Lending Club Case Study

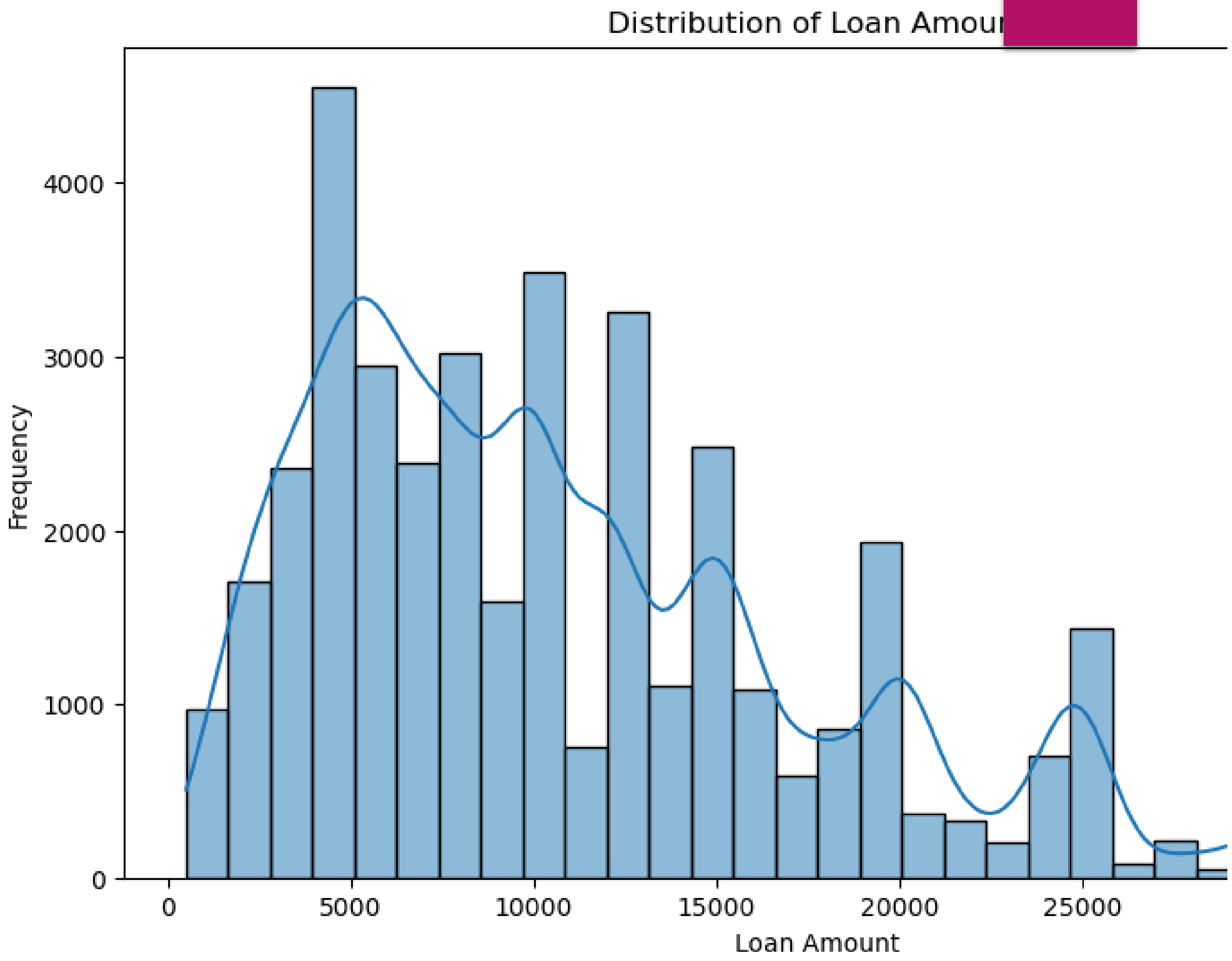
- Background
  - Lending Club is an online peer-to-peer lending platform that connects borrowers with investors.
  - It facilitates unsecured personal loans, business loans, and auto refinances loans.
  - The platform operates by offering loans at competitive rates, while investors can diversify their portfolios by investing in loans of various risk levels.
- Business Objective
  - The main objective of this study is to analyze the risk of loan defaults
  - We aim to identify key factors that drive default rates and assess the impact of borrower characteristics such as income, loan purpose, and interest rates on defaults.
  - The insights will help in making better lending decisions and managing portfolio risks.

Data Understanding

- The dataset contains records of past loans issued by Lending Club.
- It includes various attributes such as loan amount, interest rate, loan purpose, employment length, and verification status.
- The target variable is the loan status, which indicates whether a loan was fully paid, charged off, or is currently in default.
- Key variables for analysis:
    - Loan amount, interest rate, annual income, debt-to-income ratio (DTI), and loan purpose.
    - The analysis majorly focuses on these variables.

Distribution of Loan Amount

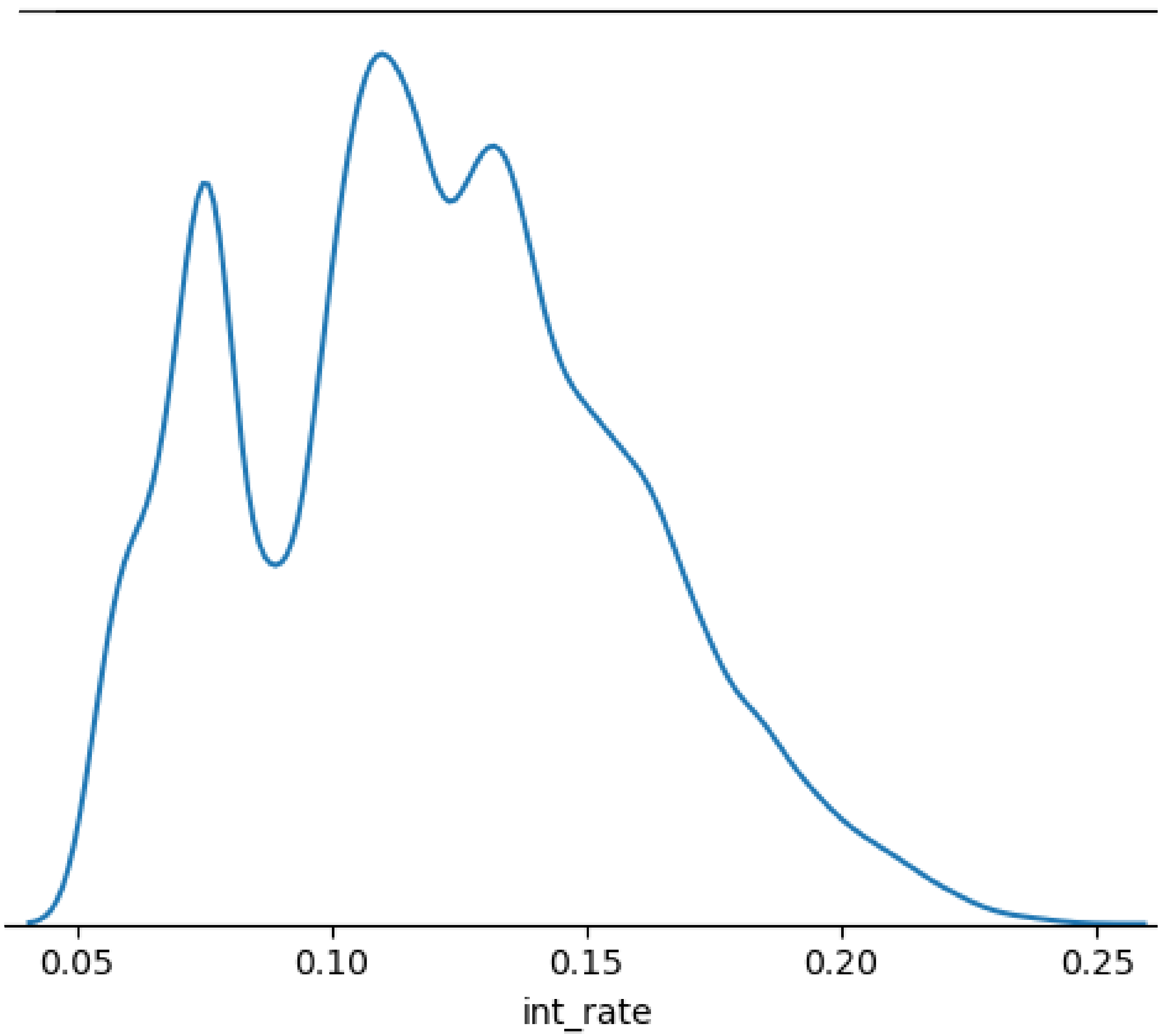
- Key Insights:
  - The distribution is right-skewed (also known as positively skewed). This means there is a tail extending to the right, indicating that a larger number of loans are for lower amounts, with fewer loans for higher amounts.
  - The highest frequency is around 4,000, indicating that this is the most common loan amount range.
  - This distribution shows that most borrowers seek loans in the mid-range, but some higher values may represent riskier loan applicants.



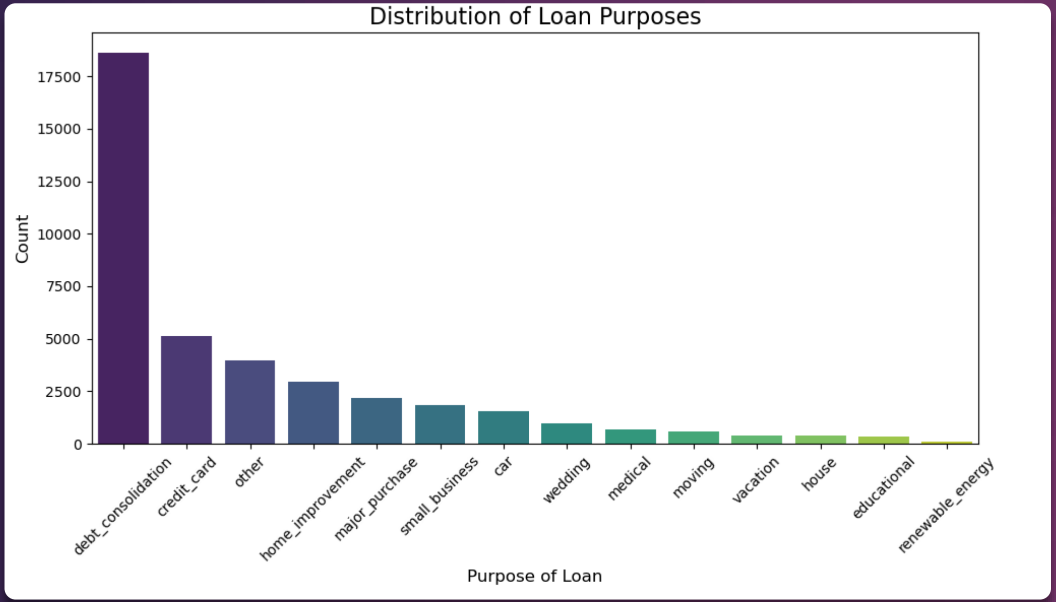
Interest Rate Distribution

Key Insights:

- Interest rates cluster around the 10%-15% range, showing this is the most common rate offered.
- The risk of default may be higher for loans with significantly higher or lower interest rates.



Distribution of Loan Purposes



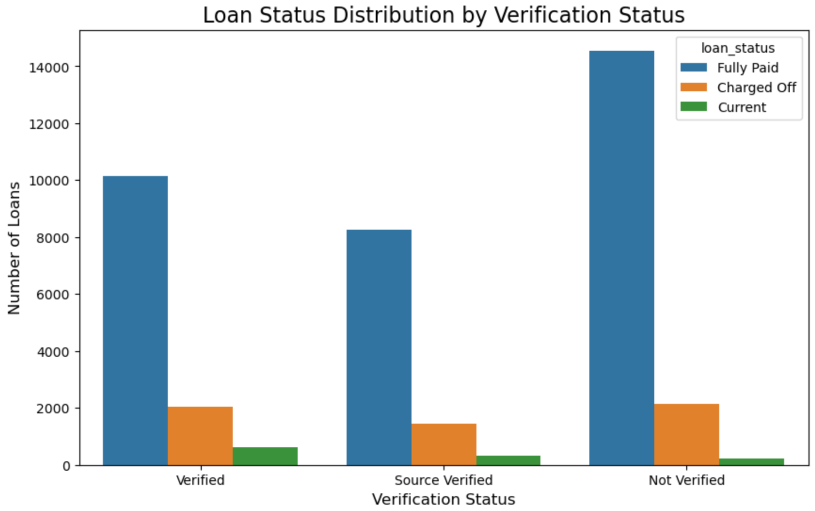
•**Key Insight:** Most loans were taken for *debt consolidation* and *credit card payments*.

•**Risk Analysis:** Borrowers seeking loans for debt consolidation and credit card payments might represent higher financial strain, which could increase their risk of default.

•**Impact:** Understanding the purpose of the loan helps in identifying potentially high-risk borrowers, allowing for better decision-making on loan approvals.

Verification Status and Loan Repayment Trends

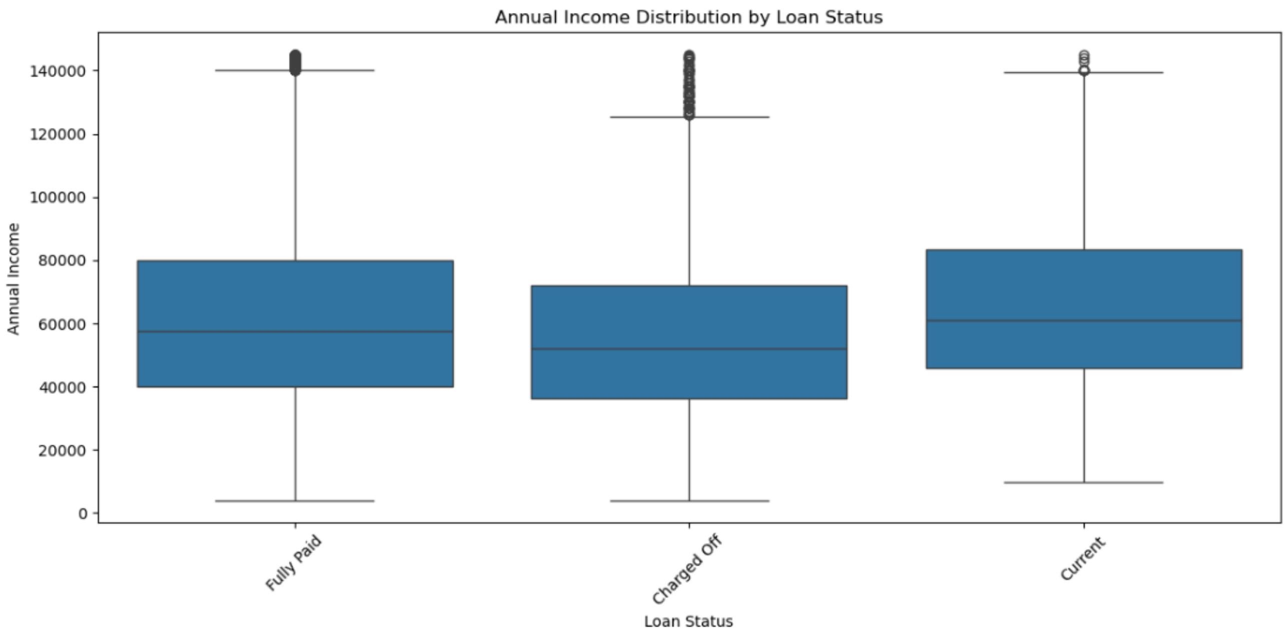
- KEY INSIGHTS:
  - The high number of 'Not Verified' loans that are 'Fully Paid' may indicate a tendency for these borrowers to prioritize repayment despite lacking formal verification.
  - Conversely, 'Verified' loans, while indicating a more stable financial background, show fewer 'Fully Paid' loans, potentially suggesting higher expectations for financial scrutiny.
- RISK IMPLICATIONS:
  - WHILE 'NOT VERIFIED' LOANS EXHIBIT A GREATER REPAYMENT RATE, THIS COULD MASK UNDERLYING RISKS, WARRANTING A DEEPER ANALYSIS INTO BORROWER BEHAVIOR.
  - THE 'VERIFIED' CATEGORY MAY REQUIRE CLOSER EXAMINATION TO UNDERSTAND WHY REPAYMENT RATES ARE LOWER, DESPITE A MORE ROBUST VERIFICATION PROCESS.



Impact of Annual Income on Loan Repayment Status

- KEY INSIGHTS:

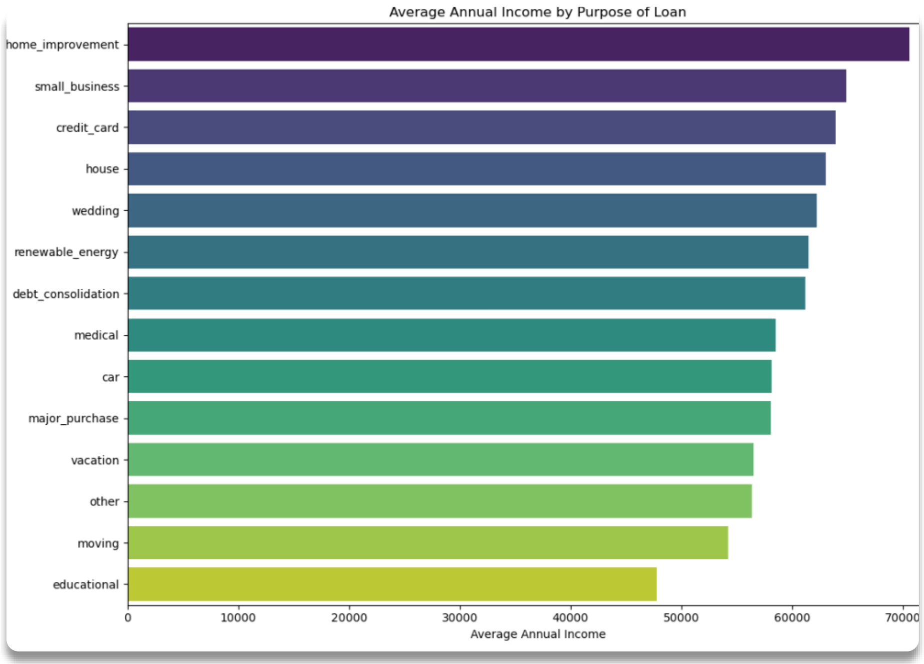
- THE MEDIAN INCOME FOR "FULLY PAID" LOANS APPEARS TO BE SLIGHTLY HIGHER THAN FOR "CHARGED OFF" AND "CURRENT" LOANS, SUGGESTING THAT HIGHER INCOME BORROWERS MIGHT HAVE A BETTER CHANCE OF REPAYING THEIR LOANS.
- THE "CHARGED OFF" CATEGORY SHOWS A SLIGHTLY WIDER INTERQUARTILE RANGE (IQR), INDICATING MORE VARIABILITY IN INCOMES AMONG BORROWERS WHO DEFAULTED ON THEIR LOANS. THIS MIGHT SUGGEST THAT INCOME LEVEL IS LESS PREDICTABLE FOR THIS GROUP.
- THE PRESENCE OF OUTLIERS, ESPECIALLY IN THE "FULLY PAID" CATEGORY, SUGGESTS THAT THERE MIGHT BE SOME CASES WHERE HIGH INCOME BORROWERS HAVE SUCCESSFULLY REPAID THEIR LOANS DESPITE OTHER FACTORS.



Analysis of Annual Income by Purpose of Loan

- KEY INSIGHTS:

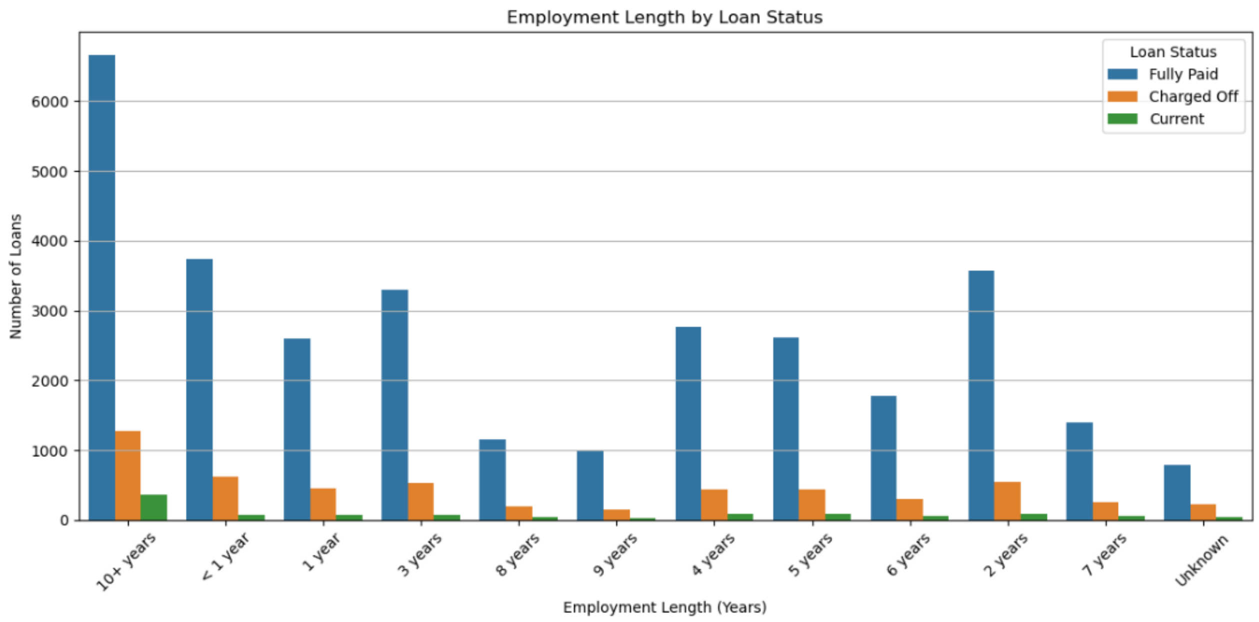
- HIGHER AVERAGE INCOMES LINKED TO HOME IMPROVEMENT AND SMALL BUSINESS LOANS SUGGEST THESE BORROWERS MAY POSE LOWER CREDIT RISKS, AS THEY ARE LIKELY TO HAVE A STRONGER FINANCIAL FOUNDATION.
- CONVERSELY, LOWER AVERAGE INCOMES ASSOCIATED WITH EDUCATIONAL AND MOVING LOANS HIGHLIGHT A POTENTIAL FOR HIGHER RISK. THESE BORROWERS MAY FACE CHALLENGES IN REPAYMENT, PARTICULARLY IF THEIR INCOME DOES NOT STABILIZE PROMPTLY AFTER TAKING OUT THE LOAN.
- LENDERS SHOULD CONSIDER NOT ONLY THE LOAN PURPOSE BUT ALSO THE BORROWER'S INCOME LEVEL WHEN ASSESSING RISK.



Analysis of Employment Length by Loan Status

- KEY INSIGHTS:

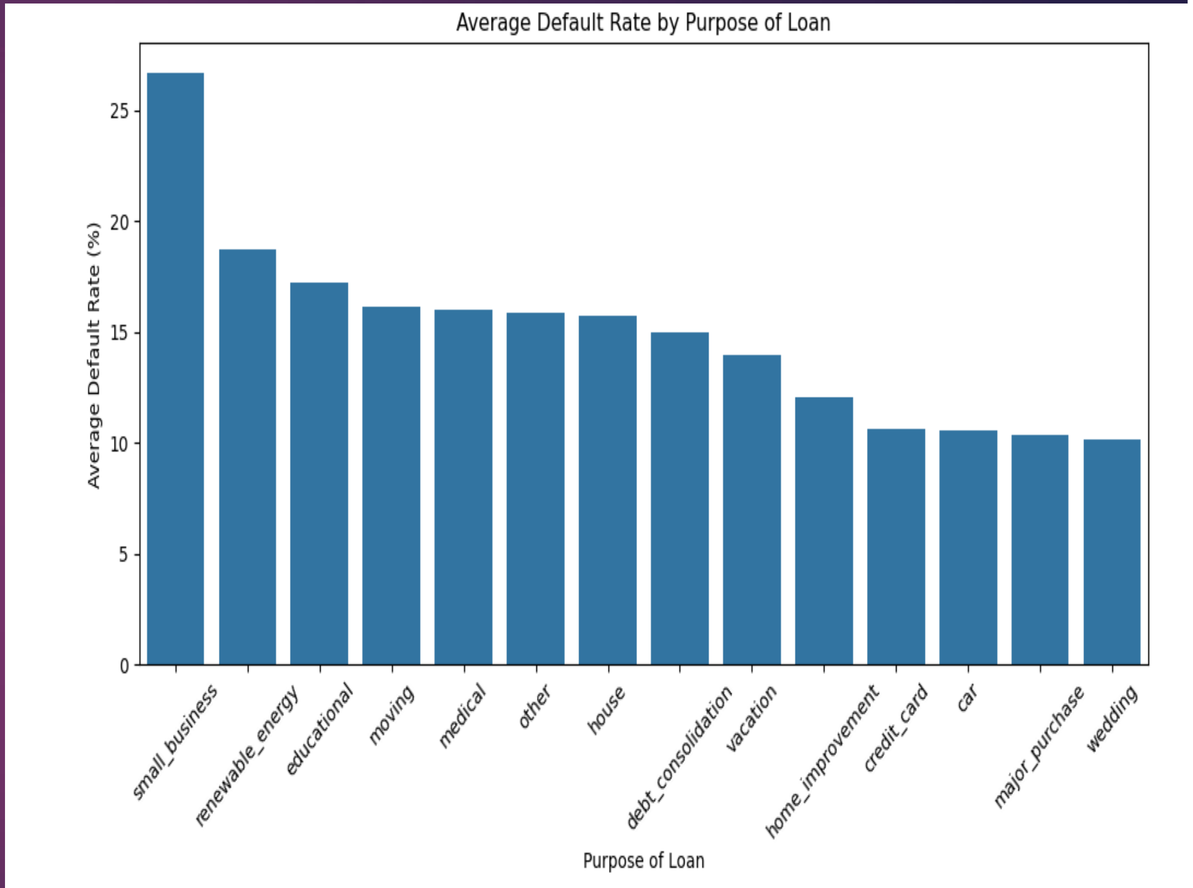
- THE CHART REVEALS A RELATIONSHIP BETWEEN EMPLOYMENT LENGTH AND LOAN STATUS. LOANS FOR INDIVIDUALS WITH LONGER TENURES (10+ YEARS) HAVE A HIGHER PROPORTION OF "FULLY PAID" STATUS COMPARED TO SHORTER TENURES.
- A SIGNIFICANT PORTION OF "CHARGED OFF" LOANS ARE FOR INDIVIDUALS WITH LESS THAN 1 YEAR OF EMPLOYMENT, SUGGESTING THAT NEWER EMPLOYEES MIGHT HAVE A HIGHER RISK OF DEFAULT.
- THE DISTRIBUTION OF "CURRENT" LOANS IS RELATIVELY SPREAD OUT ACROSS DIFFERENT EMPLOYMENT LENGTHS, INDICATING THAT EMPLOYMENT LENGTH ALONE MIGHT NOT BE A STRONG PREDICTOR OF WHETHER A LOAN WILL BE CURRENT.



Analysis of Average Default Rate by Purpose of Loan

KEY INSIGHTS:

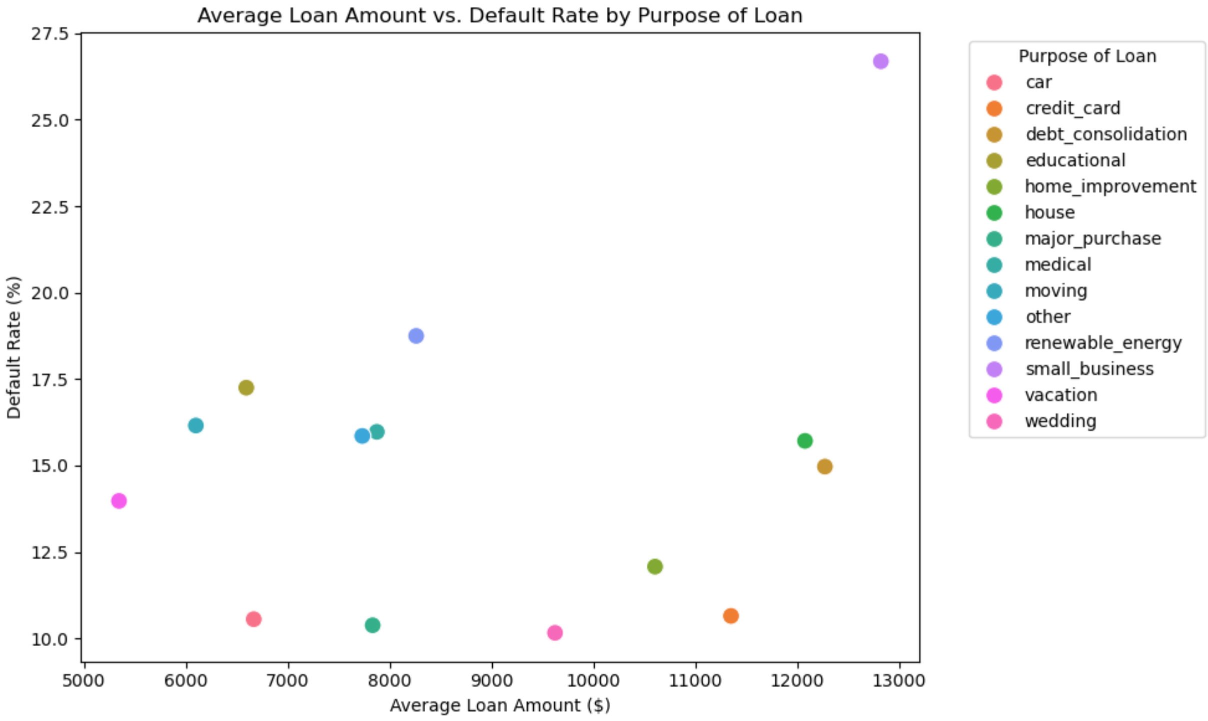
- LOANS FOR "SMALL BUSINESS" HAVE THE HIGHEST AVERAGE DEFAULT RATE, INDICATING THAT THIS CATEGORY MIGHT POSE THE GREATEST RISK.
- LOANS FOR "RENEWABLE ENERGY" AND "EDUCATIONAL" PURPOSES HAVE RELATIVELY LOW AVERAGE DEFAULT RATES, SUGGESTING THAT THESE CATEGORIES MIGHT BE CONSIDERED LOWER RISK.
- SEVERAL LOAN PURPOSES, INCLUDING "MOVING," "HOUSE," "DEBT CONSOLIDATION," AND "VACATION," HAVE SIMILAR AVERAGE DEFAULT RATES, INDICATING THAT THESE CATEGORIES MIGHT HAVE COMPARABLE RISK PROFILES.
- SOME CATEGORIES, SUCH AS "WEDDING" AND "MAJOR PURCHASE," HAVE NOTABLY LOWER DEFAULT RATES, POTENTIALLY INDICATING THAT THESE LOANS ARE ASSOCIATED WITH MORE STABLE FINANCIAL SITUATIONS OR ARE USED FOR ESSENTIAL PURPOSES.
- THE DATA SUGGESTS THAT THE PURPOSE OF A LOAN CAN BE A SIGNIFICANT FACTOR IN PREDICTING DEFAULT RISK. LENDERS MIGHT NEED TO ADJUST THEIR UNDERWRITING CRITERIA FOR DIFFERENT LOAN PURPOSES.



Average Loan Amount vs. Default Rate by Purpose of Loan

▶KEY INSIGHTS:

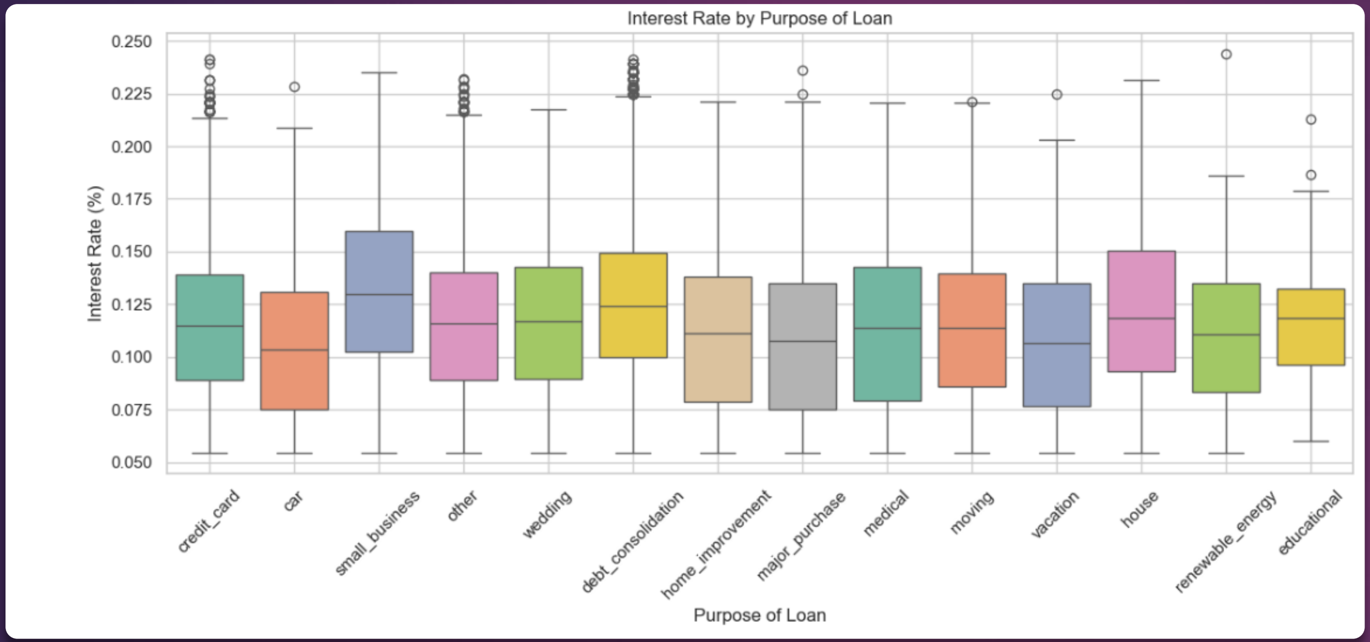
- SMALL BUSINESS LOANS SHOW THE HIGHEST DEFAULT RATE AT 26.69%, INDICATING SIGNIFICANT RISK ASSOCIATED WITH LENDING IN THIS CATEGORY.
- RENEWABLE ENERGY (18.75%) AND EDUCATIONAL LOANS (17.25%) ALSO PRESENT HIGHER DEFAULT RATES, SUGGESTING POTENTIAL RISK FACTORS IN THESE AREAS.
- LOANS FOR HOME IMPROVEMENT (12.08%), CREDIT CARDS (10.65%), AND CAR PURCHASES (10.56%) DISPLAY MODERATE DEFAULT RATES, REFLECTING A BALANCE BETWEEN RISK AND PURPOSE.
- THE AVERAGE LOAN AMOUNTS FOR HIGHER-RISK PURPOSES SUCH AS SMALL BUSINESS (12,822) AND DEBT CONSOLIDATION (12,273) ARE COMPARATIVELY SIGNIFICANT, INDICATING THAT LARGER LOANS CORRELATE WITH INCREASED RISK.
- CONVERSELY, LOAN PURPOSES LIKE VACATION (5,344) AND MOVING (6,098) HAVE LOWER AVERAGE LOAN AMOUNTS AND DEFAULT RATES, SUGGESTING LOWER RISK IN THESE CATEGORIES.



Analyzing the Loan Amount vs. Interest Rate Scatter Plot

- KEY INSIGHTS:

- THE SCATTER PLOT SHOWS A POSITIVE CORRELATION BETWEEN LOAN AMOUNT AND INTEREST RATE. THIS MEANS THAT AS THE LOAN AMOUNT INCREASES, THE INTEREST RATE TENDS TO INCREASE AS WELL.
- THE DATA POINTS ARE SOMEWHAT SCATTERED, INDICATING THAT THERE IS SOME VARIATION IN INTEREST RATES FOR LOANS OF SIMILAR AMOUNTS. THIS SUGGESTS THAT OTHER FACTORS BESIDES LOAN AMOUNT MIGHT INFLUENCE THE INTEREST RATE.
- THE RED LINE REPRESENTS A LINEAR REGRESSION MODEL FITTED TO THE DATA. THIS LINE SHOWS THE GENERAL TREND BETWEEN LOAN AMOUNT AND INTEREST RATE. A POSITIVE SLOPE OF THE LINE CONFIRMS THE POSITIVE CORRELATION.



Interest Rate Analysis by Purpose of Loan

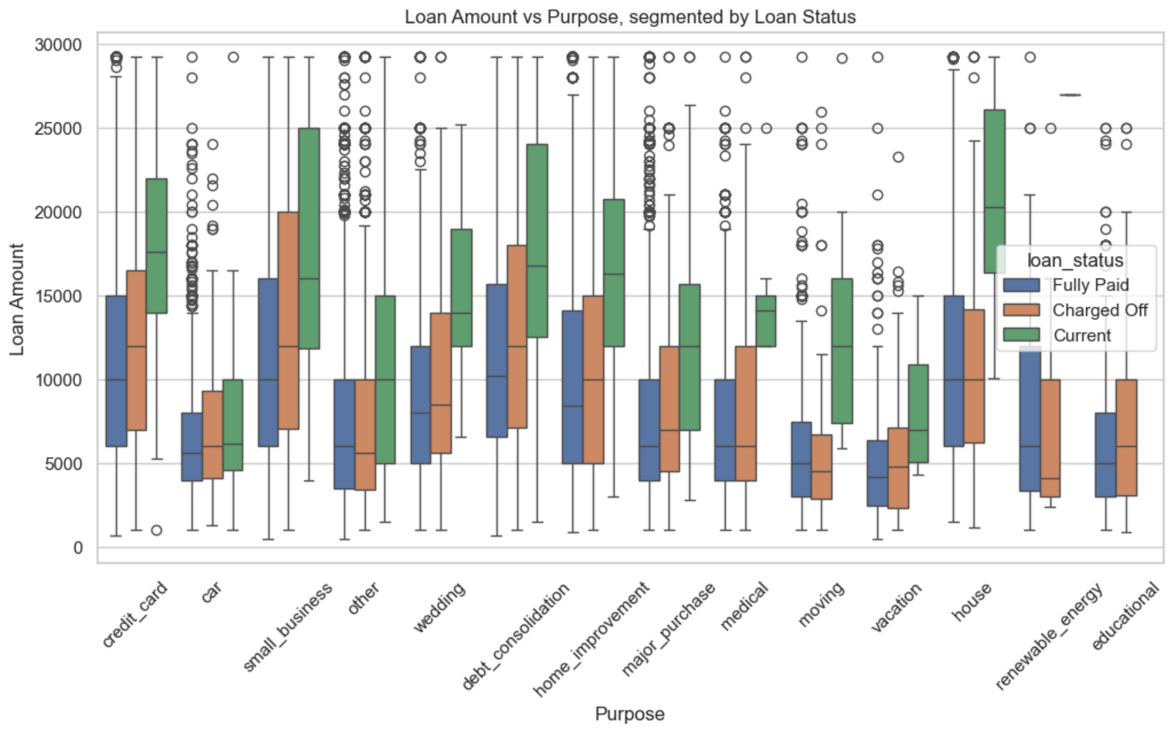
KEY INSIGHTS:

- **SMALL BUSINESS** AND **CREDIT CARD** LOANS TYPICALLY EXHIBIT HIGHER INTEREST RATES. THIS SUGGESTS THAT LENDERS CONSIDER THESE PURPOSES TO CARRY HIGHER RISK DUE TO FACTORS LIKE MARKET VOLATILITY AND BORROWER REPAYMENT CAPABILITIES.
- LOANS FOR PURPOSES SUCH AS **CAR** AND **VACATION** GENERALLY SHOW LOWER INTEREST RATES. THIS COULD INDICATE THAT LENDERS PERCEIVE THESE LOANS AS LESS RISKY, POSSIBLY DUE TO THE INTENDED USE ENHANCING PROPERTY VALUE OR DISCRETIONARY SPENDING.
- LOANS FOR **EDUCATION** PURPOSES EXHIBIT A RELATIVELY NARROW INTEREST RATE RANGE COMPARED TO OTHER PURPOSES. THIS INDICATES A CONSISTENT LENDING POLICY FOR EDUCATIONAL LOANS.

Analysis of Loan Amount vs. Purpose, Segmented by Loan Status

- KEY INSIGHTS:

- LOANS FOR "SMALL BUSINESS" HAVE A WIDE RANGE OF LOAN AMOUNTS, SUGGESTING THAT THERE MIGHT BE SIGNIFICANT VARIATION IN THE SIZE AND FINANCIAL NEEDS OF SMALL BUSINESSES.
- "HOUSE" LOANS GENERALLY HAVE HIGHER LOAN AMOUNTS, REFLECTING THE LARGER FINANCIAL COMMITMENT ASSOCIATED WITH PURCHASING A HOME.
- "CREDIT_CARD" LOANS TEND TO HAVE LOWER LOAN AMOUNTS, REFLECTING THEIR NATURE AS REVOLVING CREDIT PRODUCTS.



Analysis of Loan Amount vs. Purpose, Segmented by Loan Status

KEY INSIGHTS:

Loan Amount vs. Interest Rate:
- Higher interest rates correlate with higher loan default risk ("Charged Off").
- Lower interest rates are linked to a higher probability of full repayment ("Fully Paid").

Loan Status Distribution:
- **Fully Paid** loans (blue) are more frequent at lower interest rates.
- **Charged Off** loans (orange) increase in higher interest rate ranges (above 15%).
- **Current** loans (green) are sparsely distributed across different amounts and rates.

Debt-to-Income Ratio (DTI):
- Loan outcomes appear relatively independent of DTI size.
- Larger DTI loans are scattered across all loan statuses and interest rates.

General Trend:
- Loans with lower interest rates (5%-15%) are more likely to be fully repaid.
- Loans with higher interest rates (15%-25%) face a higher chance of being charged off.