Lending Club Case Study

Mayank Sharma & Mathavan Ramasamy

Objective

The Objective of this case study is to implement EDA technique on a real world problem and understand the insights and present in a business first manner via presentation.

Advantages of the case study:

- ➤ Showcases how EDA is applied to real-life business problems
- > Besides, it cultivates a fundamental comprehension of risk analytics in banking and financial services.
- > What steps are taken to minimize the loss of money when lending it to clients using data.
- > We gain better understanding of visualization and best charts to use for real life data.

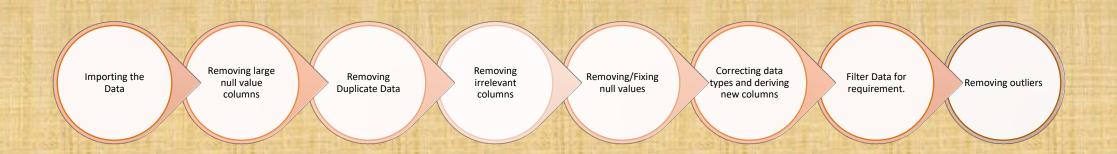
Business Insights

The business goal is to make a decision upon receiving a loan application, determining whether to approve or reject it based on specific variables.

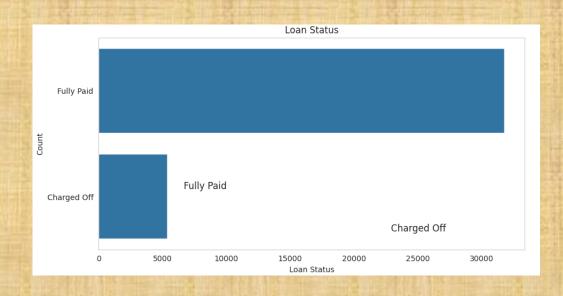
Dataset Details:

The provided data contains information about past loan applicants and whether they defaulted. The data focuses only on approved loans, excluding rejected ones, and categorizes the loans into three statuses: Fully Paid, Current, and Charged-Off.

Data Clean-up and preparation process:



Loan Status and Amount

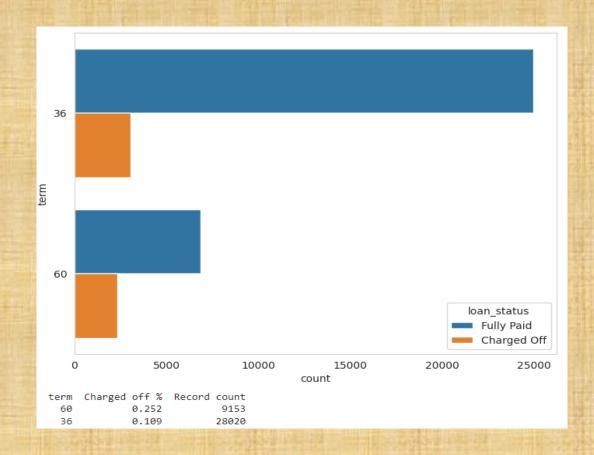


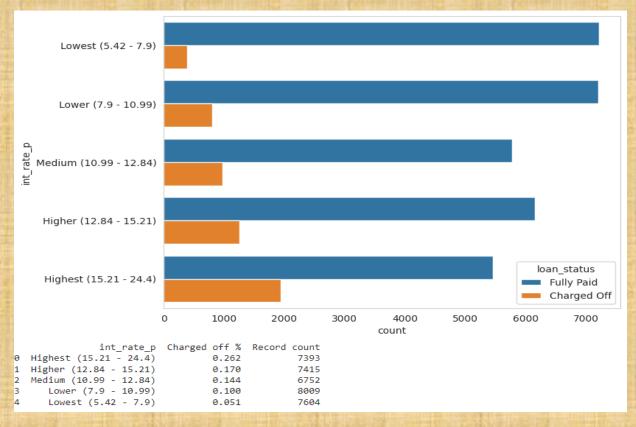
Loan Status: The number of charged off loan is much smaller(14.5%) compared to total count.



Loan Amount: It varies from 500 to 35000 with a median of 10000. The greater the loan amount, the higher the chance of the loan defaulting

Term and Interest Rate

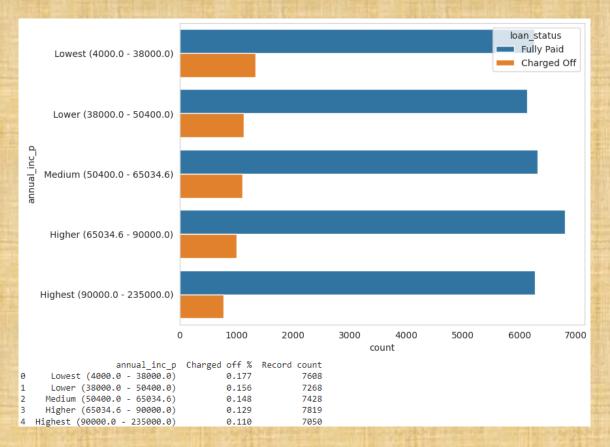


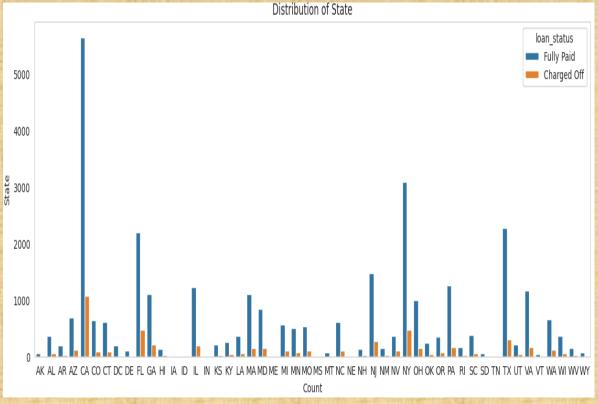


Loan Term: The length of the loan repayment term is a factor that influences the default rate

•Interest Rate: A higher interest rate results in a higher rate of defaulters.

Income and Location

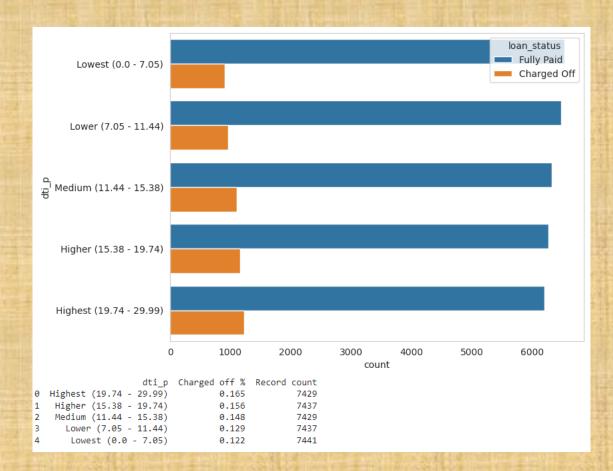




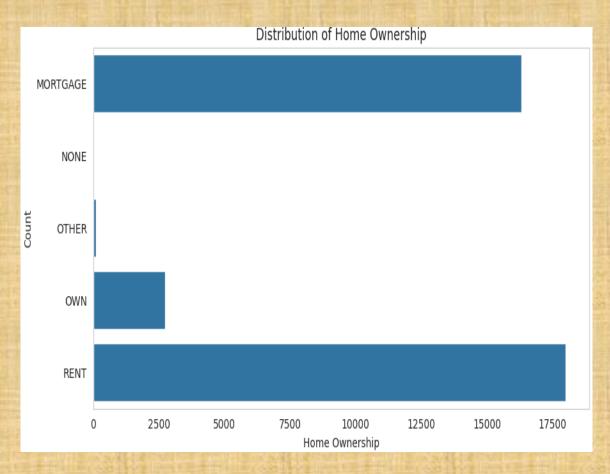
•Income: The greater the income, the greater the repayment percentage

 Location: The majority of borrowers are from large urban cities

DTI and Homeownership

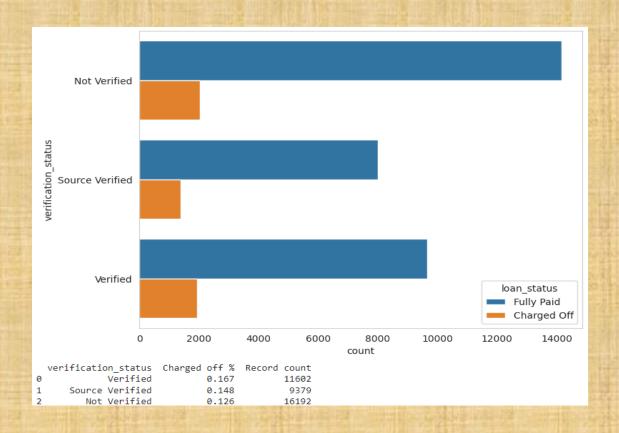


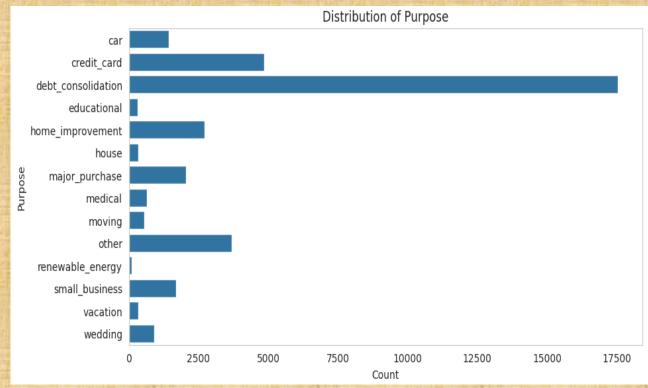
Debt-To-Income: A higher debt-to-income ratio (DTI) will lead to a higher percentage of defaulted loan.



Home Ownership: Majority of clients are lacking ownership of any property and are on rent or mortgage and have a higher chance of defaulting.

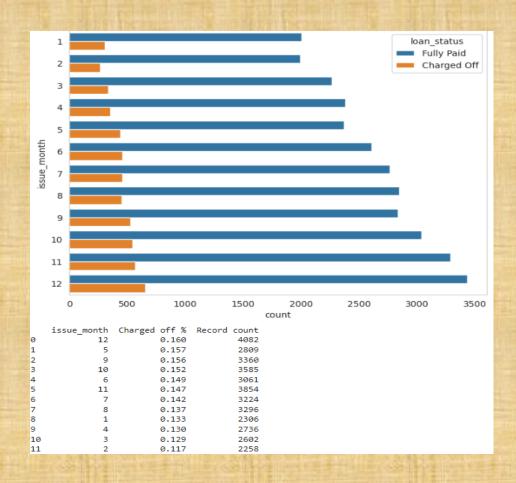
Verification & Purpose



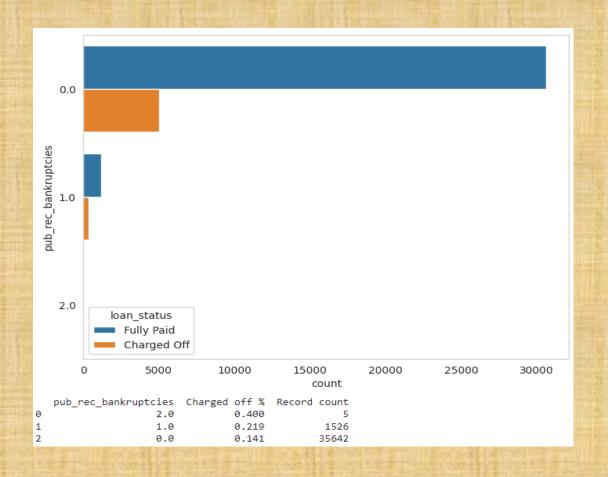


 Verification: Verified applicants have a higher default rate, indicating a potential flaw in the verification process Purpose: Loans are taken mostly for debt consolidation followed by credit card payment.

Loan request & Bankruptcy

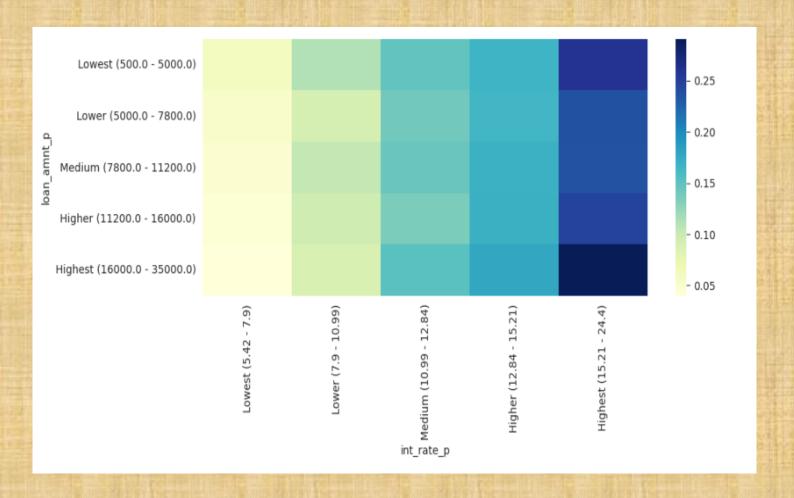


Loan Request: December sees the highest number of loan applications and the highest default ratio



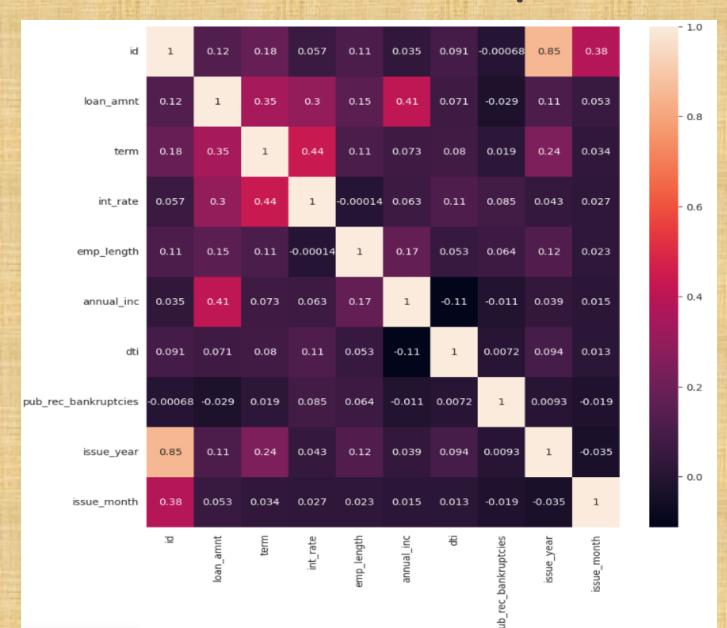
Public Recorded Bankruptcy: Having multiple bankruptcy records significantly increases the probability of loan default

Interest rate Vs Loan Amount



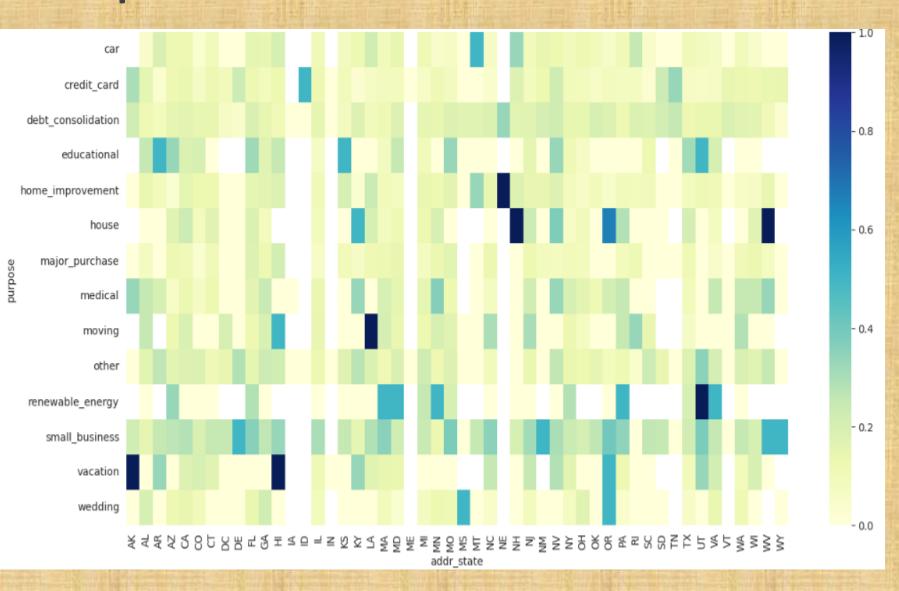
Loan default is a consequence of high loan amounts and high interest rates

Variable Relationship



Each numeric variable is completely independent of the others

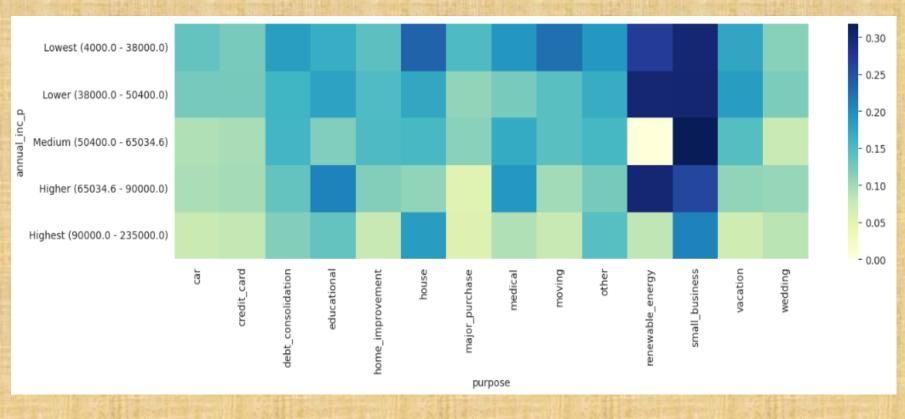
Purpose Vs Location



the darker the intersection of state has with the purpose of the loan, the risker the loan

application is

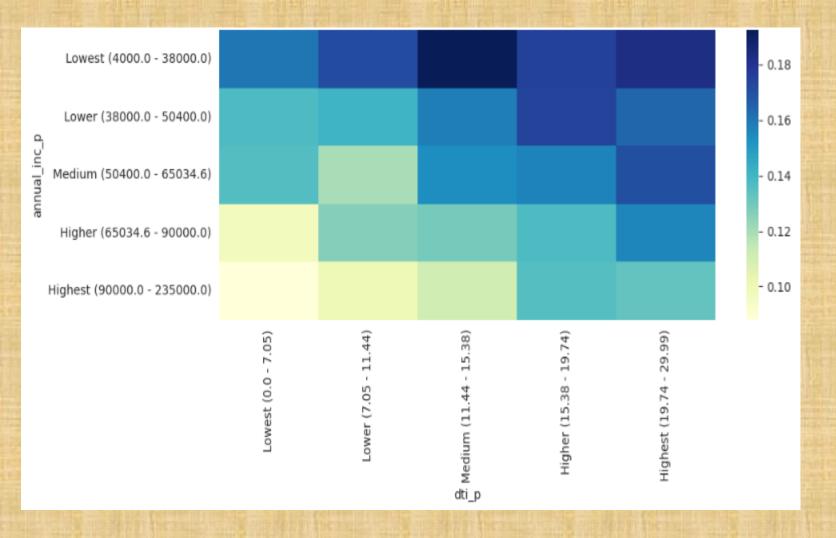
Purpose Vs Income



Loans have a risky purpose for different income groups

- 1.small business loans for lowest and medium income groups
- 2.renewable energy loans for higher income group

DTI Vs Income



The riskiest loan repayment group is the medium debt-to-income group in the lowest income range

Conclusion

Minor Impact

- .Higher loan amount
- Higher instalment amount
- Lower annual income
- Higher debt to income ratio
- Applicant's address state (NV, SD, AK, FL, etc.)
- Loan issue month (Dec, May, Sep)

Major Impact

- Higher interest rate
- Repayment term (5 years)
- Loan grade & subgrade (D to G)
- Loan purpose (small business, renewable energy, educational)
- Public bankruptcy records

Cumulative effect

- High loan amount & interest rate for lower income group
- Residential state and loan purpose
- Income group and loan purpose