Lending Club Case Study: Leveraging Data for Informed Decision Making

This case study explores the key factors influencing loan approval and default risk at Lending Club, a leading peer-to-peer lending platform. By conducting a comprehensive exploratory data analysis, we aim to uncover valuable insights that can inform strategic decision-making and risk management for lending operations.

Rohit Gotcha

Anant Joshi

Defining the Problem and Analysis Approach

1

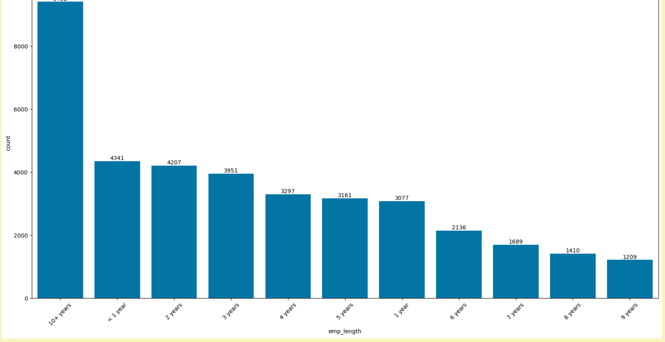
Problem Statement

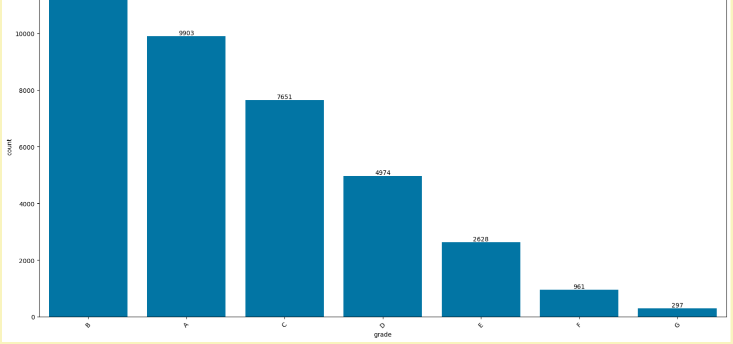
Determining the factors for loan approval and predicting loan default risk.

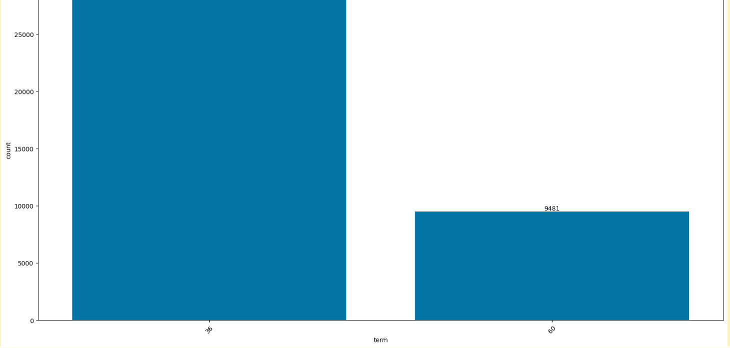
2

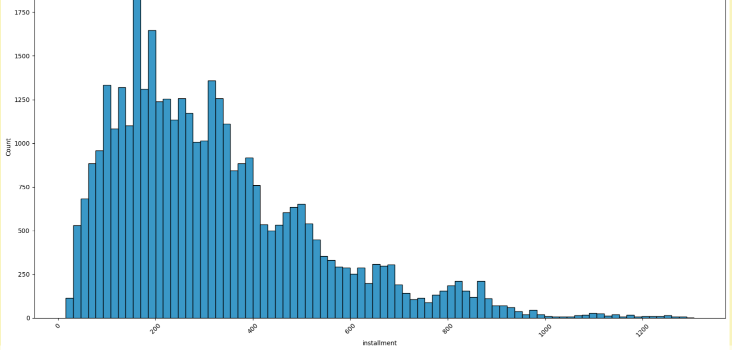
Analysis Approach

Utilizing exploratory data analysis steps like data cleaning, identify the missing values, data optimisation using standardising the data, univariate, bivariate analysis and derived metrics to identify key variables influencing loan decisions and default rates.









Unveiling Univariate Analysis Insights

1

Need for Loan

It seems, majority of loan applicants are having over 10 years of employment

2

Probable applicants that may default

Considerably less number of applicants have at least 1 public bankruptcy record

3

Applicants by Grade

There are less applicants having grades between E, F and G applied for loan compare to applicants having better grade

4

Charged off applicants

Over 5500 applicants are changed off, so they are rejected directly

5

Applicants with Leverage

More number of applicants having no assets like house or are in mortgage status

6

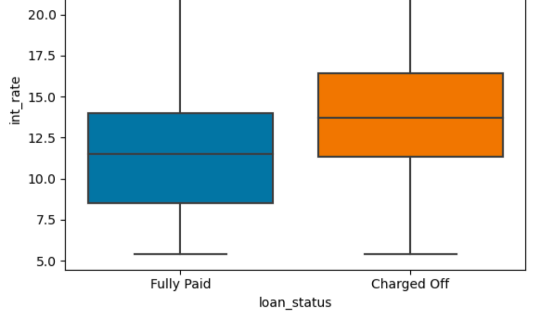
Repayment

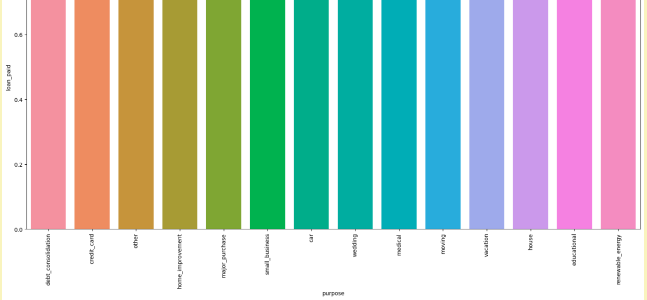
Applicants with higher premiums are likely to re-pay the loan in the shorter instalments

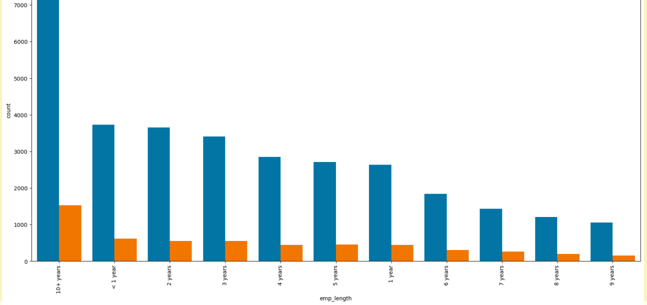
7

Loan Term Distribution

Majority of loans have terms of 36 years, affecting repayment feasibility







Unfolding Bivariate Analysis Details

By Purpose of Loan

Majority of applicants whose purpose of loan is either major purchase or home improvement have paid >80% loan

Higher Interest rates

Whoever the applicants have charged off were having higher interest rates

Annual Income vs Loan Status

Annual Income has no impact on loan status

1

2

3

4

5

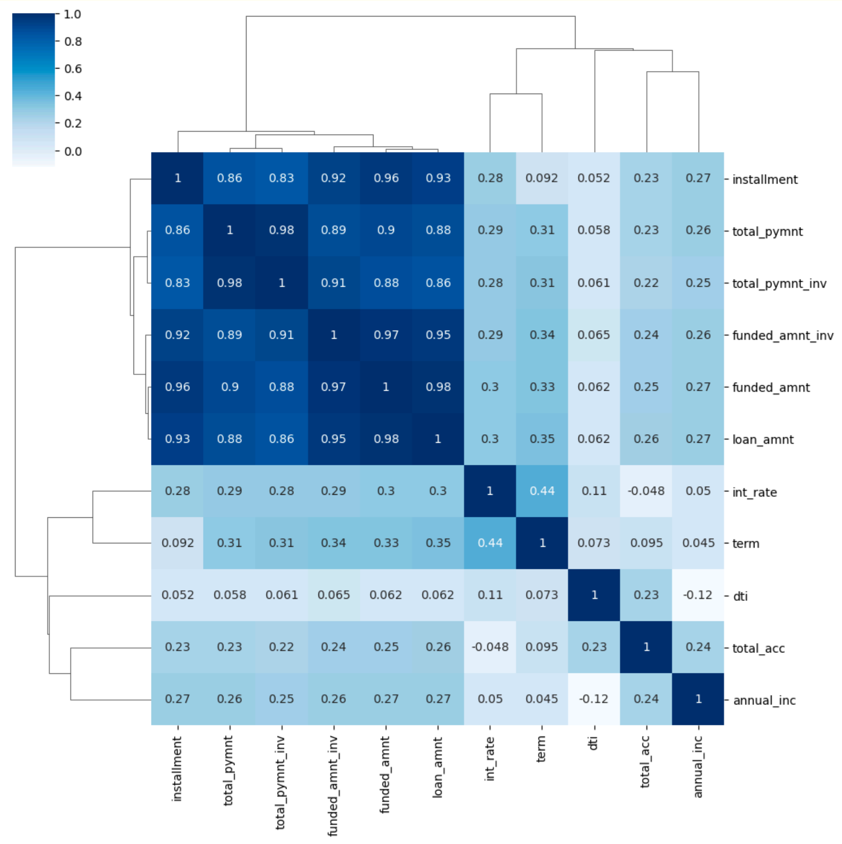
Address State

Address state has no relation with loan paid

Employee employment is positive

Employee length is criteria that directly adds value to the loan repayment

Multivariate Analysis Findings



Driving factors

1. Funded amount and total payments have strong positive correlation and are forming better cluster. So, these factors can be considered for loan approval
2. Debt to Income ratio is a strong candidate to consider the loan rejection as higher the DTI lower the loan payment capability that applicant has.

Deriving Meaningful Insights from Metrics

Debt-to-Income Ratio

Applicants with a high DTI ratio are at a higher risk of default, suggesting stricter DTI requirements.

Loan-to-Value Ratio

Higher LTV ratios correlate with increased default rates, necessitating lower LTV thresholds.

Summarizing the Key Findings

1. **Key Takeaways**

   Derived metrics like DTI and LTV provide additional insights into default risk.

2. **Recommendation**

   Stricter criteria for creditworthiness and risk assessment are recommended to mitigate defaults.

Recommendations for Lending Operations

1

Credit Scoring Models

Implement automated credit scoring models to evaluate applicant creditworthiness.

2

Income Verification

Introduce income verification procedures to ensure borrower ability to repay.

3

Loan Adjustments

Adjust loan terms and amounts based on the risk factors identified in the analysis.

Conclusion: The Power of Data-Driven Decisions

Exploratory data analysis has provided valuable insights into loan approval and default prediction, underscoring the importance of data-driven decision-making for mitigating risks and maximizing profitability in lending operations.