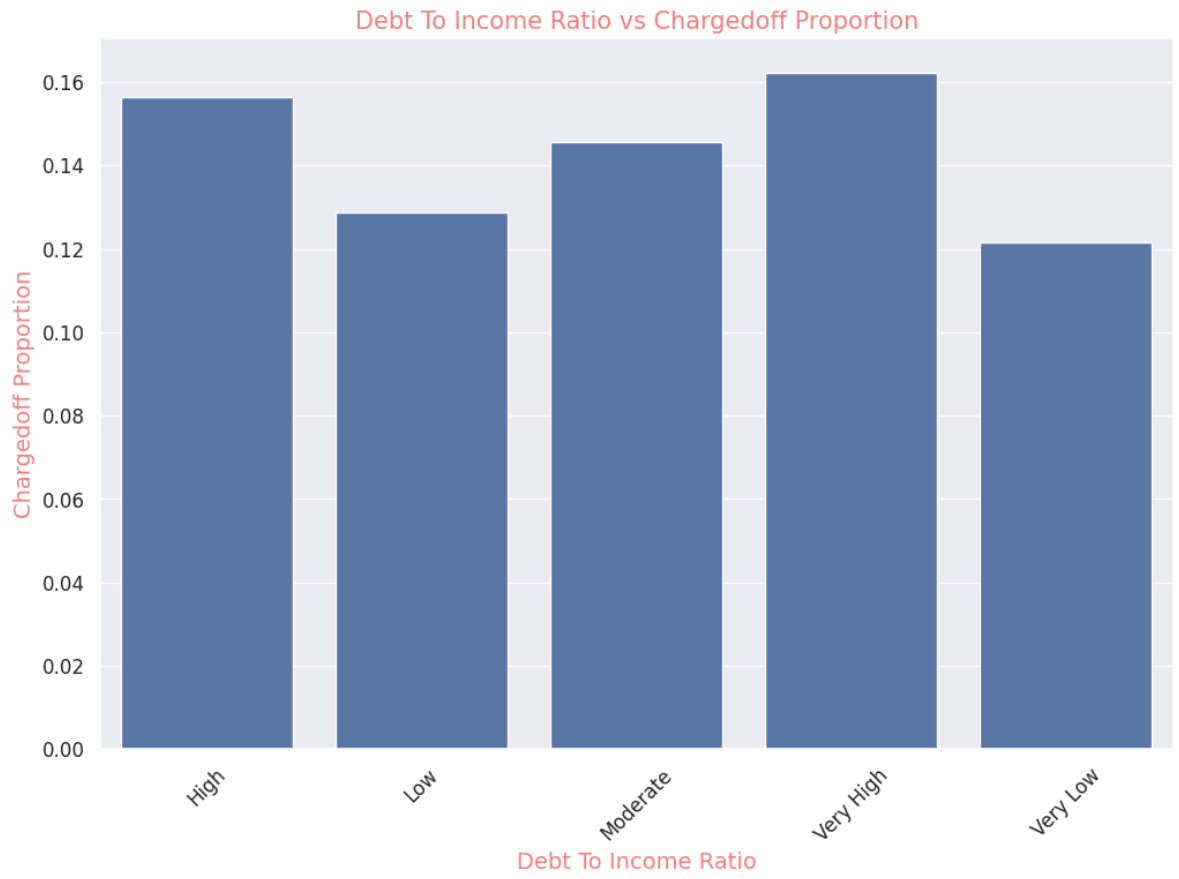
Observation

- Those applicants who is having home loan is having low chances of loan defaults.
- Those applicants having loan for small bussiness is having high chances for loan defaults.



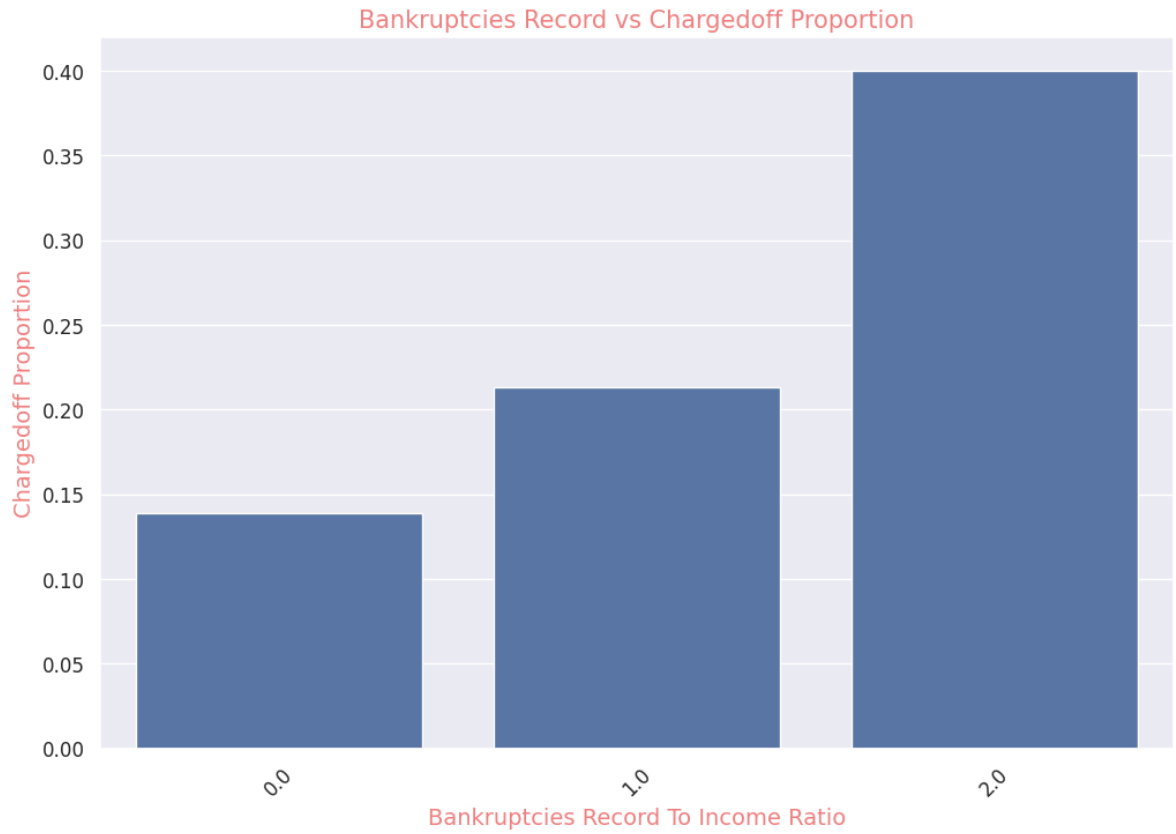
Bivariate Analysis on DTI against Chargedoff_Proportion

- As the Debt-to-Income ratio increases, the proportion of charged-off loans also tends to increase.
- This suggests that borrowers with higher DTI ratios are more likely to default, indicating a higher risk associated with lending to individuals with a larger debt burden relative to their income.



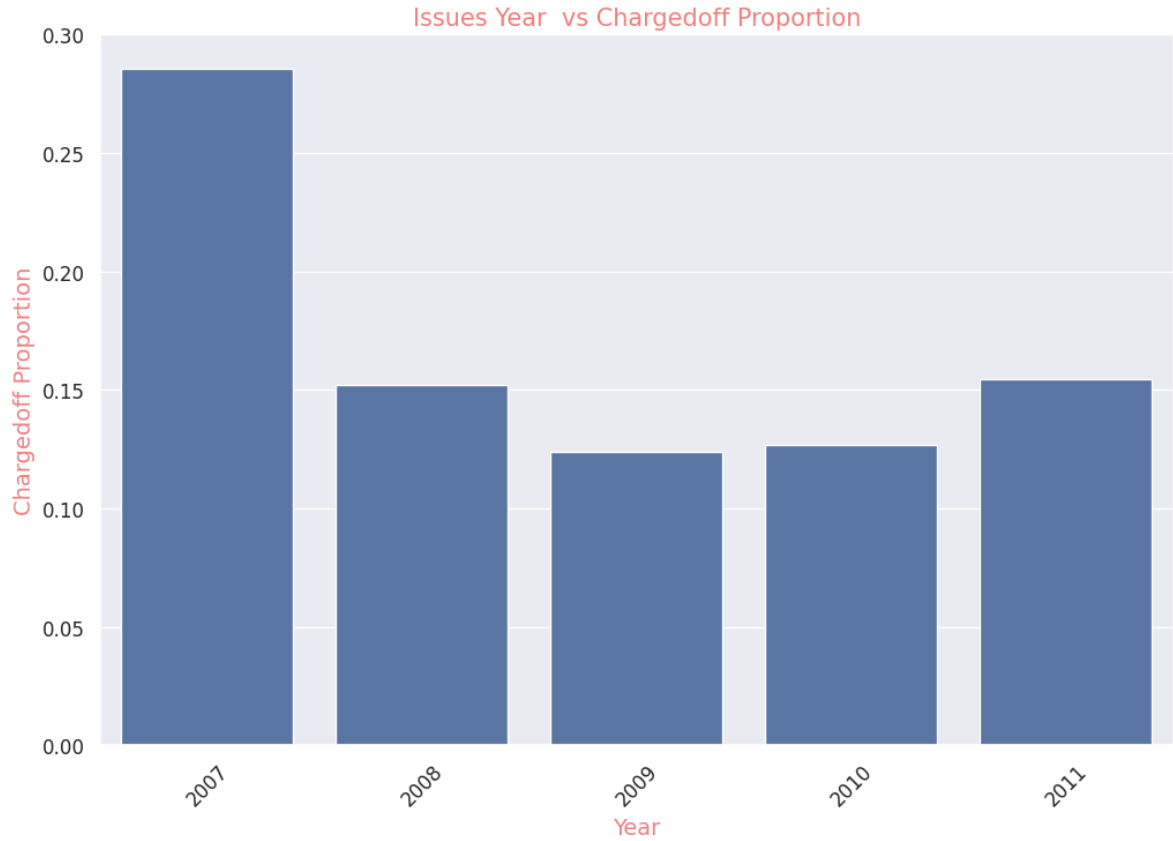
Observations

- Borrowers with a history of bankruptcies (pub_rec_bankruptcies > 0) exhibit a higher likelihood of loan default compared to those without any prior bankruptcies.
- This highlights the importance of considering past financial behavior as a key indicator of credit risk.



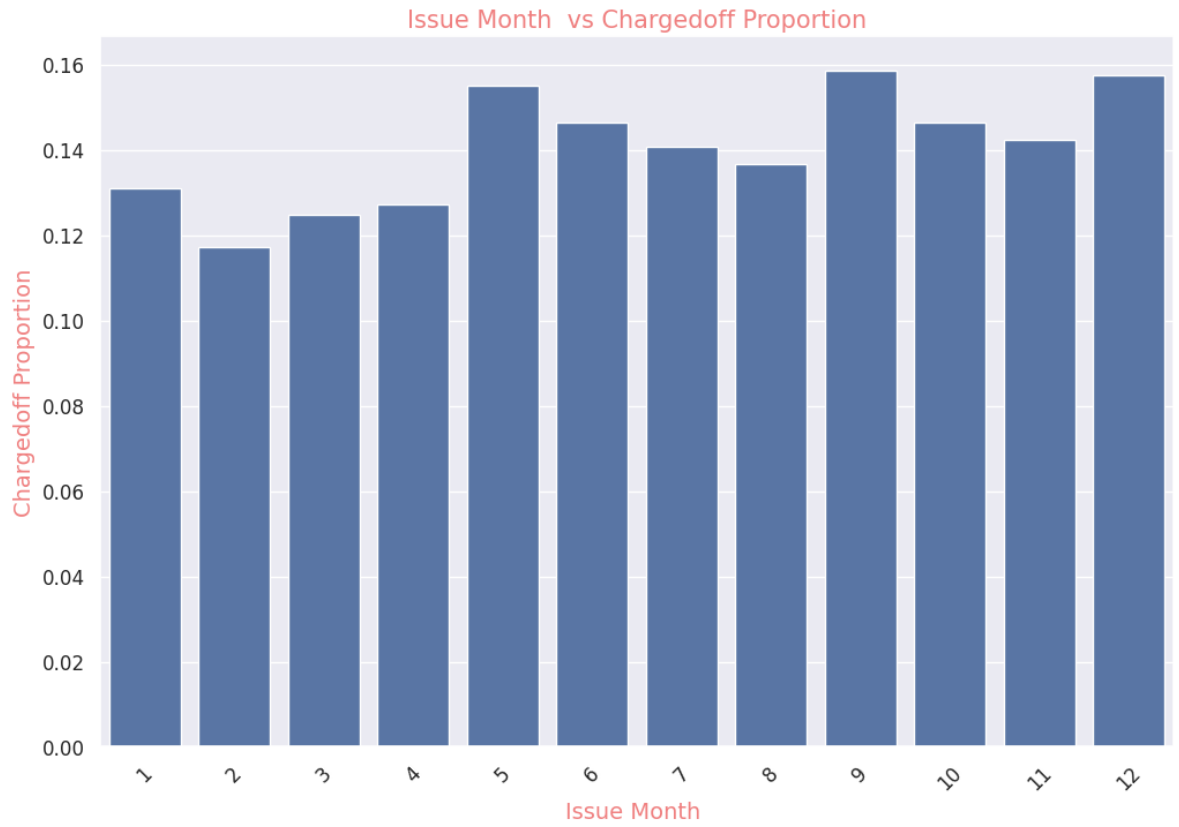
Observations

- Loans issued in earlier years (2007 and 2008) show a higher proportion of charge-offs compared to those issued in later years.
- This could be attributed to various factors, including the economic conditions prevalent during those years or changes in lending practices over time.



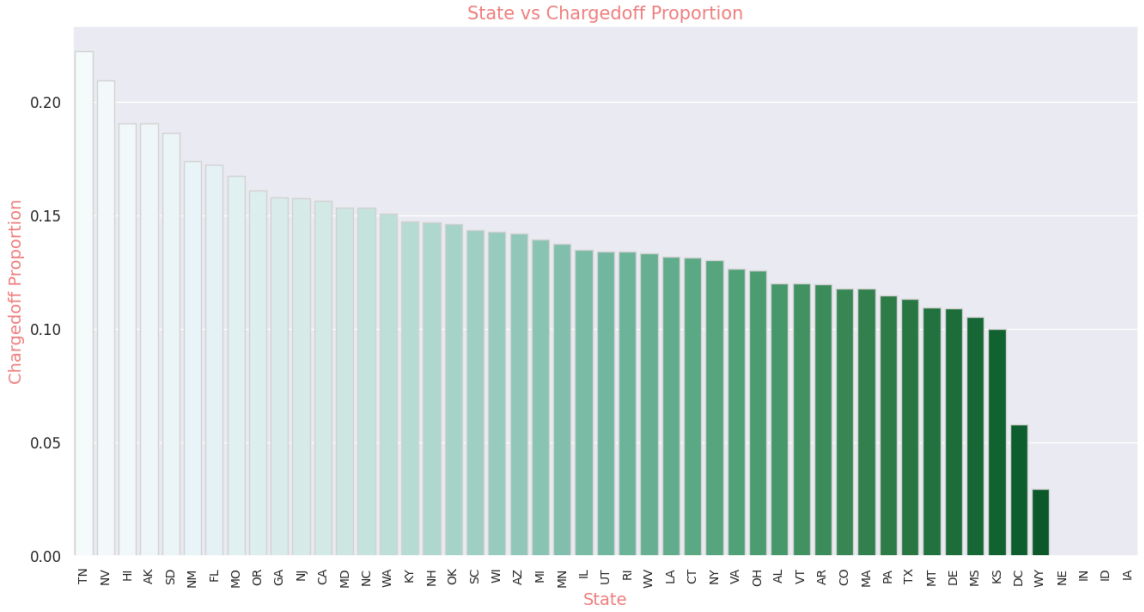
Observations

- There is no clear trend or significant variation in charge-off proportions across different months of loan issuance.
- This suggests that the month in which a loan is issued might not be a strong predictor of default risk.



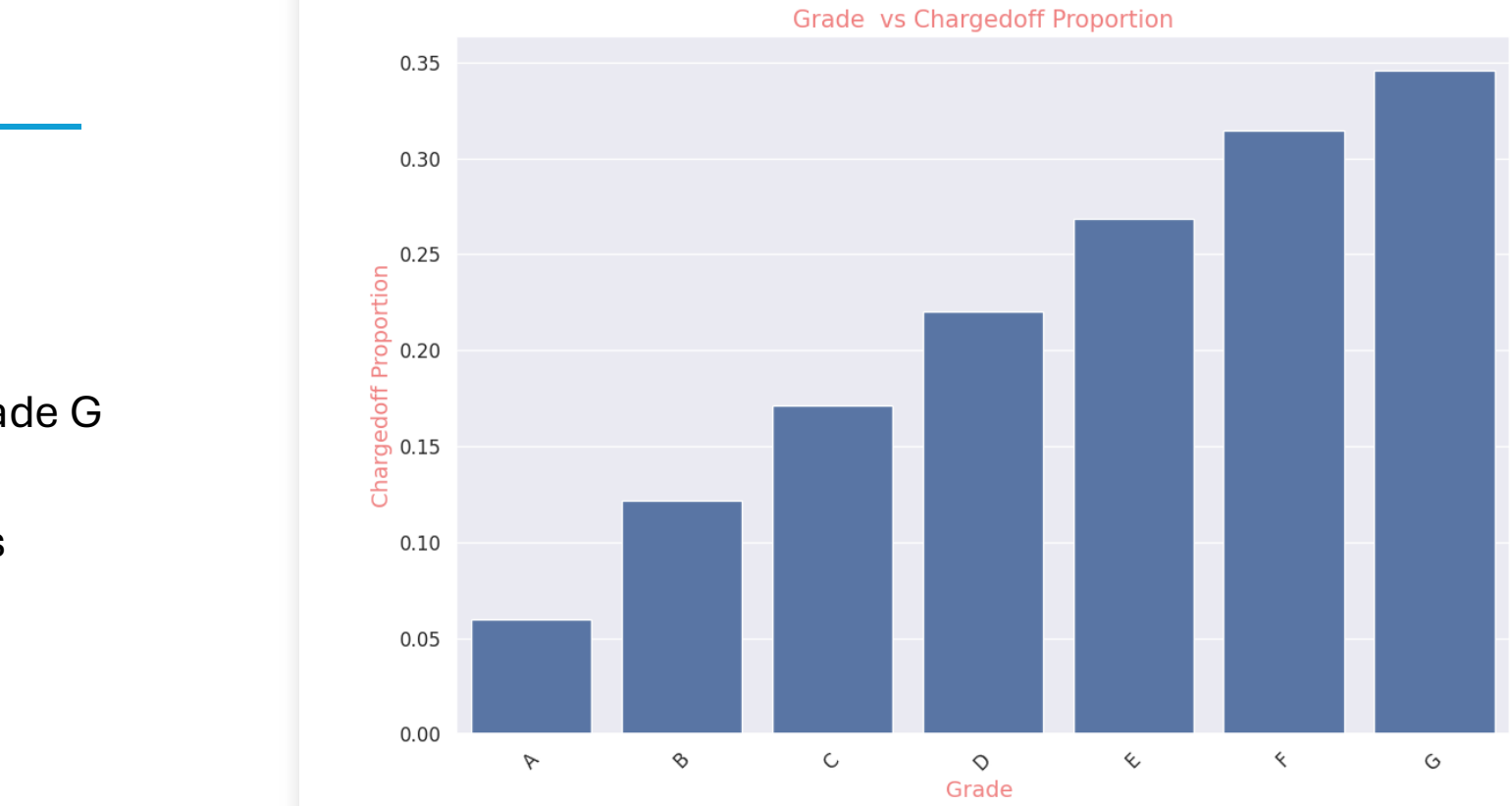
Observations

- # States like NV, AK, FL are having high chances of loan defaults.
- # States like IA, ME, ID are having low chances of loan defaults.



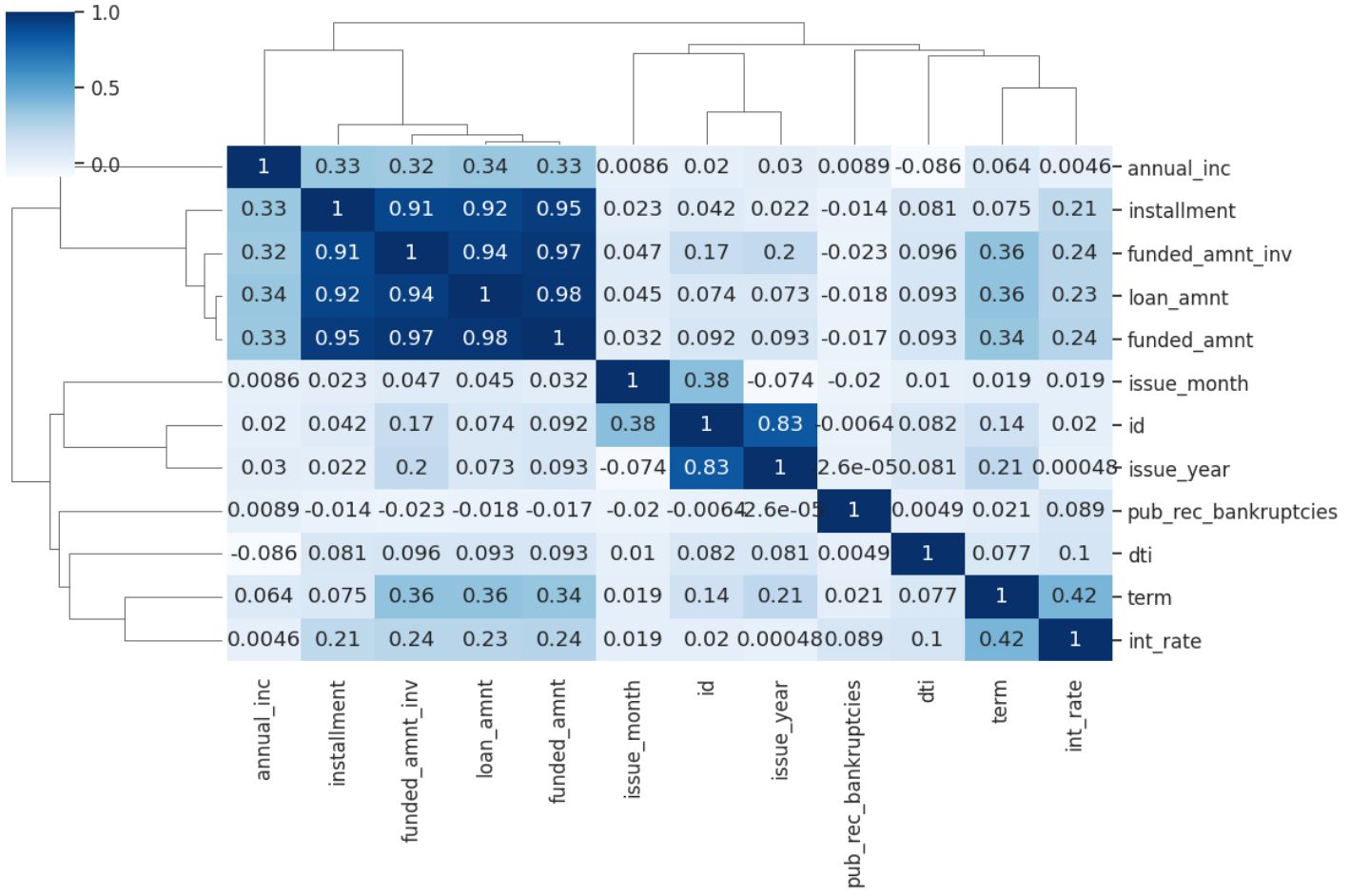
Observations

- The Loan applicants with loan Grade G is having higest Loan Defaults.
- The Loan applicants with loan A is having lowest Loan Defaults.



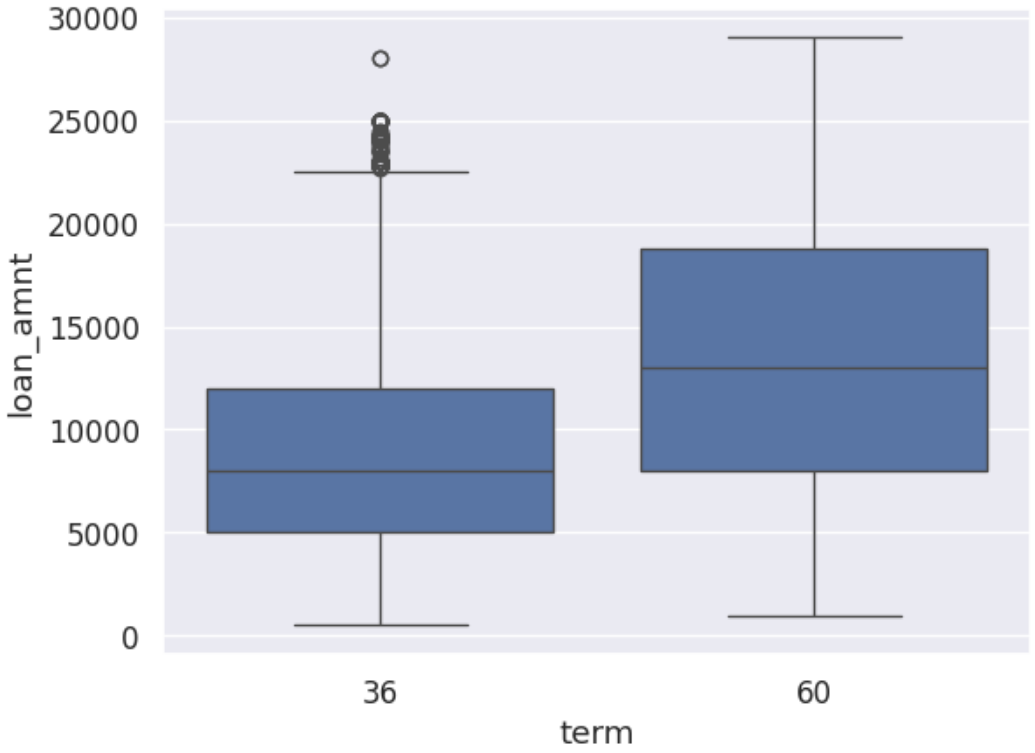
Correlation heatmap

- # Positive Correlations:
- Loan amount, funded amount, and investor-funded amount are strongly positively correlated, indicating that larger loan requests generally receive higher funding.
- Interest rate shows a moderate positive correlation with loan amount, suggesting that larger loans tend to have higher interest rates.
- Installment amount is strongly correlated with loan amount and funded amount, reflecting the direct relationship between loan size and monthly payments.
- # Negative Correlations:
- Public record bankruptcies exhibit a weak negative correlation with annual income, implying that individuals with higher incomes are less likely to have a history of bankruptcies.
- # Overall:
- The heatmap reveals the interdependencies among various numerical variables in the loan dataset.
- # - These correlations can be leveraged to gain insights into factors influencing loan approval, interest rates, and the likelihood of default.



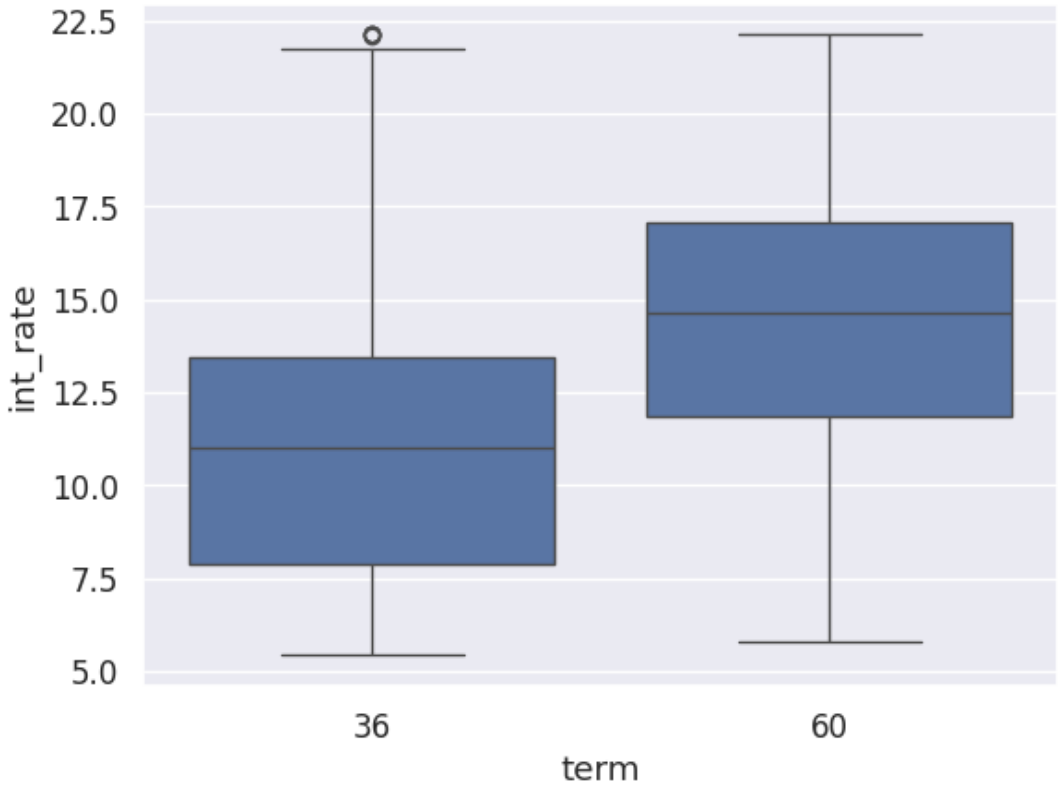
Bivariate analysis of the relationship between loan term ('term') and loan amount ('loan_amnt')

- If term is increasing loan amount is also increasing



Bivariate analysis of the relationship between loan term ('term') and interest rate

- Loans with longer terms (60 months) generally have higher interest rates than those with shorter terms (36 months).
- This reflects the increased risk associated with longer-term lending, as the lender is exposed to potential default for a more extended period.



Summary

- Loan applicants with lower income and higher interest rates are more likely to default.
- Borrowers who don't own homes and have a history of bankruptcies also pose higher risk.
- Loan purpose influences default rates, with small business loans being riskier.
- Loans issued in earlier years (2007-2008) show higher charge-off rates.
- States like NV, AK, FL exhibit higher default rates compared to IA, ME, ID.