Lending Club Case Study

- Chandra Sekhar Kommuri
- Vikrant Yadav

Business Understanding

- Lending Club is a consumer finance company which specializes in lending various types of loans to urban customers.

- When the company receives a loan application, the company has to make a decision for loan approval based on the applicant's profile.

- Two types of risks are associated with the bank's decision:
  - If the applicant is likely to repay the loan, then not approving the loan results in a **loss of business** to the company
  - If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a **financial loss** for the company

Business Objectives

Minimize Credit Loss

Identify and reduce the number of loans issued to high-risk applicants.

Improve Risk Assessment

Utilize EDA to uncover patterns and consumer attributes that are strong indicators of default.

Strategic Loan Management

Implement data-driven strategies to mitigate potential losses, thereby enhancing the company's financial stability. For example: *Denying loans, Reducing loan amounts, Modifying interest rates and Adjusting loan term*

Data Understanding

Customer Information

- Annual Income
- Employment length
- Home Ownership
- Address State

Loan Information

- Purpose
- Loan Amount
- Funded Amount
- Interest Rate
- Installment
- Loan Term
- Grade/Sub-Grade
- Loan Status

Credit History & Delinquency Metrics

- Delinquents in 2 years
- Open Account
- Total Account
- Revolving Utilization
- Credit enquires in last 6 months

Approach

Data Acquisition

- ***Data Source***
  - Loan dataset from Lending Club.
- ***Data Description***
  - Original dataset shape - (39,717 rows, 111 columns)
  - Data is provided for 5 years from 2007 to 2011 in CSV format.

Data Preparation

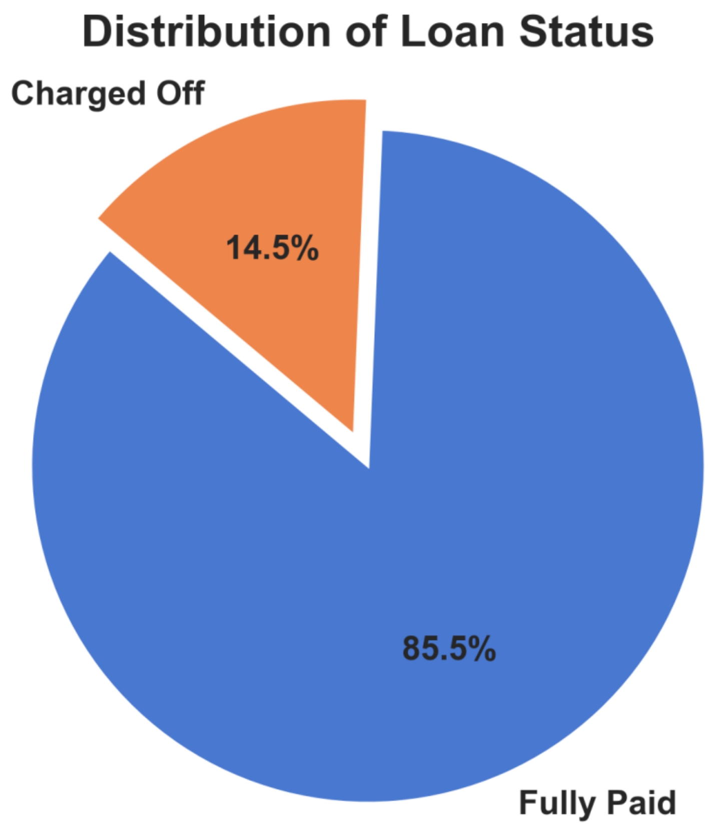
- ***Data Cleaning***
  - Dropping rows/columns that are completely null
  - Missing values treatment
  - Handling outliers
  - Filtering data
- ***Data Transformation***
  - Standardize numbers
  - Derive new columns

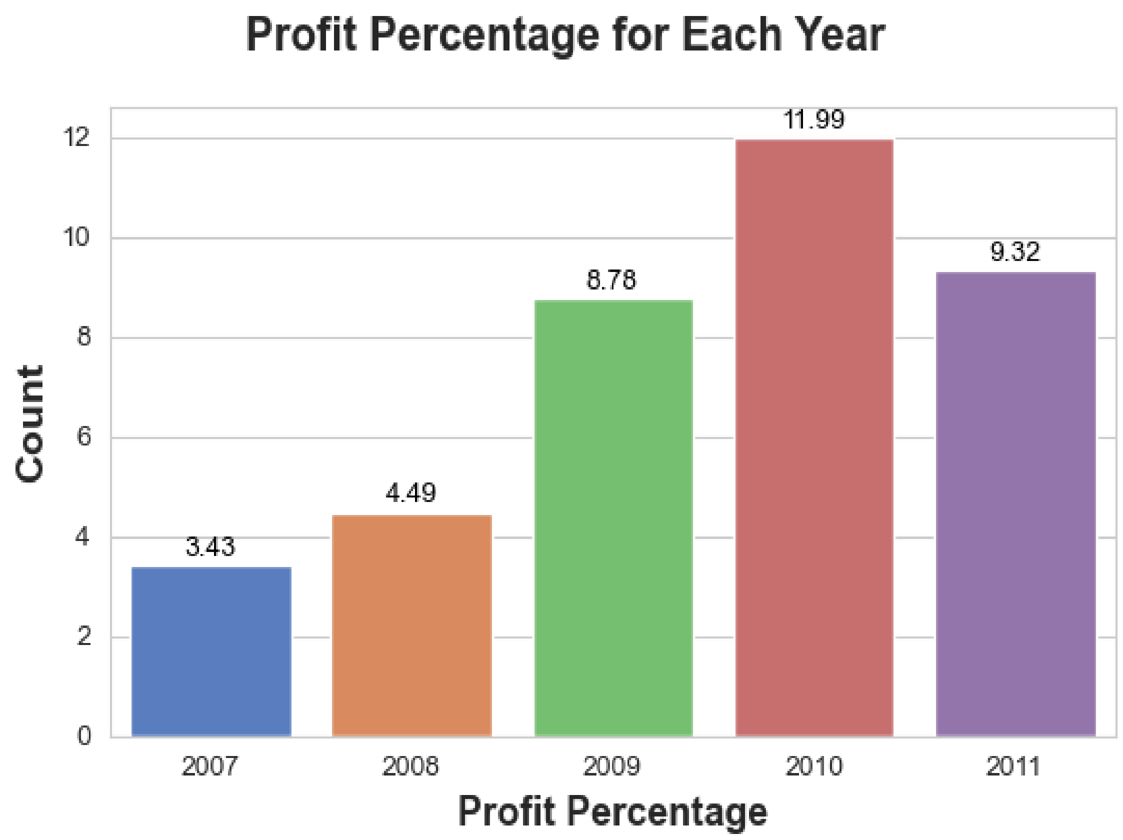
Exploratory Data Analysis(EDA)

- ***Univariate Analysis***
  - Categorical
  - Numerical
- ***Segmented Univariate Analysis***
- ***Bivariate Analysis***
- ***Multivariate Analysis***
- ***Inferences and Recommendations***

- Though we have analyzed all the relevant attributes in the loan data, this presentation includes analysis of only attributes that are most relevant to the business.
- Analysis does not include ongoing loans. i.e, loan_status = 'Current'

Summary of Analysis





Loan Distribution / Univariate Analysis

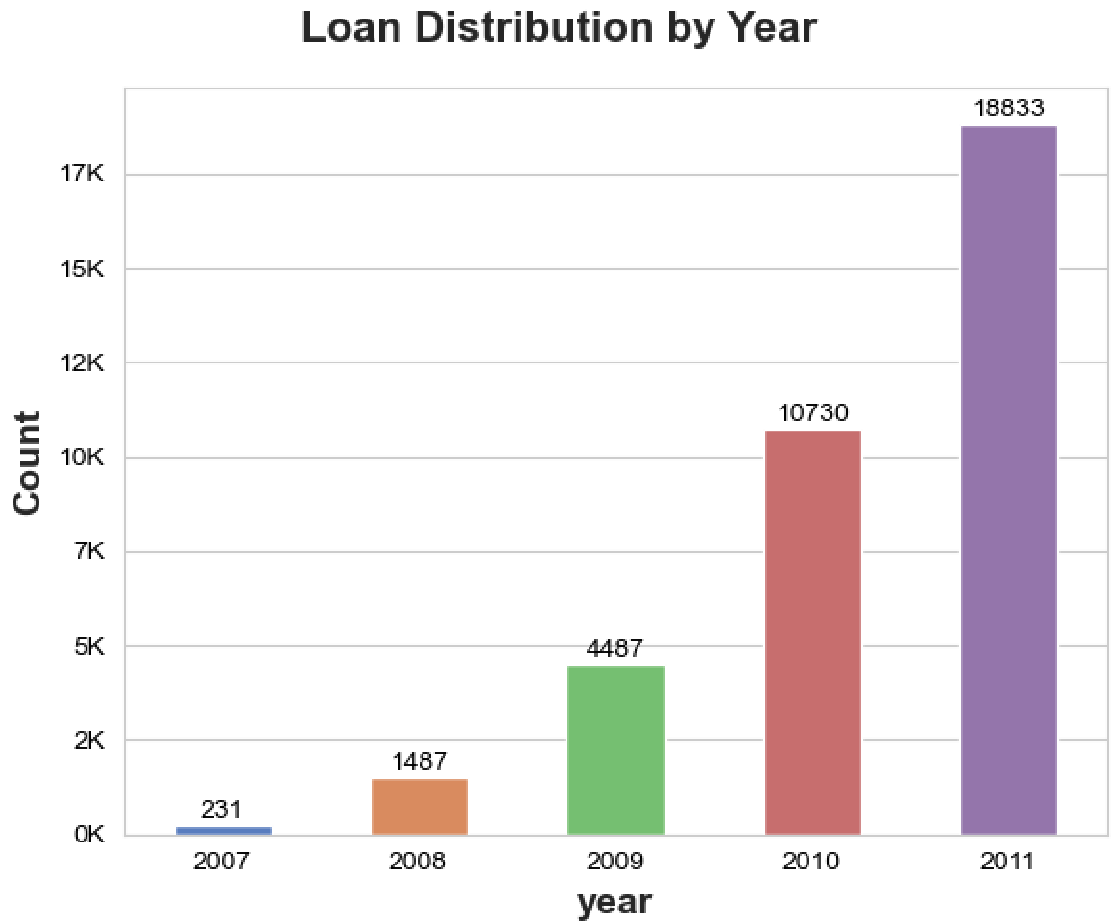
- ❖ Univariate analysis is a type of statistical analysis that involves examining one variable or characteristic at a time.

- ❖ This type of analysis is often used to describe patterns or distributions within the data and to understand the characteristics of the variable.

- ❖ We used this technique to understand the loan distributions across various attributes.

Year

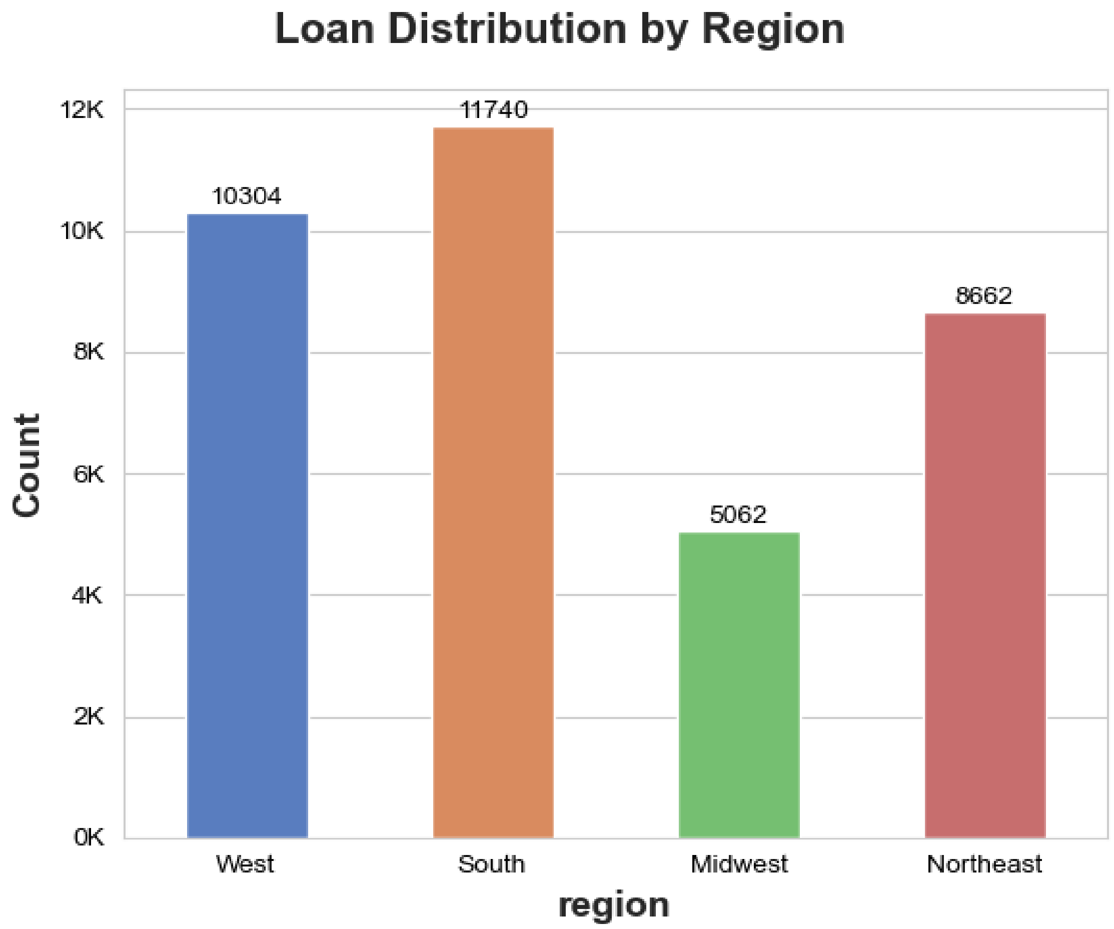
- ❖ There is an exponential increase in loan distribution for every year.



Region

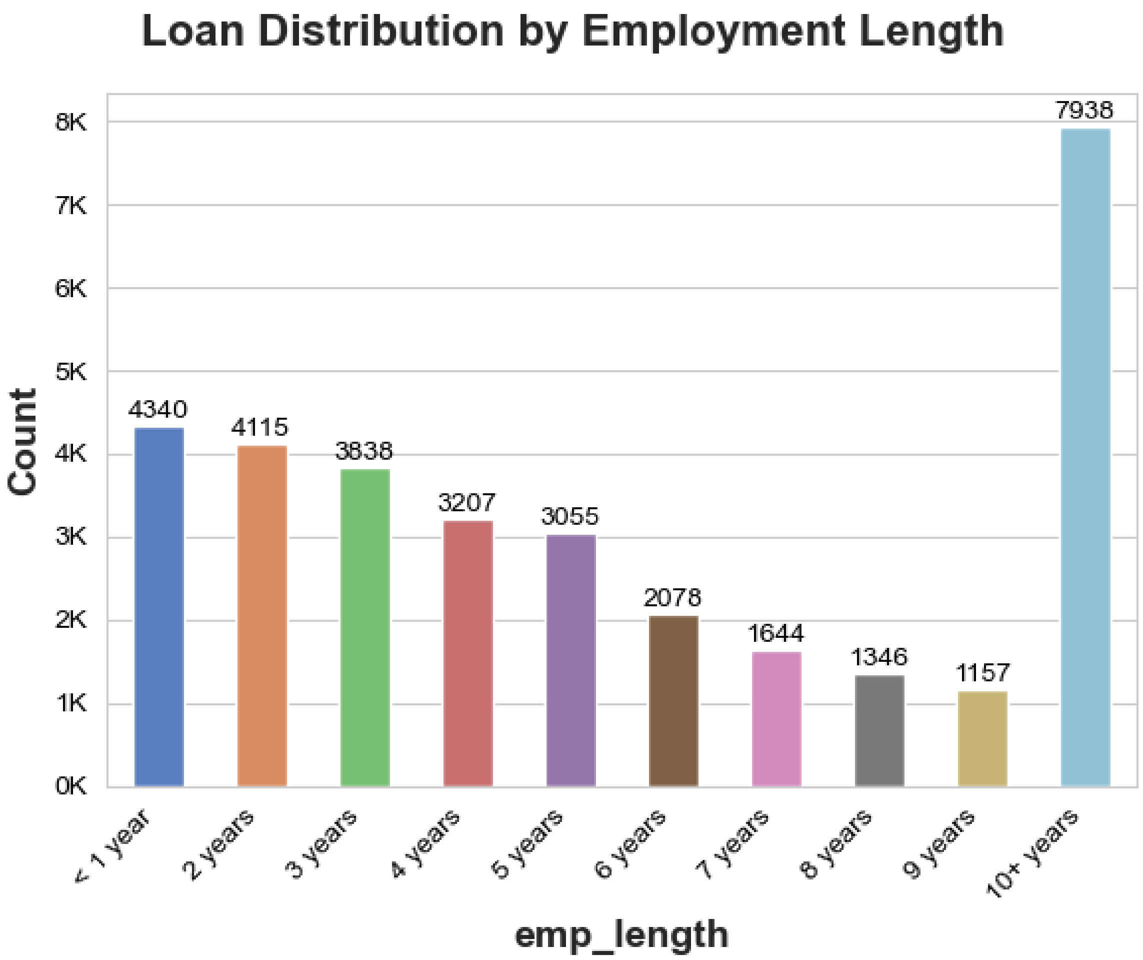
❖ Derived a new attribute **region** by grouping the states based on their geographic locations.

❖ Majority of the loans are offered to *South* region customers.



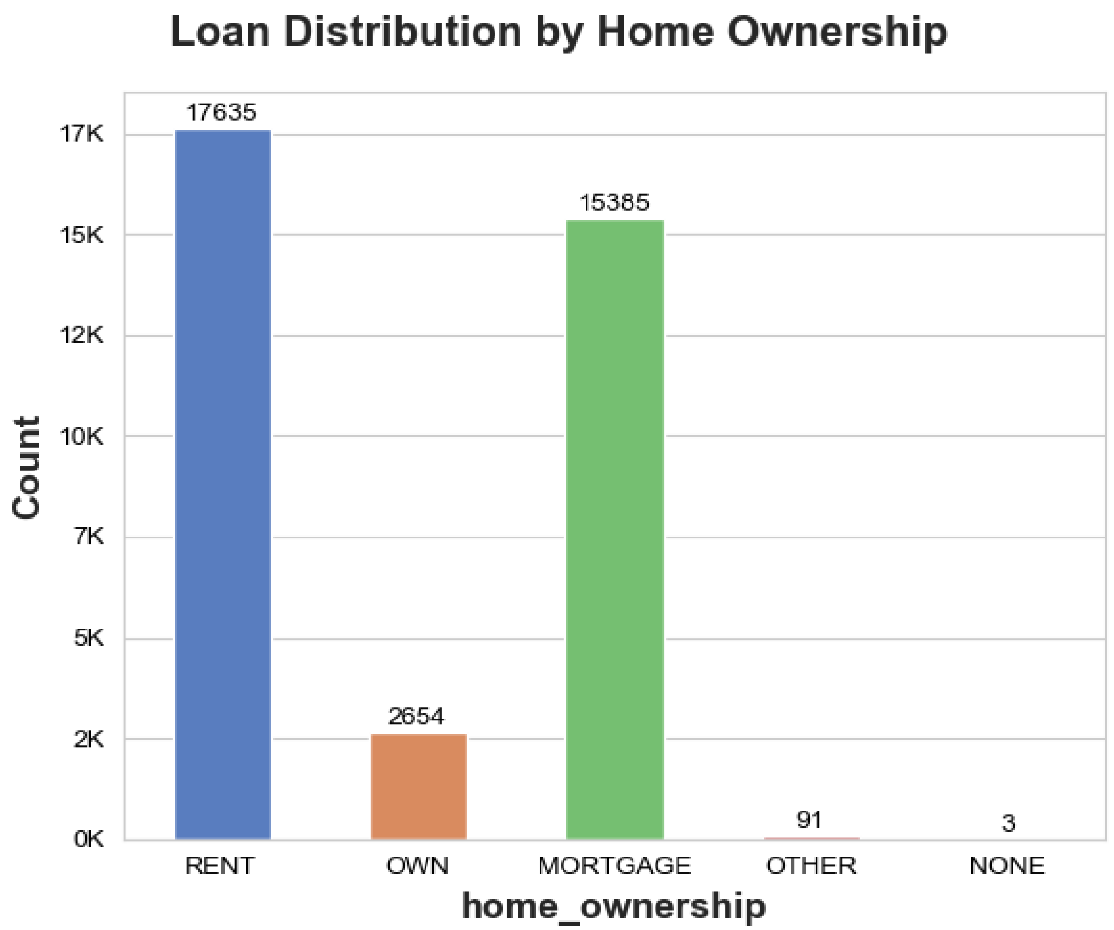
Employment Length

❖ Customers with *10+ years* of experience were offered more loans.



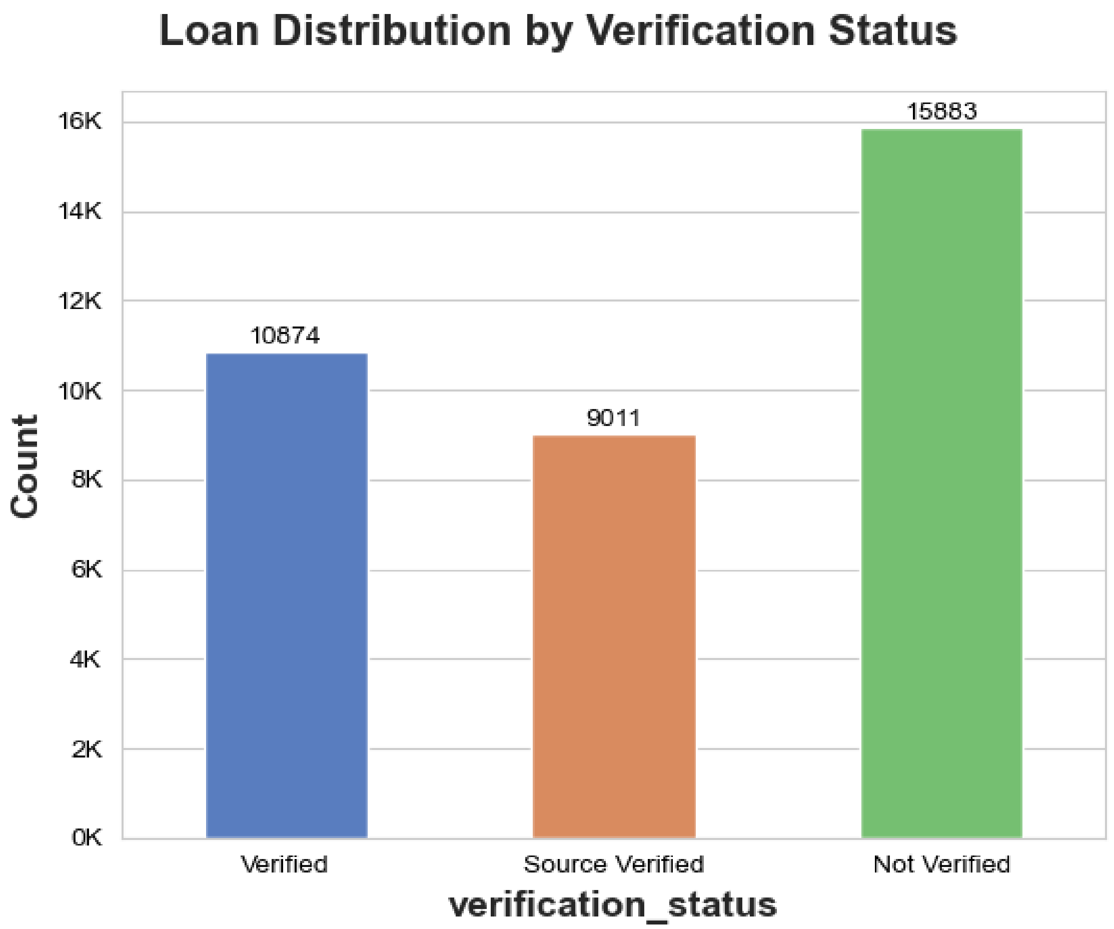
Home Ownership

❖ Customers living in *RENT / MORTGAGE* houses are taking more loans.



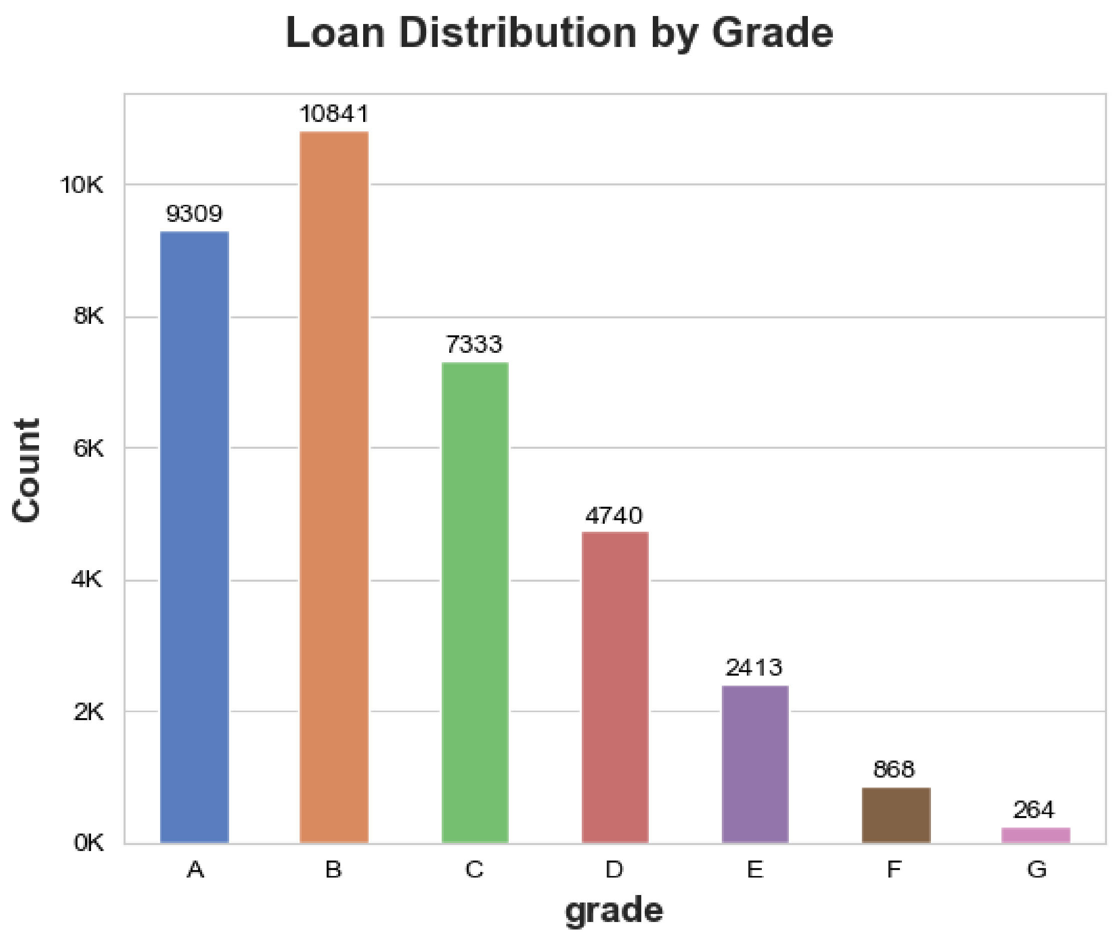
Verification Status

❖ Company offered more loans to customers though they are *Not Verified*.



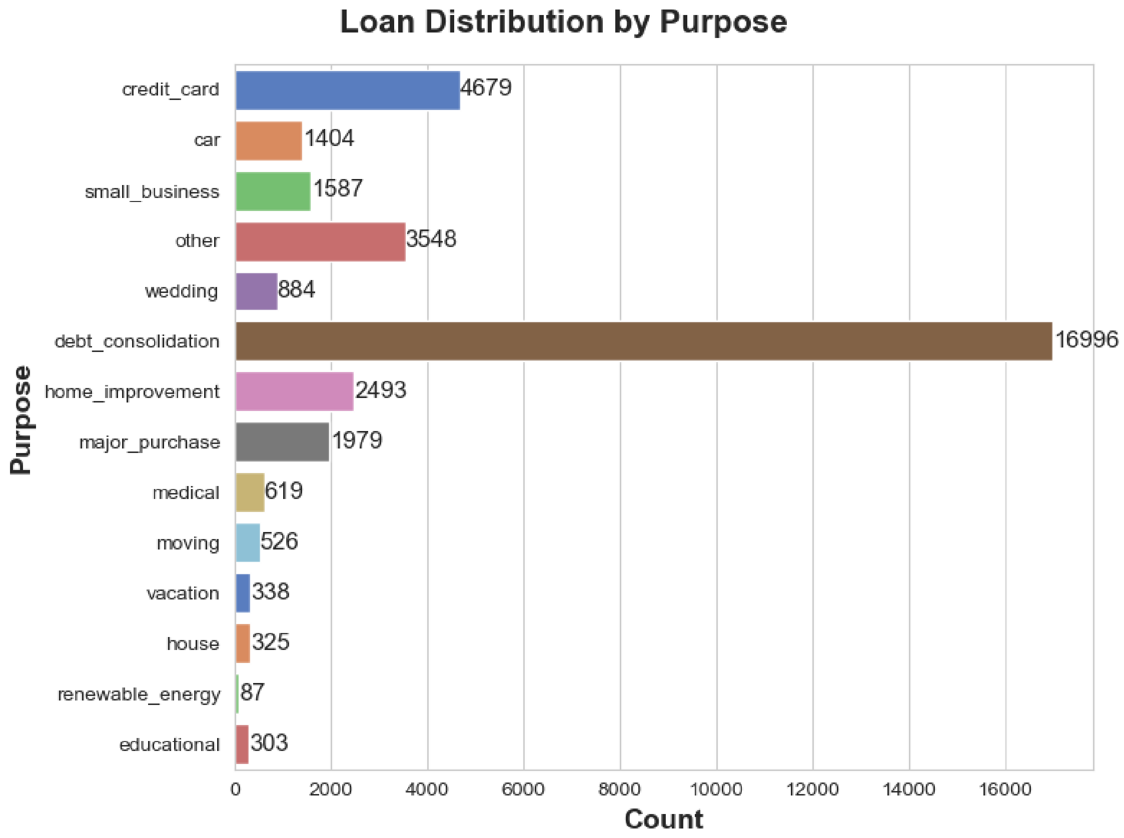
Grade

- ❖Grade is the attributed assigned to each loan by Lending Club.
- ❖*A* being low risk and *G* high risk



Purpose

❖ Majority of the loans are taken for *debt_consolidation*



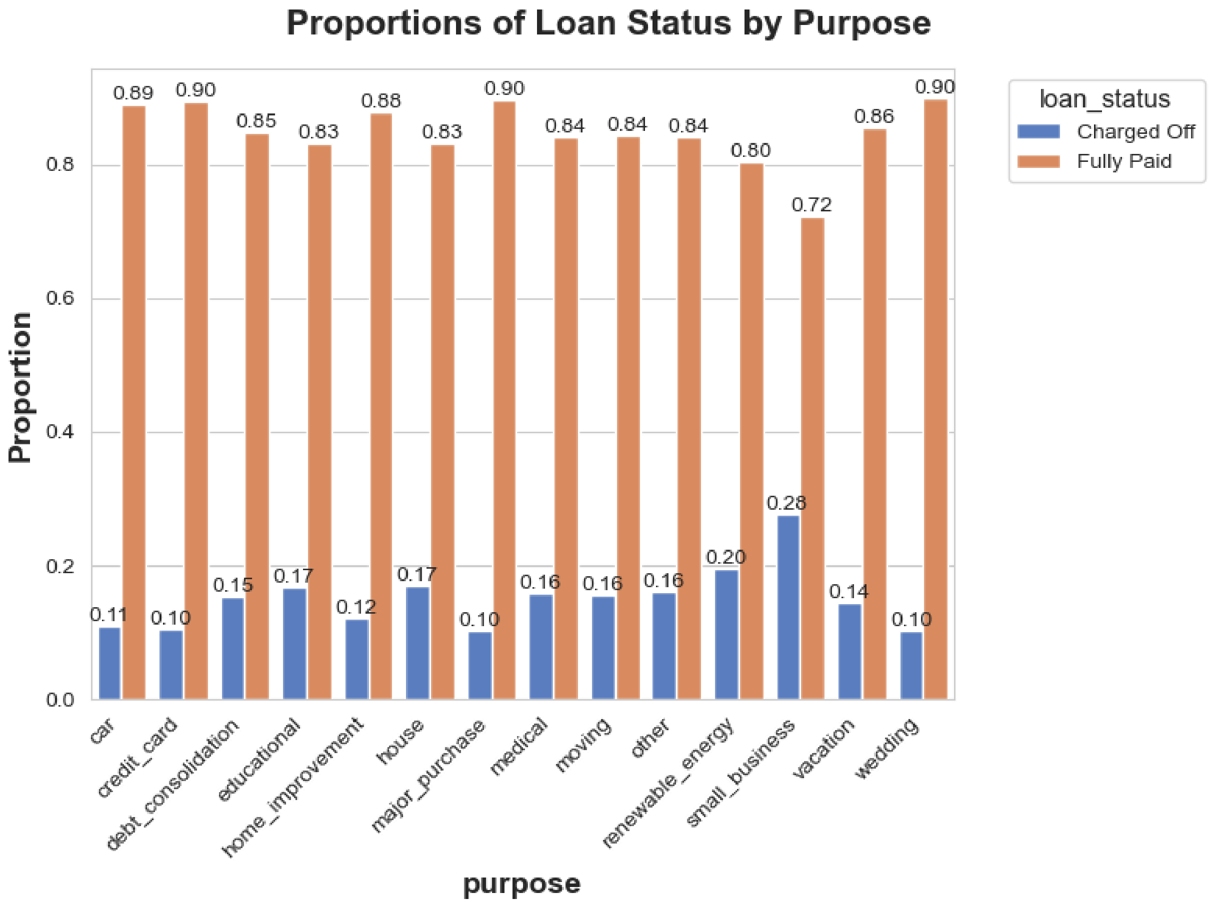
Segmented Univariate Analysis

- Segmented univariate analysis is a type of statistical analysis in which you examine one variable at a time (univariate analysis), but separately within different segments or subgroups of your dataset.

- In this case study **Loan Status (loan_status)** attribute is analyzed within various segments.

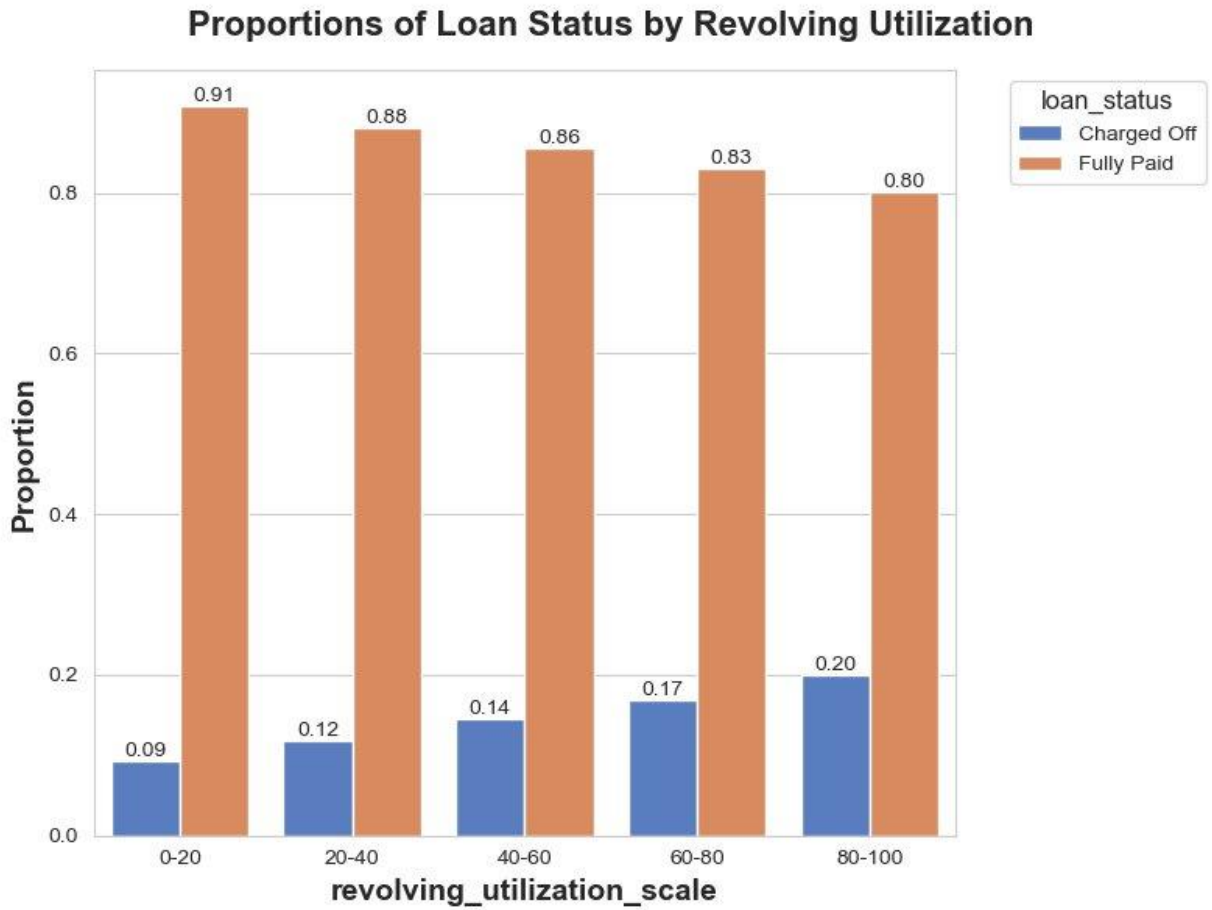
Purpose

- ❖ Loans taken for small_business and renewable_energy have high default rate.



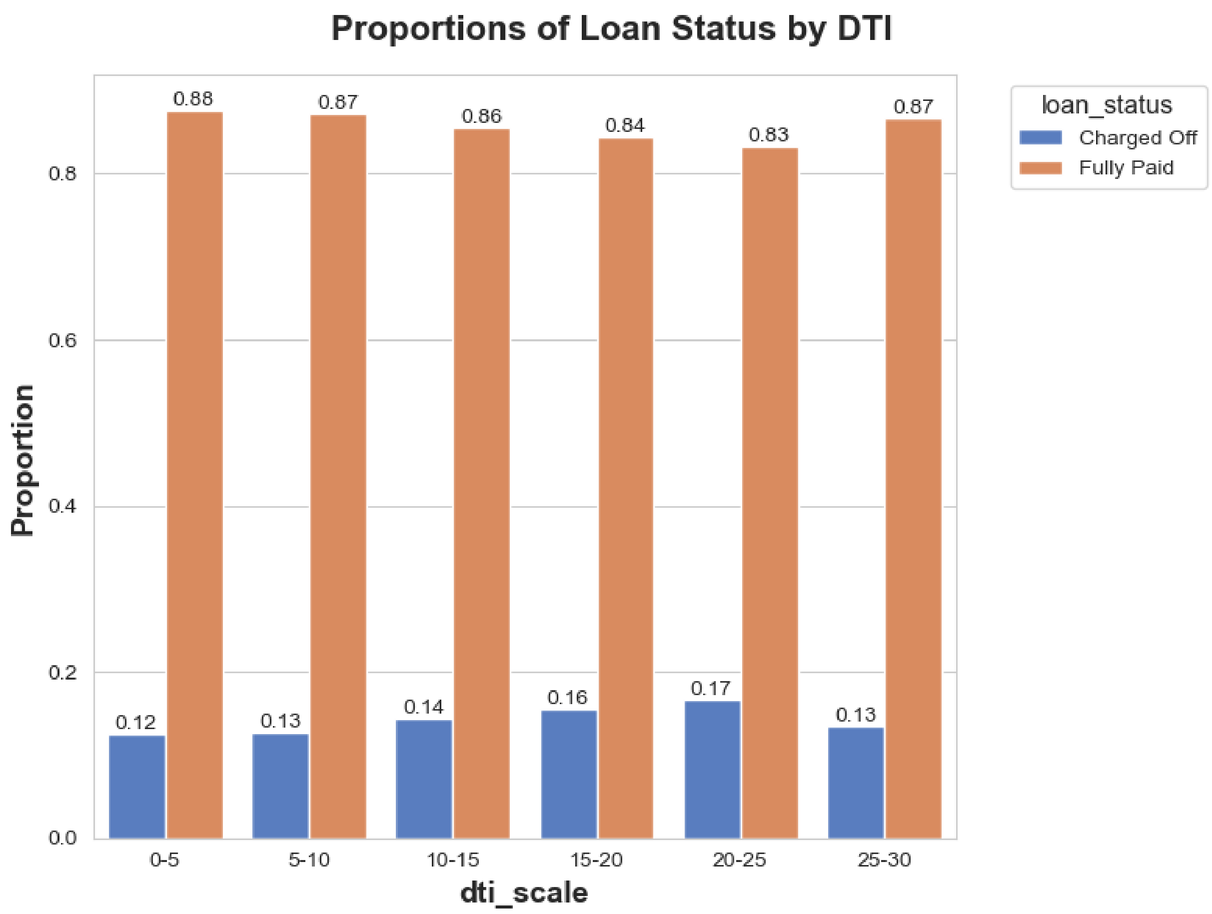
Revolving Utilization

- ❖ Charged-off rate increases with utilization.



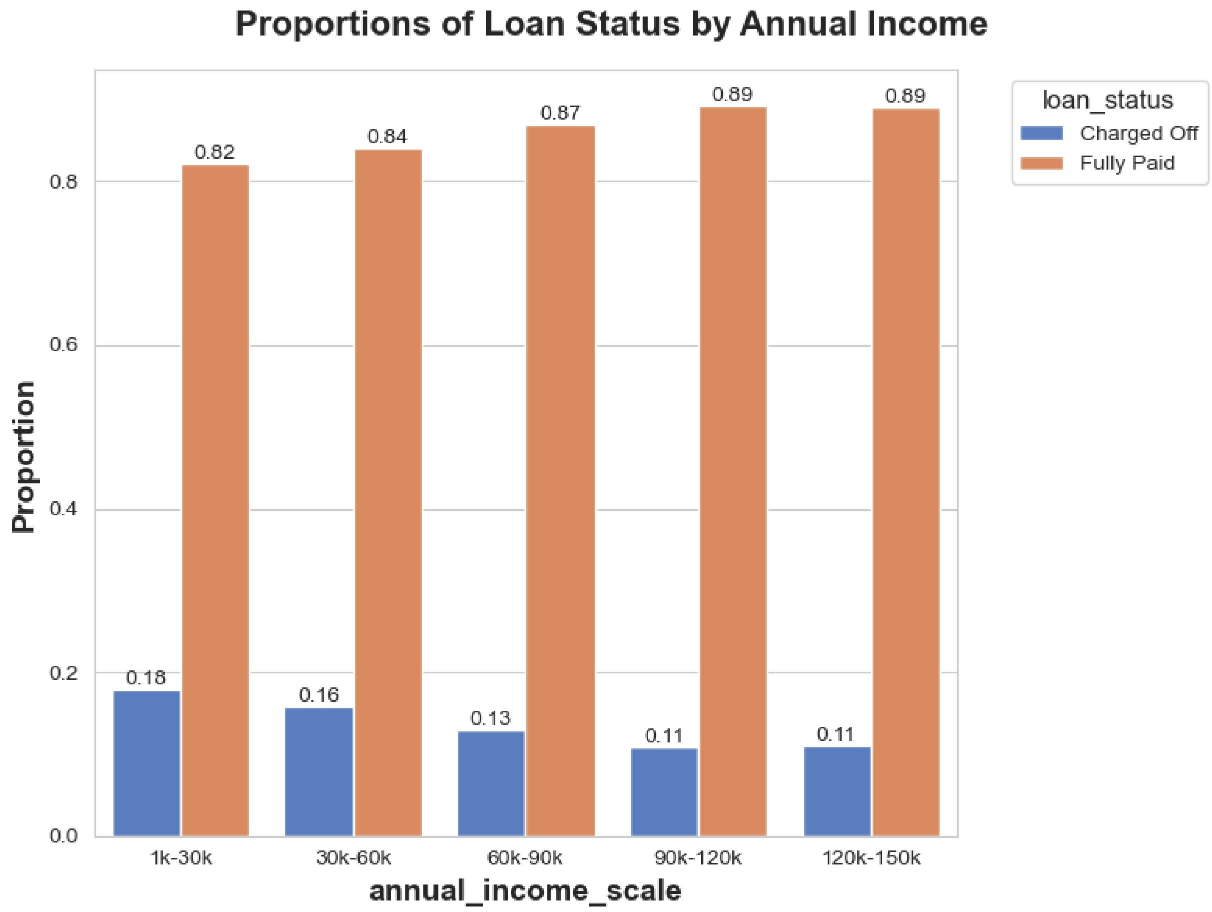
Debt-to-Income Ratio (DTI)

❖ 20-25% segment has high default rate



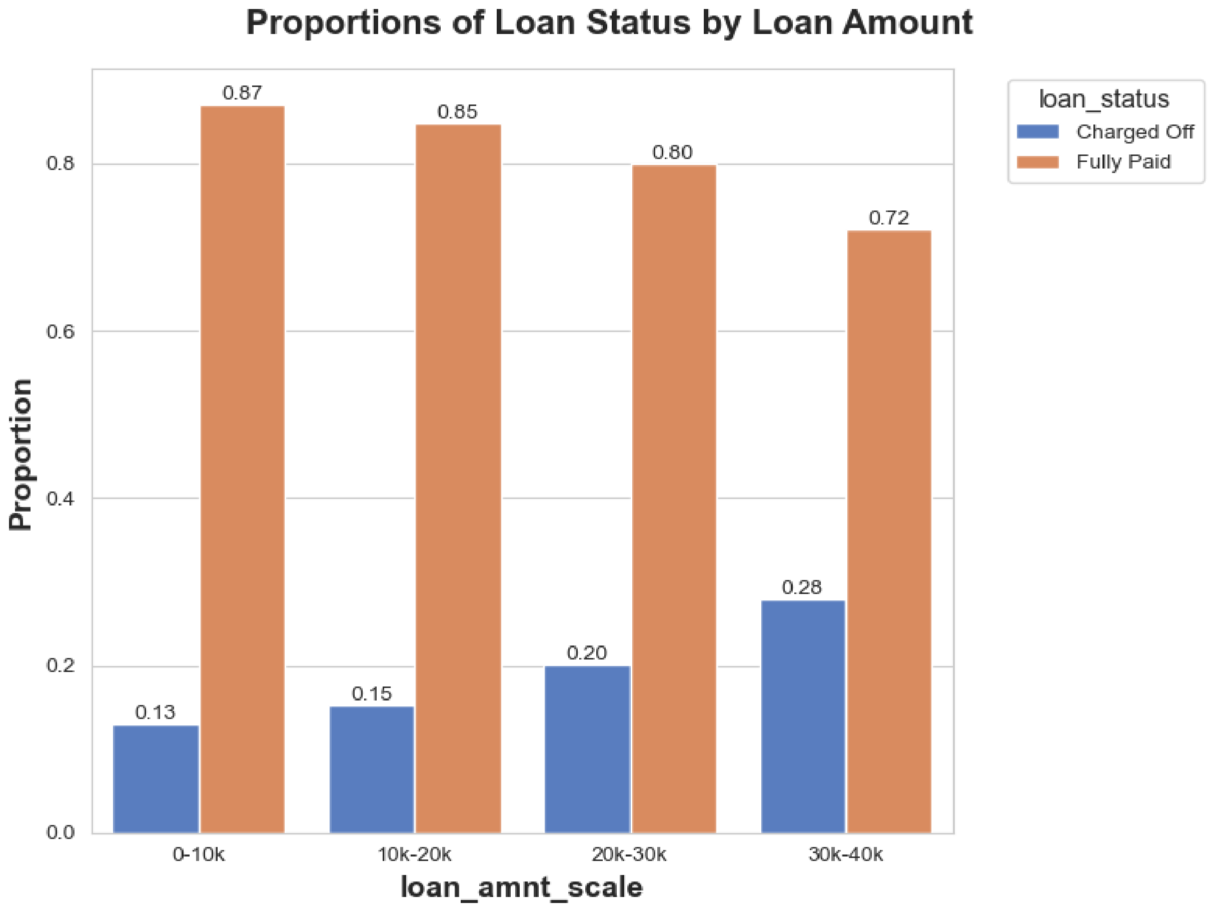
Annual Income

- ❖ Charged off rate decreases when annual income increases.



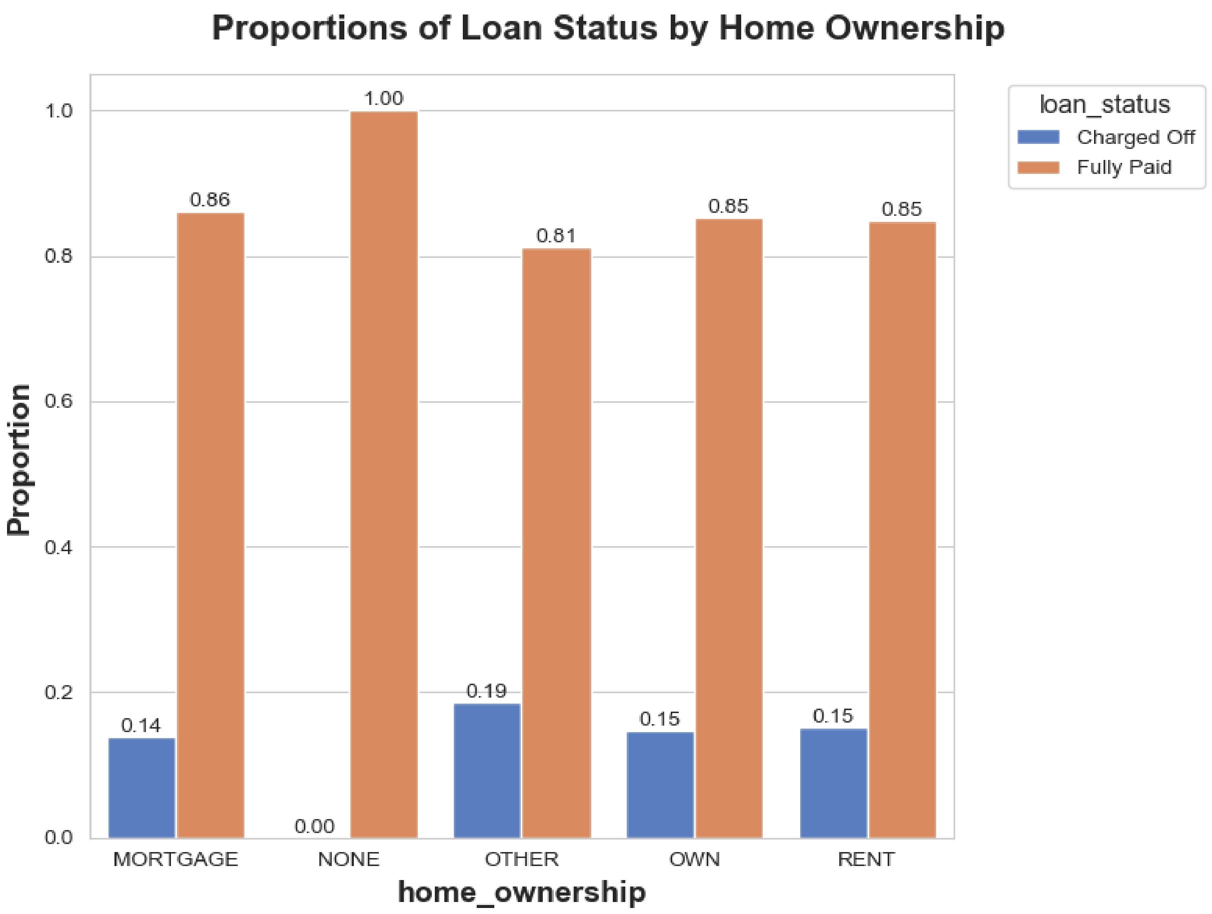
Loan Amount

❖ Risk of default increases with loan amount.



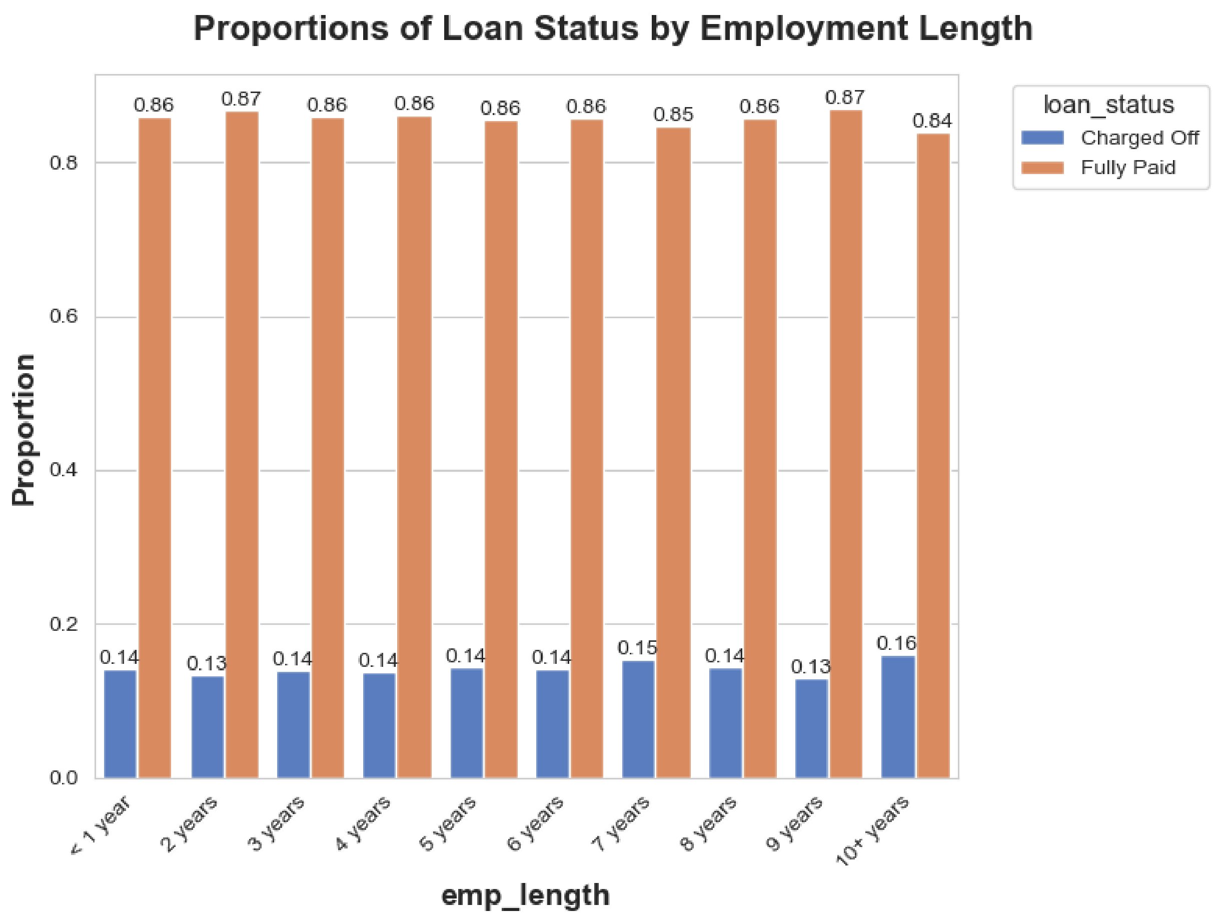
Home Ownership

❖ OTHER has highest default rate.



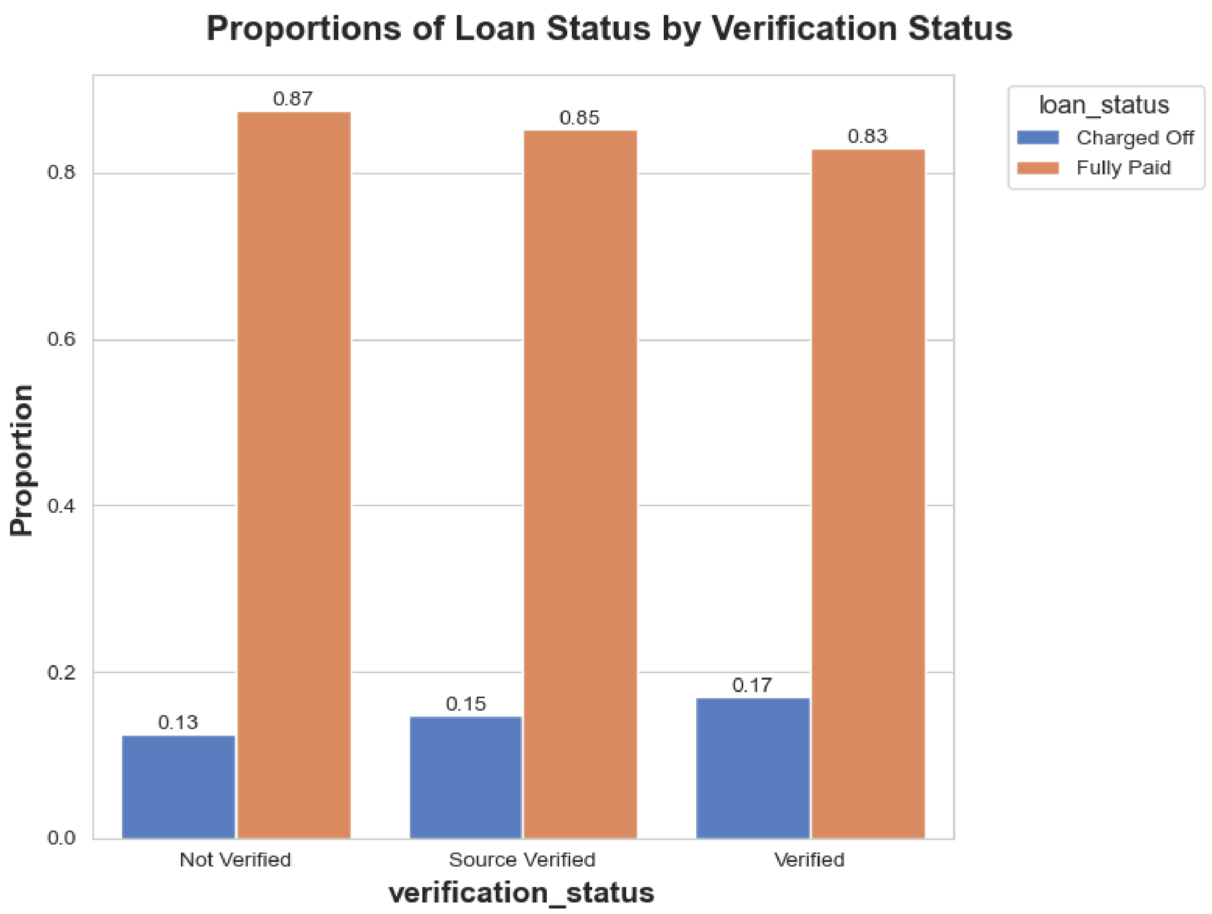
Employment Length

❖ Charged off rate is almost same in all experience levels.



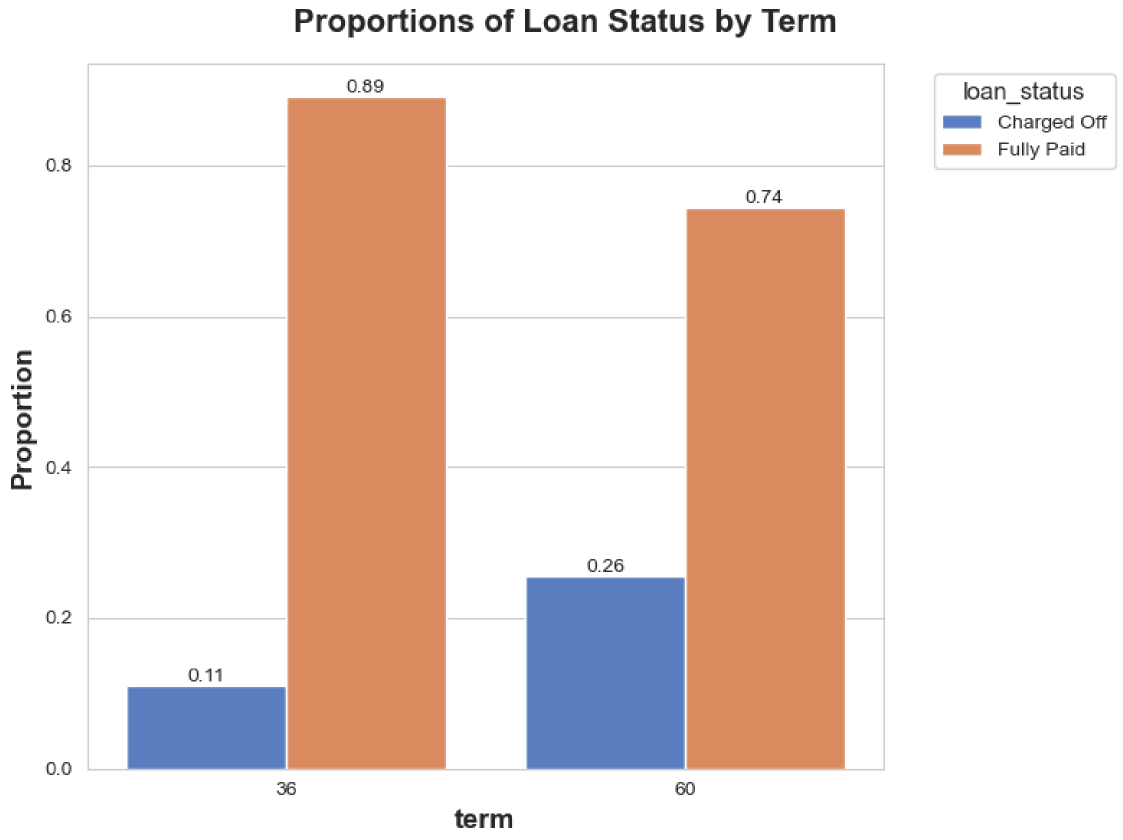
Verification Status

❖ Surprisingly *Verified* customers are defaulting more.



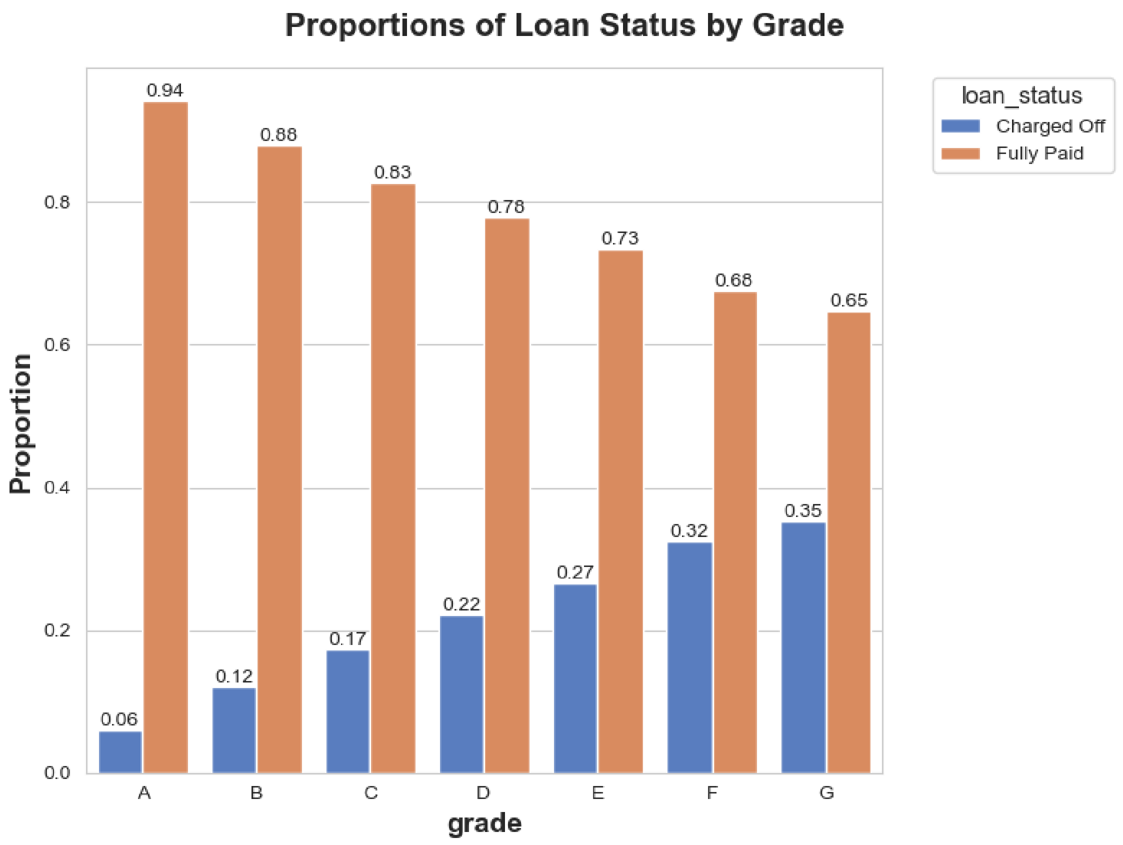
Term in months

❖ 60 months term has high default rate.



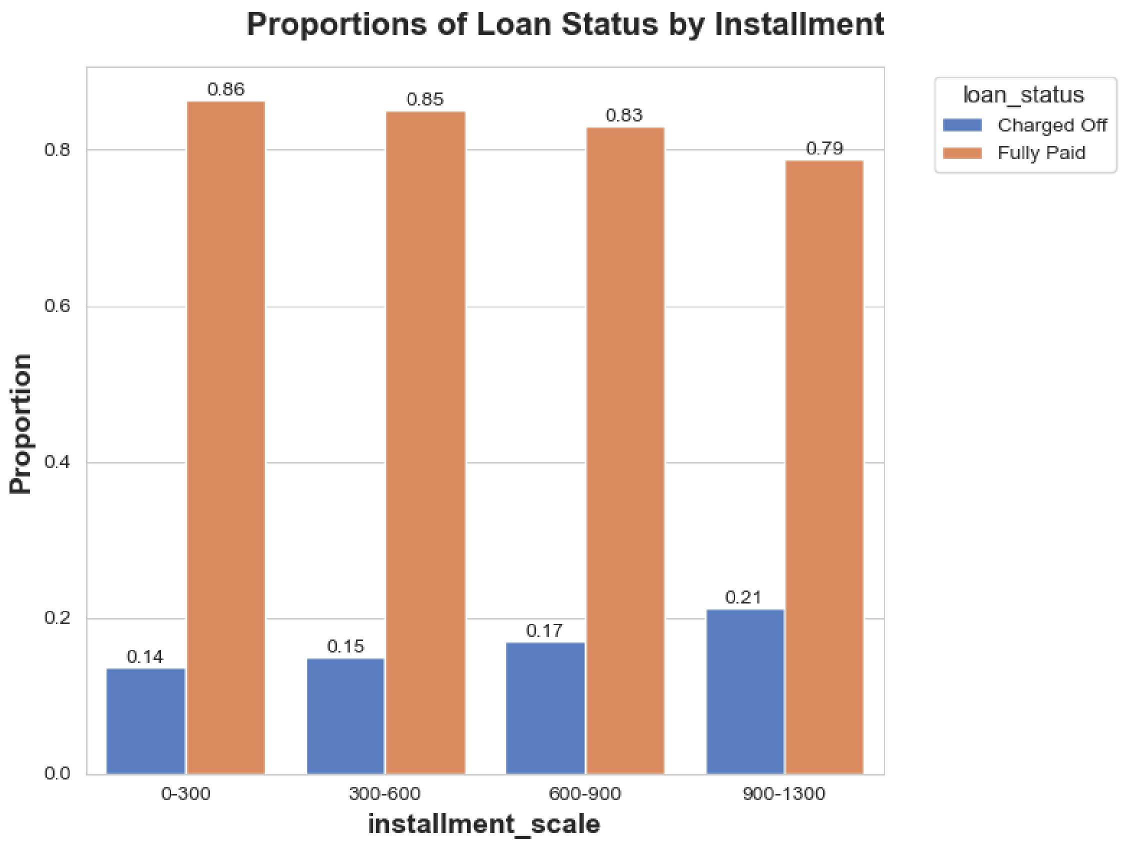
Grade

❖ As classified by Lending Club, default rate increases when loan grade decreases.



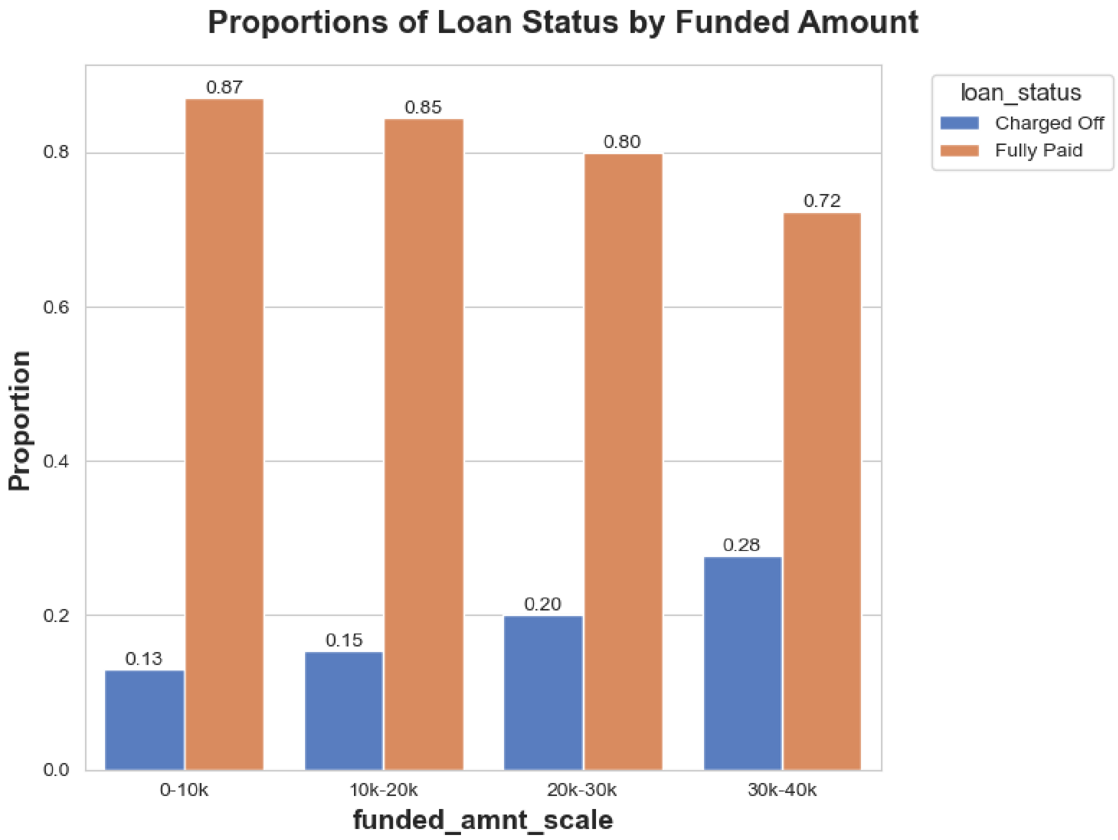
Installment

- ❖ Charged-off rate increases with installment amount.



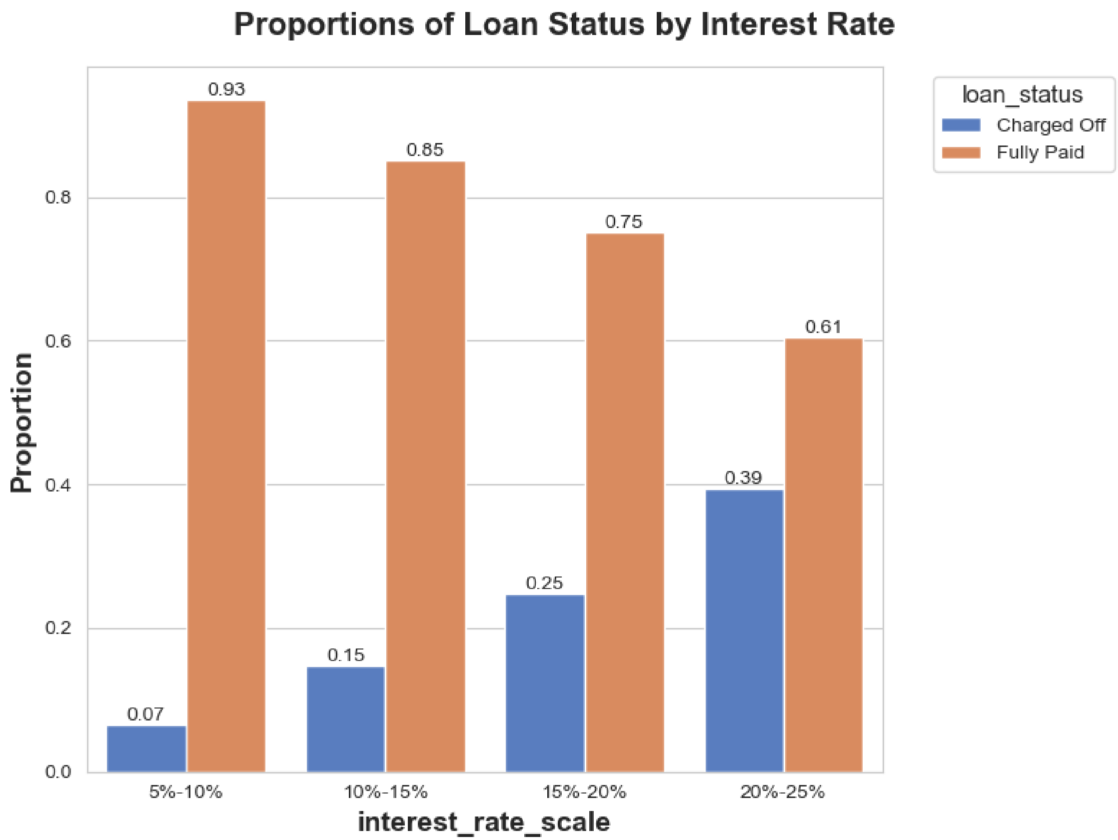
Funded Amount

❖ Charged-off rate increases with funded amount.



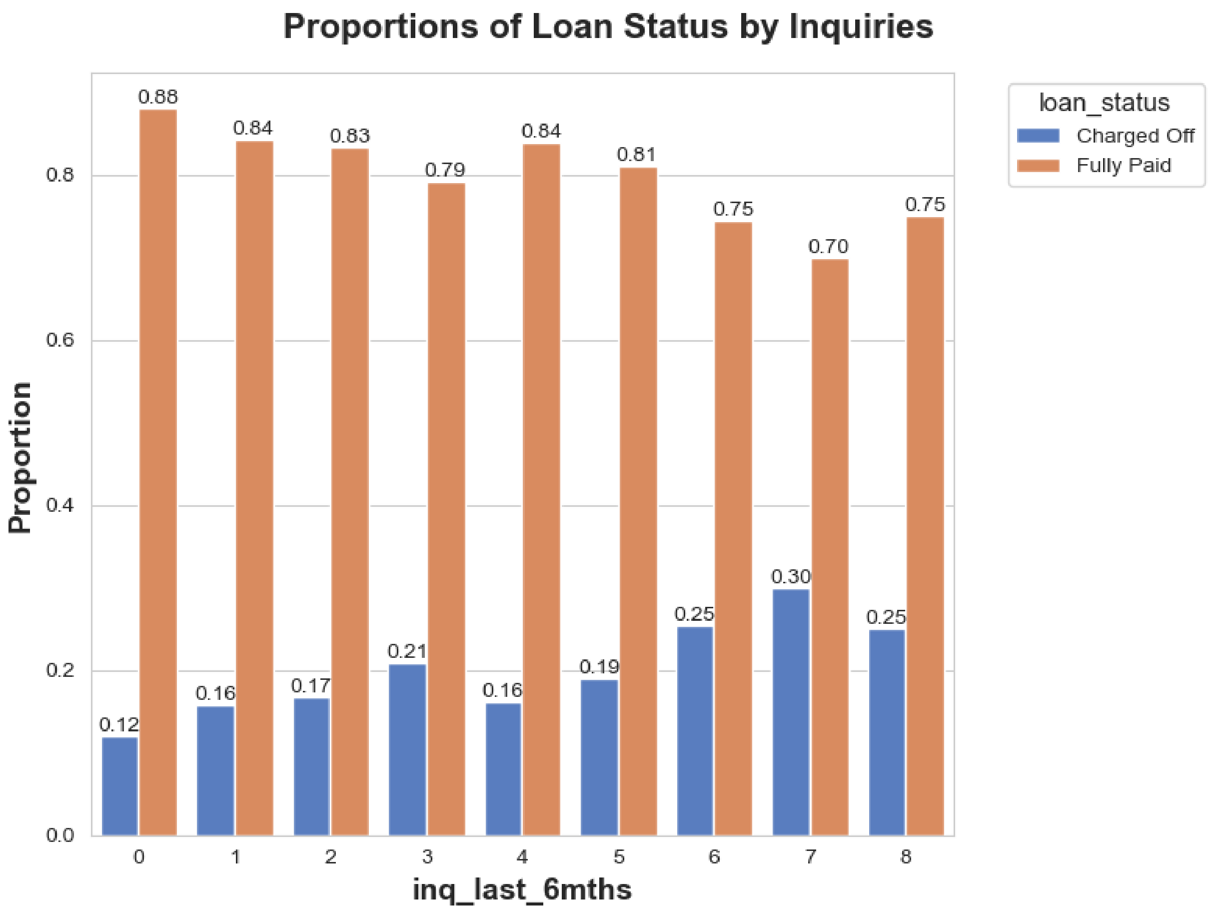
Interest Rate

❖ Charged-off rate increases with interest rate.



Inquiries in last 6 months

- ❖ It represents the customers' credit appetite.
- ❖ Charged-off rate increases with inquiries.



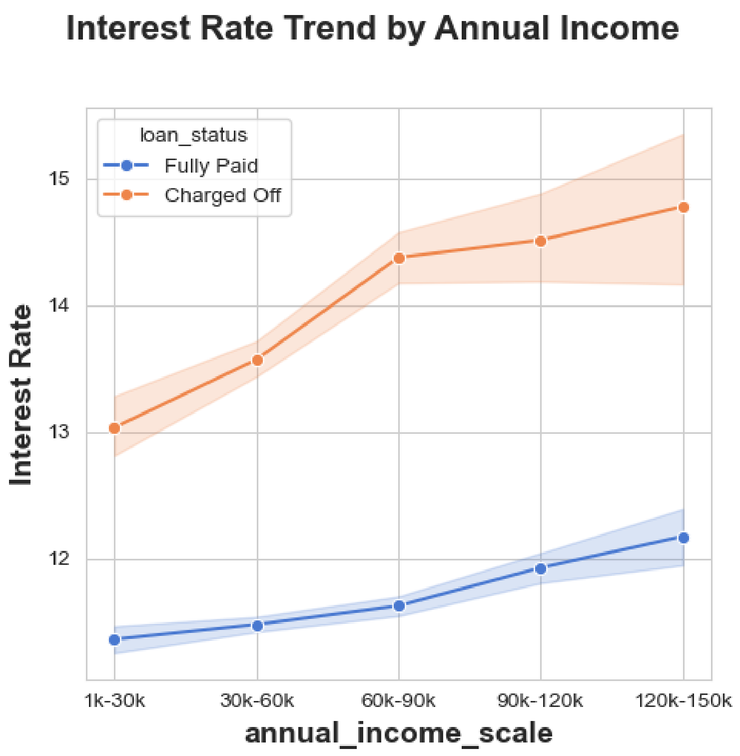
Bi-Variate Analysis

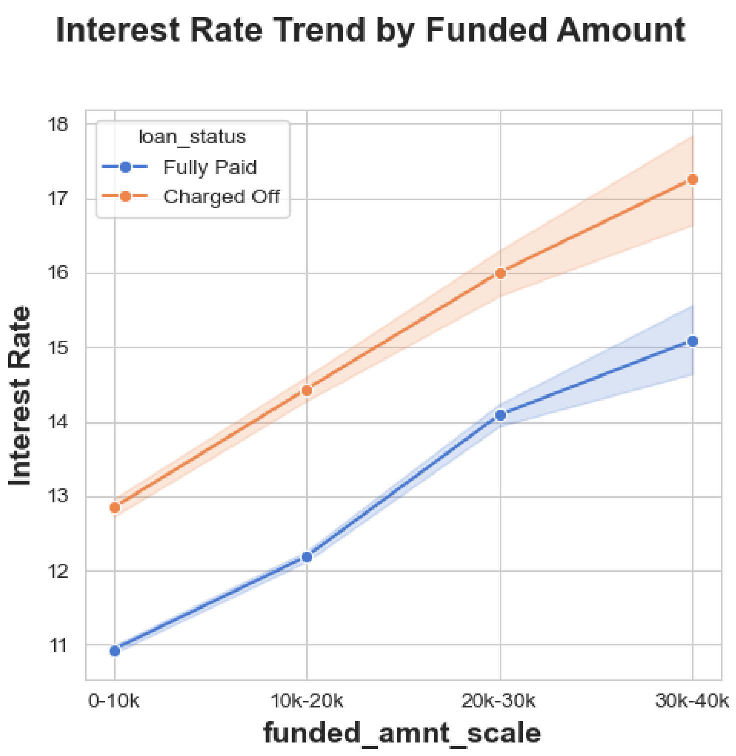
- Bivariate analysis is a type of statistical analysis that investigates the relationship between two variables.

- It involves comparing two sets of data with the goal of exploring a potential association.

Impact of Financial Factors on Interest Rates

- Average Interest Rate Increases with Higher Funded Amount and Annual Income.
- Average Interest Rate for Charged-Off Loans is Higher Than for Fully Paid Loans.





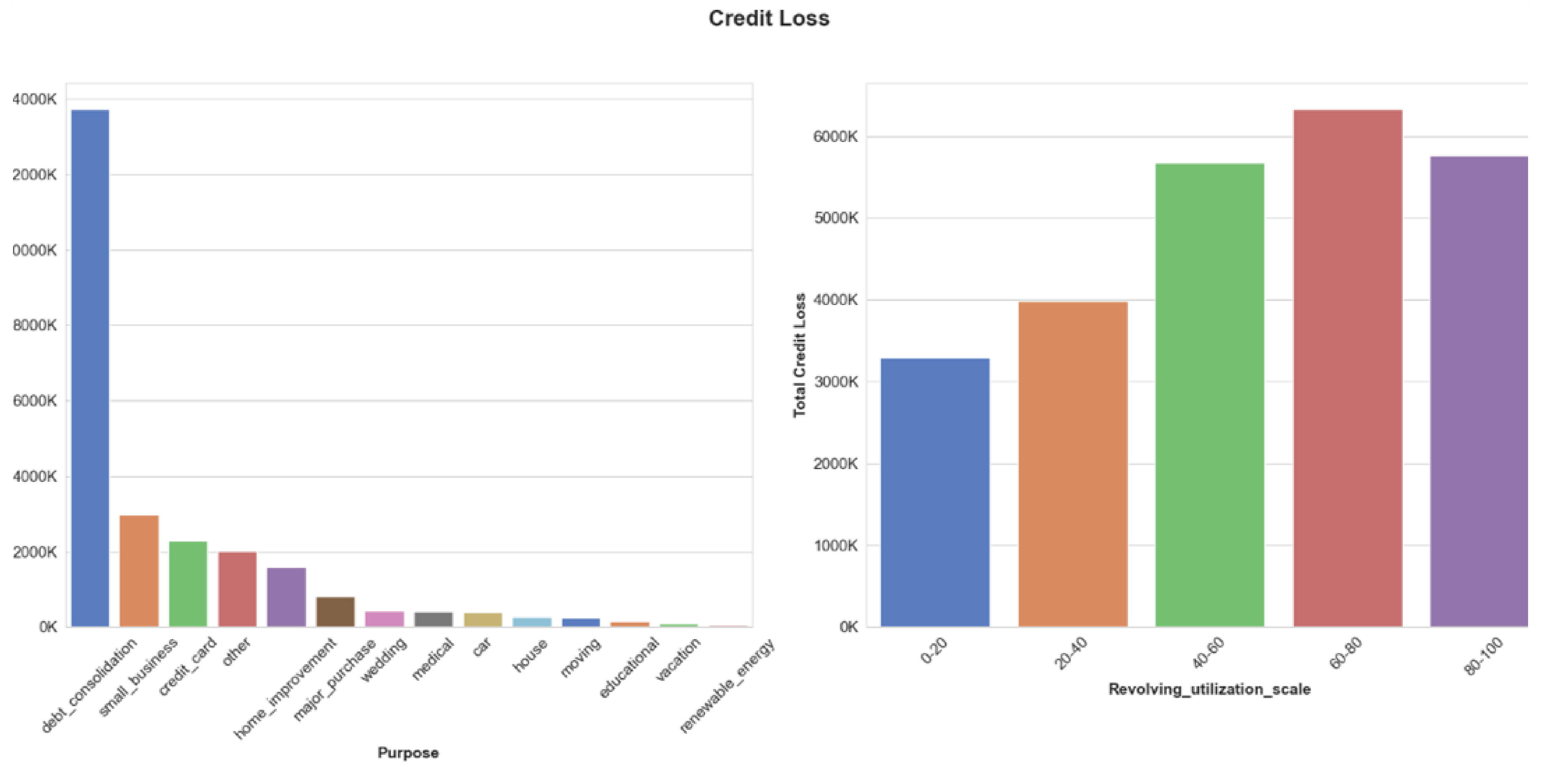
Credit Loss

- Credit loss refers to the amount of money the company loses when a borrower defaults, meaning the borrower fails to pay back the loan as agreed.

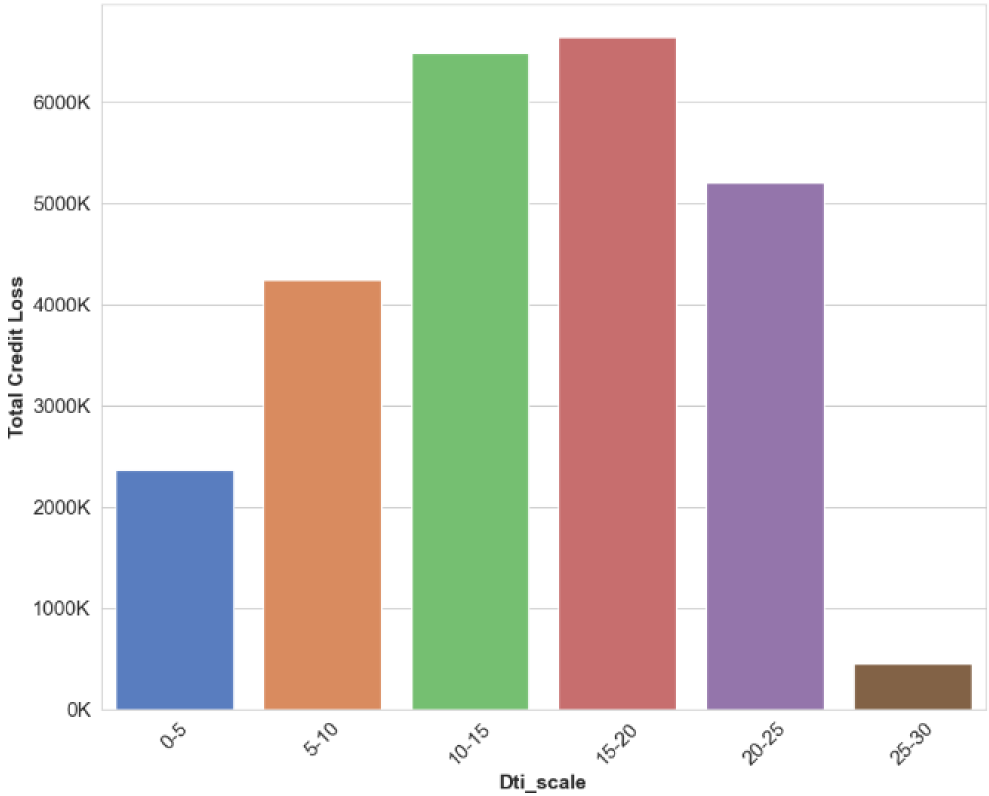
```
Credit Loss = Funded Amount - Total Payment
```

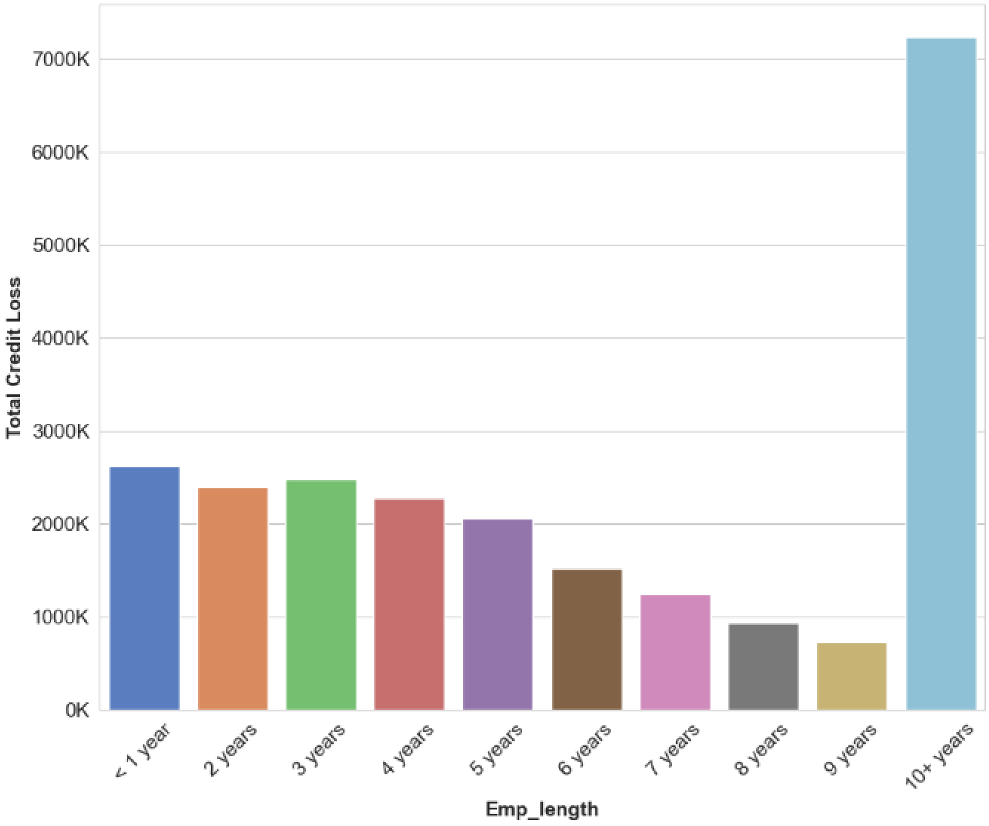
- For Calculating credit loss, only data with *'Charged Off'* status is considered.
- This approach ensures that we accurately measure the loss from defaulted loans.

- Credit loss increases with increase in Revolving Utilization
- Debt Consolidation has more credit loss as volume of loans in that is also high



❖ Credit loss increases with Debt-to-Income Ratio increases.
❖ Credit loss is highest for employment length of 10+ Years.





Conclusions

- ✓ **Purpose:** The purpose of loans taken for *small_business* and *renewable_energy* are more risky. Default rates are **28%** and **20%** respectively. While the other purposes are relatively less risky, we observed more credit loss for *debt_consolidation* purpose which is ~ $14M because volume is high in this purpose.

- ✓ **Revolving Utilization:** Higher utilization rates correlate with higher default rates. The highest charged-off rate in this segment is **20%** for the customers who utilize their available credit 80%-100%. Credit loses are also proportional to the utilization rates.

- ✓ **DTI (Debt-to-Income):** Higher DTI ratios lead to higher default rates. Charged-off rate increases starting from 13% to 17%. Higher DTI ratios lead to higher charged-off rates and increased credit losses.

- ✓ **Annual Income:** Higher income correlates with lower default rates. Charge-off rate decreases from 18% (1k-30k) to 11% (90k-150k). Credit loss is highest for income range 30k-60k ~ $11M.

Conclusions

- ✓ **Loan Amount:** Larger loan amounts correlate with higher default rates. Charge-off rates rising from 13% for loan amount between 0-10k to 28% for loan amount between 30k-40k.
- ✓ **Home Ownership:** Category OTHER has higher default rate 19%.
- ✓ **Employment Length:** Default rate almost same across all levels approximately ~15%. Out of all 10+ years people are defaulting slightly more than others with 16%. But, there is huge credit loss in this category though the difference is just 1% because of volume.
- ✓ **Verification Status** Even verified customers are more likely to default at the rate of 17% and credit loss is ~ $12.5M.
- ✓ **Term:** 60 months loan term has more default rate compared to 36 months. Charged-off percentage is 26% for 60 months and 11% for 36 months.
- ✓ **Grade:** Lower loan **grade**s have higher default rate. The risk level increases from grade A(6%) to grade G(35%).

Conclusions

- ✓ **Installment:** Higher monthly installments correlate with higher default rates. Charged-off rate increases from 14% (0-300) to 21% (900-1300).

- ✓ **Funded Amount:** Larger funded amounts correspond to higher default rates. The charged-off rate increases from 13% (for amounts between 0-10k) to 28% (for amounts between 30k-40k).

- ✓ **Interest Rate:** Interest rate is positively correlated with default rate. Charged-off rate is 7% for 5-10% and 39% for 20-25%.

- ✓ **Total Accounts:** Total accounts show stable default rates with slight variations. Charge-off rate is relatively stable around 14-15% across most ranges.

- ✓ **Public Records:** People with derogatory/bankruptcy history are more likely to default. But, loans offered to such customers is also very less.

Recommendations to Reduce Defaults

- Enhance Risk Assessment for High-Risk Loan Purposes:
  - Implement stricter credit checks and higher eligibility criteria for loan purposes that exhibit higher default rates, particularly for small business and renewable energy loans.
  - Consider offering financial advisory services to potential borrowers in these categories to improve their business plans and financial stability, thereby reducing the likelihood of default.

- Limit High Revolving Utilization Borrowers:
  - Introduce policies to limit loan approvals for borrowers with high revolving utilization rates (80%-100%).
  - Offer financial management workshops and tools to help borrowers manage their credit utilization more effectively, which can lead to lower default rates.

- Target High DTI Ratio Borrowers with Financial Counseling:
  - Implement mandatory financial counseling sessions for borrowers with high debt-to-income ratios to help them better manage their debts and improve their financial health.
  - Consider introducing debt consolidation programs that provide better terms and lower interest rates to help reduce their financial burden and risk of default.

Recommendations to Reduce Defaults

- Tailor Loan Amounts Based on Risk Assessment:
  - Adjust loan amounts based on comprehensive risk assessments, particularly for higher loan amounts which show increased default rates.
  - Offer staggered or phased loan disbursement plans for higher loan amounts, contingent on the borrower's performance and repayment ability.

- Promote Loans to Higher Income and Low-Risk Borrowers:
  - Focus marketing efforts on attracting higher income borrowers who demonstrate lower default rates.
  - Develop loyalty programs and incentives for low-risk borrowers (e.g., those with lower DTI ratios, higher annual incomes, and lower revolving utilization rates) to encourage timely repayments and reduce overall default rates.