Lending Club Case Study

Study Group

Business Understanding

Lending Club is a finance company that specializes in lending to urban customers.

When receiving a loan application, the decision to approve hinges on the applicant's profile.

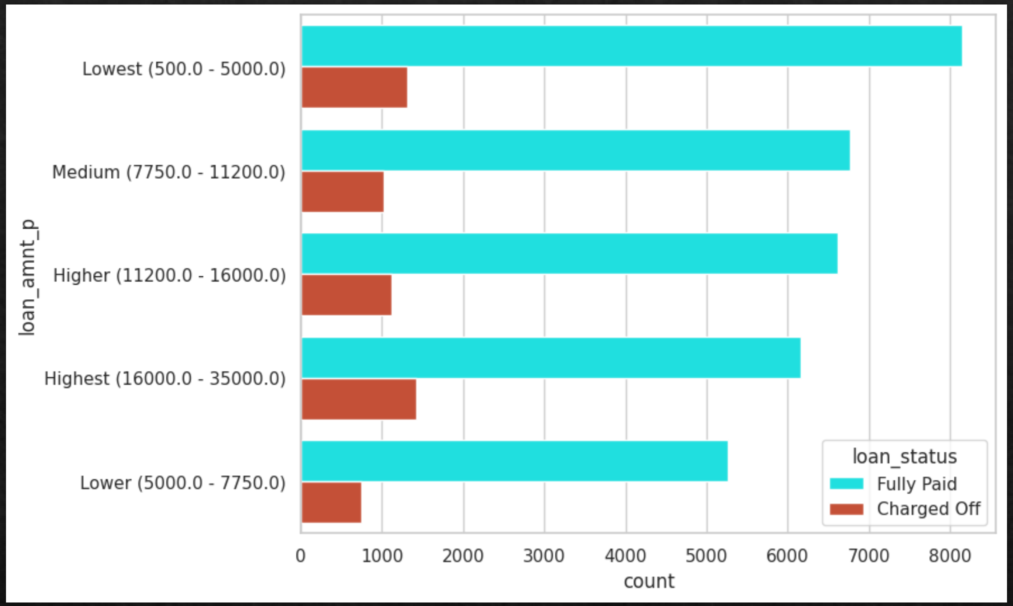
The bank faces two risk scenarios in the decision-making process:

- If the applicant is expected to repay, denying the loan results in lost business for the company.
- If the applicant is likely to default, approving the loan could lead to financial loss for the company.

Objective

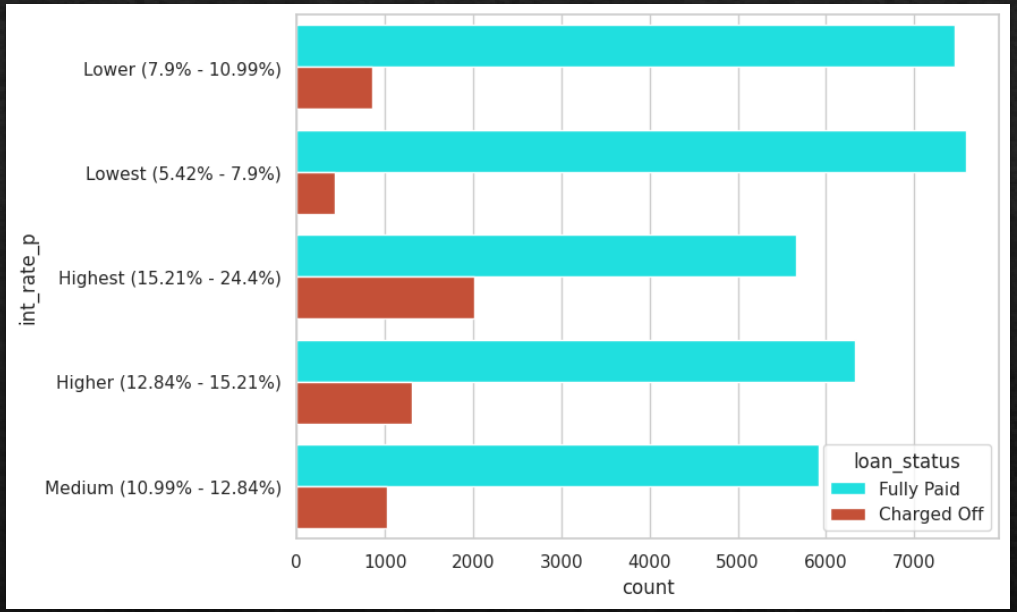
Find the driving factors which lead to the defaulted loans which are major source of loss for the company.

Loan Amount



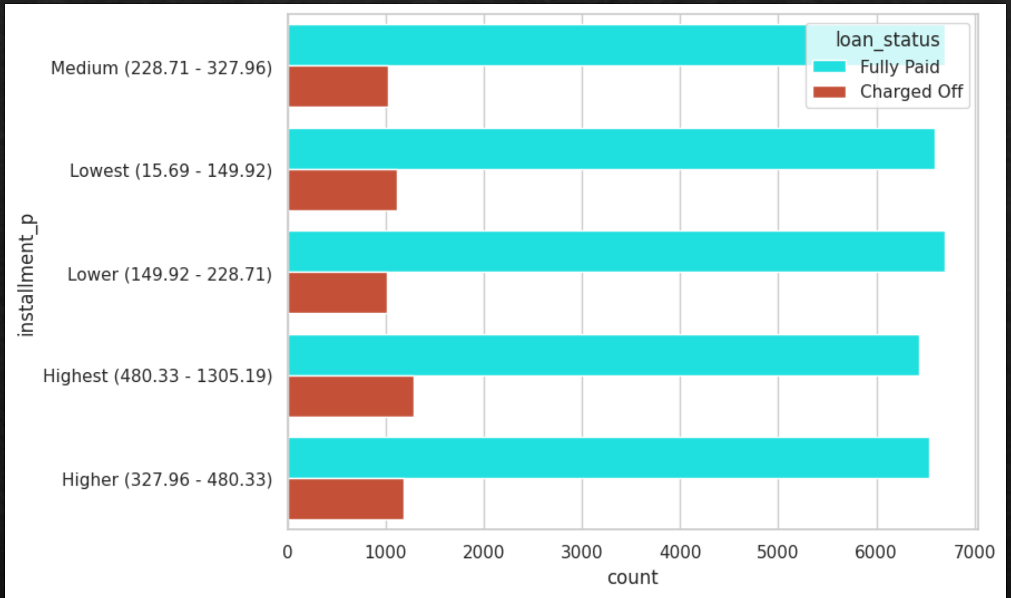
The probability of a loan defaulting increases with the loan amount.

Interest rate



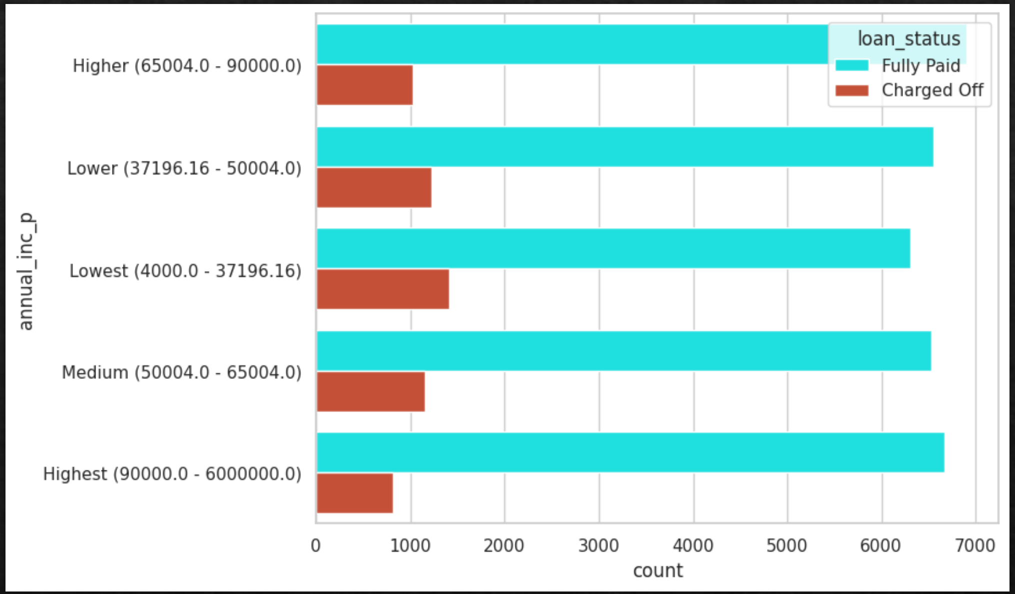
An increase in the interest rate correlates with a higher percentage of charged-off loans.

Instalment Amount



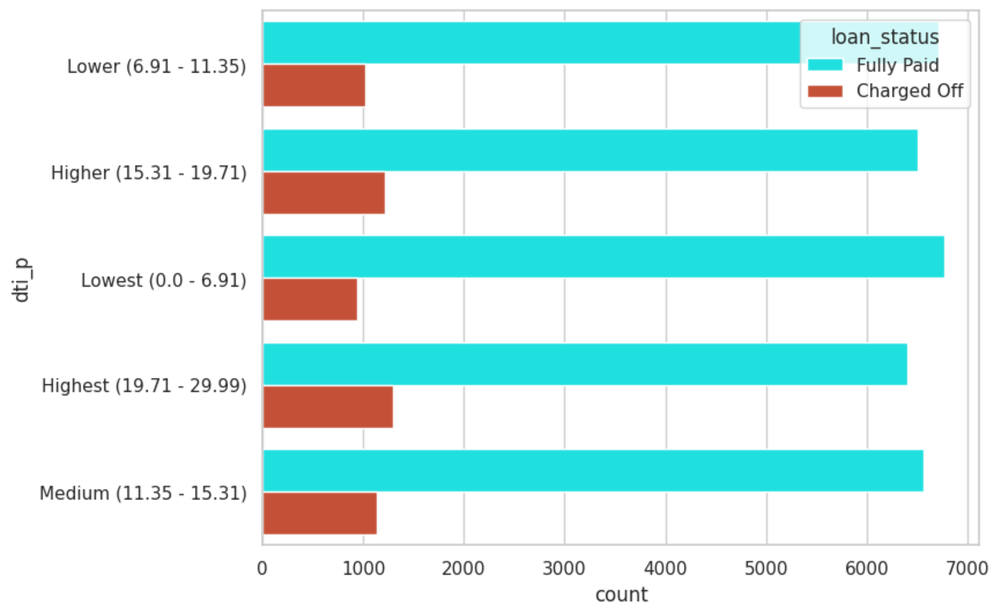
Higher installment amounts shows higher default percentages.

Annual Income



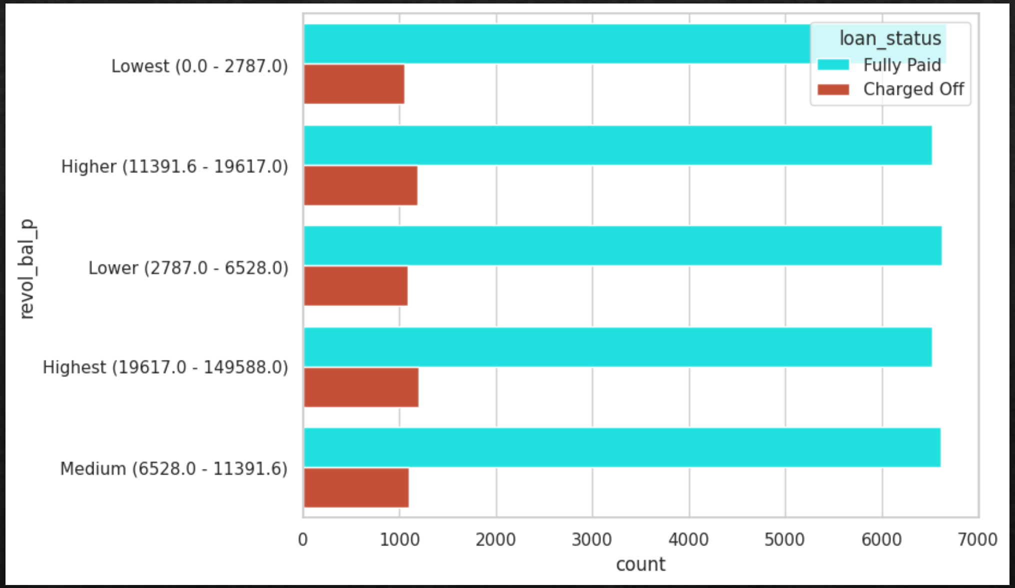
Higher the income higher the repayment %

Debt to Income Ratio



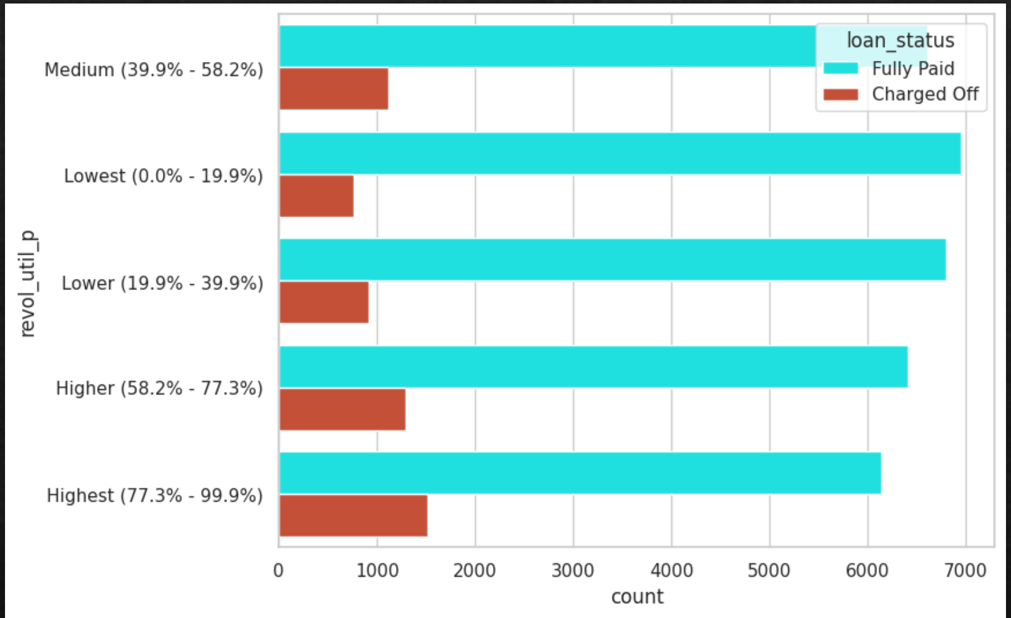
Higher DTI (debt to income ratio) will lead to higher charged off %

Revolving Balance



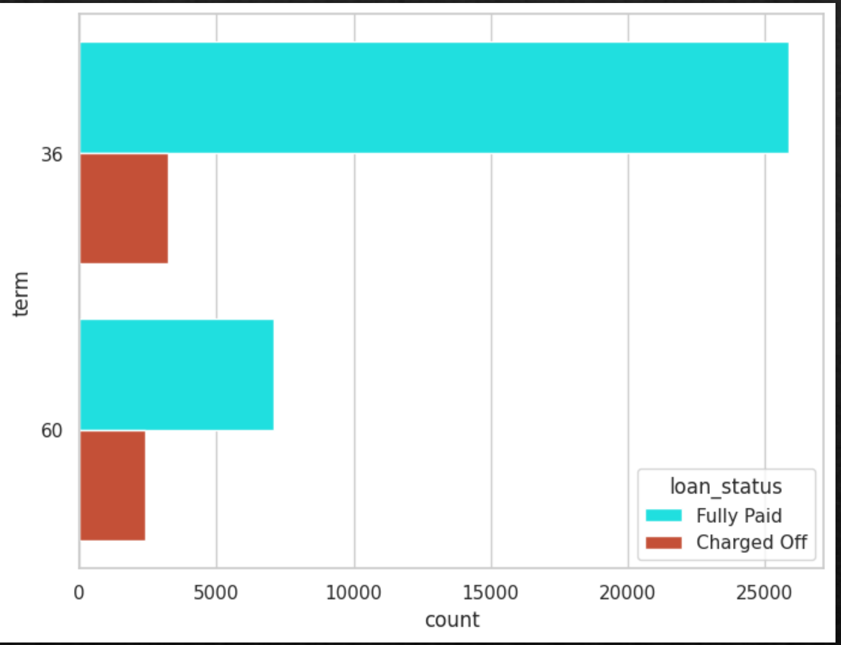
This shows the total credit revolving balances slightly influence the default percentage. Higher the revolving balance, bigger the chance of the loan getting defaulted.

Revolving Line Utilization Rate



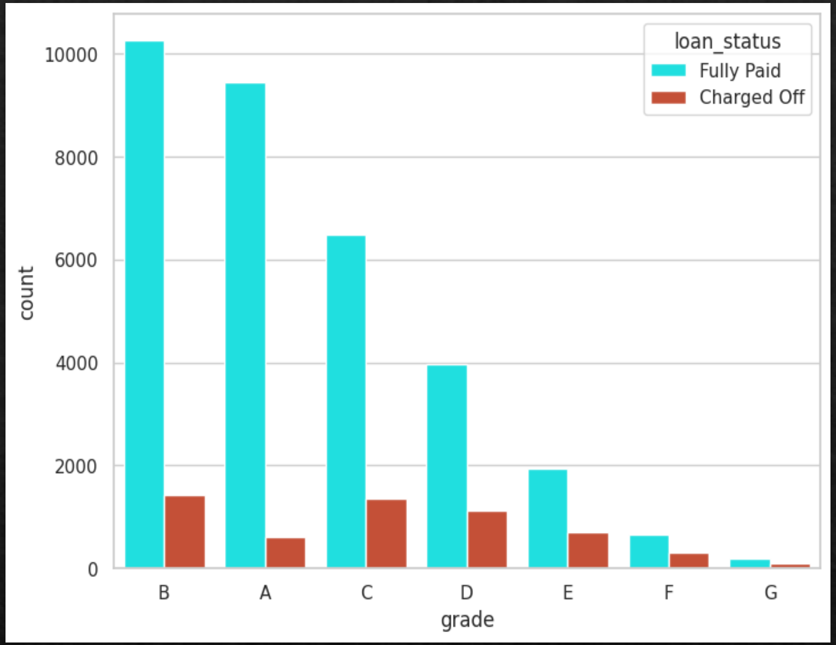
This data demonstrates a significant correlation between the revolving line utilization rate and the default percentage; as the utilization rate increases, so does the charged-off percentage.

Terms



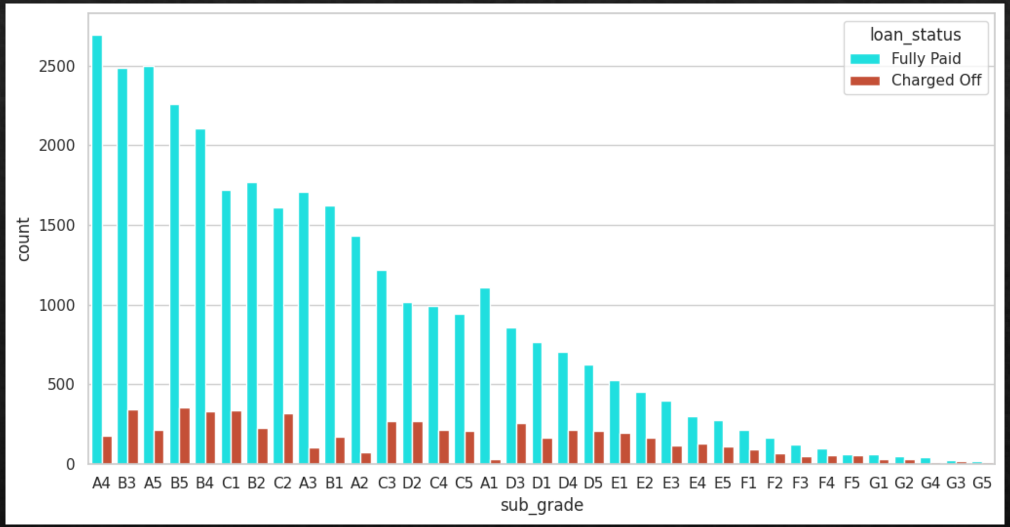
For loans with 5 year repayment term, the default percent is 25%. And for 3 year loan repayment term, the default is only for 11% of the cases. Therefore, loan repayment term plays a factor in judging the default rate.

Grade



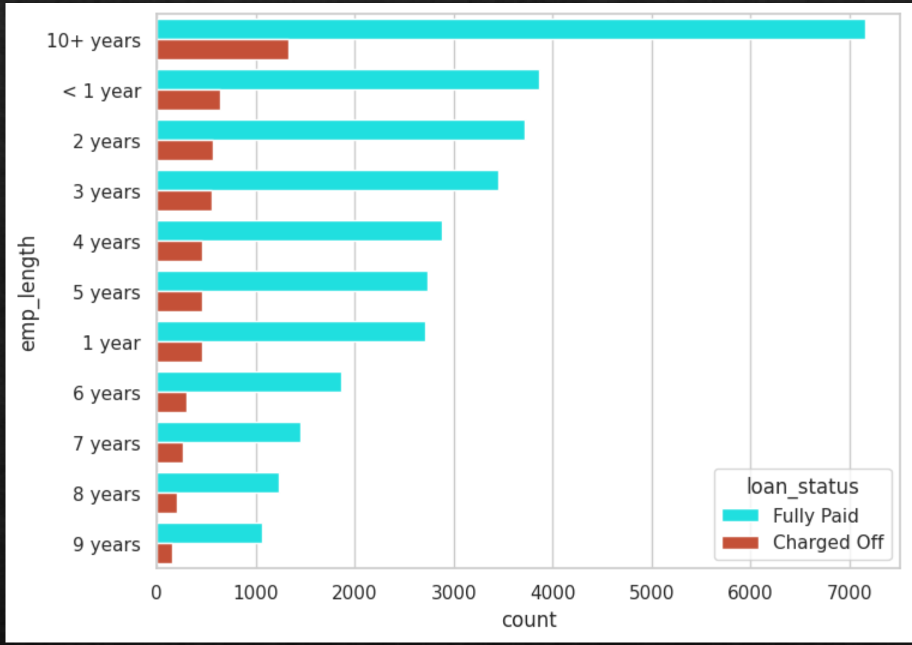
We can clearly see that loan grades having highest default percentages. G, F, E and D form grades where default rate is much higher than others.

Sub Grade



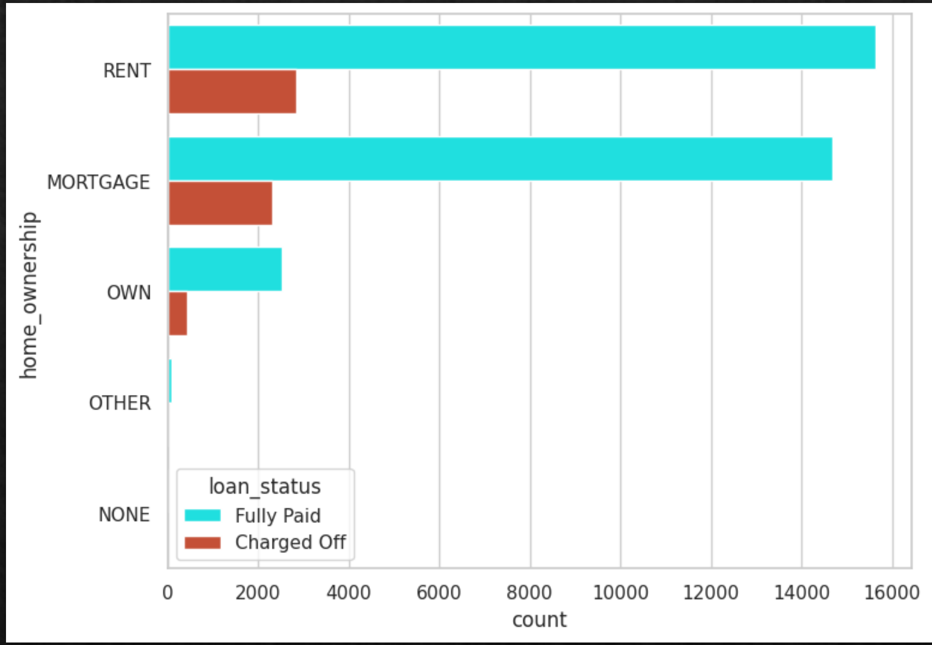
The table illustrates the relationship between loan sub-grade and default percentage. Sub-categories G3 and F5 exhibit default rates exceeding 40%, making this field a robust indicator of default likelihood.

Employment Length



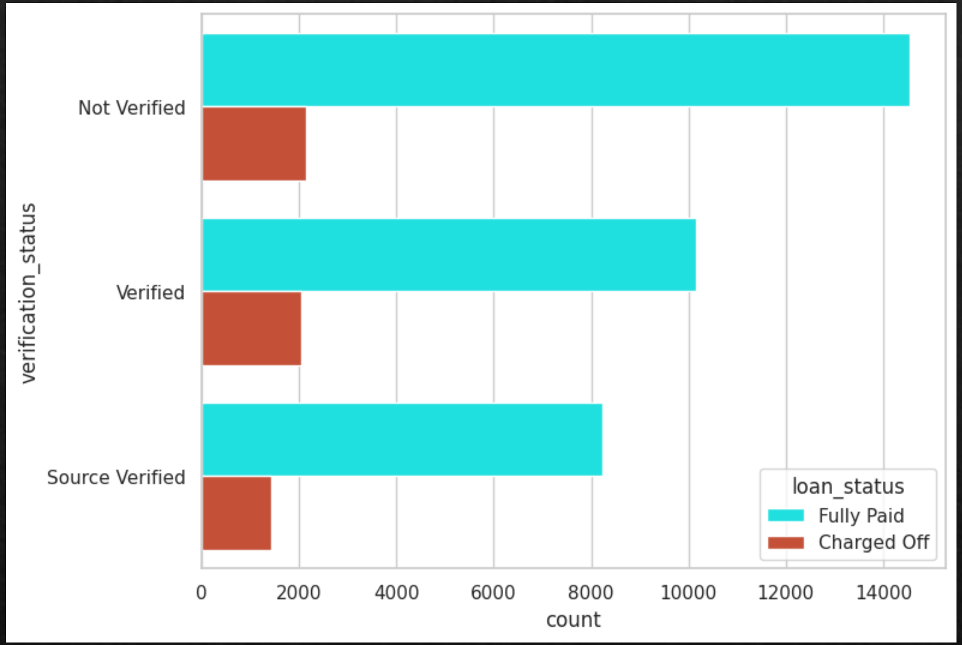
There is nothing much to conclude from this data, we can see having no employment will lead to Defaulted payment.

Home Ownership



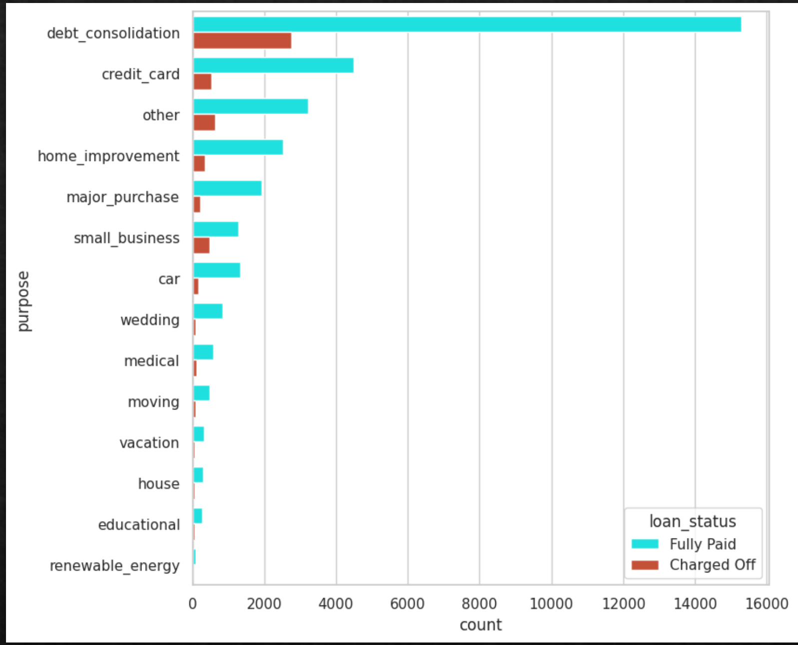
We have not much info to conclude as it says "OTHER",

Verification Status



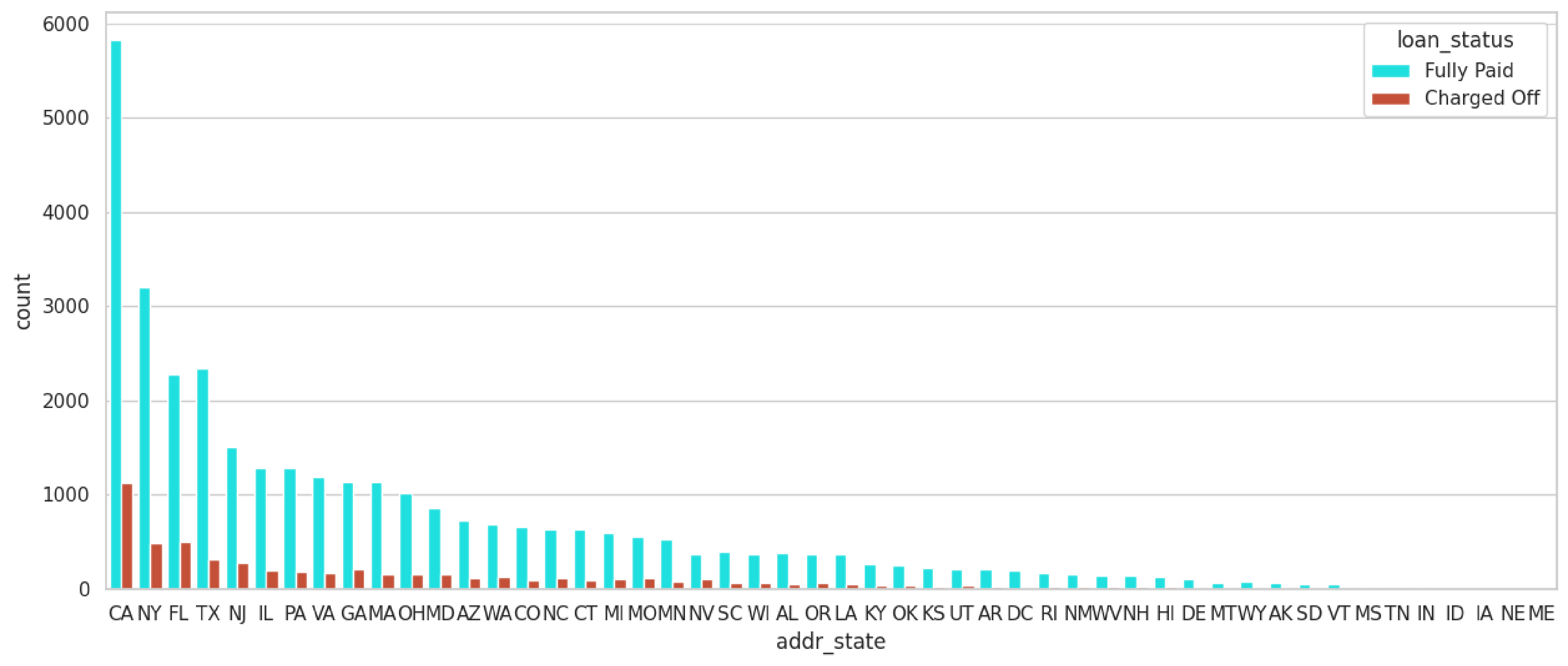
Interestingly verified applicants have more defaulted % which means the verification process is not correct.

Purpose



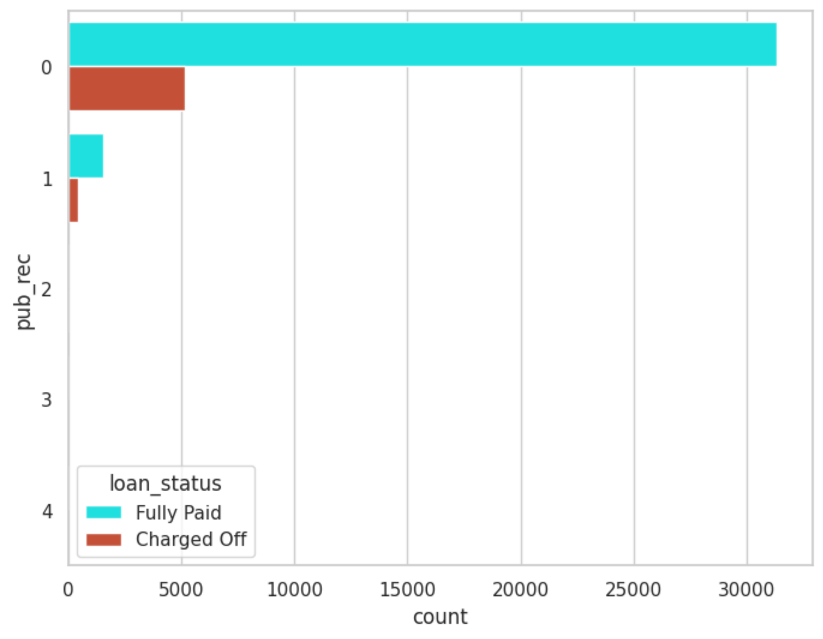
From the above analysis it is evident that the loans taken for small business, renewable energy and educational are the riskier ones.

Address States



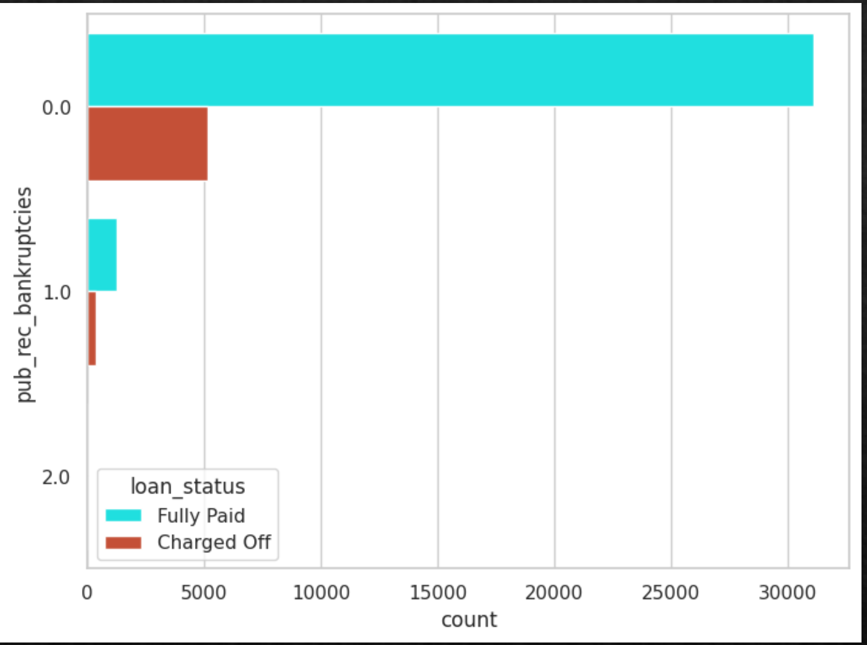
From the above table we can see that the some of the risky states are NE, NV, SD, AK, FL, MO

Public Records



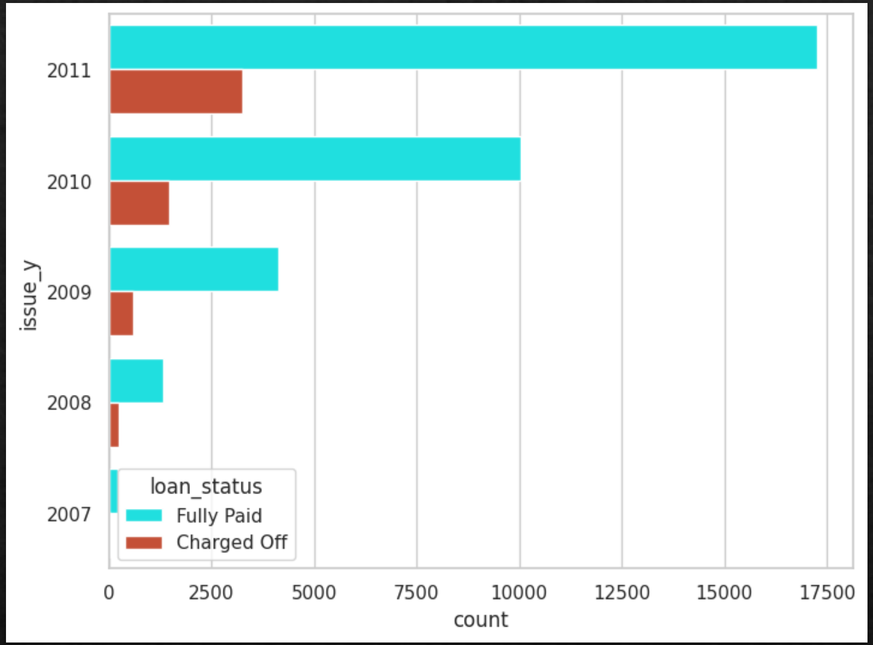
Data clearly shows that individuals with non-zero derogatory public records have higher chances of charged off.

Public Bankruptcy Records



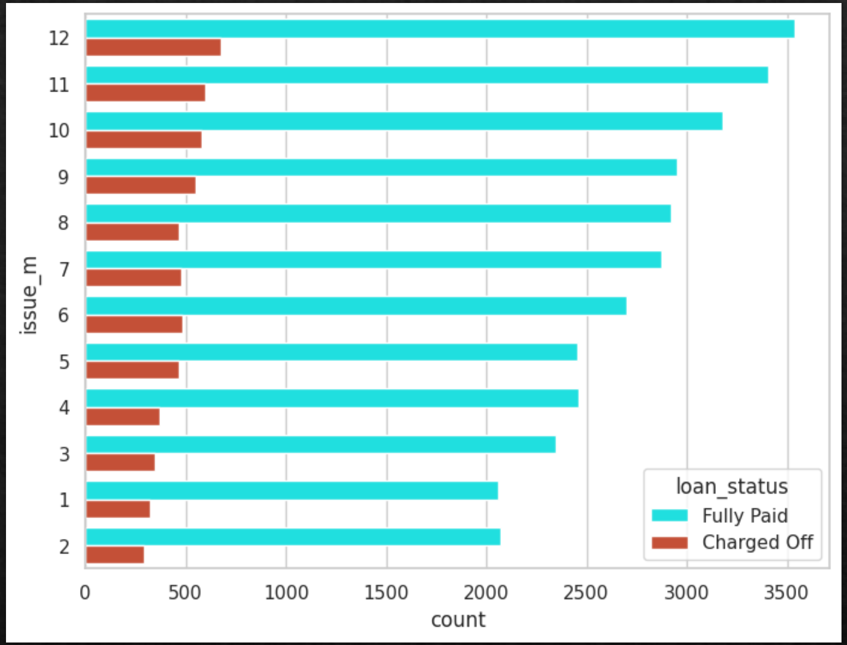
Higher the number of public bankruptcy records, bigger the chance of defaulting the loan.

Issue Year



With each passing year loan taken are increasing exponentially which indicate we are seeing large increase in DTI ratio and decrease in defaulting rate.

Issue Month



We see a gradual increase in loan taken through the year, with lesser defaulting rate in April ,August, December quarter wise and better more late in year

Annual Income Vs Instalments

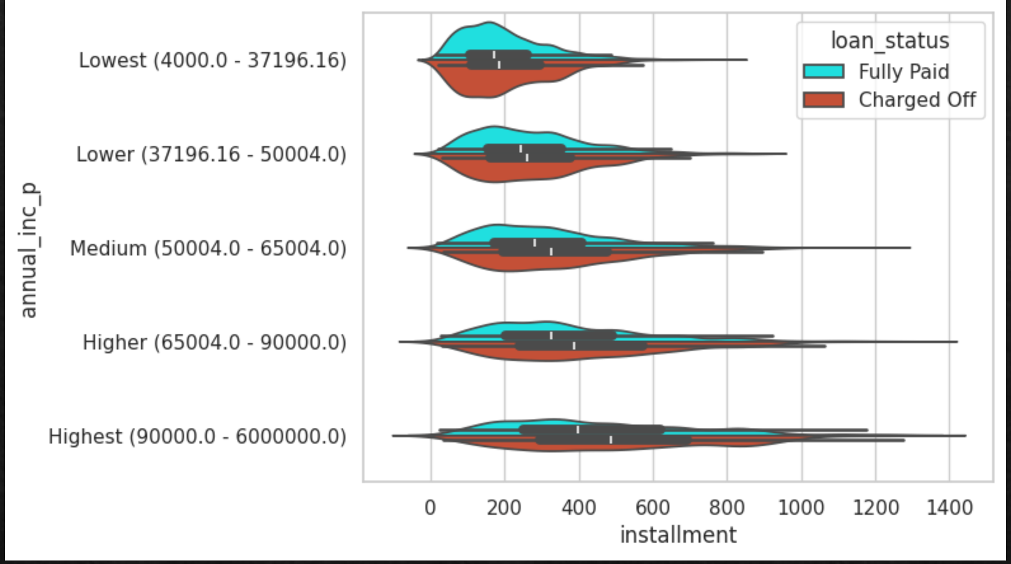
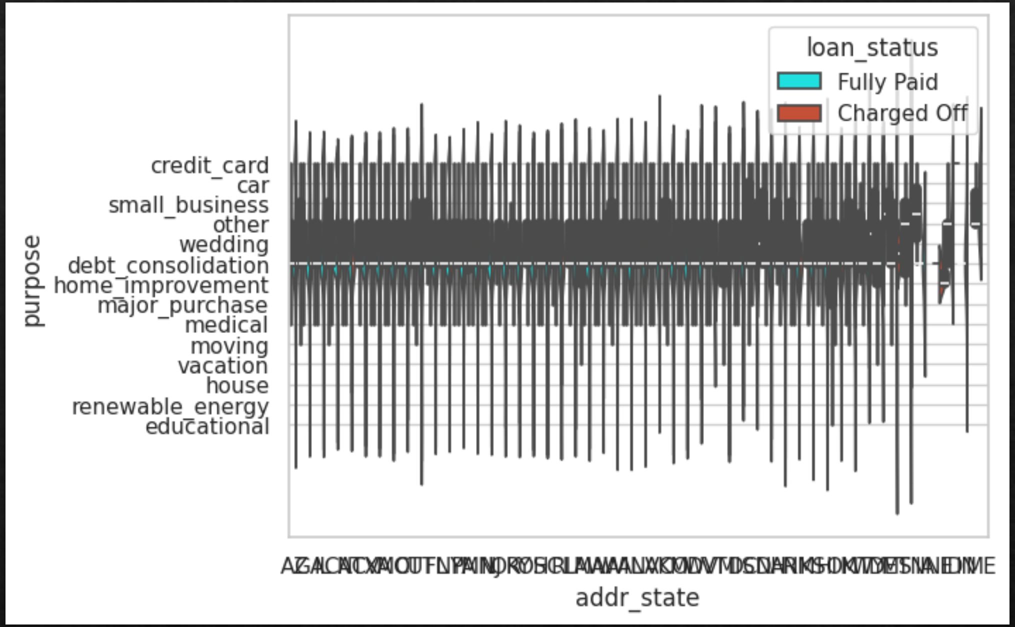


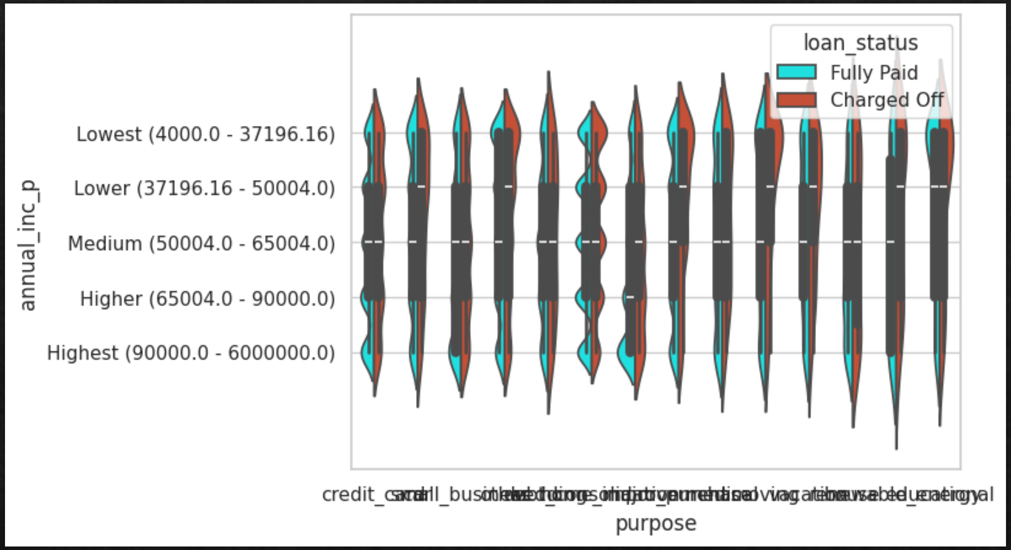
Figure shows that for higher installments for any income group have more number of defaults.

Purpose Vs Address States



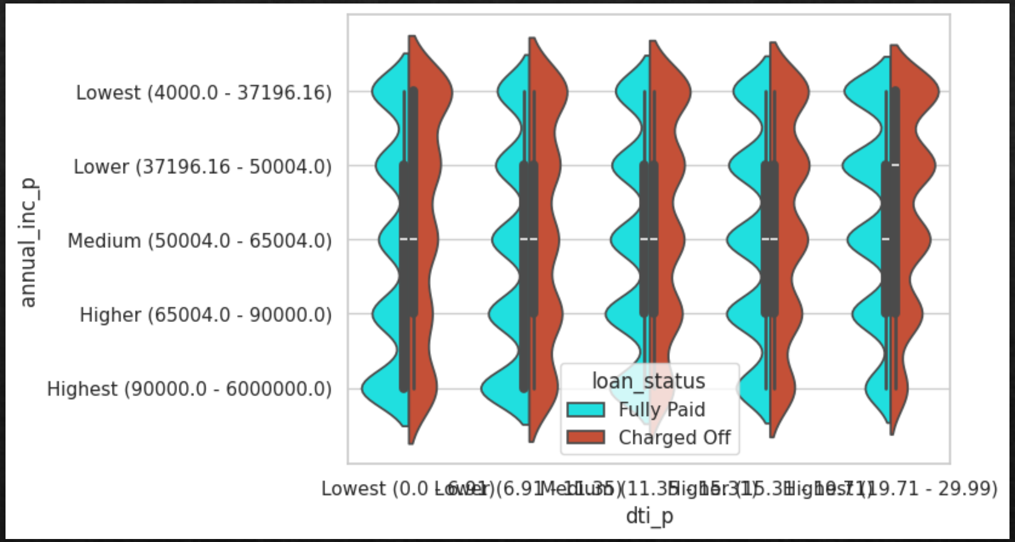
Based on the plot analysis, states with darker shading at the intersection with loan purposes indicate higher risk for loan applications.

Annual Income VS Purpose



Small business loans being riskier for the lowest and medium income groups.

Annual Income VS Debt-to-income



Medium debt-to-income group in the lowest income range is the most risky when it comes to loan repayment.

Recommendations

Certainly! Here's a more formal presentation of the key findings from the exploratory data analysis (EDA) of the loan dataset, categorizing factors by their impact on loan default:

Factors with Minor Impact

- Higher loan amount (above $16,000)
- Higher installment amount (above $327)
- Lower annual income (below $37,000)
- Higher debt-to-income ratio (above 15%)
- Applicant's address state (e.g., NV, SD, AK, FL)
- Loan issue month (Dec, May, Sep)

Factors with Heavy Impact

- Higher interest rate (above 13%)
- Higher revolving line utilization rate (above 58%)
- Repayment term of 5 years
- Loan grade & sub-grade (D to G)
- Missing employment record
- Loan purpose (small business, renewable energy, educational)
- Derogatory public records (1 or 2)
- Public bankruptcy records (1 or 2)

Factors with Combined Impact

- High loan amount & interest rate for lower income group
- High installment and longer repayment term
- Home ownership (other) and loan purpose (car, moving, or small business)
- Residential state and loan purpose
- Income group and loan purpose

This EDA identifies these factors as the most impactful drivers for loan default, providing valuable insights into risk assessment and decision-making processes within the lending industry.

“

Thank You

”

Harsh Kumar Dewangan & Sarvjeet Luthra