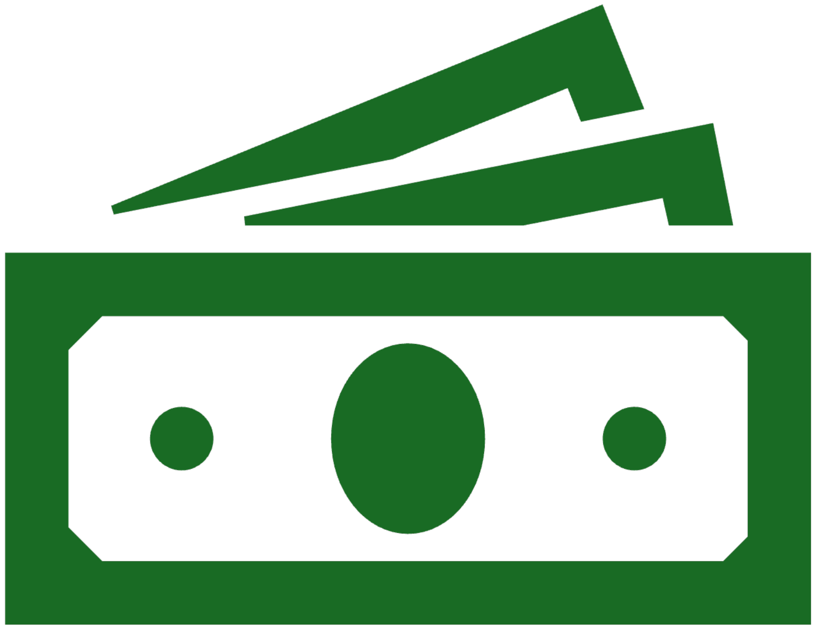

Nithiesh Kumar M

&

Adila Parveen Hafeezulla

Lending Club Case Study

Problem Statement

The company offers personal, business and medical procedure financing.

The company provides low-interest loans via online platform.

The greatest financial risk is to lend the loan to applicants who are default on their loans, which would result in significant credit loss.

The goal is to decrease credit loss by employing exploratory data analysis (EDA) techniques to identify the primary factors contributing to loan default.

Approach

To mitigate credit loss, we will conduct EDA to uncover the key drivers behind loan defaults.

By analyzing various data points and patterns within our borrower profiles.

We aim to identify factors strongly correlated with default behavior.

The analysis will guide our strategic decisions in risk assessment, enabling us to refine our lending practices and minimize exposure to default risk.

Univariate Analysis

- Use histograms to plot the graphs for numerical columns.
- Use count plot to plot the graphs for categorial columns.

Insights from Univariate Analysis

- More than half of the loans taken have a term of 36 months compared to 60 months.
- The interest rate is clustered more around 5-10% and 10-15%, with a slight drop near 10%.
- A significant portion of loans are graded 'A' and 'B', indicating a predominance of high-grade loans.
- Within grades 'A' and 'B', there is a concentration of loans with lower subgrades, consistent with the overall grade distribution.
- A majority of borrowers have work experience exceeding 10 years.
- Most borrowers do not own property and are either on mortgage or rent.
- Approximately 50% of borrowers are verified by the company or have their sources verified.

Insights from Univariate Analysis cont'd

- Annual income follows a left-skewed normal distribution, suggesting that the majority of borrowers have relatively low annual income.
- A large percentage of loans are taken for debt consolidation, followed by credit card usage.
- The majority of borrowers hail from large urban cities like California, New York, Texas, Florida, etc.
- Most borrowers have significant debt compared to their registered income, with a concentration in the 10-15 DTI ratio range.
- The bulk of loans are approved in the last quarter of the year.
- The number of approved loans increases exponentially over time, indicating a rising loan approval rate.
- The quantity of defaulted loans is comparatively lower than that of fully paid ones.

Bi-variate Analysis

- Use box plot to plot the graphs for categorial columns.

Insights from Bi-variate Analysis

- Grade A, which represents the lowest risk, also exhibits the lowest DTI ratio, suggesting that higher-grade loans have a lower rate of default.
- Grades represent risk levels, indicating that interest rates increase with risk.
- Employees with 10+ years of experience are more likely to default but also have a higher chance of fully paying the loan.
- Defaulted loan amounts increase with interest rates and show a decline after a 17.5% interest rate.
- Borrowers with less than $50,000 in annual income are more likely to default, while those with higher annual incomes are less likely to default.
- Defaulted loans are fewer for borrowers who own their property compared to those on mortgage or rent.

Insights from Bi-variate Analysis cont'd

- Loan status varies with DTI ratio; loans in the 10-15 DTI ratio range have a higher number of defaulted loans, but higher DTI ratios have a higher chance of defaulting.
- Loans with a 36-month term mainly consist of grade A and B loans, while loans with a 60-month term mostly consist of grade B, C, and D loans.
- The 60-month term has a higher chance of defaulting than the 36-month term, whereas the 36-month term has a higher chance of fully paid loans.
- The mean and 25th percentile are the same for both, but a larger 75th percentile is observed in defaulted loans, indicating that a larger loan amount has a higher chance of defaulting.
- Debt consolidation is the most popular loan purpose and has the highest number of fully paid and defaulted loans.
- Fully paid loans are increasing exponentially over time compared to defaulted loans.

Multi-variate Analysis

- Use heatmaps to plot the graphs for all the columns.

Conclusion

- Overall, the analysis suggests several key factors that significantly influence the probability of loan default. These factors include the purpose of the loan, income level, interest rates, loan amount, borrower's employment length, verification status, debt-to-income ratio (DTI), loan grade, property ownership status, geographical location, bankruptcy history, and time of loan approval.
- Loan Characteristics: Loans for certain purposes, such as home improvement or small business, loans with higher interest rates, and larger loan amounts, are more likely to default. Additionally, loans with longer terms, particularly 60-month terms, have a higher chance of default compared to 36-month terms.
- Borrower Profile: Borrowers with lower annual incomes, higher DTI ratios, and longer employment histories (10+ years) are more likely to default. Furthermore, borrowers with lower loan grades (e.g., E, F, G) are at a higher risk of default.
- Verification and Ownership Status: Loans that are verified by the company or have a verified source are associated with lower default rates. Moreover, borrowers who own their property have a lower default rate compared to those who are on mortgage or rent.
- Geographical and Economic Factors: Borrowers from large urban cities like California, New York, Texas, and Florida are more likely to default. This might indicate regional economic factors influencing default rates.

Conclusion

- Recommendations: To mitigate default risks and minimize credit losses, it's essential to focus on assessing and managing the following factors:
- DTI and Loan Grades: Monitor and assess the debt-to-income ratio and loan grades to better gauge the borrower's ability to repay the loan.
- Verification and Income: Ensure thorough verification processes, especially for borrowers with higher default risk characteristics. Consider income levels as a key determinant of repayment capacity.
- Geographical Considerations: Consider regional economic factors and borrower demographics when assessing default risk.
- Proactive Monitoring: Continuously monitor loan performance and adjust lending practices accordingly to minimize default risks.

Thank You ☺