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

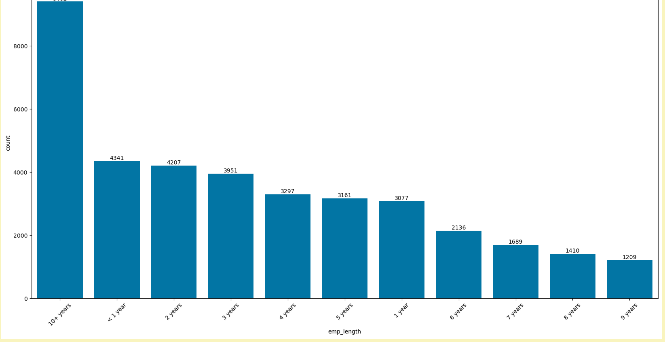
Problem Statement

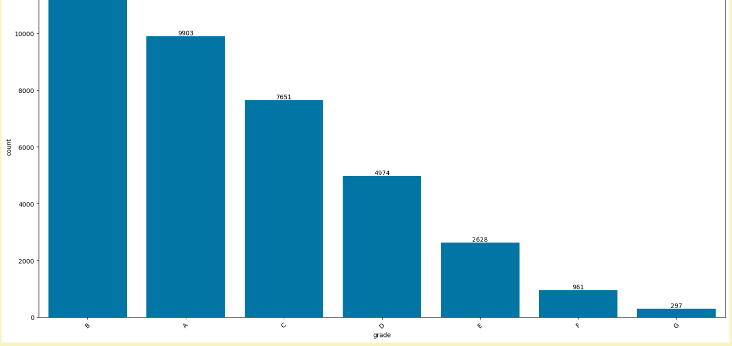
Determining the factors for loan approval and predicting loan default risk.

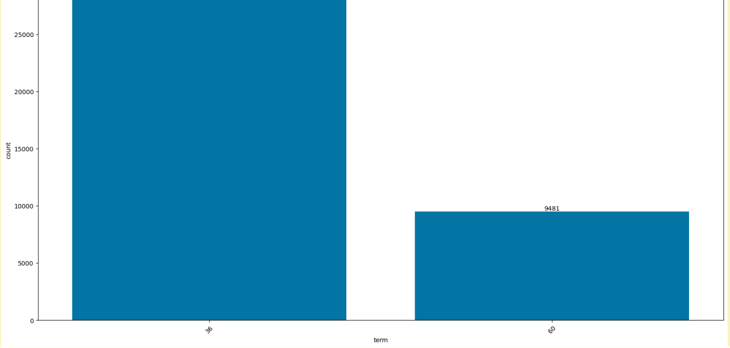
2

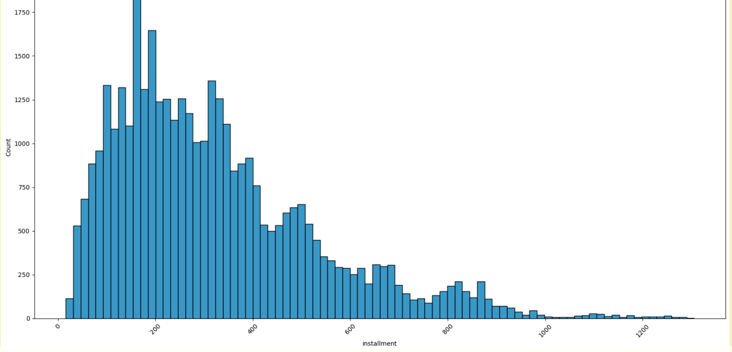
Analysis Approach

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

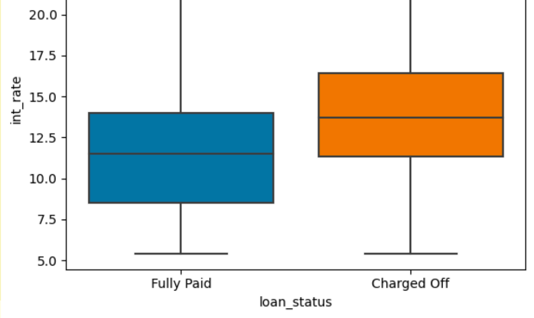
7

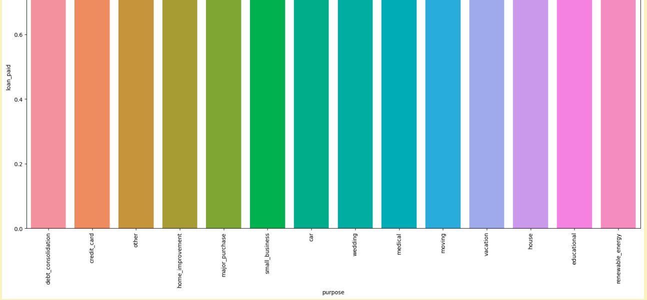
Loan Term Distribution

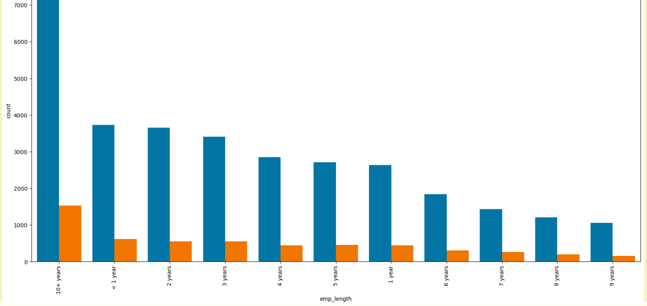
Majority of loans have terms of 36 years, affecting repayment feasibility

Summary of Univariate Analysis

1. The majority of the loans fall within the range of 5000 to 15000 dollars.
2. The majority of loan interest rates fall between 9% and 14%.
3. Most of the individuals with "Charged off" loans resided in rental properties. The lending company needs to evaluate the financial situation of individuals residing in rental properties, as they could be more vulnerable to changes in the economy.
4. The majority of the borrower's annual income falls between 40,000 and 75,000.
5. The majority of the loans were used for consolidating debts and paying off credit card bills. The number of loans that have been charged off is also high for these loans.
6. The majority of them reside in either rented accommodations or have mortgages on their homes. There is a high number of applicants from these categories, leading to a high rate of charge-offs.
7. It is observed that the number of loan applications is increasing each year. Therefore, the rise in the number of loan applications is contributing to an increase in the number of charged off applications.
8. Applicants who had borrowed money to be repaid over a 60-month period had a higher percentage of charge-offs compared to those who had taken a loan for 36 months.







Unfolding Bivariate Analysis Details

1. **By Purpose of Loan**

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

2. **Address State**

   Address state has no relation with loan paid

3. **Higher Interest rates**

   Whoever the applicants have charged off were having higher interest rates

4. **Employee employment is positive**

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

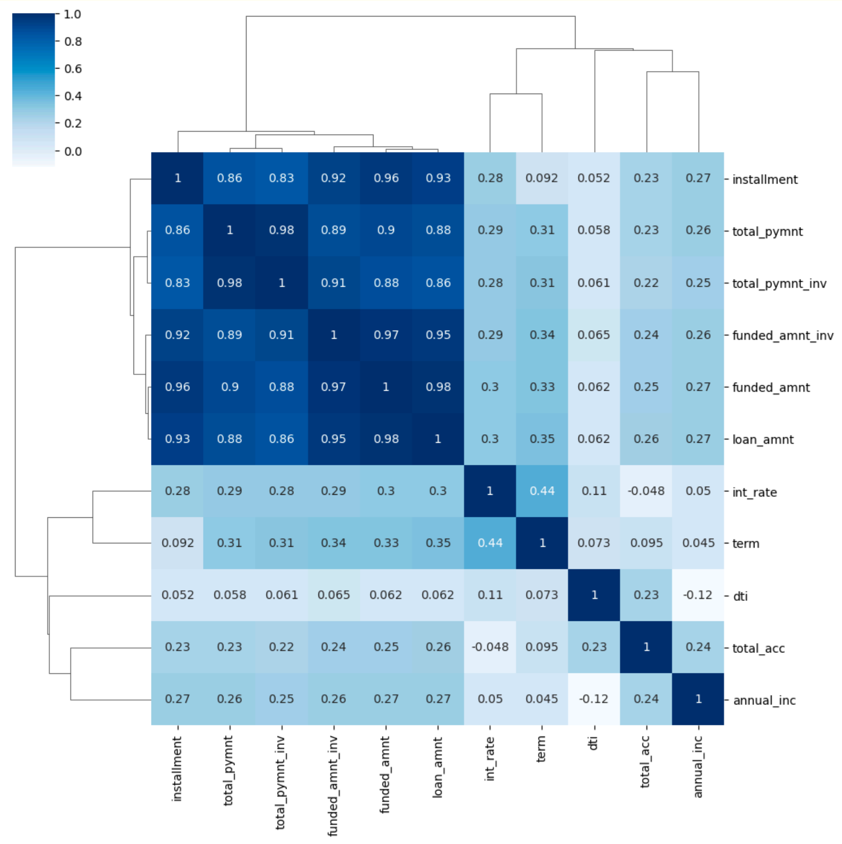
5. **Annual Income vs Loan Status**

   Annual Income has no impact on loan status

Summary of Bivariate Analysis

1. Most borrowers have been in the workforce for ten years or more, and they are also the most likely to default on their loans.
2. Most of the loans that are labeled as "Charged Off" come from loan applicants in Grades B, C, and D.
3. Grade "A" has very less chances of charged off.
4. Applicants with Sub Grades B3, B4, C1, C2 and B5 are at a higher risk of defaulting on their loans.
5. sub Grades of "A1” has very less chances of charged off
6. December tends to be the most popular month for borrowing money, likely because of the holidays.
7. The amount of people applying for loans has consistently risen from 2007 to 2011, suggesting a promising trend for the years ahead.
8. Borrowers who apply for loans with a 60-month term are at a higher risk of default compared to those who opt for a 36-month term loan.
9. Loan applicants who have been verified are defaulting at a higher rate compared to those who have not been verified.
10. Borrowers from California, Florida, and New York have the highest probability of failing to repay their loans.
11. Individuals who reside in homes that are rented or under a mortgage are at a higher risk of failing to repay their loans.
12. Debt consolidation is the category that sees the highest number of loans being issued, and also the highest rate of defaults among borrowers in that category.
13. Most of the borrowers who defaulted on their loans had very high Debt-to-Income (DTI) ratios.
14. Most of the borrowers who failed to repay their loans had borrowed $15,500 or more.
15. A large number of borrowers who failed to repay their loans had taken out loans with interest rates between 13% and 17%.
16. Charged off proportion is increasing with higher interest rates.
17. interest rate less than 10% has very less chances of charged off.

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.

Summary of Multivariate Analysis

Insights and observations from the Multivariate analysis , Correlation Analysis

Conclusions drawn from correlation measurements.

High correlation

- There is a significant correlation between annual income and loan amount.
- The term is closely linked to the interest rate.
- There is a significant relationship between installment and funded amount, loan amount, and funded amount invested.

Low correlation

- The length of employment has a low correlation with the majority of the variables.
- dti shows a low correlation with the majority of the fields.
- The annual income is inversely related to the Debt-to-Income ratio (DTI).
- The debt-to-income ratio refers to the portion of a person's monthly gross income that is used to cover debt payments.
- This indicates that when annual income is low, the debt-to-income ratio (DTI) is high and conversely.

Opposite relationship

- There is a negative correlation between annual income and dti ratio.
- pub_rec_bankrupticies has a strong negative relationship with nearly all variables.

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.

Recommendations for Lending Operations

Loan Amount and Interest Rate Analysis

Given that the majority of loans fall within the range of $5000 to $15000 and most interest rates range between 9% and 14%, the lending company should consider offering loans within these ranges to minimize the risk of defaults. Loans falling outside these ranges could be subject to closer scrutiny or higher interest rates.

Borrower Profile Analysis

Since individuals with "Charged off" loans predominantly reside in rental properties, the lending company should assess the financial stability of applicants residing in rental properties more thoroughly. This could involve additional verification steps or adjustments to loan terms based on the applicant's housing situation.

Term and Grade Analysis

Applicants opting for longer loan terms, particularly 60-month terms, exhibit a higher percentage of charge-offs. Additionally, loans from Grades B, C, and D have a higher likelihood of being charged off. Therefore, the lending company should exercise caution when approving loans with longer terms or from lower-grade applicants, implementing stricter criteria or higher interest rates for such cases

Thank You