

LENDING CLUB CASE STUDY

AMIT KUMAR

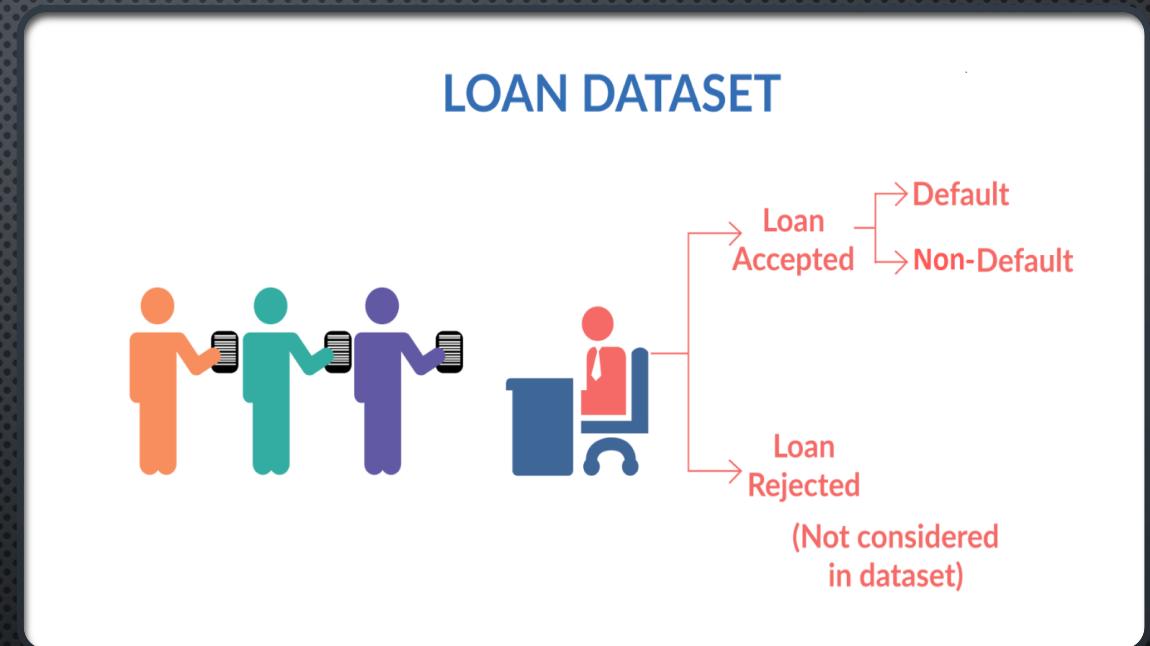
KANHAIYA CHOUDHARY

ML AUG 2024

PROBLEM STATEMENT

- TO UNDERSTAND THE **DRIVING FACTORS (OR DRIVER VARIABLES)** BEHIND LOAN DEFAULT, I.E. THE VARIABLES WHICH ARE STRONG INDICATORS OF DEFAULT.
- THE COMPANY CAN UTILISE THIS KNOWLEDGE FOR ITS PORTFOLIO AND RISK ASSESSMENT.

- **FULLY PAID:** APPLICANT HAS FULLY PAID THE LOAN (THE PRINCIPAL AND THE INTEREST RATE)
- **CURRENT:** APPLICANT IS IN THE PROCESS OF PAYING THE INSTALMENTS, I.E. THE TENURE OF THE LOAN IS NOT YET COMPLETED. THESE CANDIDATES ARE NOT LABELLED AS 'DEFAULTED'.
- **CHARGED-OFF:** APPLICANT HAS NOT PAID THE INSTALMENTS IN DUE TIME FOR A LONG PERIOD OF TIME, I.E. HE/SHE HAS DEFAULTED ON THE LOAN



PROBLEM SOLVING METHODOLOGY

- TO UNDERSTAND HOW **CONSUMER ATTRIBUTES** AND **LOAN ATTRIBUTES** INFLUENCE THE TENDENCY OF DEFAULT.
- USING EDA FOR THIS ANALYSIS
- STEPS
 - DATA UNDERSTANDING
 - DATA CLEANING
 - UNIVARIATE ANALYSIS
 - SEGMENTED UNIVARIATE ANALYSIS
 - BIVARIATE ANALYSIS
 - RECOMMENDATION

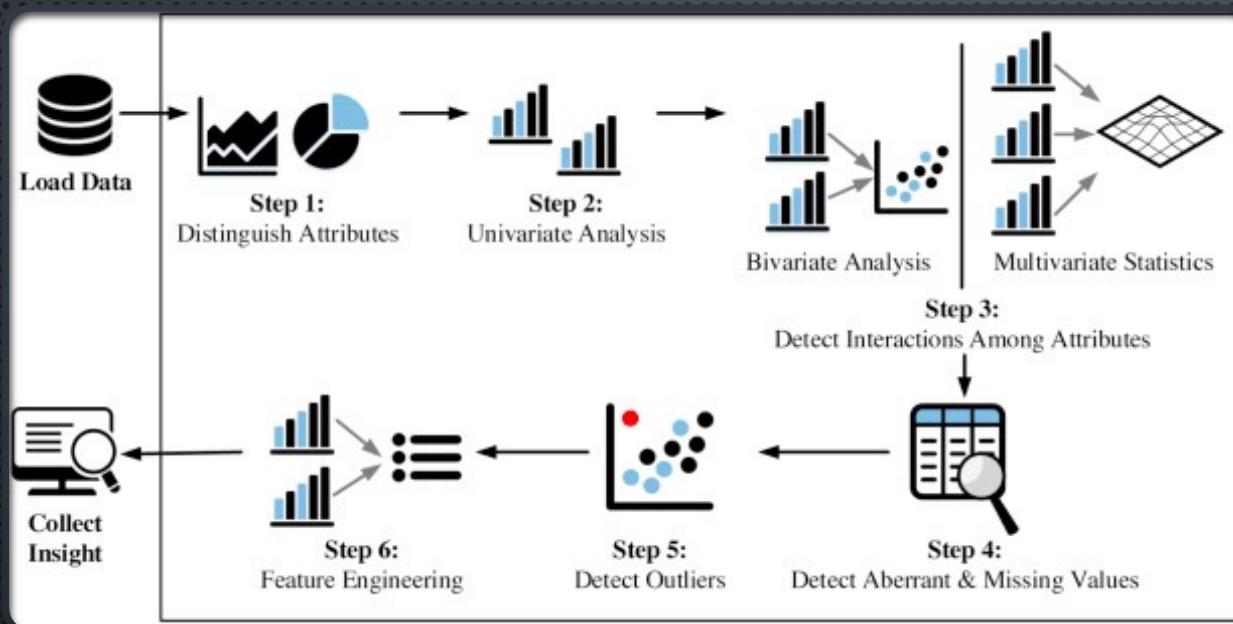
DATA CLEANING

- CHECK THE PERCENTAGE OF MISSING VALUES
- REMOVE ALL THOSE WITH VERY HIGH MISSING PERCENTAGE
- FOR COLUMNS WITH LESS MISSING PERCENTAGE: PERFORM IMPUTATIONS
- DROP ROWS WHERE THE MISSING PERCENTAGE IS QUITE HIGH

DATA ANALYSIS

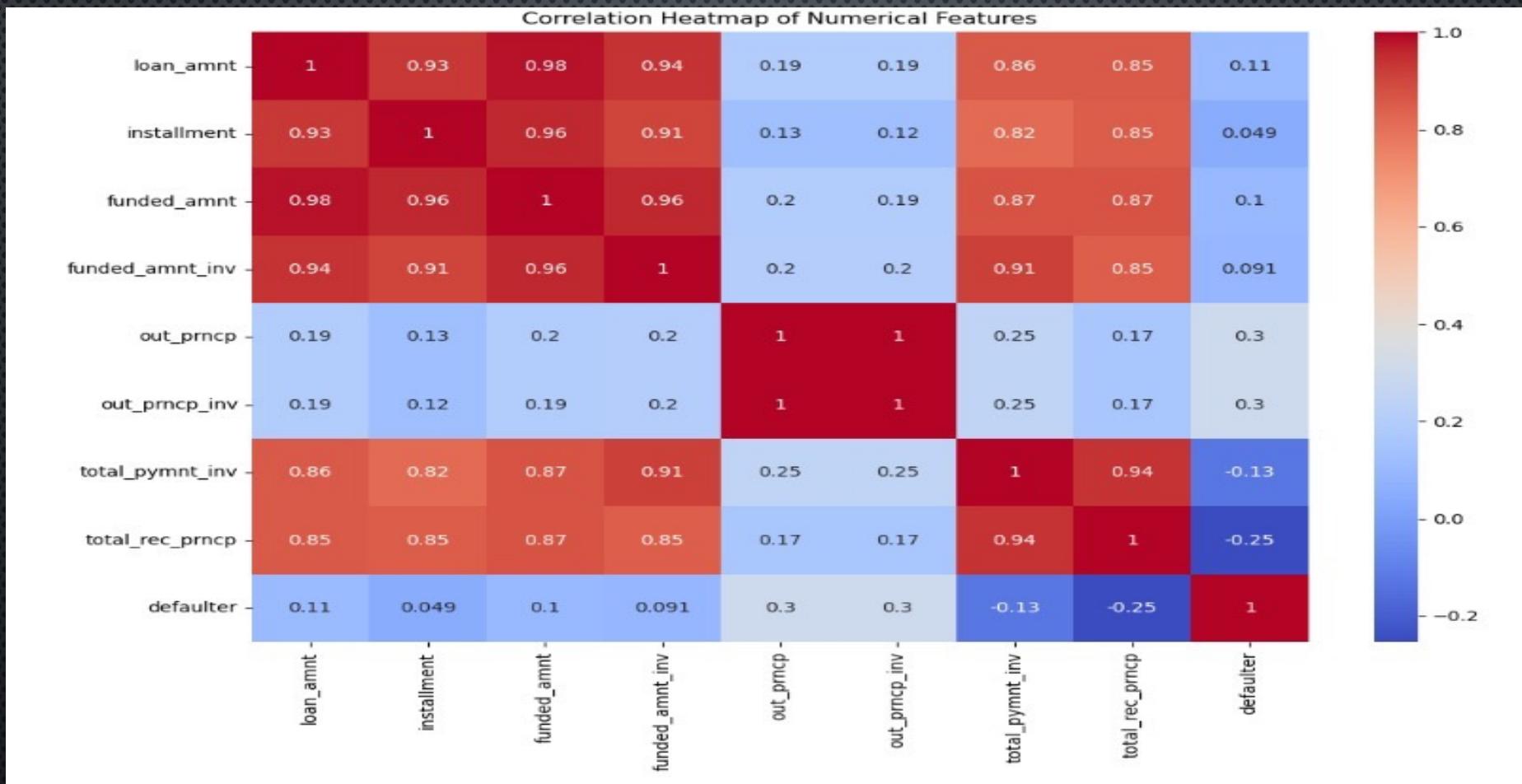
- THE OBJECTIVE IS TO IDENTIFY PREDICTORS OF DEFAULT SO THAT AT THE TIME OF LOAN APPLICATION, WE CAN USE THOSE VARIABLES FOR APPROVAL/REJECTION OF THE LOAN.
 - THERE ARE BROADLY THREE TYPES OF VARIABLES –
 - THOSE WHICH ARE RELATED TO THE APPLICANT
 - LOAN CHARACTERISTICS
 - CUSTOMER BEHAVIOR VARIABLES

EDA PROCESS

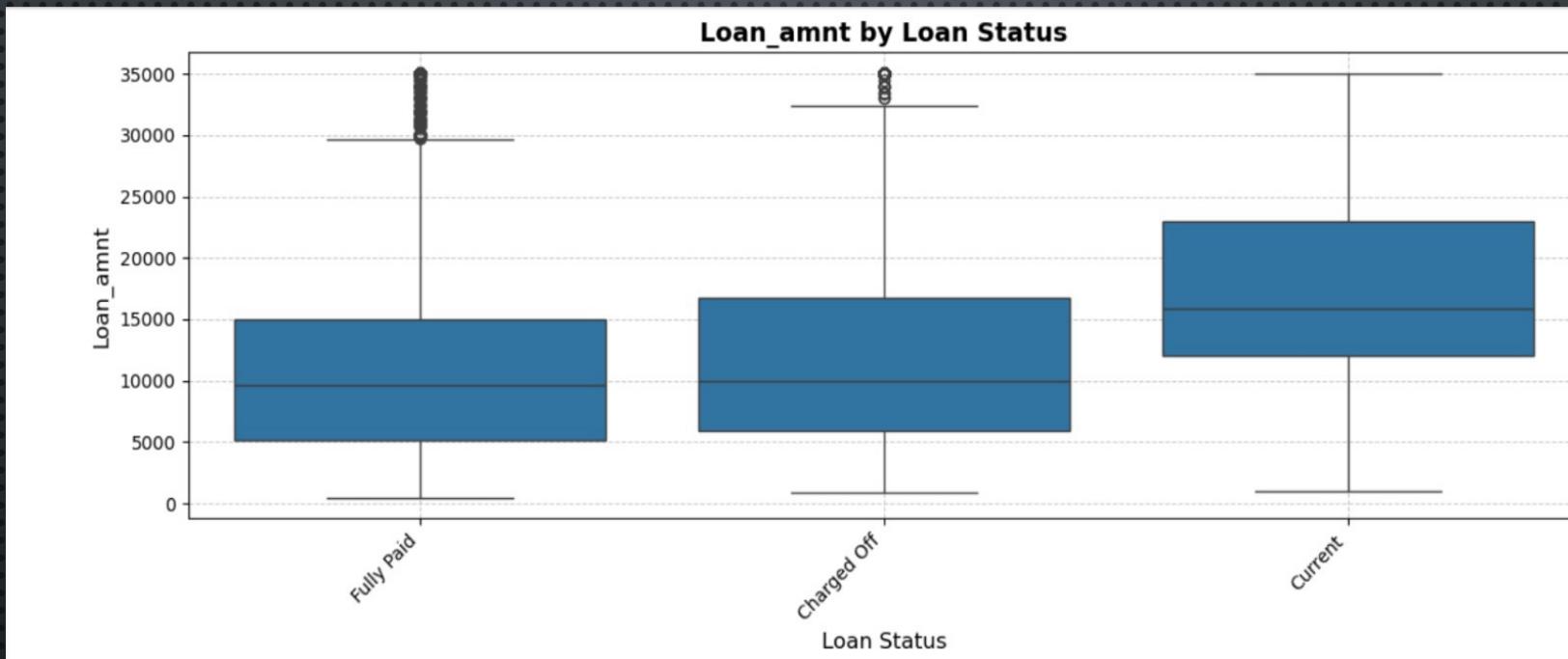


ANALYSIS

CORRELATION HEATMAP

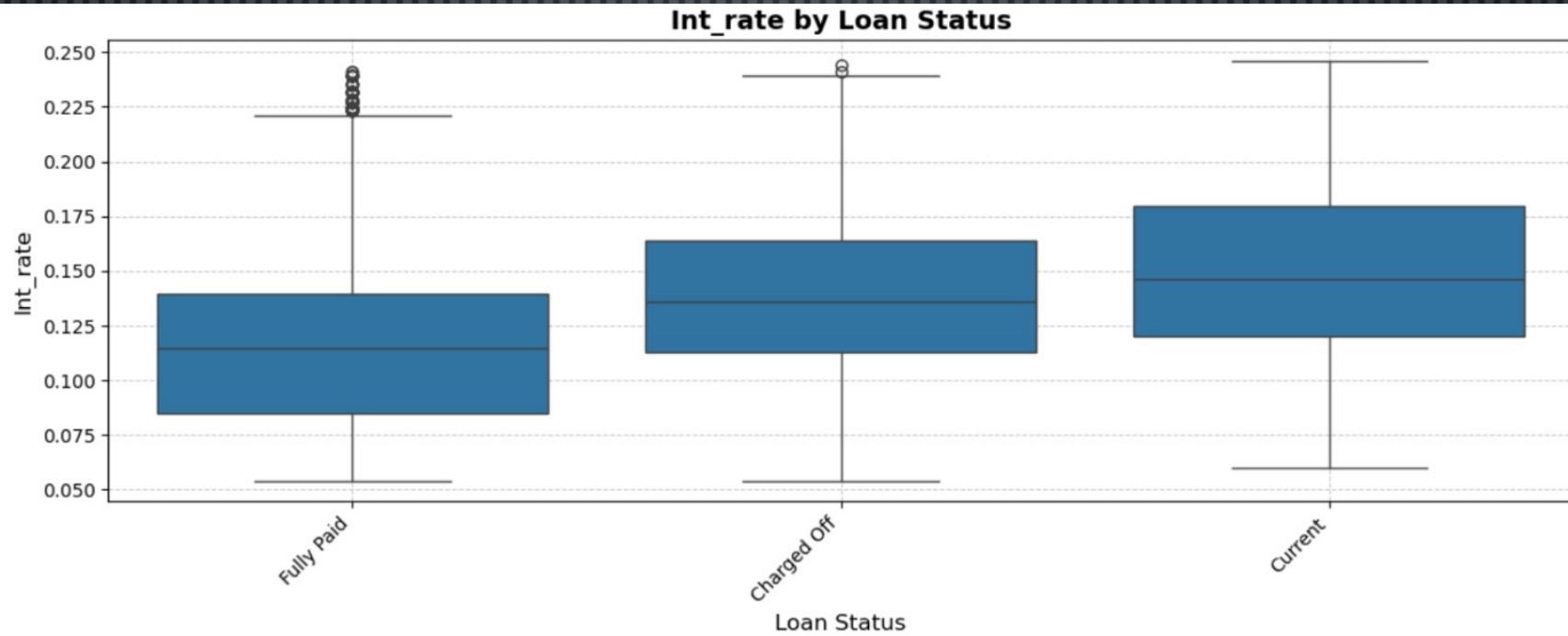


IMPACT OF LOAN AMOUNT VARIABLE ON CHARGED OFF



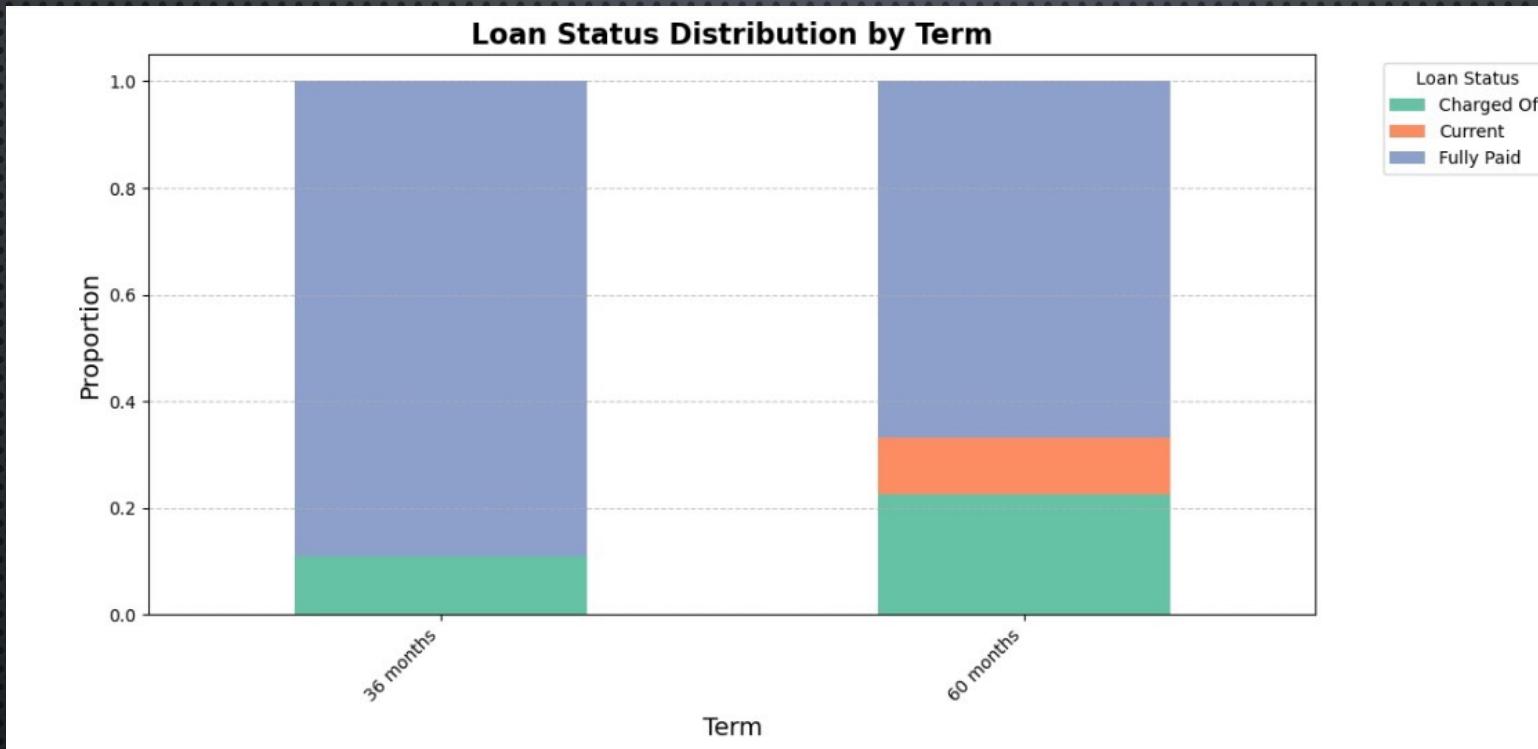
Higher loan amounts are associated with an **increased likelihood of default**. This suggests that borrowers who take larger loans may face more financial strain, leading to a higher risk of non-repayment.

IMPACT OF VARIABLE INTEREST RATE ON CHARGED OFF



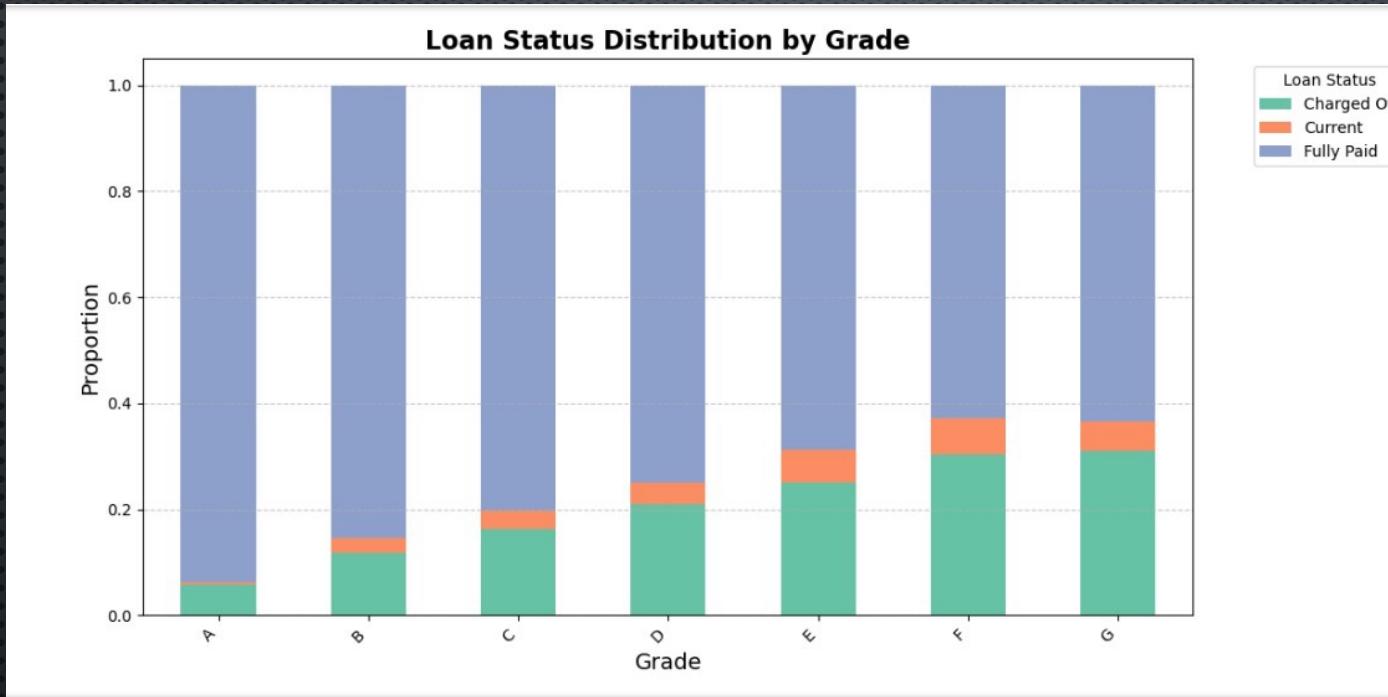
Loans with **higher interest rates** are more prone to default. Borrowers with higher interest rates may have a lower creditworthiness, leading to higher risks for the company.

IMPACT OF TERM VARIABLE ON CHARGED OFF



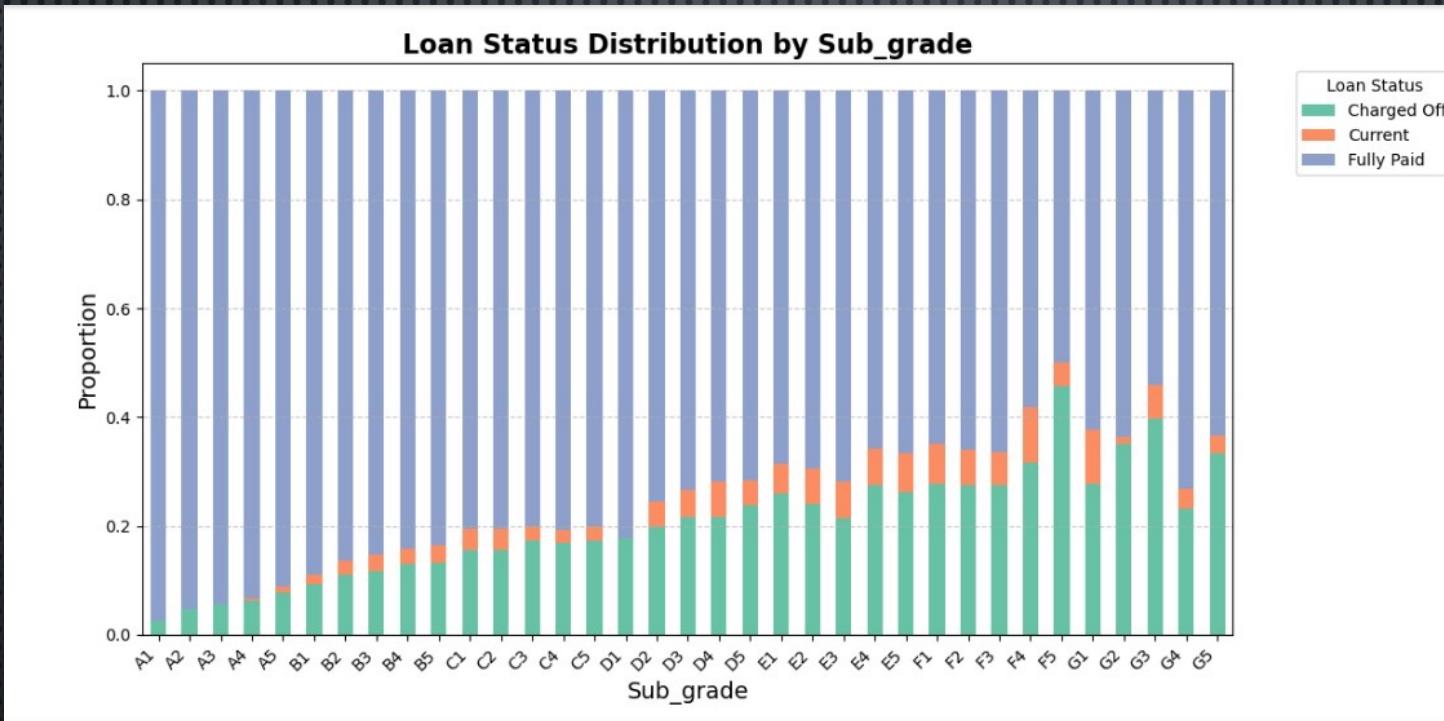
Loans with a **60-month term** are at a **higher risk of default** compared to loans with a 36-month term. Longer repayment periods may create more opportunities for financial distress over time.

IMPACT OF GRADE VARIABLE ON CHARGED OFF



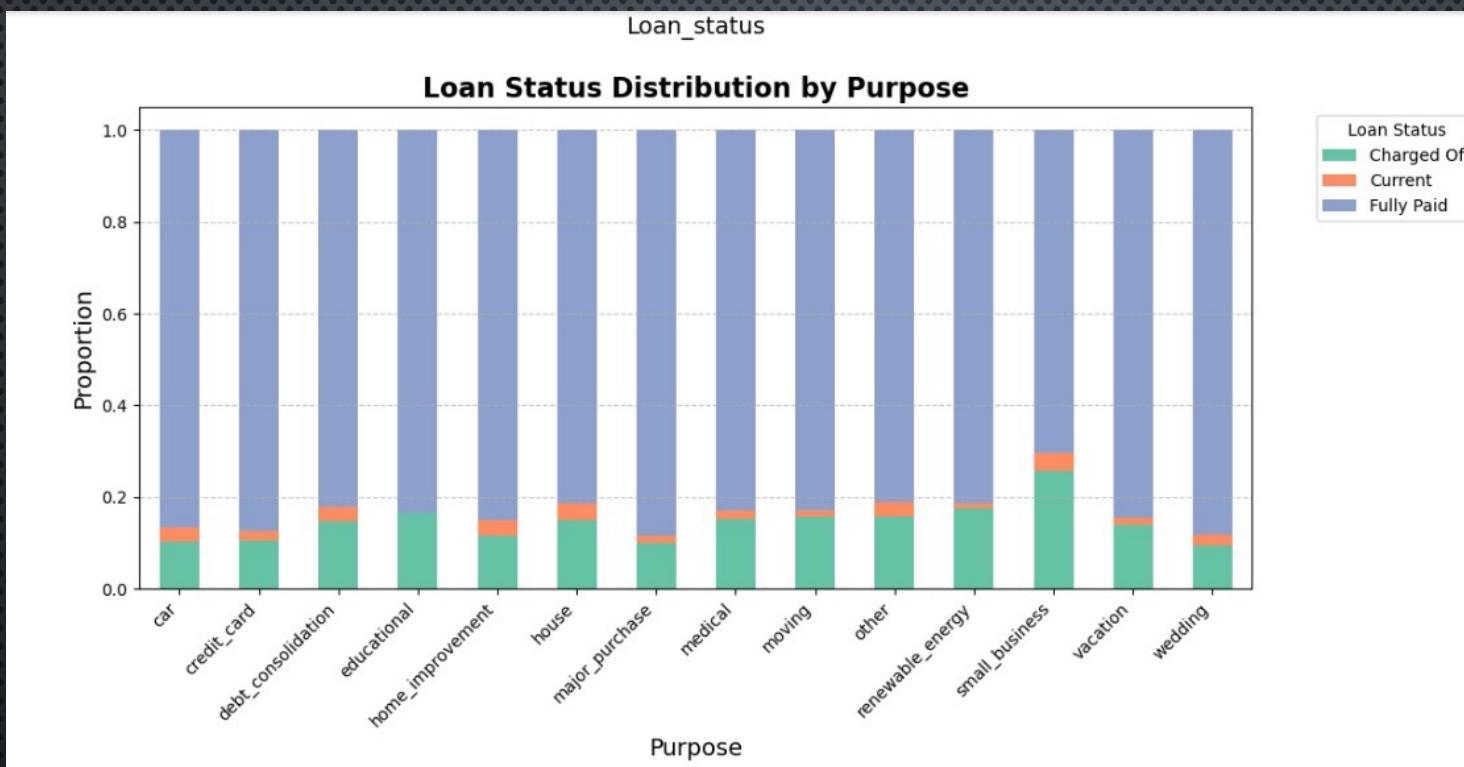
Loan defaults gradually increase as the grade worsens from **A to G**. Loans with higher grades (closer to A) tend to perform better, whereas loans with grades closer to G are riskier.

IMPACT OF SUB GRADE VARIABLE ON CHARGED OFF



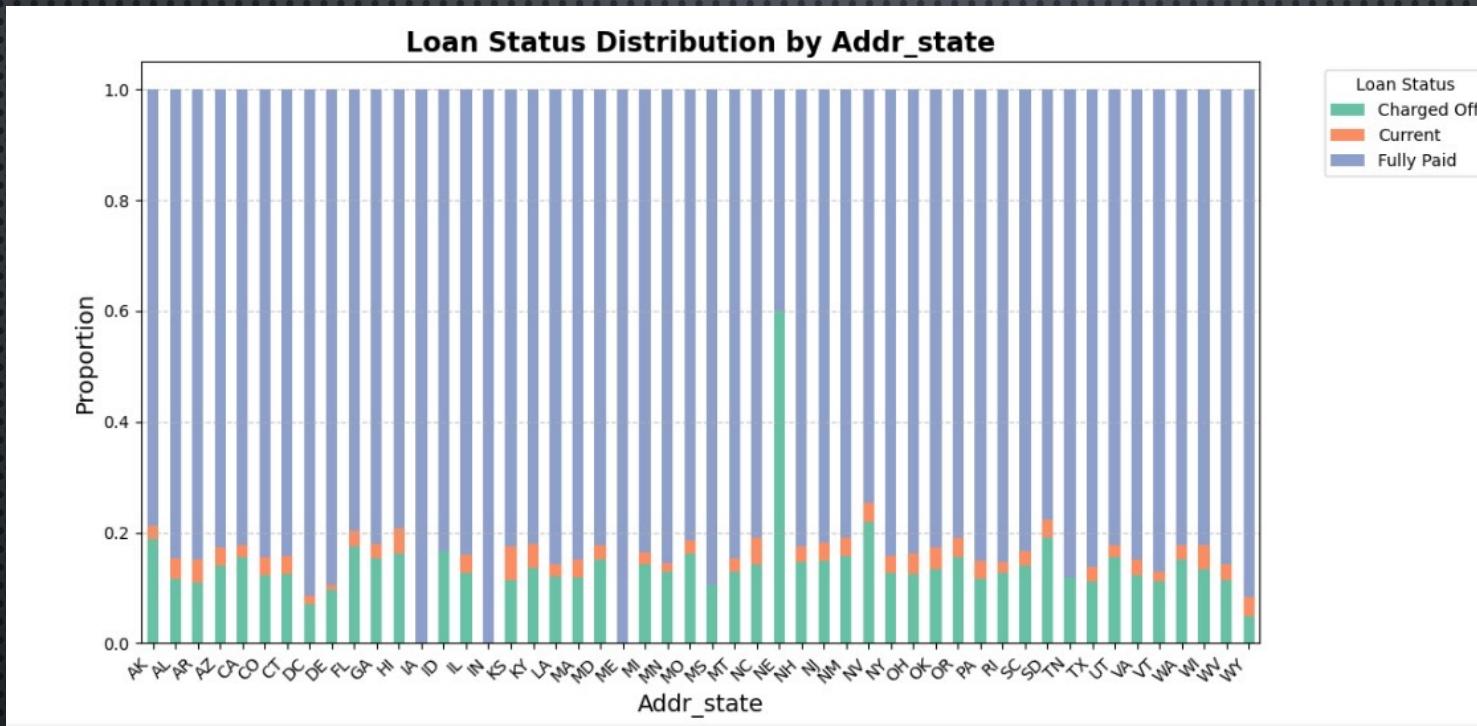
Similar to the overall grade, **sub-grade** analysis reveals that **loan defaults increase** as the sub-grade worsens from **A1 to G5**, with some exceptions. This provides a more granular view of creditworthiness.

IMPACT OF PURPOSE VARIABLE ON CHARGED OFF



Certain loan purposes, specifically **small business loans**, exhibit a **higher risk of default**. This could be due to the volatility and risk associated with small businesses.

IMPACT OF ADDRESS STATE VARIABLE ON CHARGED OFF



NE state loans (loans issued to borrowers in Northeastern U.S. states) appear to be at a **higher risk of default** compared to loans in other regions. This may point to region-specific economic factors affecting borrowers' ability to repay.

CONCLUSION

- BASED ON THE ANALYSIS OF THE KEY DRIVER VARIABLES, WE CAN CONCLUDE THAT CERTAIN FACTORS SIGNIFICANTLY INFLUENCE THE LIKELIHOOD OF LOAN DEFAULTS. THESE FINDINGS CAN HELP THE COMPANY IN MAKING DATA-DRIVEN DECISIONS TO MITIGATE RISKS ASSOCIATED WITH LOAN DEFAULTS:
1. **LOAN AMOUNT:**
 1. HIGHER LOAN AMOUNTS ARE ASSOCIATED WITH AN **INCREASED LIKELIHOOD OF DEFAULT**. THIS SUGGESTS THAT BORROWERS WHO TAKE LARGER LOANS MAY FACE MORE FINANCIAL STRAIN, LEADING TO A HIGHER RISK OF NON-REPAYMENT.
 2. **INTEREST RATE:**
 1. LOANS WITH **HIGHER INTEREST RATES** ARE MORE PRONE TO DEFAULT. BORROWERS WITH HIGHER INTEREST RATES MAY HAVE A LOWER CREDITWORTHINESS, LEADING TO HIGHER RISKS FOR THE COMPANY.
 3. **LOAN TERM:**
 1. LOANS WITH A **60-MONTH TERM** ARE AT A **HIGHER RISK OF DEFAULT** COMPARED TO LOANS WITH A 36-MONTH TERM. LONGER REPAYMENT PERIODS MAY CREATE MORE OPPORTUNITIES FOR FINANCIAL DISTRESS OVER TIME.
 4. **GRADE:**
 1. **LOAN DEFAULTS GRADUALLY INCREASE** AS THE GRADE WORSENS FROM **A TO G**. LOANS WITH HIGHER GRADES (CLOSER TO A) TEND TO PERFORM BETTER, WHEREAS LOANS WITH GRADES CLOSER TO G ARE RISKIER.
 5. **SUB-GRADE:**
 1. SIMILAR TO THE OVERALL GRADE, **SUB-GRADE** ANALYSIS REVEALS THAT **LOAN DEFAULTS INCREASE** AS THE SUB-GRADE WORSENS FROM **A1 TO G5**, WITH SOME EXCEPTIONS. THIS PROVIDES A MORE GRANULAR VIEW OF CREDITWORTHINESS.
 6. **LOAN PURPOSE:**
 1. CERTAIN LOAN PURPOSES, SPECIFICALLY **SMALL BUSINESS LOANS**, EXHIBIT A **HIGHER RISK OF DEFAULT**. THIS COULD BE DUE TO THE VOLATILITY AND RISK ASSOCIATED WITH SMALL BUSINESSES.
 7. **ADDRESS STATE:**
 1. **NE STATE LOANS** (LOANS ISSUED TO BORROWERS IN NORTHEASTERN U.S. STATES) APPEAR TO BE AT A **HIGHER RISK OF DEFAULT** COMPARED TO LOANS IN OTHER REGIONS. THIS MAY POINT TO REGION-SPECIFIC ECONOMIC FACTORS AFFECTING BORROWERS' ABILITY TO REPAY.
- IN SUMMARY, BY CONSIDERING THESE DRIVER VARIABLES IN THE LOAN APPROVAL PROCESS, THE COMPANY CAN IMPLEMENT MORE EFFECTIVE RISK MITIGATION STRATEGIES, SUCH AS ADJUSTING LOAN TERMS, INTEREST RATES, OR EVEN DENYING LOANS IN CASES OF HIGH RISK. THESE ACTIONS WILL HELP MINIMIZE FINANCIAL LOSSES AND OPTIMIZE LOAN PORTFOLIO PERFORMANCE.

THANK YOU