Lending Club Case Study

Group Members:

1. Ratnakar Maurya
2. Meghna Majumder

The problem

Company

Lending Club stands as the most prominent digital platform for loans, enabling individuals to secure personal and business loans, as well as funds for medical treatments.

Through a swift online portal, borrowers can conveniently obtain loans at reduced interest rates

Context

Lending Club seeks to comprehend the key elements that contribute to loan defaults, namely the influential factors that strongly signal the likelihood of default.

By harnessing this knowledge, the company can effectively employ it in portfolio management and risk evaluation..

Problem statement

In the capacity of a data scientist at Lending Club, the individual is assigned the duty of analyzing a dataset encompassing details concerning previous loan applicants. Through the application of exploratory data analysis (EDA), insights are sought into the manner in which consumer attributes and loan-specific characteristics intersect, ultimately shaping the propensity for loan default.

Analysis Approach

> Drop columns with null values, all random values or single category value
> Convert values to proper int, float, date representations

> Analyze variables against segments of other variables
> Create derived variables

Publish insights and observations

Clean Data → Univariate Analysis → Segmented Univariate Analysis → Bivariate Analysis → Summarize Results

> Check distributions and frequencies of various numerical and categorical variables
> Create derived variables

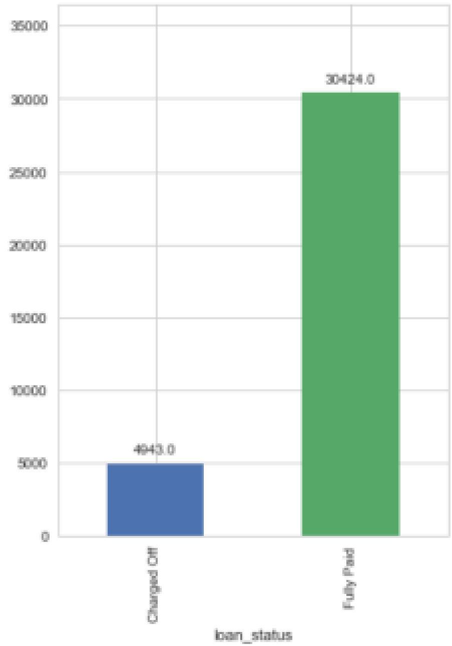
> Do correlation analysis
Check how two variables affect each other or a third variable
> Analyze joint distributions

Analysis - Overall Loan Status

Total Loans

Approximately 14% of loans are defaulted

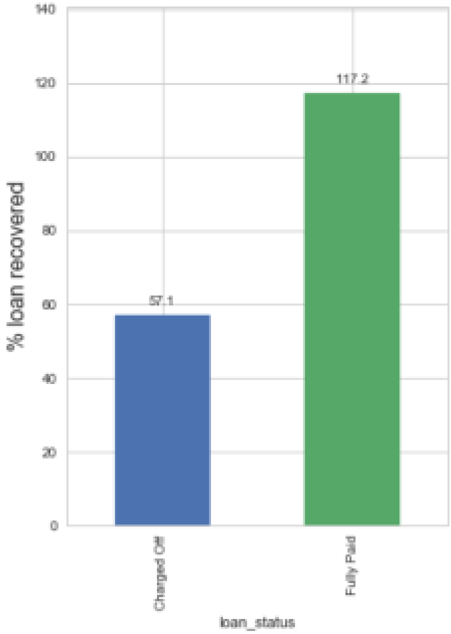
A variable that raises the default percentage beyond 16.5% ought to be regarded as a potential business risk.



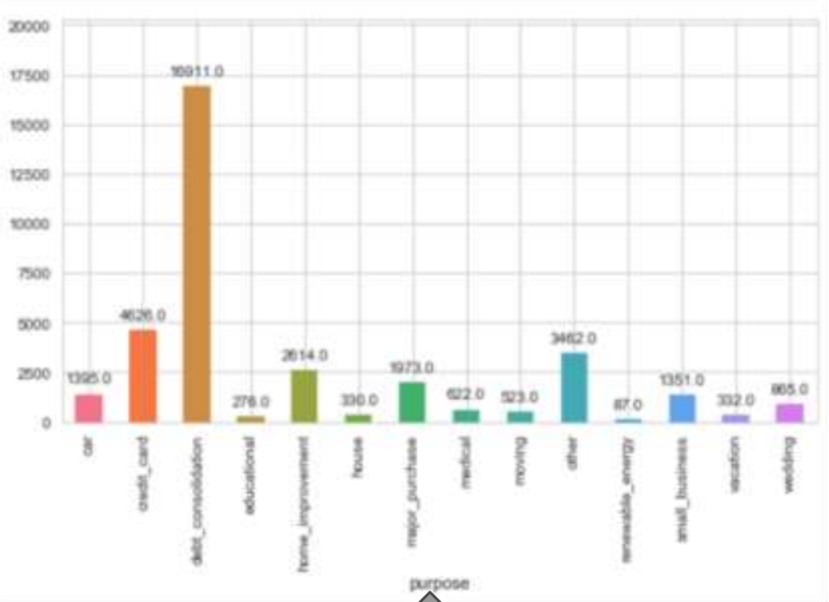
Total Money Earned

Lending Club only recovers 57% of the loan amount when loans are **defaulted**.

On fully paid up loans, the company makes 17% profit.

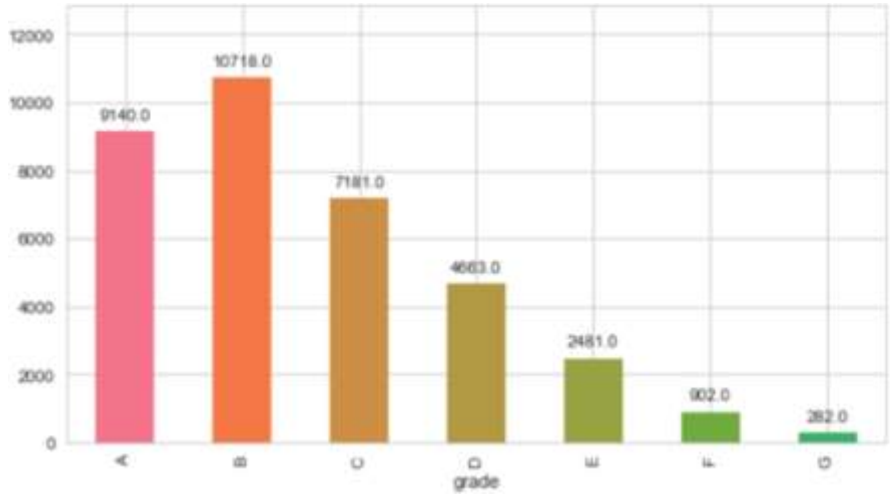


Analyzing Loans - Gaining Insight



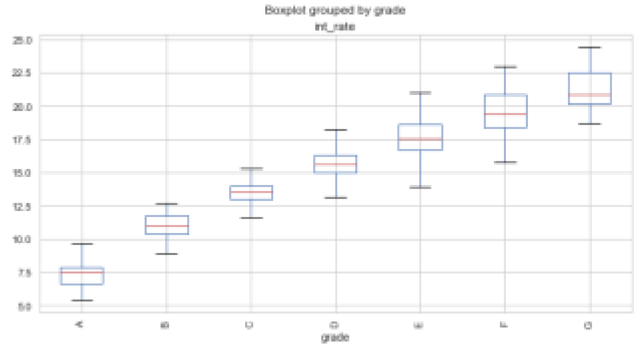
Maximum number of loans are for debt consolidation, followed by credit card

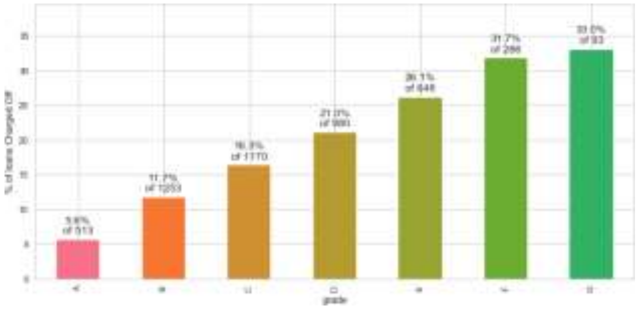
Most loans are of high quality, falling under grade A or B.



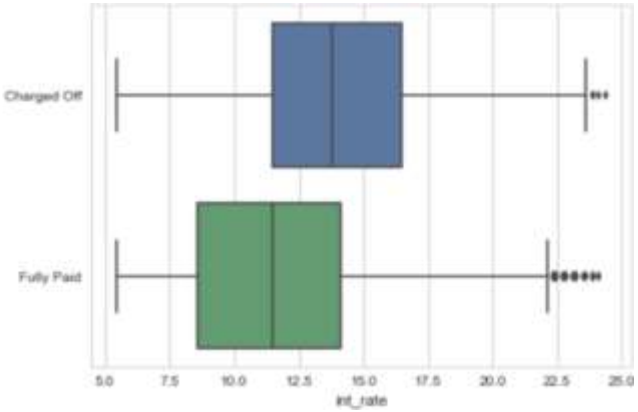
Analyzing Loans - Gaining Insight Continued

- Loans with lower grades experience a greater frequency of defaults, indicating the effectiveness of the grading system. Lending Club applies elevated interest rates as loan grades decline.
- However, as demonstrated in the upcoming slide, the primary factor driving defaults is the higher interest rate.

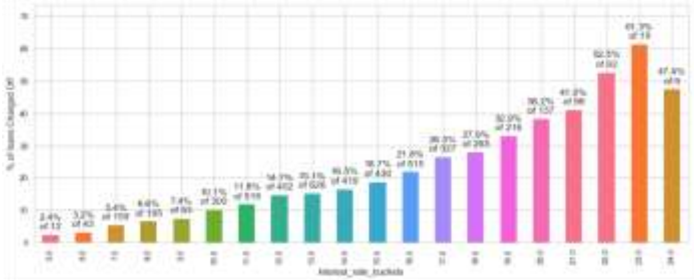


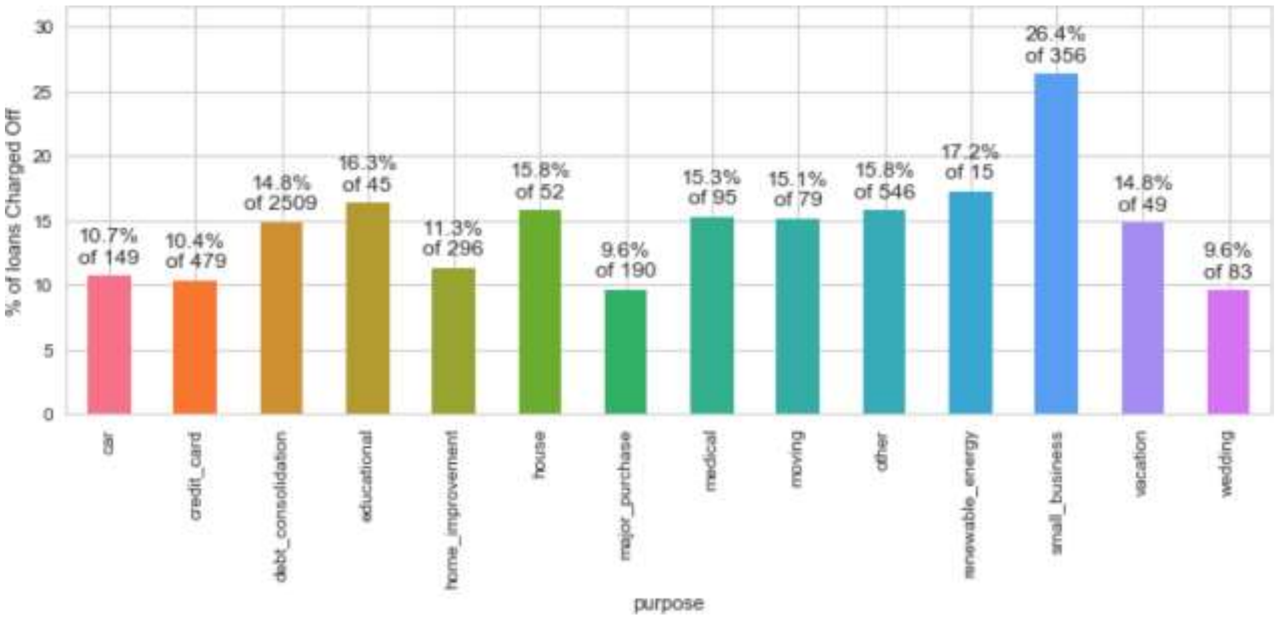


Analysis - Defaults by Interest Rate



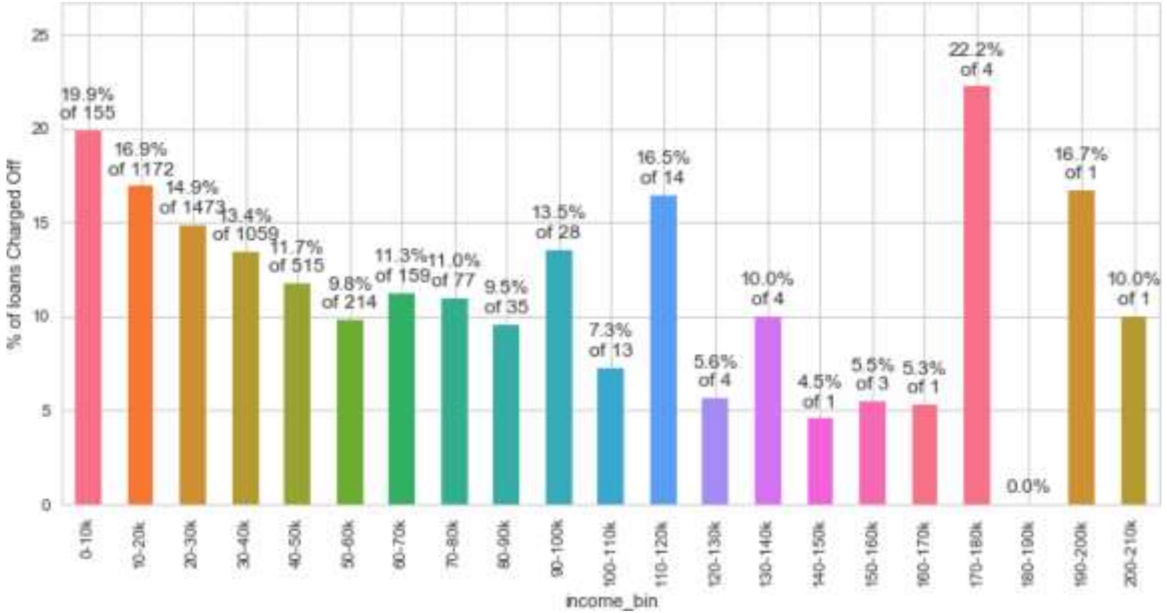
- The proportion of defaults consistently rises with increasing interest rates. For rates exceeding 19%, over 33% of loans are categorized as Charged Off.





Examining Defaults Based on Loan Purposes

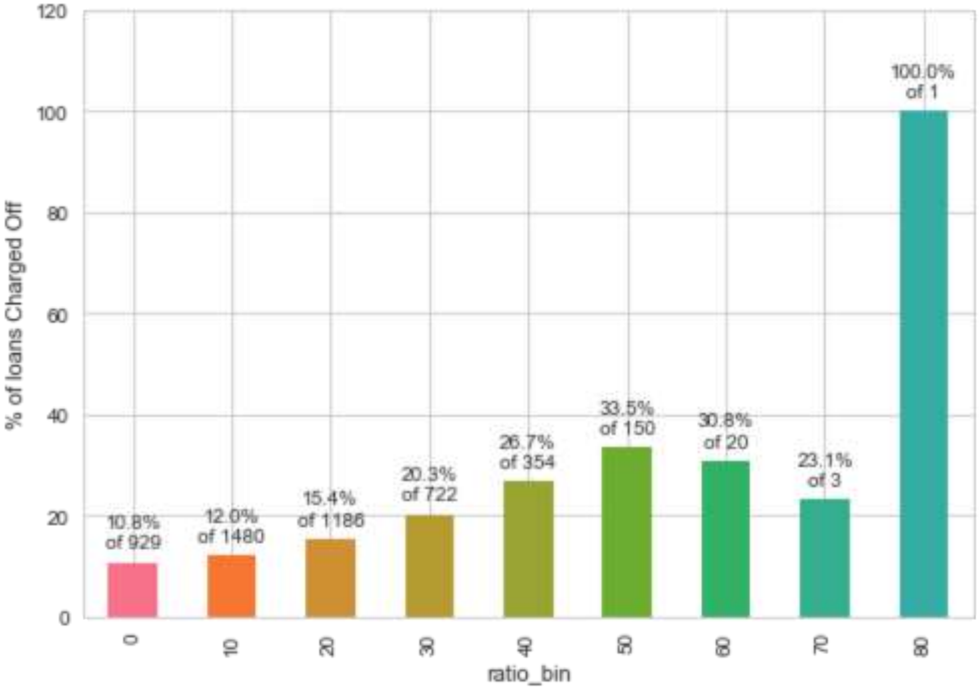
- Default rates surpass 25% for loans acquired with the intention of funding small business operations



Exploring Defaults in Relation to Borrower Income

- Borrowers with annual incomes below $20,000 exhibit notably higher loan default rates. Default rates decrease as annual income increases.
- However, as illustrated in the forthcoming slide, the ratio of loan amount to income proves to be more crucial.

Analysis - Defaults by ratio of amount to income



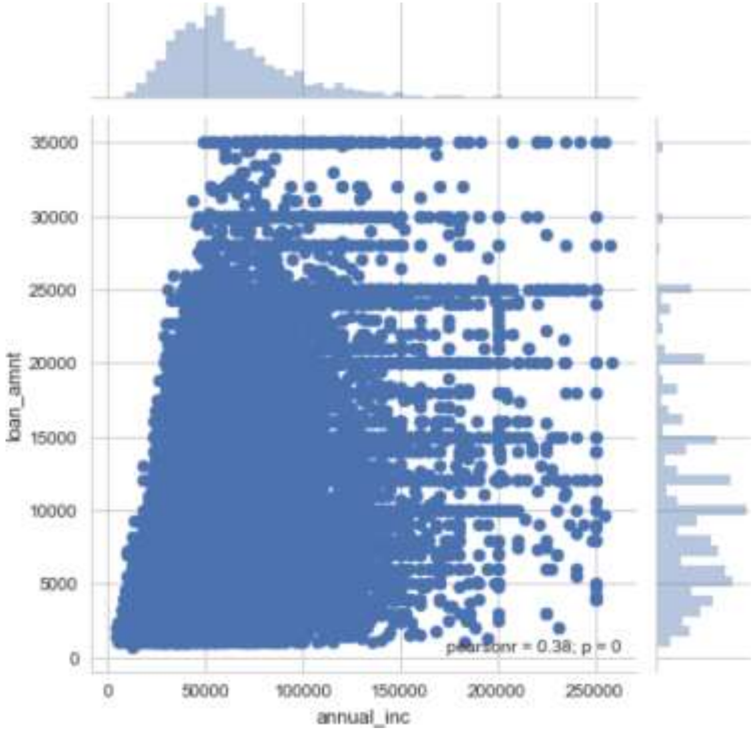
The X-axis represents the percentage of the loan amount relative to the borrower's annual income

Default rates remain low as long as the loan amount constitutes less than 20% of the borrower's annual income.

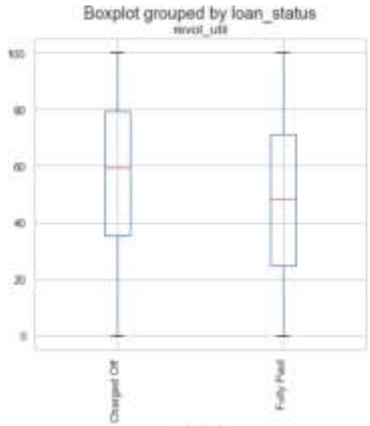
Loan amounts equivalent to or exceeding 30% of the borrower's annual income are associated with a significantly elevated default rate

Analysis - Defaults by ratio of amount to income Continued

- Lending Club has extended high-value loans to individuals with low income.
- Instances exist where individuals with an income of $50,000 or less have received loans of $25,000 or more.
- This practice of providing large loans to low-income individuals should be limited.

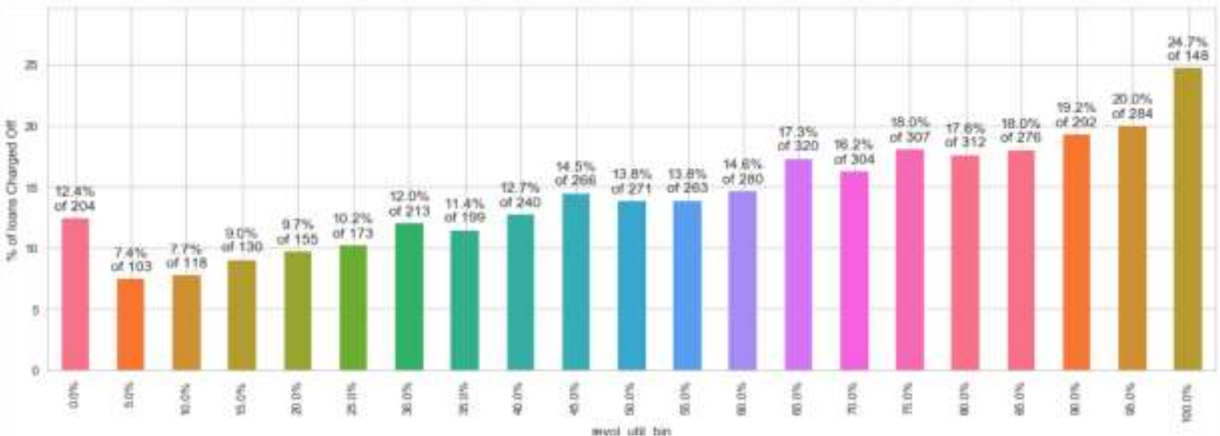


Examining Defaults Based on Revolving Line Utilization Rate

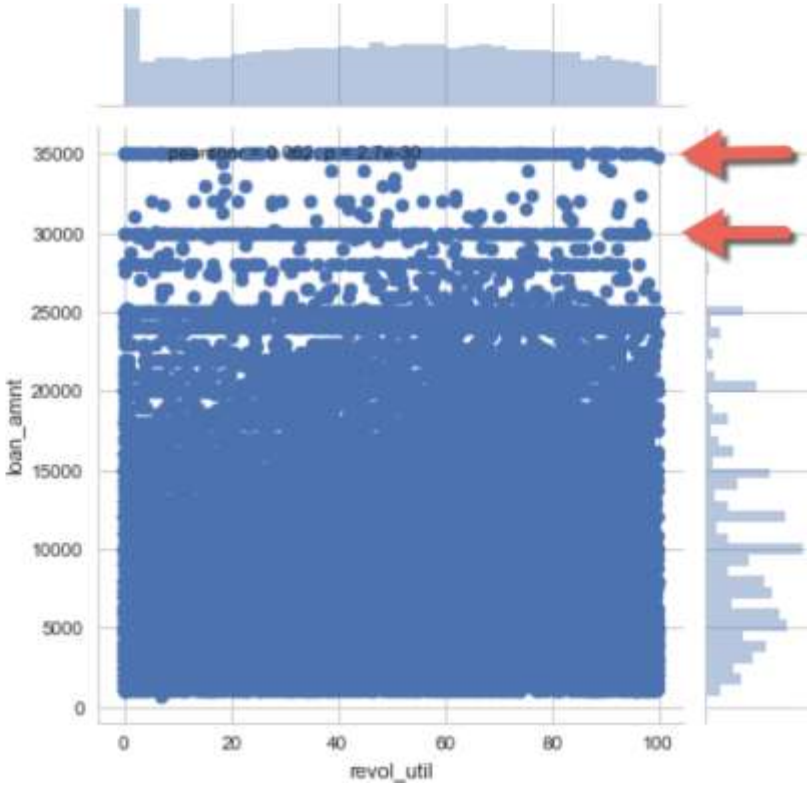


Borrowers with high utilization of Revolving Line of Credit at the loan initiation stage exhibit a higher likelihood of default.

Loans associated with a utilization rate exceeding 75% are considered risky in terms of default potential.



Analysis - Defaults by Rev Util Continued

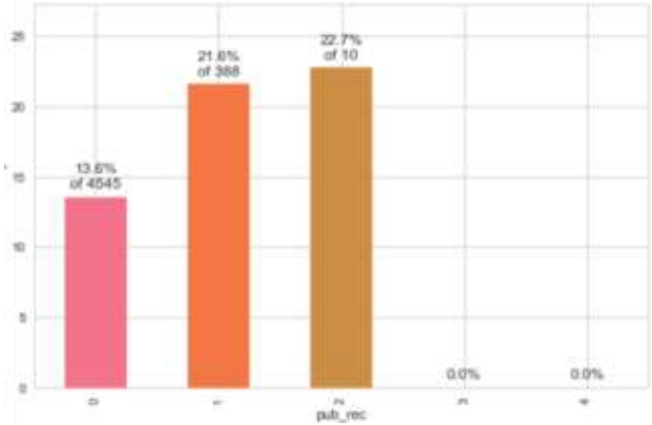


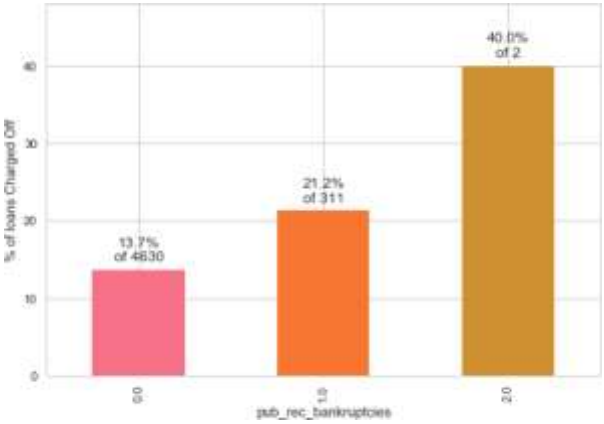
- Notably, there are instances where substantial loans have been granted to borrowers with a revolving line utilization rate surpassing 75%.
- This practice of extending such high-value loans to borrowers with elevated utilization rates should be discontinued.
- Moreover, there exists a concentrated occurrence of low-value loans. Consideration should be given to reducing the frequency of approval for these lower value loans

Analyzing Defaults Based on Prior Bad Records

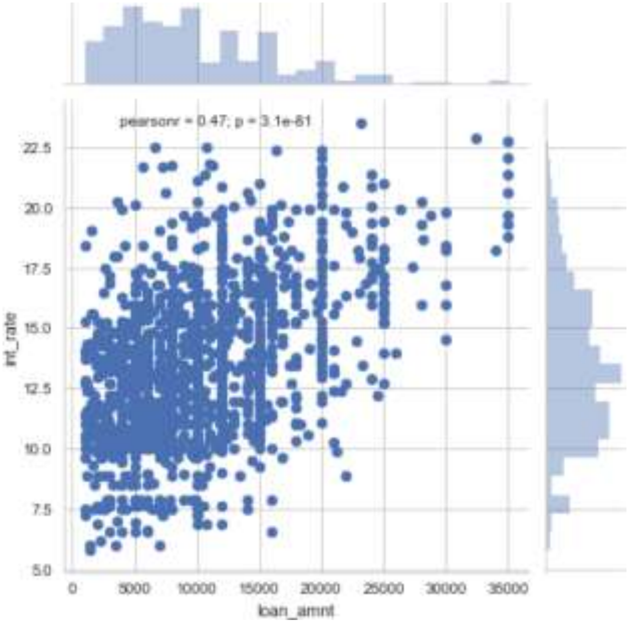
- A significant majority, approximately 94%, of borrowers have no records of public derogatory incidents.
- Even the presence of a single derogatory record considerably amplifies the likelihood of Charge Off.

- A substantial majority, around 96%, of borrowers do not have any bankruptcy records.
- The presence of just one bankruptcy record substantially elevates the probability of Charge Off.
- Notably, there exists an 83% correlation between Public Derogatory Records and Public Bankruptcy records. Either of these variables can be used effectively for analysis.





Analyzing Defaults Based on Prior Bad Records-Continued



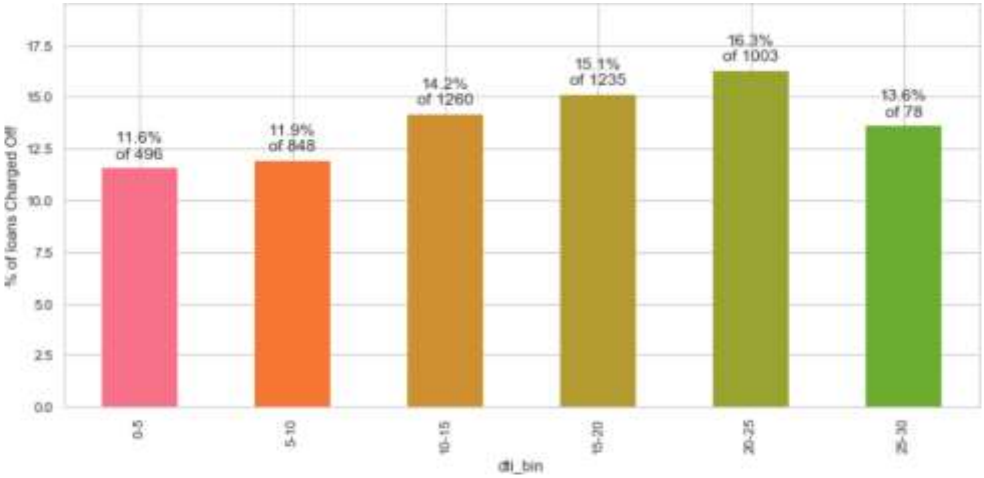
Data of people with >0 bad records

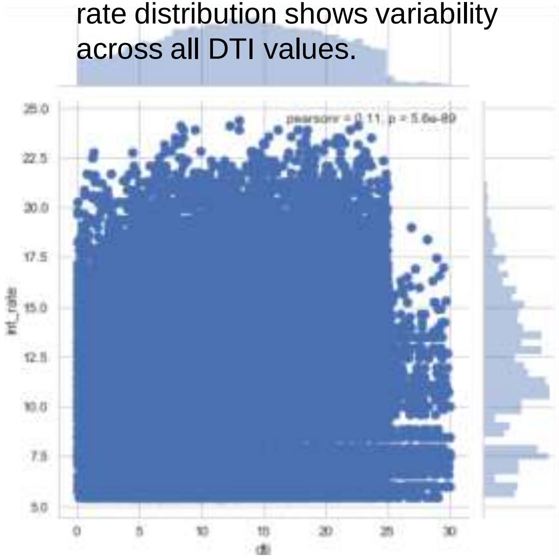
- The practice of extending both high-value loans and low-interest loans to individuals with prior public derogatory records has been observed.
- Ceasing this practice could lead to enhanced business metrics and outcomes.

Analyzing Defaults Based on Debt-to-Income Ratio

- The percentage of defaults experiences an increase in tandem with the Debt-to-Income (DTI) ratio.
- As the DTI ratio surpasses 20, loans begin to exhibit elevated risk levels.

It would be advisable to apply higher interest rates for borrowers with higher DTI ratios. However, the observed interest rate distribution shows variability across all DTI values.





Suggested Steps: Attracting Higher-Quality Borrowers

1. Cease approvals for loans where the loan amount as a percentage of income exceeds 30%.
2. Decrease the number of loan approvals specifically for small business purposes.
3. Discontinue approvals for high-value loans when the revolving line utilization rate surpasses 75%.
4. Refrain from approving loans for individuals with prior derogatory records, or at the very least, avoid high-value loans for such individuals.
5. Initiate the practice of charging higher interest rates for loans with Debt-to-Income (DTI) ratios exceeding 20%.