Identifying Risky Loan Applicants: An Exploratory Data Analysis

A Case Study on Predicting Loan Defaults

Introduction to Loan Risk Analysis
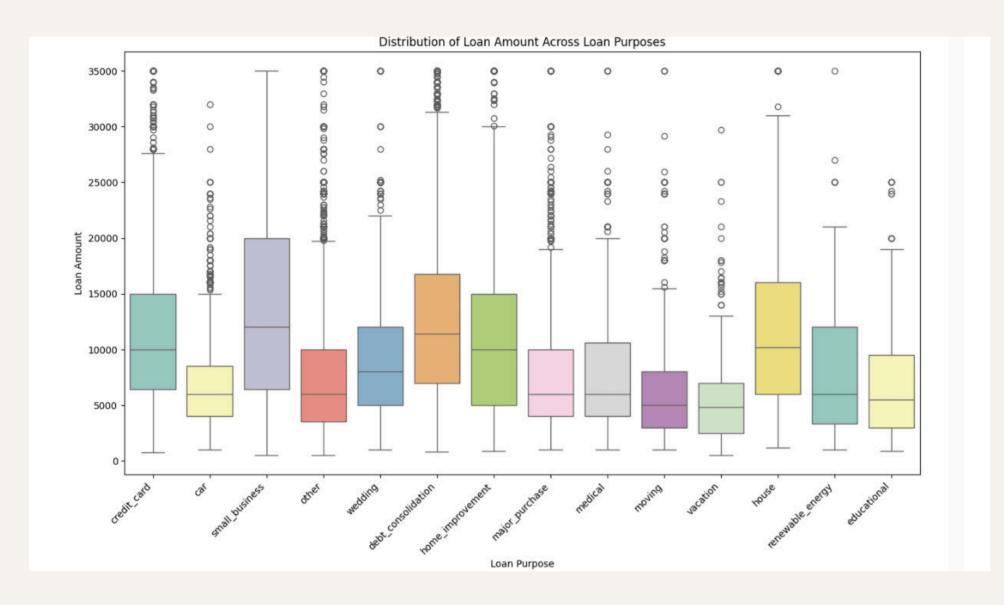

Exploratory Data Analysis (EDA) is crucial for understanding the **risk factors** associated with loan applicants. This presentation will delve into various **applicant profiles**, highlighting patterns and trends that can inform lending decisions and risk management strategies.

Dataset Overview

- **Description:**
  Dataset Size: The dataset consists of 39717 loan records, each representing an individual loan application.
  Time Period: The data covers loans issued from 2007 to 2010.
- **Key Variables**: Loan Amount, Interest Rate, Term, Employment Length, Home Ownership, Annual Income, Loan Status
- **Target Variable**: Loan Status: This is the primary variable used to determine whether a loan was successful or ended in default. A 'Charged Off' status indicates that the borrower failed to repay the loan, making it a critical indicator of default. This status serves as the dependent variable in our analysis to identify patterns associated with loan defaults.

Key Findings - Financial Insights

Distribution of Loan Amount Across Loan Purposes

This analysis with respect to the charged off loans for each variable suggests the following. There is a more probability of defaulting when :

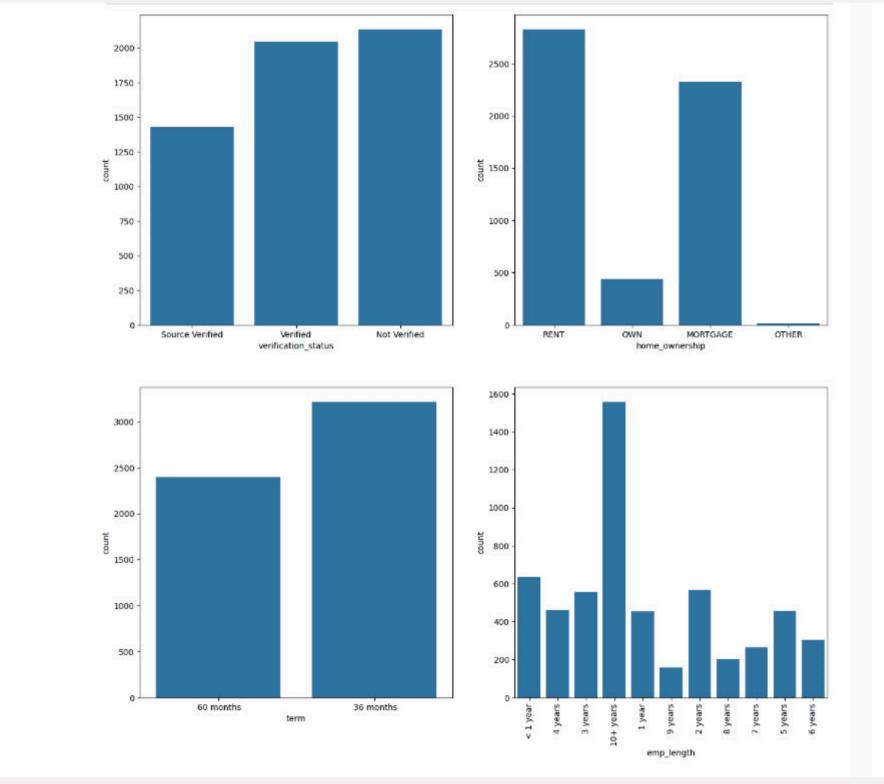- Applicants taking loan for 'home improvement' and have income of 60k -70k

Key Findings - Financial Insights

This analysis with respect to the charged off loans for each variable suggests the following. There is a more probability of defaulting when :

- Applicants having house_ownership as 'RENT'
- Applicants with employement length of 10
- When the loan status is Not verified
- Term of 36 months

Conclusion

- Applicants having house_ownership as 'RENT',employement length greater than 10, loan status which is Not verified, are strongly correlated with an increased likelihood of loan default.
- Employment length plays a significant role in default risk, with shorter employment lengths correlating with higher default rates.
- Other Reasons can be as below
  - Applicants who have taken a loan for small business and the loan amount is greater than 14k
  - Applicants taking loan for 'home improvement' and have income of 60k-70k
  -For grade G and interest rate above 20%

Thanks!

Contributors to the Case Study: Omkar Satapathy and Shahbaaz Shaikh.