Lending Club Case Study Data Analysis

Sneh Valani

Business Understanding
Analyzing Loan Default Risk for Informed Lending Decisions

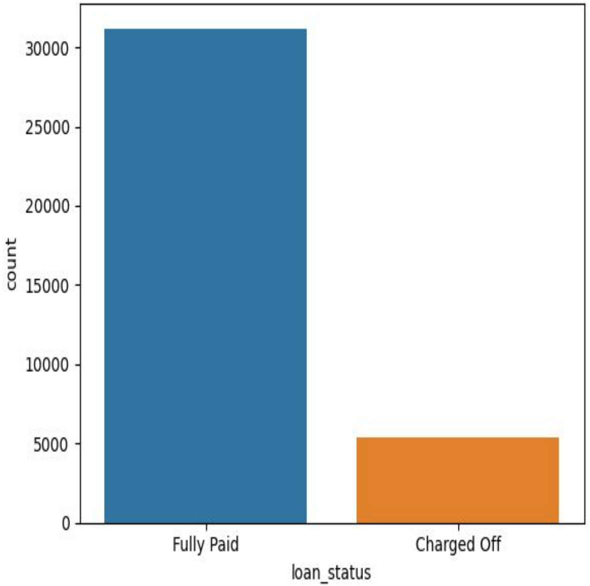
- Company Profile: You work for a consumer finance company specializing in lending various types of loans to urban customers.
- Loan Approval Decision: When evaluating loan applications, the company faces two types of risks:
- Approving a loan to someone likely to repay results in a loss of business if denied.
- Approving a loan to someone likely to default may lead to financial losses.
- Data Purpose: The provided data contains information about past loan applicants and their default status, aiming to identify patterns indicating the likelihood of default.
- EDA Objective: The case study's objective is to use Exploratory Data Analysis (EDA) to understand how consumer attributes and loan attributes influence the tendency to default.
- Loan Decision Outcomes: When a person applies for a loan, there are two possible decisions:
- Loan Accepted: This can result in three scenarios:
- Fully paid: The applicant has paid both the principal and interest rate as agreed.
- Current: The applicant is in the process of paying the installments and has not defaulted.
- Charged-off: The applicant has defaulted on the loan by not paying installments as agreed.
- Loan Rejected: The company rejected the loan application, and such data is not available in the dataset.

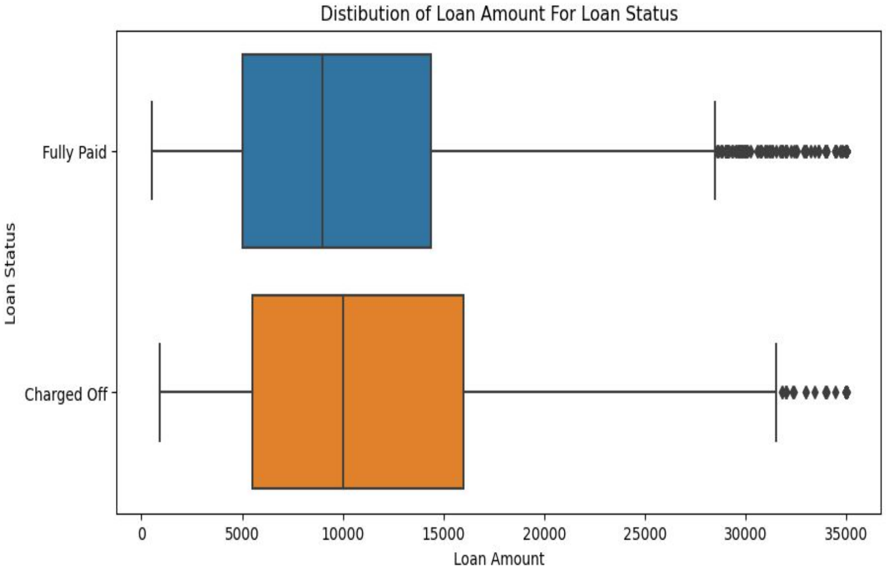
Objective: Identifying Risky Loan Applicants for Credit Loss Reduction

- ➔ The company operates as a prominent online loan marketplace, providing personal loans, business loans, and medical procedure financing.
- ➔ Credit loss, primarily caused by loan defaults, is a significant financial challenge for lending companies.
- ➔ The primary focus of this case study is to use Exploratory Data Analysis (EDA) to identify 'charged-off' customers, who represent loan defaulters, and understand the driving factors (driver variables) behind loan defaults.
- ➔ By pinpointing these critical factors, the company aims to improve its risk assessment and portfolio management, reducing credit losses.
- ➔ This objective aligns with the broader field of risk analytics, highlighting the importance of variables in managing and mitigating credit loss.

Examining Loan Default Rates: Univariate Insights

Relationship Between Loan Status and Loan Amount

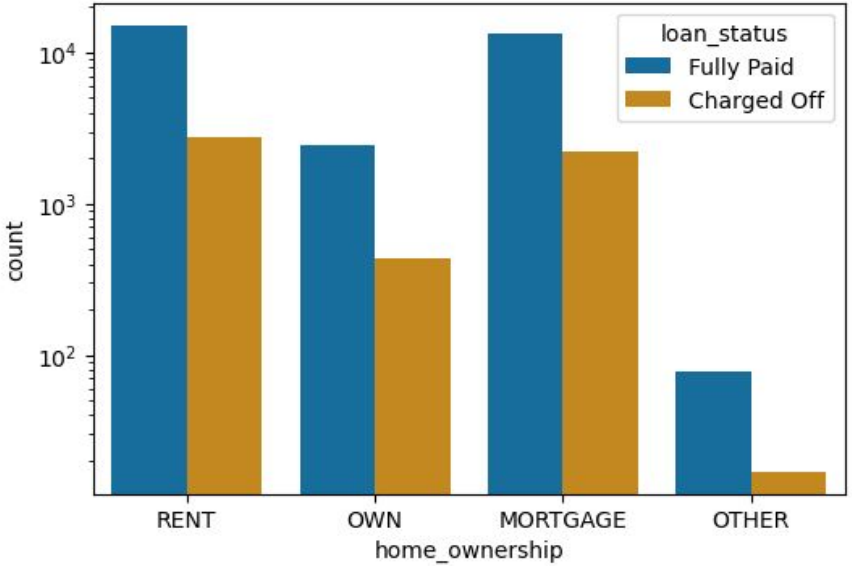


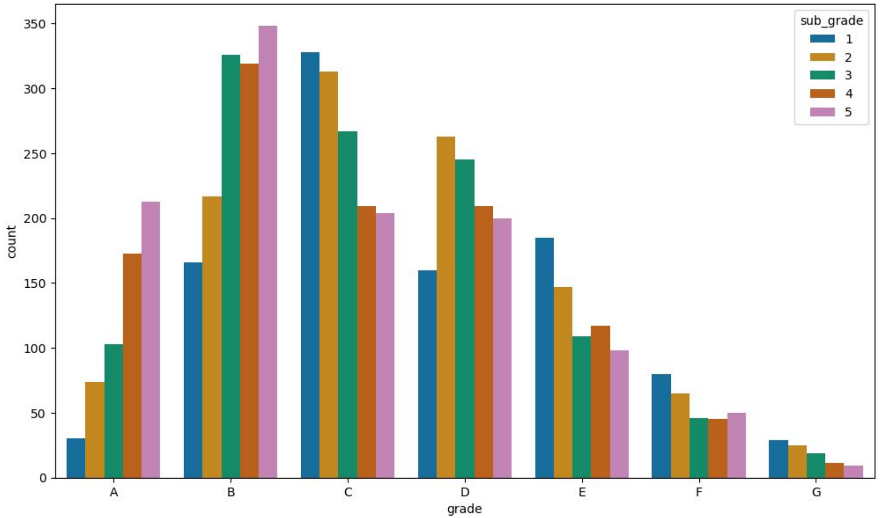


The dataset predominantly consists of fully paid loans, surpassing the number of charged-off loans, signifying that most borrowers fulfill their loan repayment commitments.

The box plot analysis reveals that borrowers who fully repay their loans tend to borrow larger amounts of money compared to those who default on their loans. Additionally, there is more variability in loan amounts for fully paid loans.

Analysis of the 'home_ownership' variable and Sub-Grade Variable

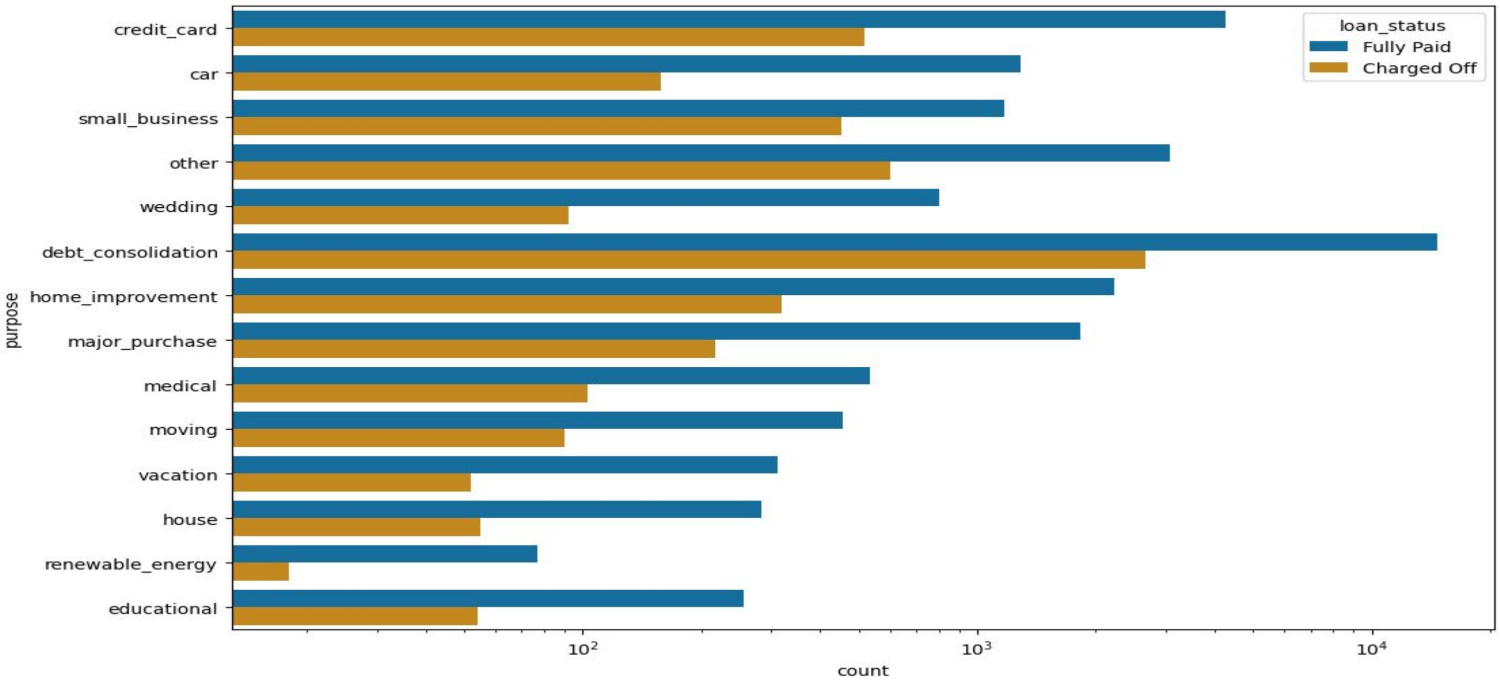




Borrowers with 'RENT' as their home ownership status are at a higher risk of loan default.

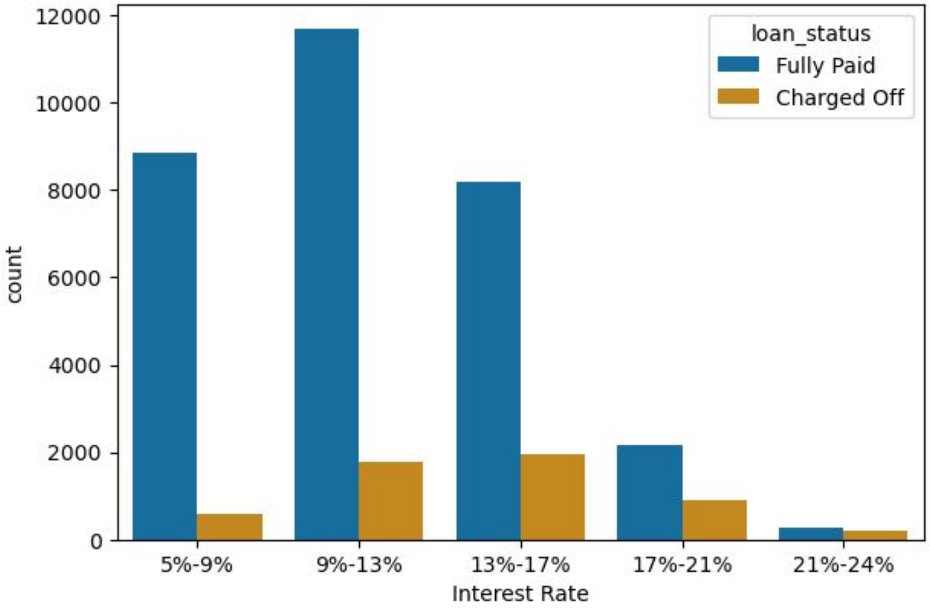
Defaults increase with loan grade. The steepest rise is between grades B and C, indicating higher default risk for grade B borrowers.
Default risk also increases for grades A, D, and E.

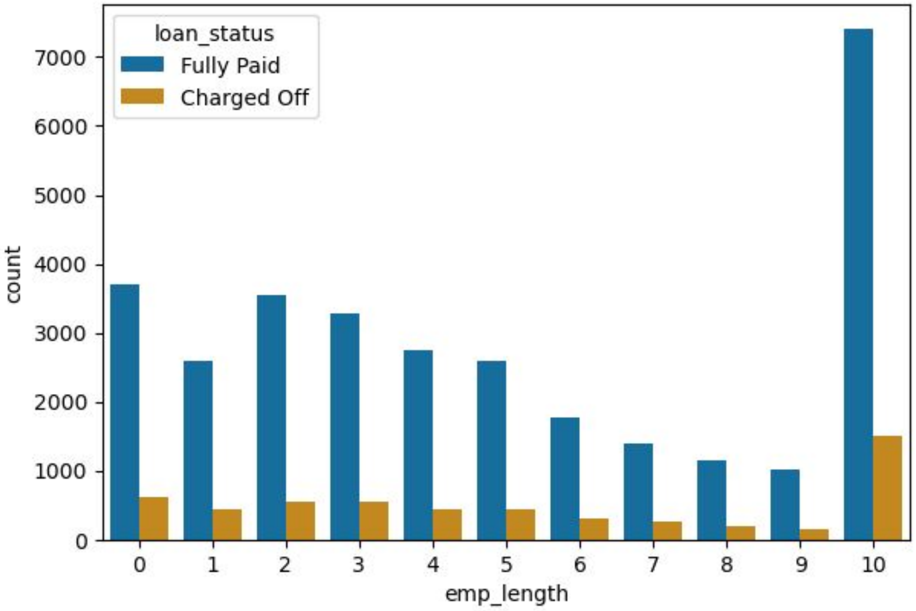
Exploring Loan Utilization and Default Rates



Borrowers frequently opt for loans to merge and manage their existing debts.
It's noteworthy that the default rate is notably high for loans earmarked for debt consolidation and small business ventures.

Loan Count Analysis: Impact of Employment Length and Interest Rates

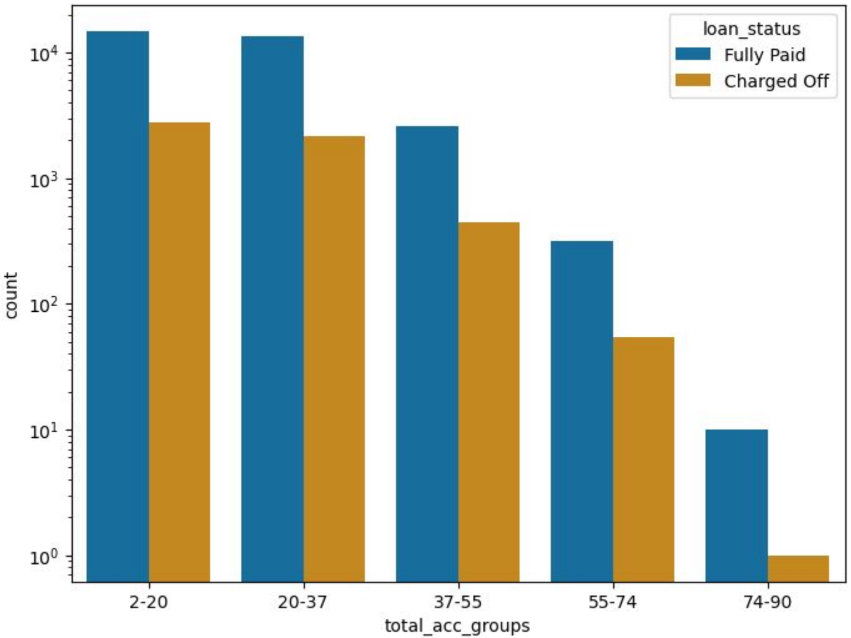


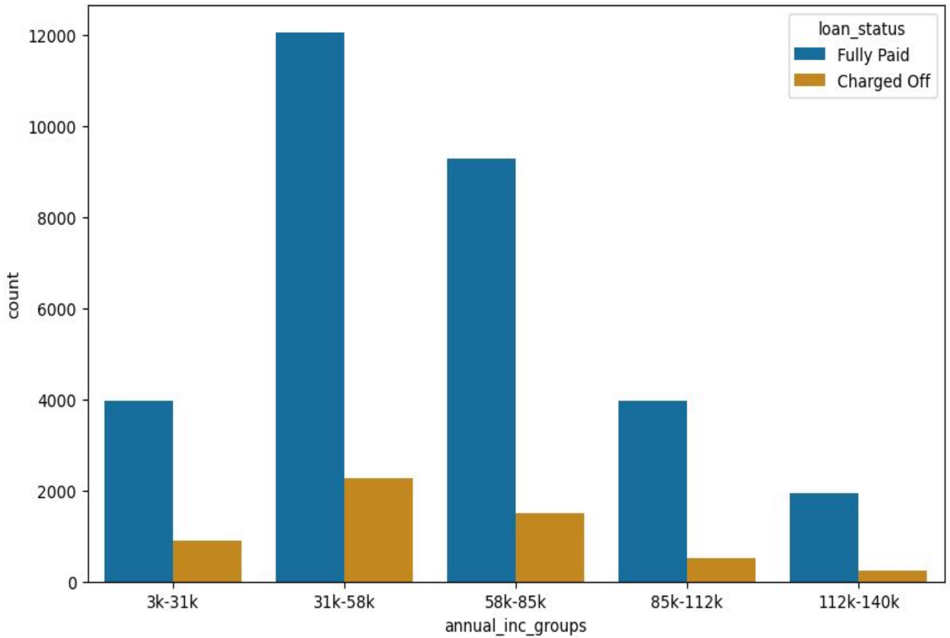


Loans with interest rates falling within the 13%-17% range exhibit a higher default rate.
Likewise, loans in the 9%-13% interest rate bracket also experience elevated default rates.

Borrowers with extensive employment experience (over 10 years) and those with minimal experience (less than 1 year) have a higher likelihood of loan default.

Analyzing Borrower Risk: Credit Lines, Income, and Loan Defaults

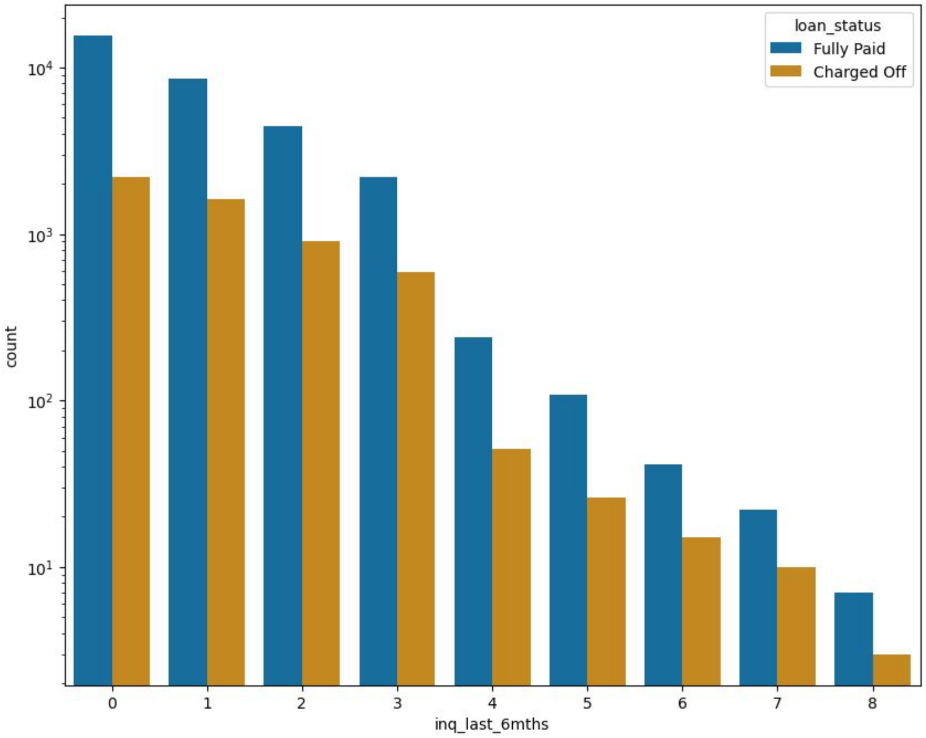




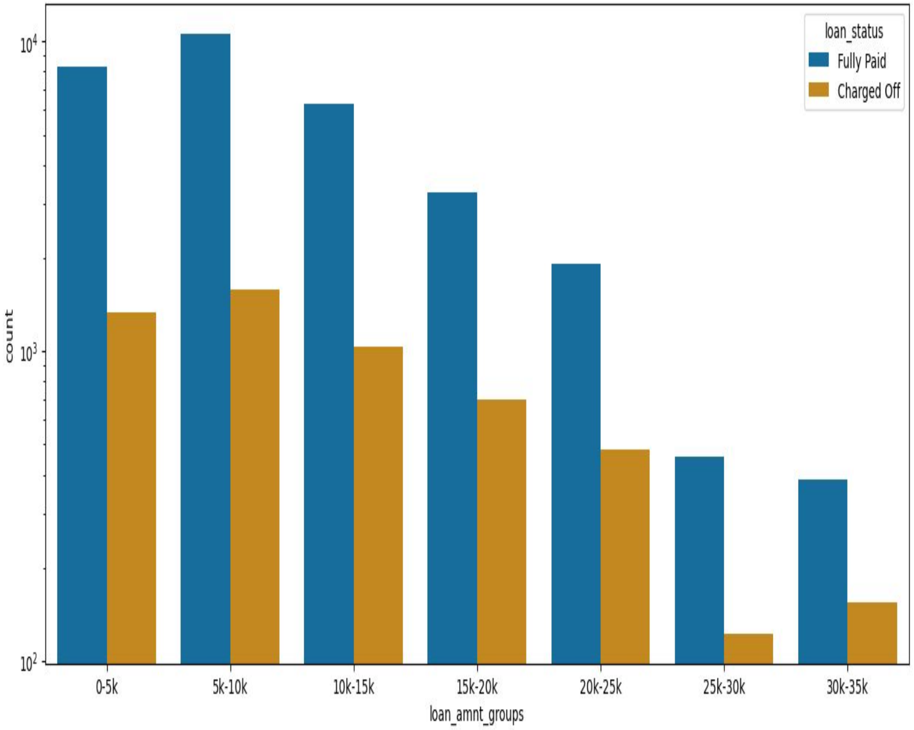
Borrowers with 2-20, 20-37, and 37-55 credit lines in their credit files display progressively higher default rates, with the highest risk observed in the latter group.

Borrowers with annual incomes in the 31k-58k and 58k-85k brackets display elevated default rates, with the lower income group, 31k-58k, having the highest default rate.

Analyzing Borrower Risk: Credit Inquiries, Loan Amount, and Default Rates



Borrowers who have made fewer than 4 credit inquiries within the last 6 months show a notably higher rate of loan defaults.



Borrowers applying for loans in the range of 0-25k exhibit a higher default rate compared to those seeking loan amounts exceeding 25k. Notably, the highest default rate is observed among borrowers applying for loans in the 5k-10k range.

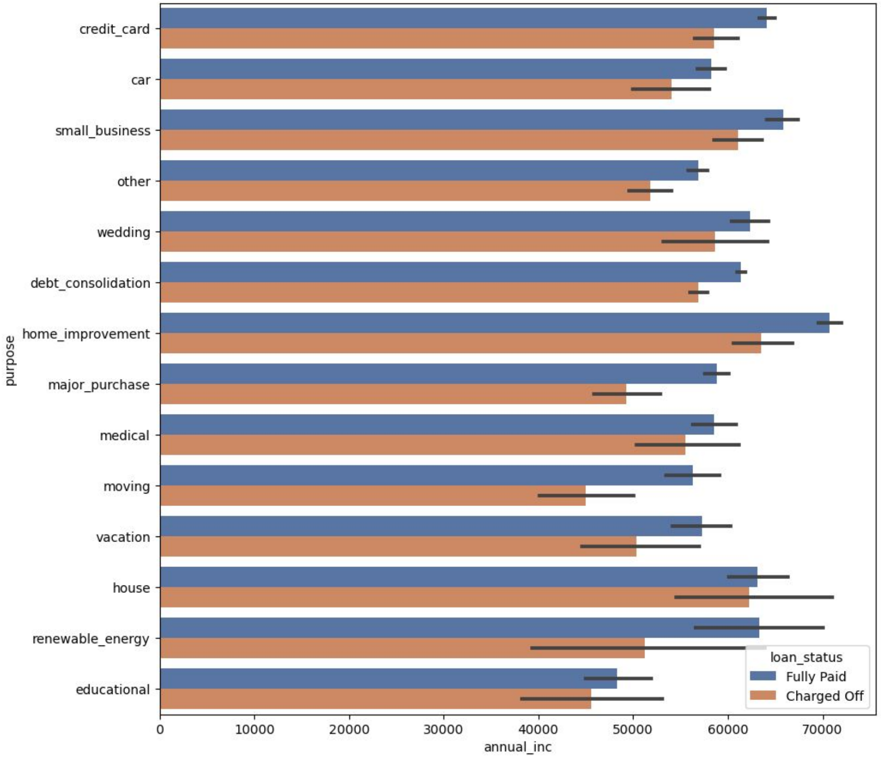
Key Observations on Loan Default in Univariate Analysis

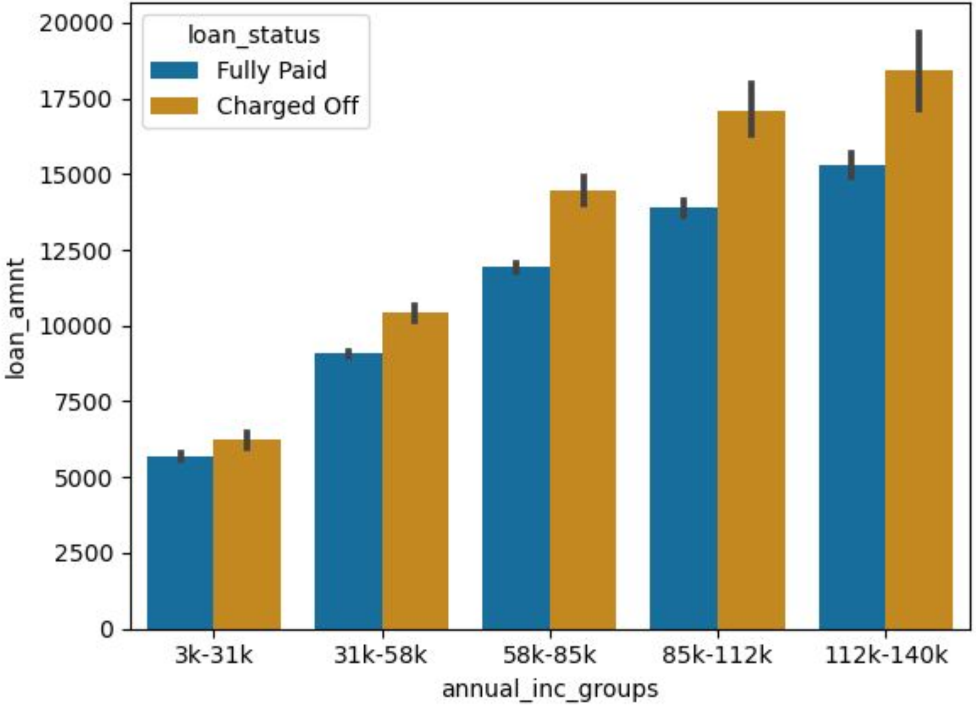
The loan analysis highlights an elevated risk of default associated with specific borrower characteristics and conditions, including

- Homeownership status as 'RENT.'
- Loan purpose for debt consolidation.
- Interest rates in the 13-17% range.
- Annual incomes ranging from 31,201 - 58,402.
- Having 20-37 open accounts.
- Employment history of 10 years.
- Investment-funded amounts between 5,000 and 10,000.
- Loan amounts between 5,429 and 10,357.
- Debt-to-income ratios (DTI) within 12-18.
- Monthly installments between 145-274.
- Loan terms set at 36 months.
- Loan status categorized as 'Not Verified.'
- Absence of recent inquiries in the last 6 months.
- No derogatory public records.
- Loan purpose classified as 'debt consolidation.'
- Loan grade of 'B.'
- A specific subgrade level of 'B5.'
- Loans issued in later months of the year have a higher default risk, adding a temporal dimension to the analysis.

Examining Loan Default Rates: Bivariate Insights

Analyzing Annual Income in Relation to loan purpose and

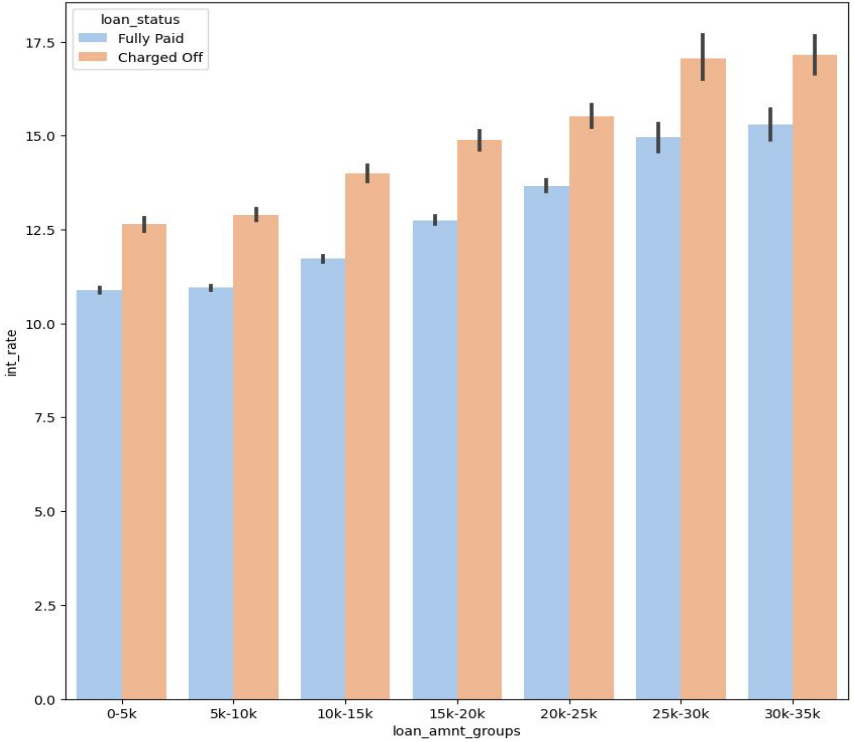


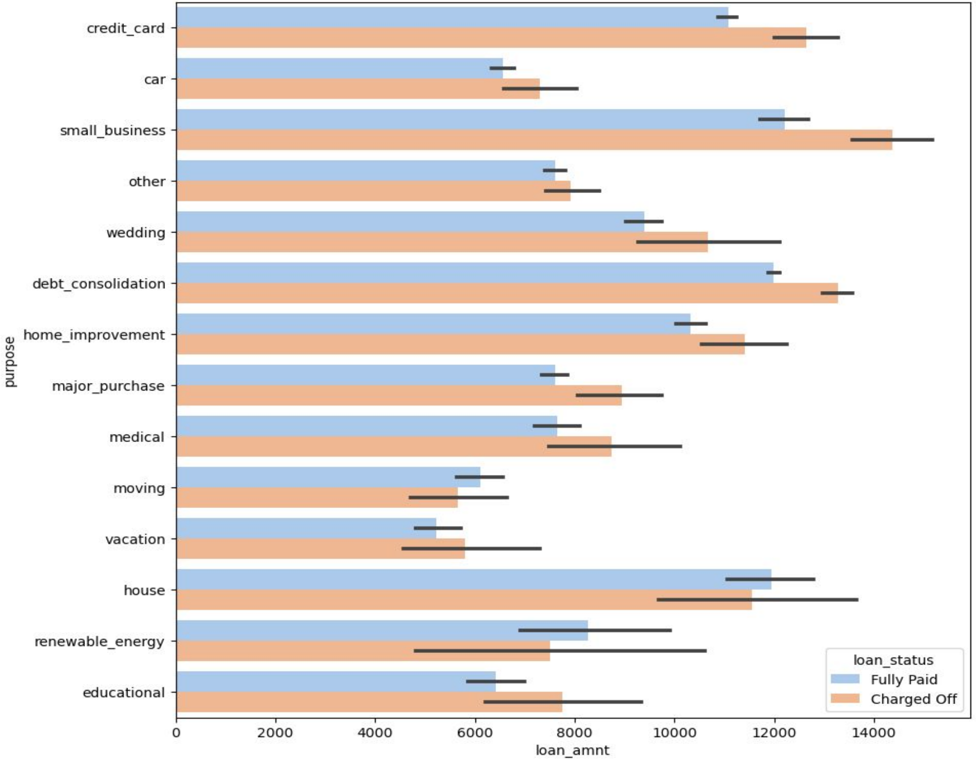


While 'debt consolidation' experiences the highest number of loan applications and defaults, it's not linked to the highest annual incomes. Borrowers with higher salaries gravitate towards 'home improvement,' 'house,' 'renewable energy,' and 'small business' loan applications.

Irrespective of their income levels, it's noteworthy that borrowers who defaulted on loans consistently held higher loan amounts. This trend underscores that loan amount, rather than income, is a significant contributing factor to loan defaults.

Exploring Loan Amount in Relation to interest rate and purposes of loan

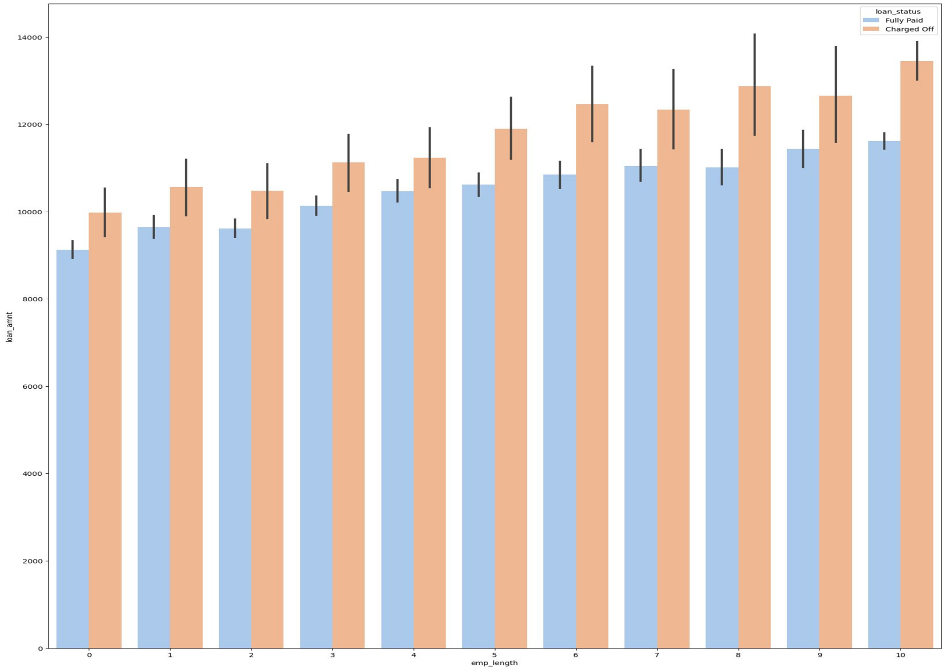




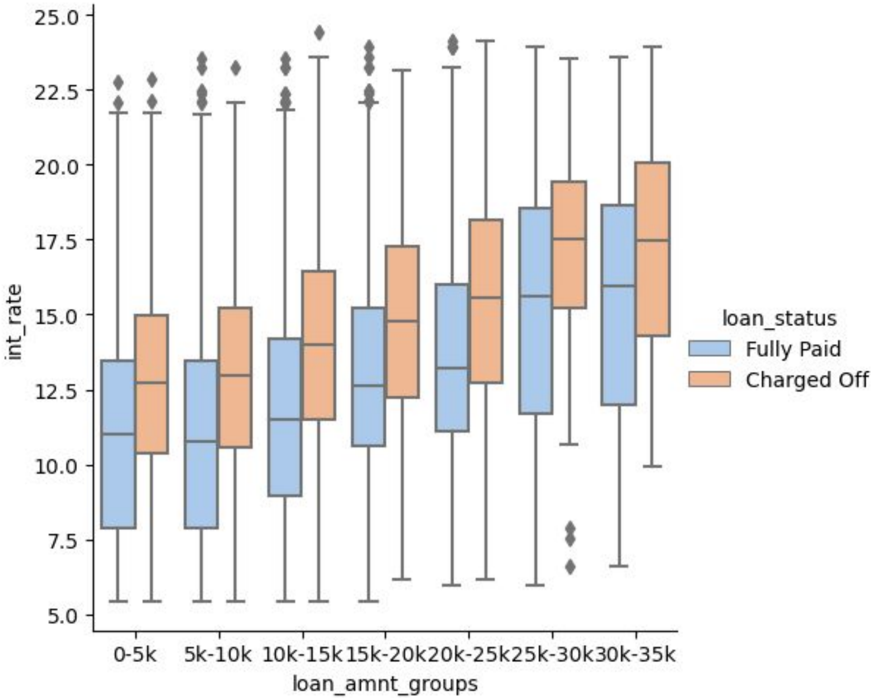
Higher interest rates correspond to a decrease in fully paid loans and an increase in charged-off loans across all loan amounts. Moreover, larger loan amounts are linked to a similar trend, with a higher likelihood of charge-offs and fewer loans being fully paid.

When considering loans for various purposes, it's evident that those allocated for small businesses, debt consolidation, and home improvement especially when obtained in larger amounts, exhibit a higher incidence of defaults as compared to loans with different purposes

Exploring Loan Amount in Relation to Employment length and Interest Rate of loan



It's observed that applicants with a more extended employment history are more likely to secure higher loan approval amounts. This indicates a positive correlation between employment tenure and loan approval, with financial institutions often favoring individuals with established work histories when considering loan amounts.



Across various loan amount categories, it's evident that interest rates for charged-off loans consistently outstrip those for fully paid loans. This marked interest rate differential is a crucial factor contributing to loan defaults. Higher interest rates place added financial burden on borrowers, making repayment more challenging and elevating the risk of loan default.

Key Observations on Loan Default in Bivariate Analysis

In our bivariate analysis of loans, we identify specific conditions that are strongly associated with loan outcomes

- Borrowers opting for 'home improvement' loans with incomes in the 60k - 70k range.
- Applicants with 'MORTGAGE' as their home ownership status and incomes of 60k - 70k.
- Borrowers facing interest rates ranging from 21-24% with incomes of 70k - 80k.
- Borrowers obtaining loans in the 30k - 35k range with interest rates between 15-17.5%.
- Applicants seeking loans for small businesses with amounts exceeding 14k.
- Applicants with 'MORTGAGE' as their home ownership status and loan amounts in the 14k - 16k range.
- Loans graded as 'F' within the 15k - 20k loan amount bracket.
- Applicants with 10 years of employment history and loan amounts between 12k - 14k.
- Verified loans exceeding 16k in amount.
- Loans graded as 'G' with interest rates exceeding 20%.