Lending Case Study

Ravi P & Revanth NKPS

Loan Risk Assessment Case Study

Analyzing loan risk based on applicant income and loan purpose

Loan risk assessment evaluates a borrower's likelihood of repaying a loan by analyzing factors like credit history and financial stability. It helps lenders minimize potential losses and set appropriate loan terms. Modern tools like AI and data analytics enhance the precision of this process.

Assess the risk associated with different loan applications based on income levels and loan purposes

The case study aims to help understand which combinations of income levels and loan purposes have higher or lower risk.

Introduction

Technologies used

- Python library - version 3.7
- NumPy
- Pandas
- Seaborn
- Matplotlib

Data OverView

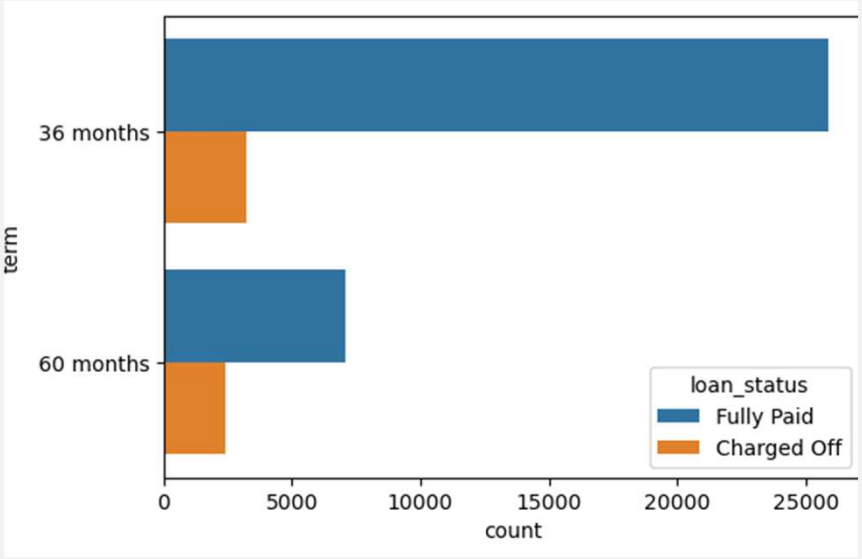
- Dataset contains 39,717 records and 111 columns.
- Types of data: Loan details, borrower information, and loan status.
- Columns with more than 10,000 missing values were removed
- The dataset was reduced from 111 to 53 columns by removing these columns.
- Dropped columns with unique values and those not relevant to loan approval and the dataset was further reduced to 39717 rows and 28 columns.
- Loan status 'Current' was removed as it does not help in decision-making.
- Added new columns for issue year and month by splitting the issue date.

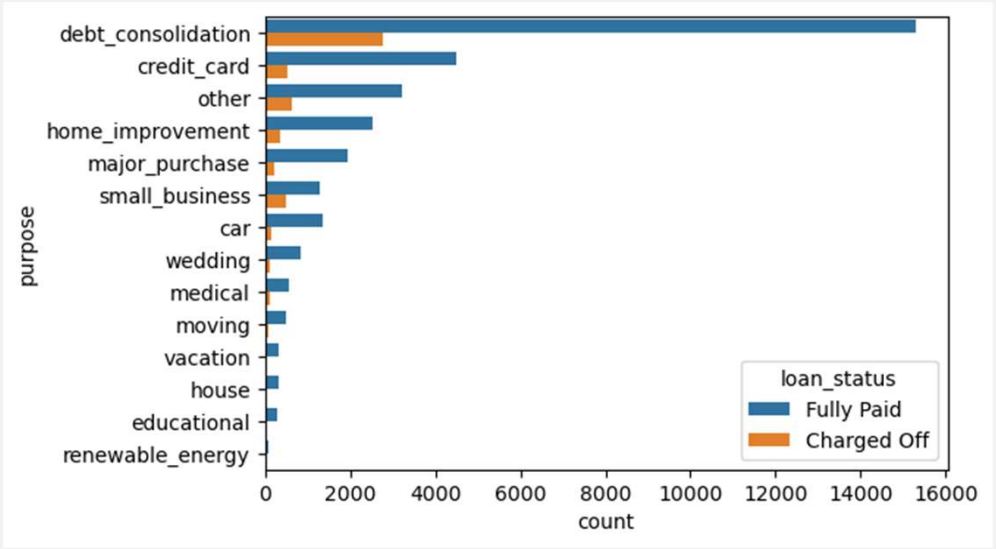
Data Analysis

- The analysis aims to identify the risk levels associated with different combinations of loan purpose and income:
- Medium-income applicants for small business loans are high risk.
- High-income applicants for car, credit card, major purchases, home improvement, vacation, and wedding loans are lower risk.

Key insights

- Loans for small businesses by medium-income applicants show higher risk.
- Loans for high-income applicants for purposes like car purchase or wedding show lower risk.

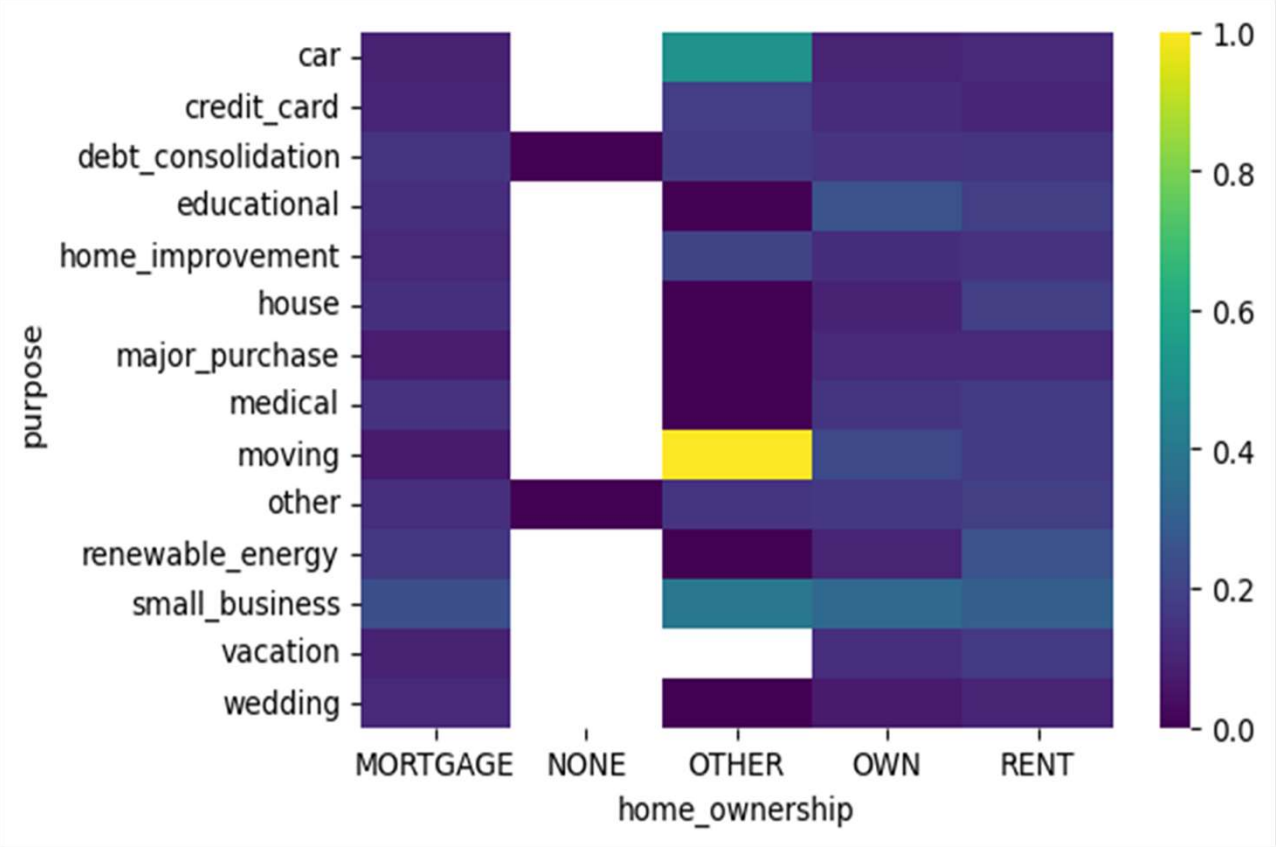




Univariate Analysis

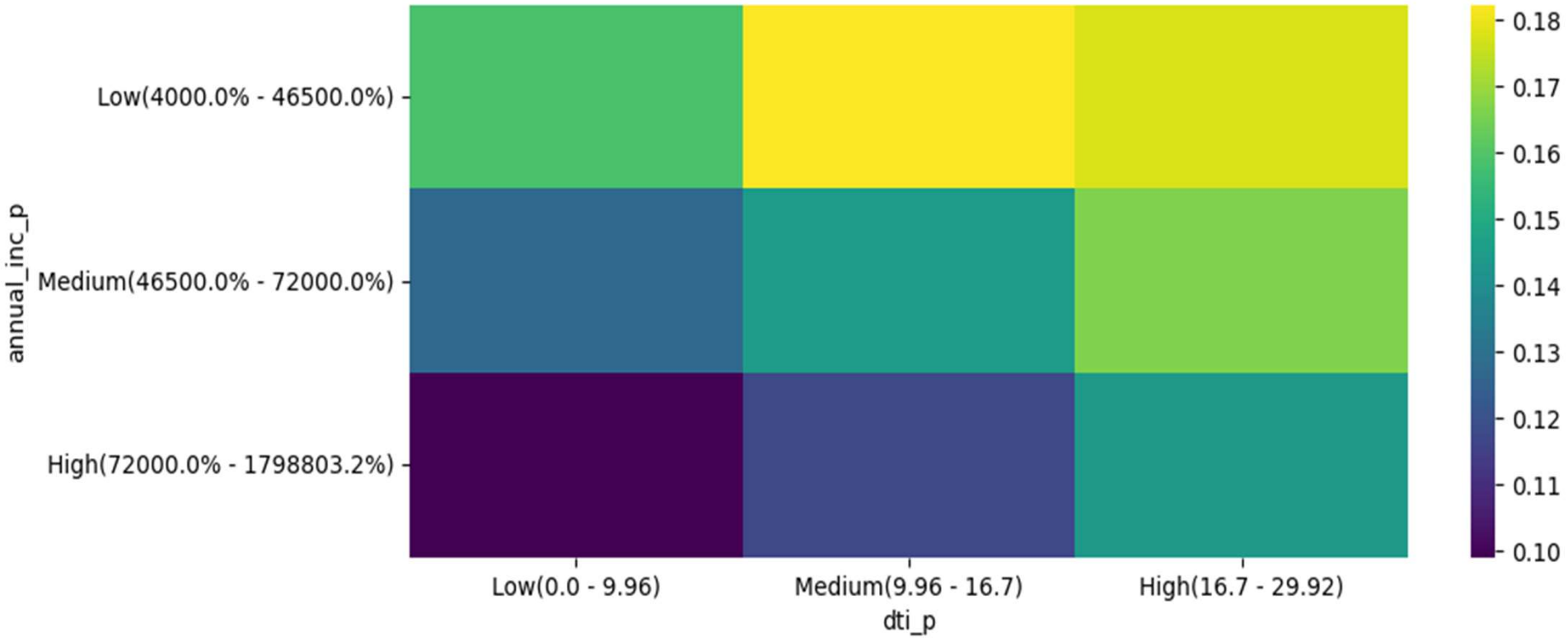
- From the bar plot graph we can conclude that the ratio of charged-off is higher if the loan tenure is of 60months.
- Approving loans to small business is turning out to be more risk.

Bi-variate Analysis

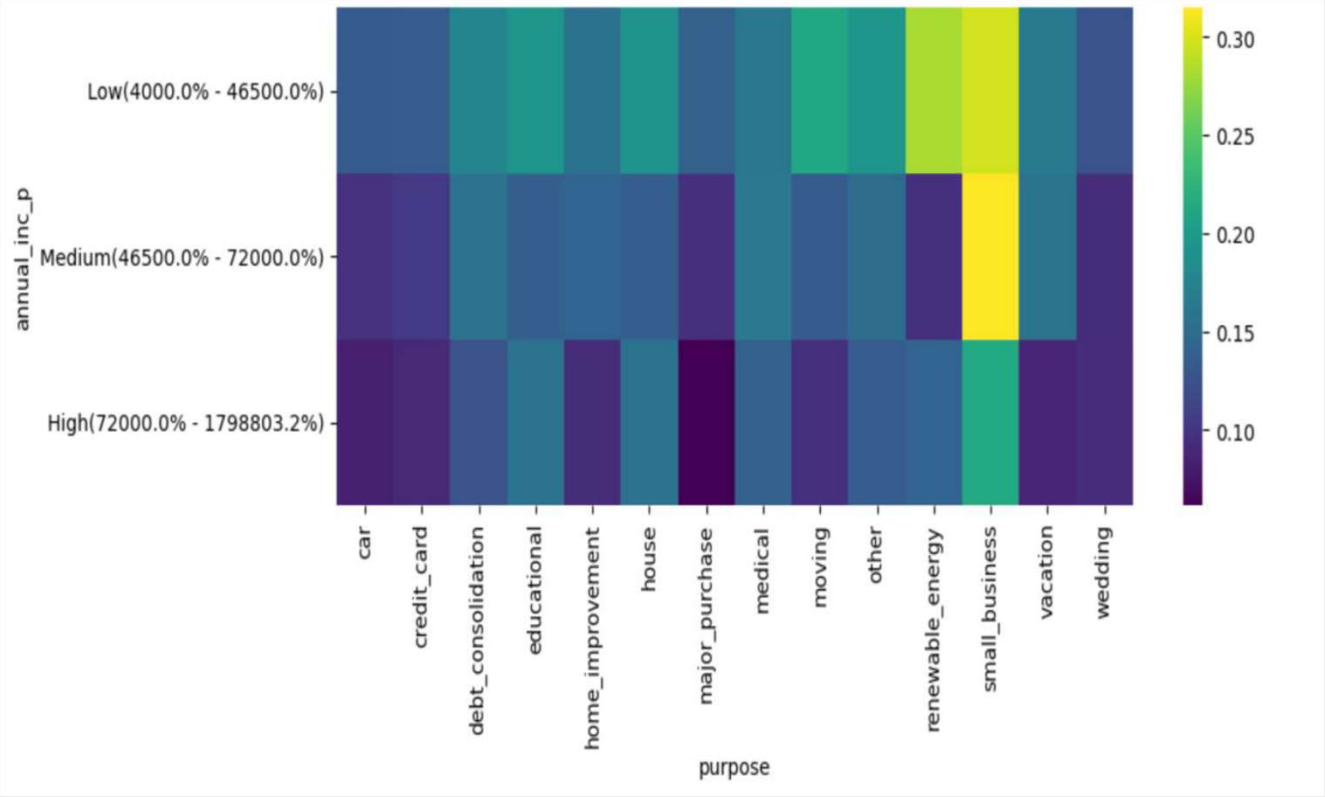


- The intersection of home ownership (None) and purpose (Other) has a darker color and represents riskier
- The intersection of home ownership (None) and purpose (debt consolidation) has the darker color representing riskier
- OTHER with small business ,moving, car purpose is risker

Bi-variate Analysis : Medium to High debt to income vs lower the annual income risk is higher



Bi-variate Analysis



- Applicant with medium income applying loan for small business purpose is very risky.
- Applicant with high income applying loan for car, credit_card, major_purchase, home_improvement, vacation, wedding purpose less risky.

Conclusion

- Through Exploratory Data Analysis (EDA) of the provided loan dataset, we were able to identify key characteristics of customers who are more likely to repay their loans, such as stable income, lower debt-to-income ratios, and positive credit histories.

- These insights can help the company refine its loan approval process, reducing the risk of approving high-risk loans and improving overall profitability.

- Implementing data-driven decision-making will allow the company to make more informed choices, ultimately minimizing financial losses from charged-off loans.

ThankYou

Ravi P
Revanth NKPS
ravipanavirs@gmail.com and revanth.nkps@gmail.com