Identifying Loan Default Risk Using Exploratory Data Analysis (EDA)

A Data-Driven Approach for Financial Risk Minimization

Group Members:
Ravikanth Nagaraj
Geetika Kapoor

8/19/2024

Problem Statement

The task is to identify patterns in loan default behavior using EDA. Two risks are involved in loan approval:

- Loss of business if a repayable loan is denied.
- Financial loss if a defaulting loan is approved

Business Understanding

Lending companies face two types of risks:

- Loan accepted: Three scenarios: fully paid, current, or charged-off (defaulted).
- Loan rejected: No transactional history of rejected applicants is available in the dataset.

Loan Approval Scenarios

When a loan is accepted, three possible outcomes:

- Fully Paid: The loan is fully paid.
- Current: The loan is still being paid.
- Charged-Off: The loan is defaulted.

Business Objective

The aim is to minimize credit loss by identifying risky applicants and defaulters using EDA. The focus is on recognizing the driving factors behind loan default and preventing financial losses.

Approach to the Problem

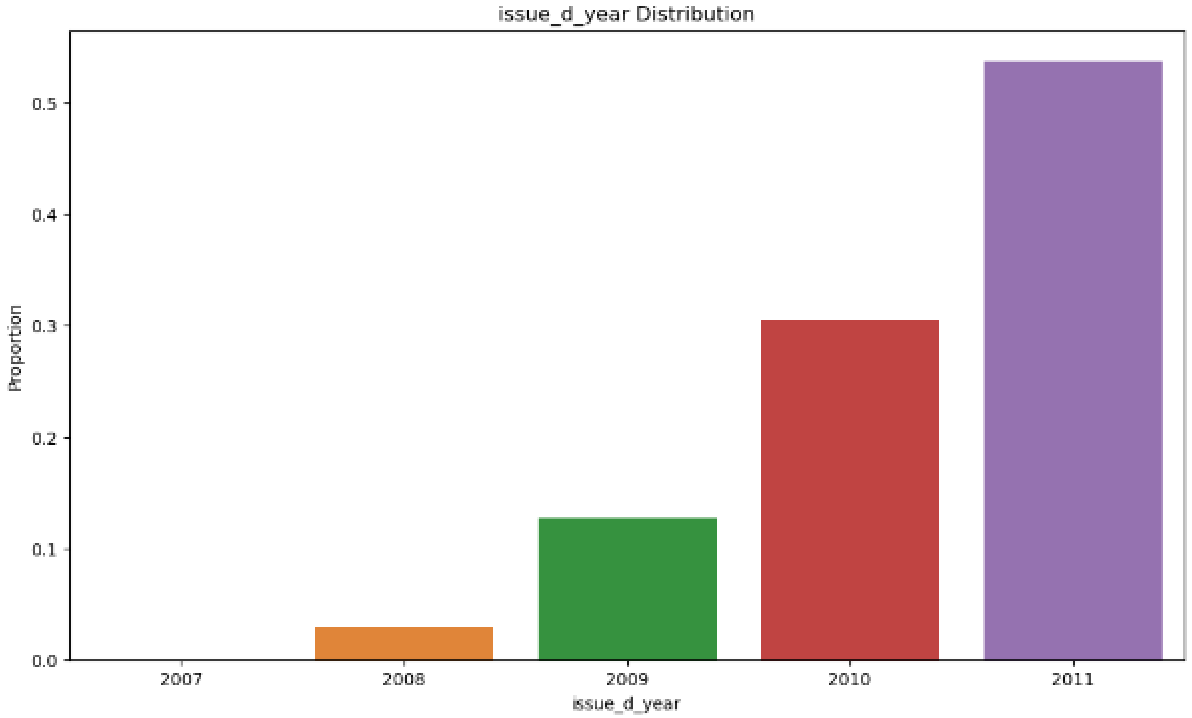
- Using EDA, we will explore the dataset to uncover insights about:
- Consumer attributes.
- Loan characteristics that influence default risk.

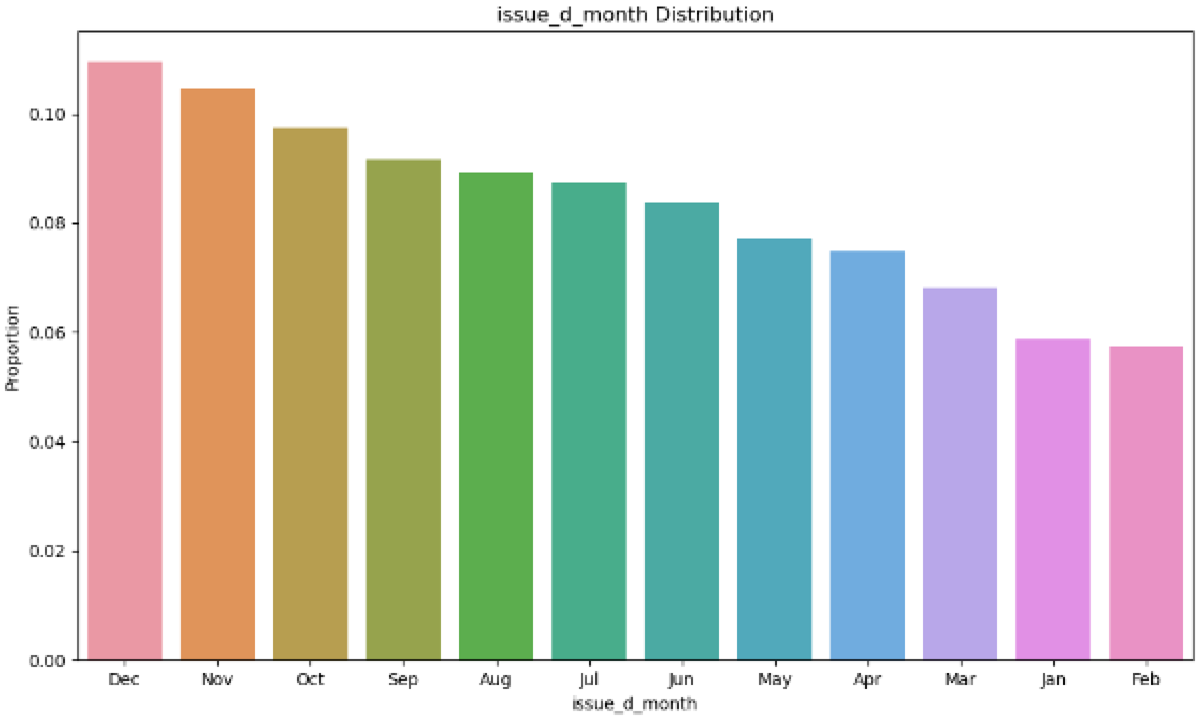
Data Understanding

Dataset Overview:

- Key driver variables: int_rate ,dti, loan_amnt, grade, annual_inc,
- Target variable: loan_status (fully paid, charged-off).

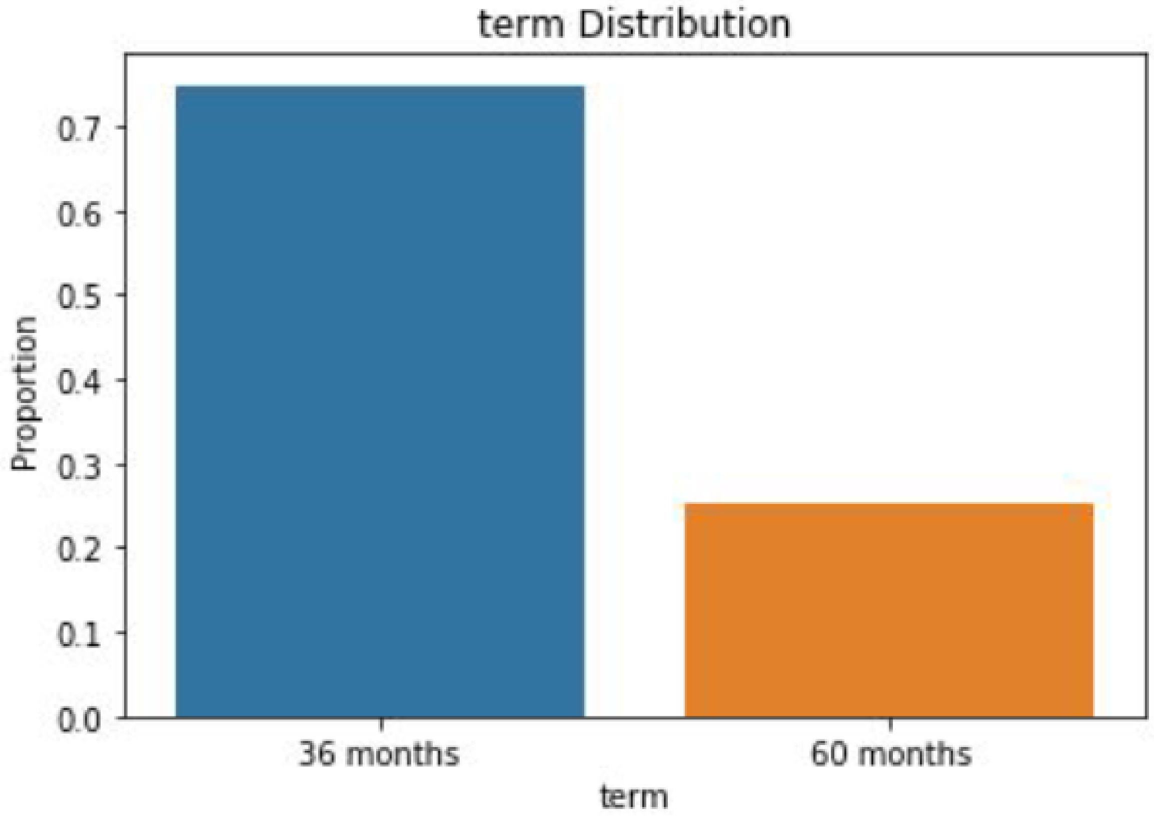
Analysis- Issued Year, Month

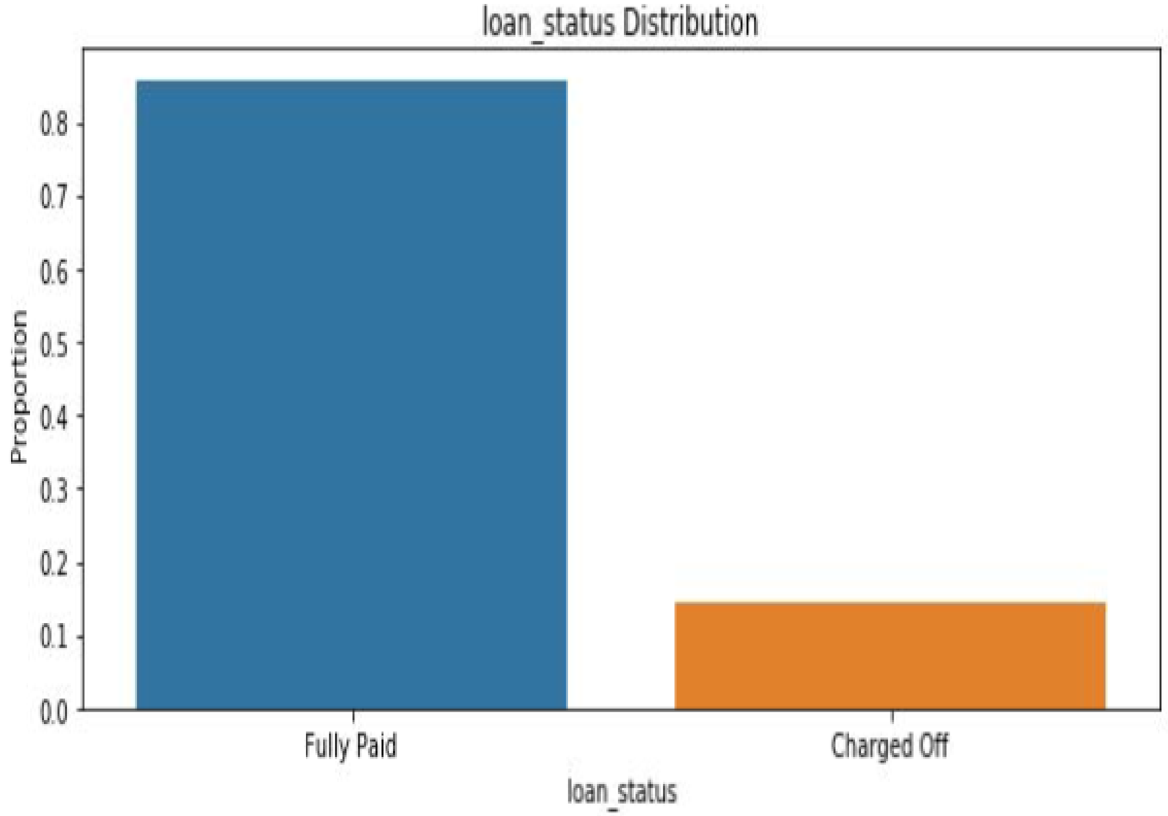




Lending Club (LC) has been doubling its loan issuance each year. Additionally, there is a spike in loan issues during the last three months of the year—October, November, and December.

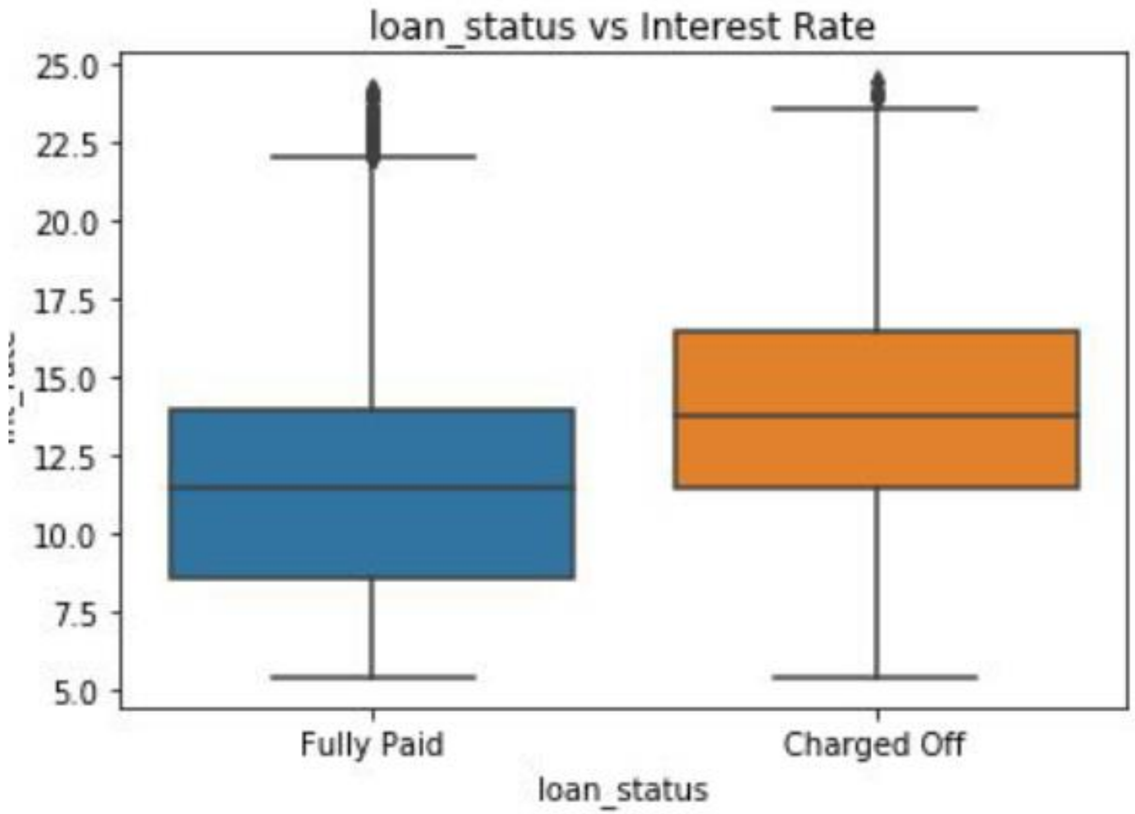
Analysis- Term, Loan Status





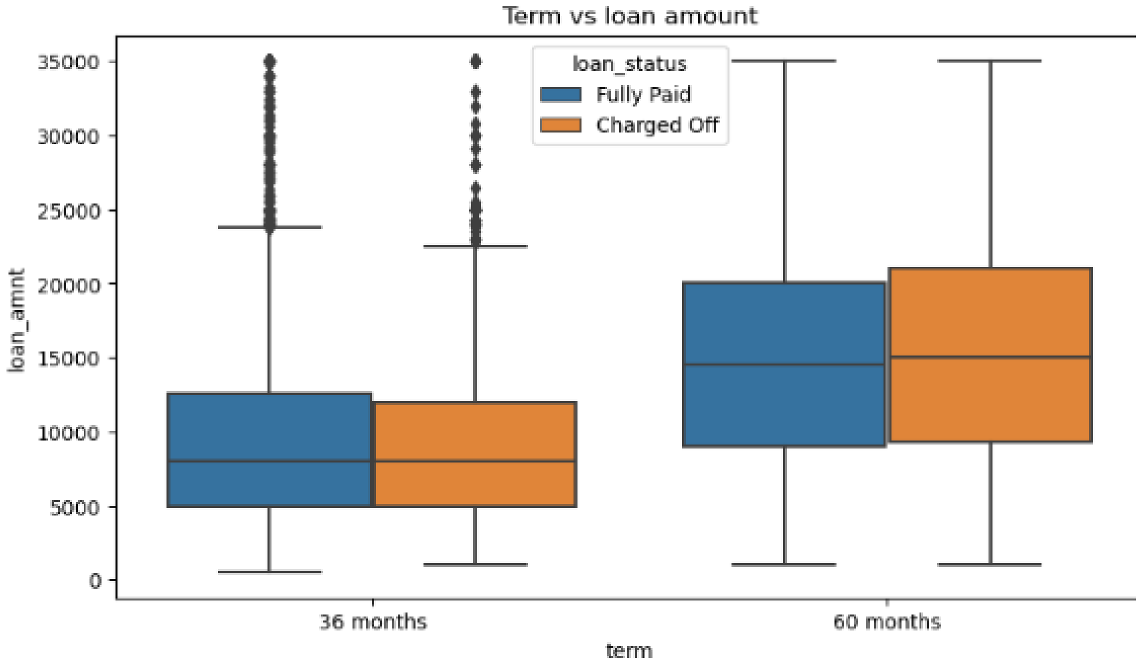
Around 75% borrowers took loans with 36 months term
The charged off borrowers are around 15% and fully paid is around 85%

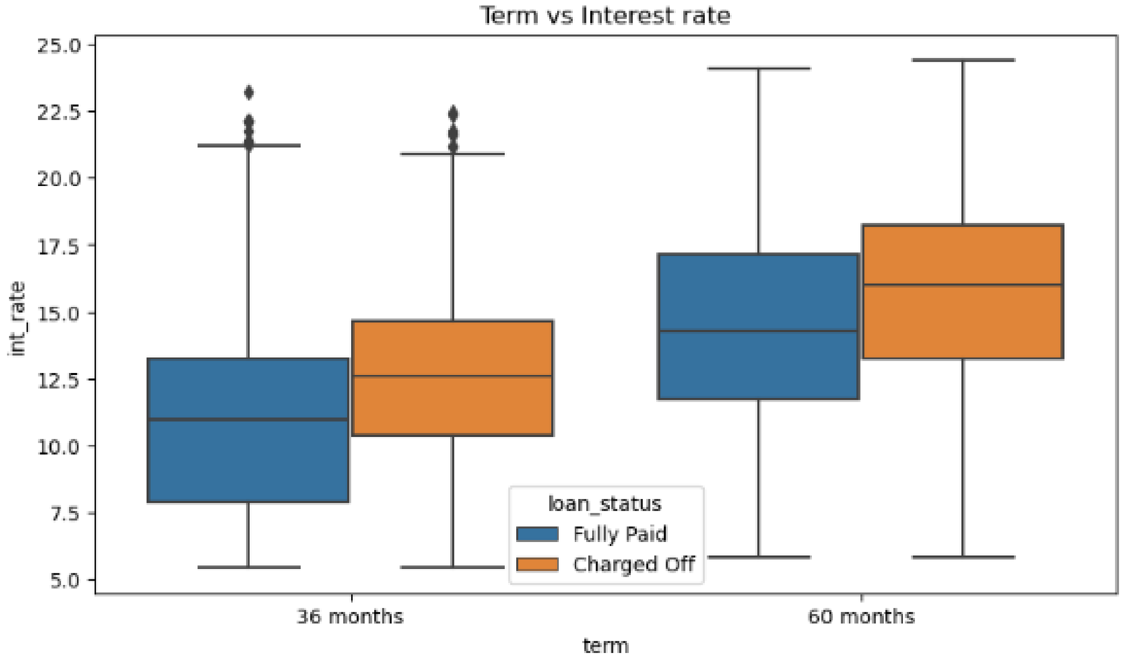
Analysis- Loan Status, Interest rate



Charged-off loans have higher amounts compared to fully paid loans.

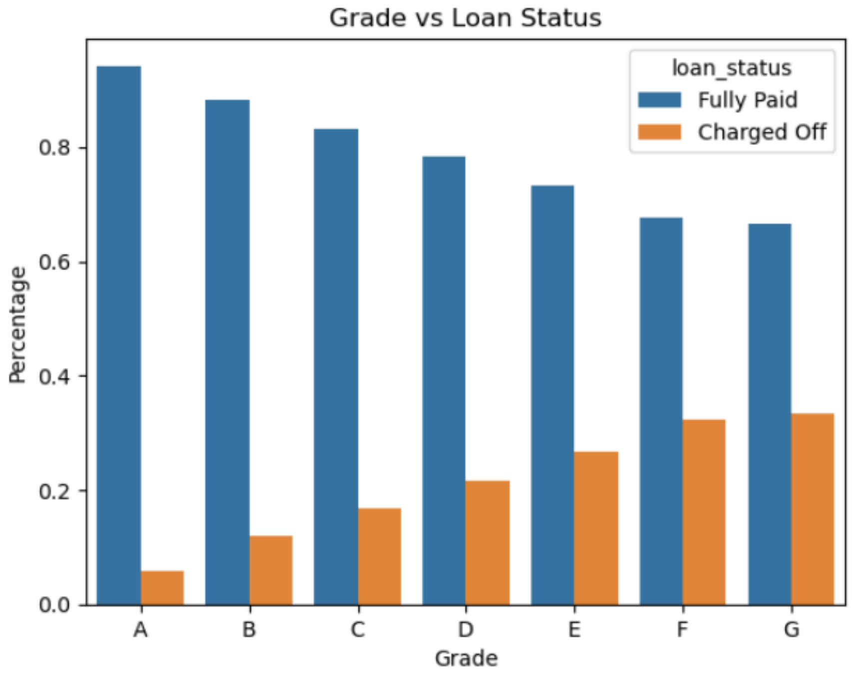
Analysis- Term, Loan amount, Interest rate

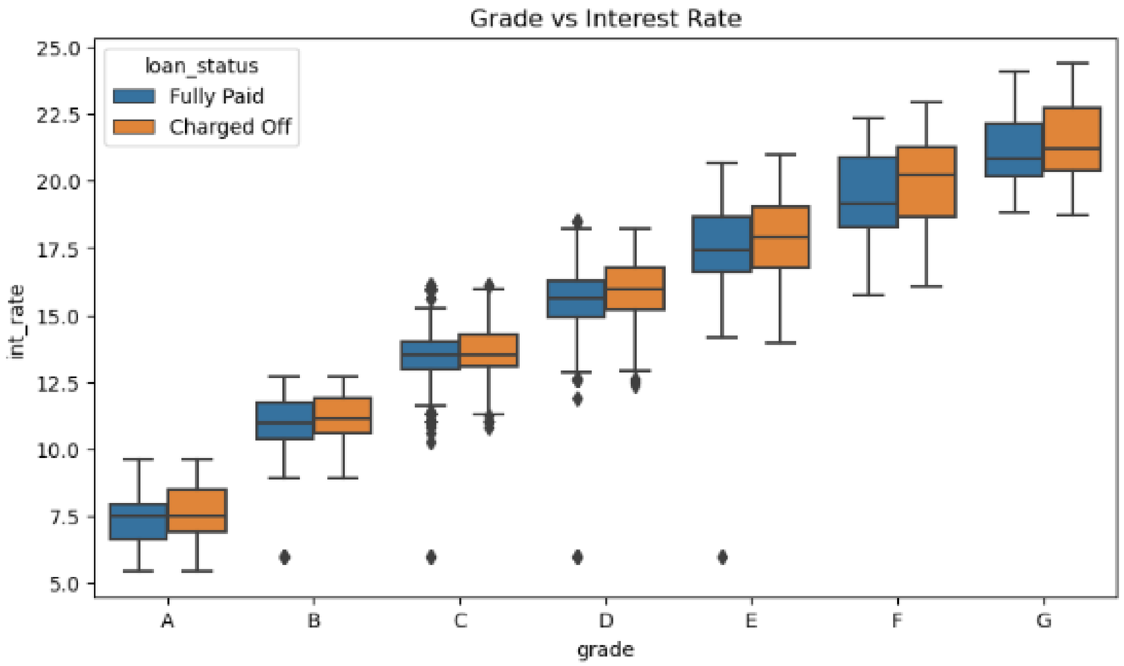




- Loan amount does not significantly influence defaults for either the 36-month or 60-month tenures. Borrowers are evenly distributed between default and non-default statuses across both loan tenures.
- Higher interest rates are associated with a higher default rate for both 36-month and 60-month tenures.

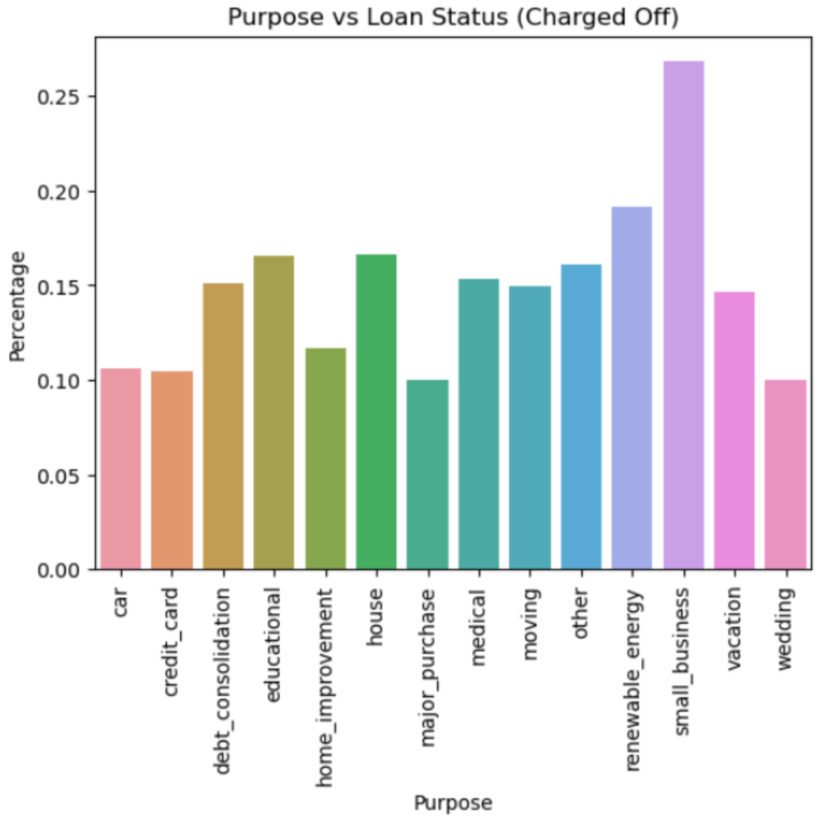
Analysis- Grade, Loan status, Interest rate





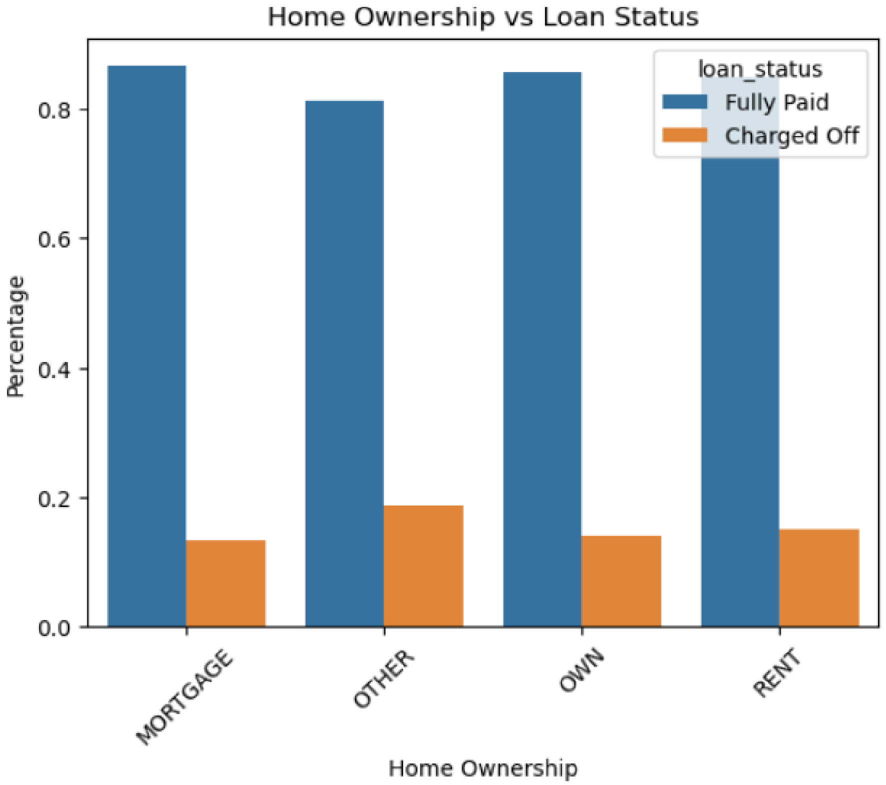
- The graph clearly indicates that the number of charged-off loans increases as credit grades decrease.
- As credit grades decrease, interest rates gradually increase, and borrowers with lower grades become more prone to defaulting on their loans.

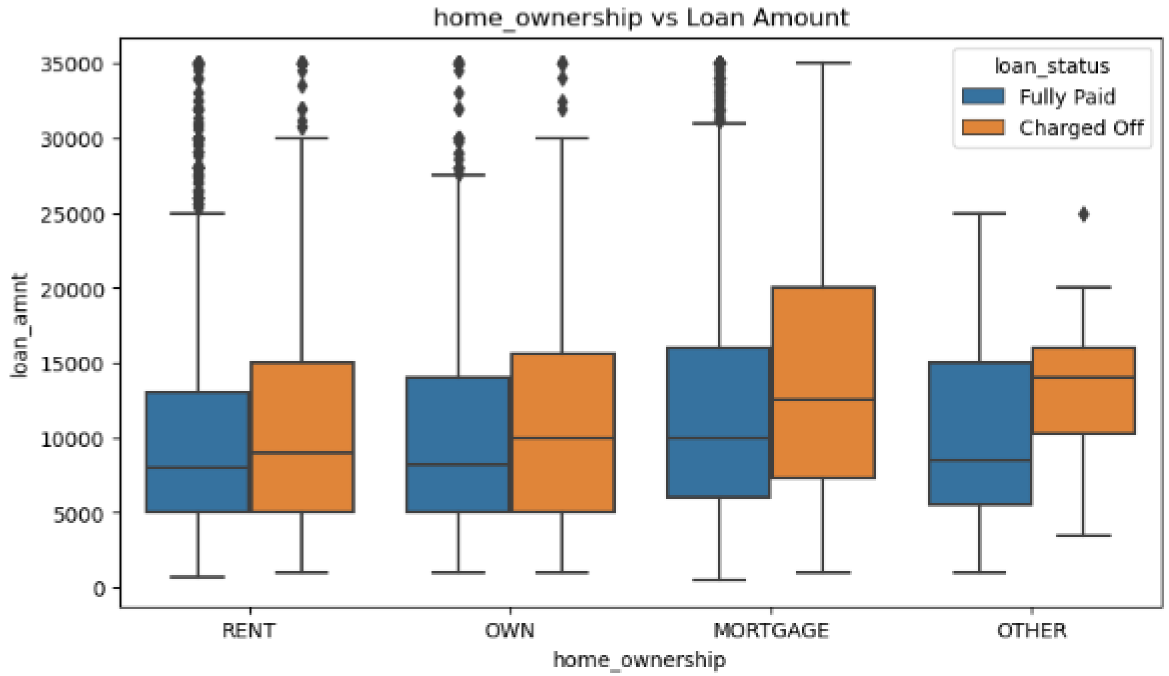
Analysis-Purpose vs Loan Status



- Small business purpose category have defaulted more

Analysis-Purpose vs Loan Status





- There is around 20% chance of loan default in each home ownership category.
- From the 2nd plot we can see the people with higher loan amounts in mortgage home ownership has high default rate than others

Insights from EDA

Key Insights:

Interest Rates and Defaults:

- Higher interest rates are strongly associated with higher default rates, irrespective of loan tenure and verification status.
- Borrowers with high revolving credit utilization face higher interest rates and tend to default more frequently.

Loan Amount and Defaults:

- Larger loan amounts, particularly in small business and home loan categories, are associated with higher default rates.
- High loan amounts in small business loans show a strong correlation with defaults.

Credit Grades and Defaults:

- Lower credit grades (F and G) are associated with higher default rates and larger loan amounts. Borrowers in these grades have higher public derogatory records.
- A-grade borrowers have fewer derogatory records and a lower default rate.

Home Ownership and DTI:

- Borrowers with mortgage ownership have higher incomes but also higher loan amounts, which correlates with higher default rates.
- "Other" home ownership categories have a lower DTI ratio, but an equal likelihood of default compared to other categories.

Insights from EDA contd..

Key Insights:

Public Records:

- Public derogatory records are highly correlated with public bankruptcy records, indicating that higher derogatory records are a strong predictor of default.
- Borrowers with 4 public derogatory records are predominantly in the mortgage ownership category and "Not Verified" status.

Verification Status:

- Verified loans tend to have higher loan amounts, and there is a slight increase in loan amounts for defaulted verified and source-verified loans.

Recommendations to Lending Club

Recommendations to reduce defaults:

Adjust Interest Rates:

- Implement a more nuanced interest rate model that considers credit grades, loan amounts, and public derogatory records to better assess risk.
- Consider capping interest rates for high-risk borrowers or adjusting rates based on their credit utilization and derogatory records.

Loan Amount Controls:

- Set stricter guidelines for loan amounts, especially for small business and high-risk categories. Consider limiting loan amounts for borrowers with lower credit grades or higher DTI ratios.

Enhance Credit Assessments:

- Strengthen credit assessments by incorporating more detailed analyses of public records and credit utilization.
- Implement additional checks and balances for borrowers with high public derogatory records.

Recommendations to Lending Club.

Refine Home Ownership Criteria:

- Adjust lending criteria for different home ownership categories based on observed default rates and DTI ratios.
- Consider offering tailored financial products that align with the risk profile of borrowers in different home ownership categories.

Improve Verification Processes:

- Increase focus on thorough verification processes to ensure that loans are given to verified borrowers, as they tend to have more reliable repayment behavior.
- Explore options for improving verification for categories with higher default rates, such as the "Not Verified" category.

Monitor and Analyze Trends:

- Continuously monitor default rates and interest rate impacts across different loan types, credit grades, and home ownership categories.