

# Lending Club Case Study

Submitted by:

Sumanth Bharadwaj Prasannakumar  
Swagat Kumar Mishra

# Background

- Lending club is the largest peer-to-peer marketplace, facilitating personal loans, business loans and financing medical procedures.
- Application is processed online and decision is made based on the data available.
- There is always risk of default associated with lending.

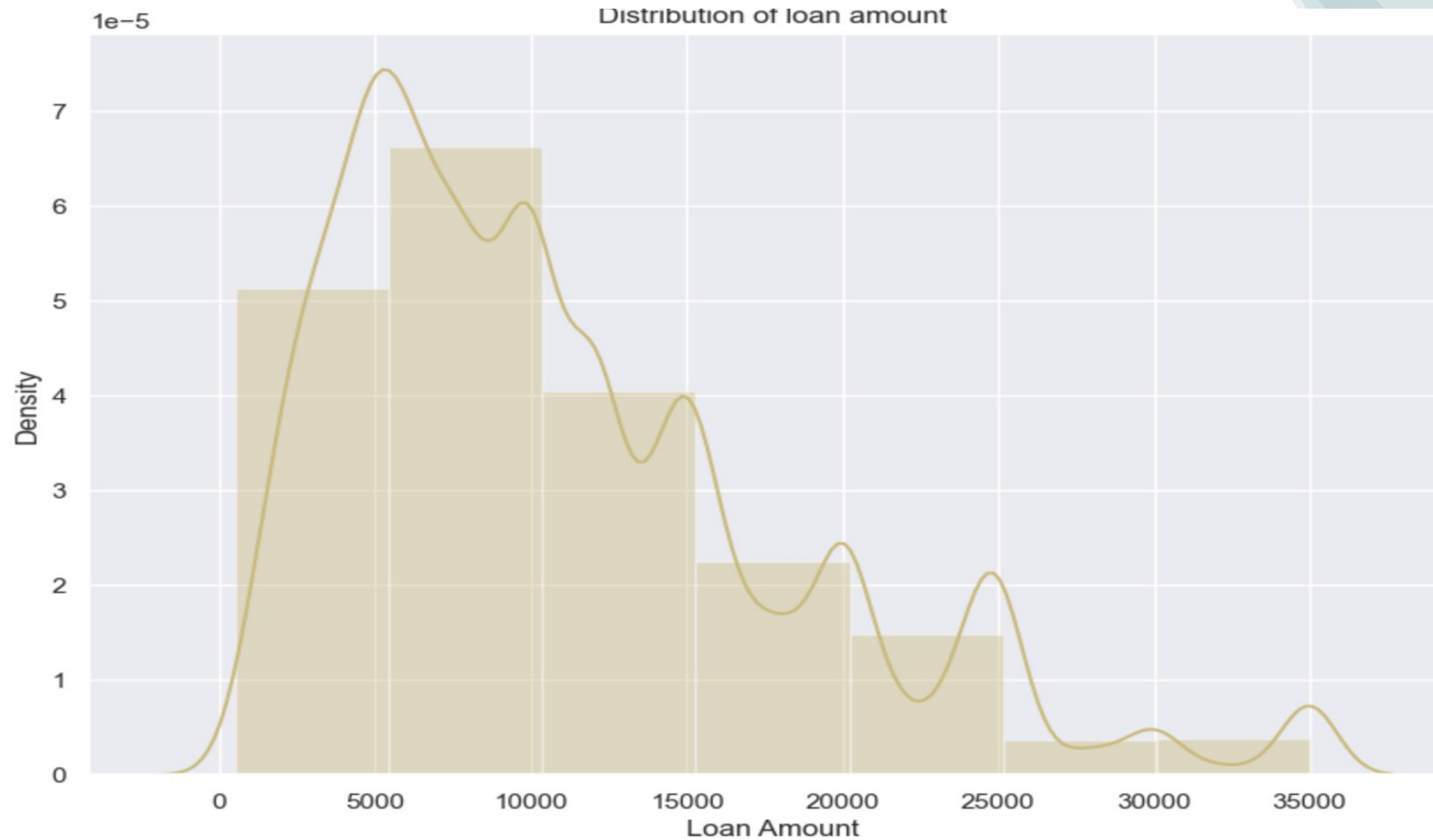
# Business Objective

- Identify key factors driving loan defaults
- Enable data-driven decisions on whether to approve or reject loan applications based on borrower profiles.

# Approach

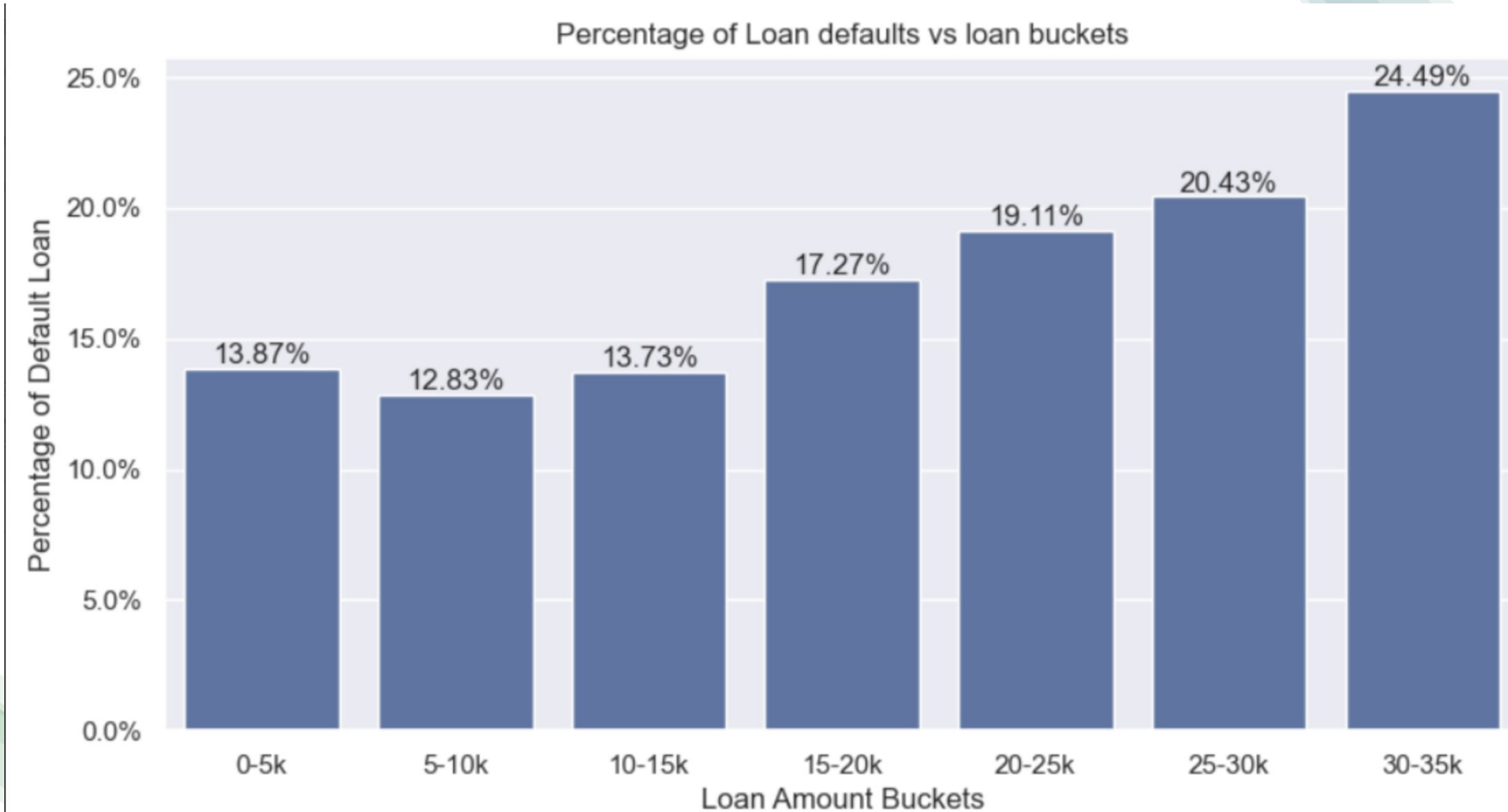
- Data Preparation
  - Data Removal
    - Drop columns having
      - Only null values
      - Only one value
      - All unique values (like `id`)
    - Drop any empty rows
  - Data Cleaning
    - Convert values to correct data type int, float, date representation
    - Fill in missing values with appropriate information wherever required
    - Remove or replace any outliers if we think that will impact the analysis
  - Data Derivation
    - Derive data from the existing data to perform more analysis.
- Distribution Analysis
  - Analyse distribution patterns and frequencies of important data fields in the dataset. Identify any striking feature in that.
- Relationship Analysis
  - Analyse the default behaviour by taking multiple datasets into consideration.
  - Identify any particular combination of data fields that results in higher default rate.
- Key Insights Summary
  - Summarize critical findings and observations to inform decision-making

# Loan Amount



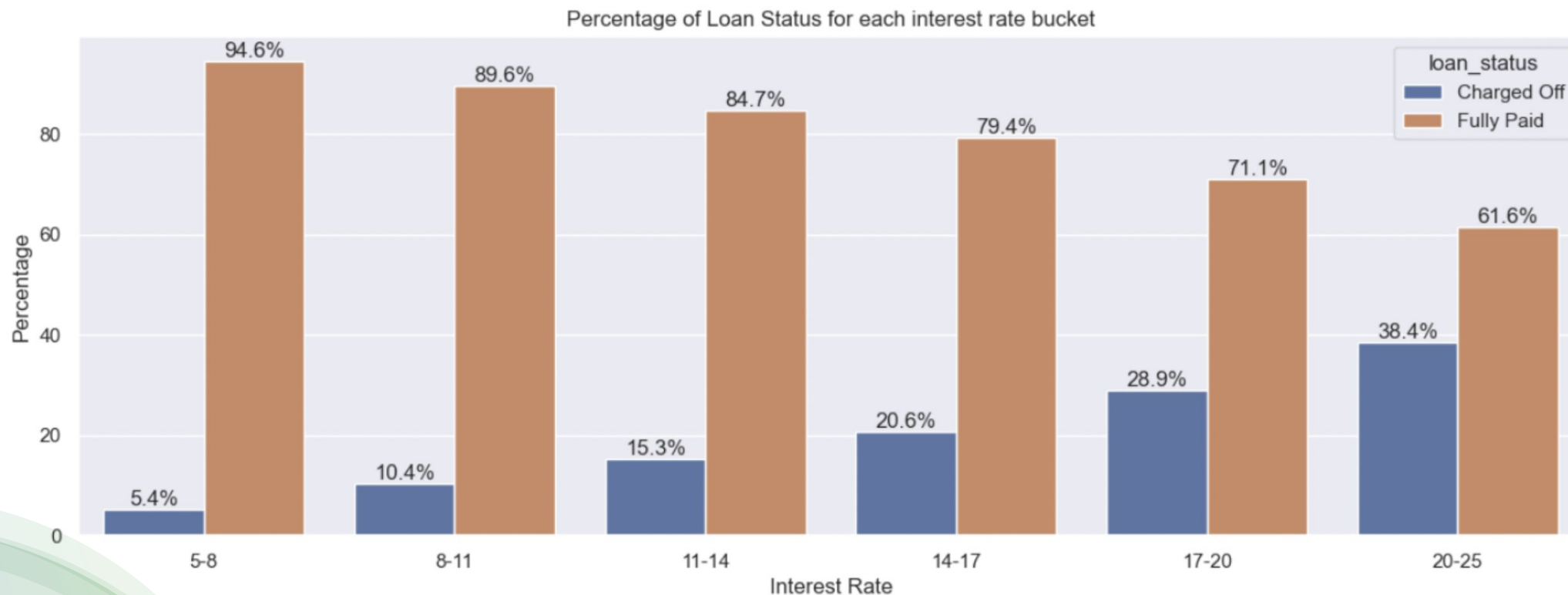
- The maximum number of loan profiles have loan amount ranging between 5K-10K.
- Average loan amount is 11K

# Percentage of Default vs Loan Amount Buckets



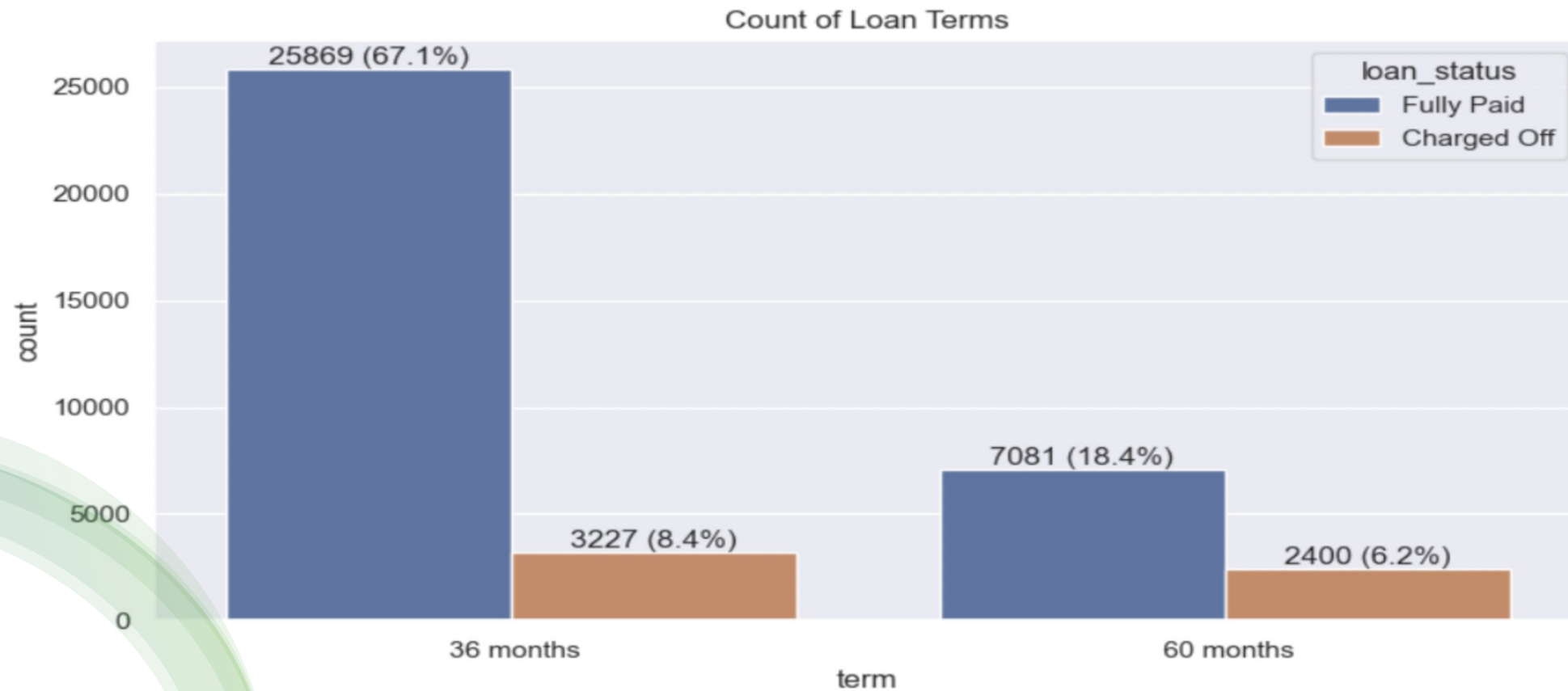
- The loan amount higher than 15K has higher default rate.

# Relation Of Defaulters with Interest Rate



- Average interest rate is around 12%.
- Most of applications have interest rate ranging between 9%-14%.
- As the interest rate increases the default rate increases.

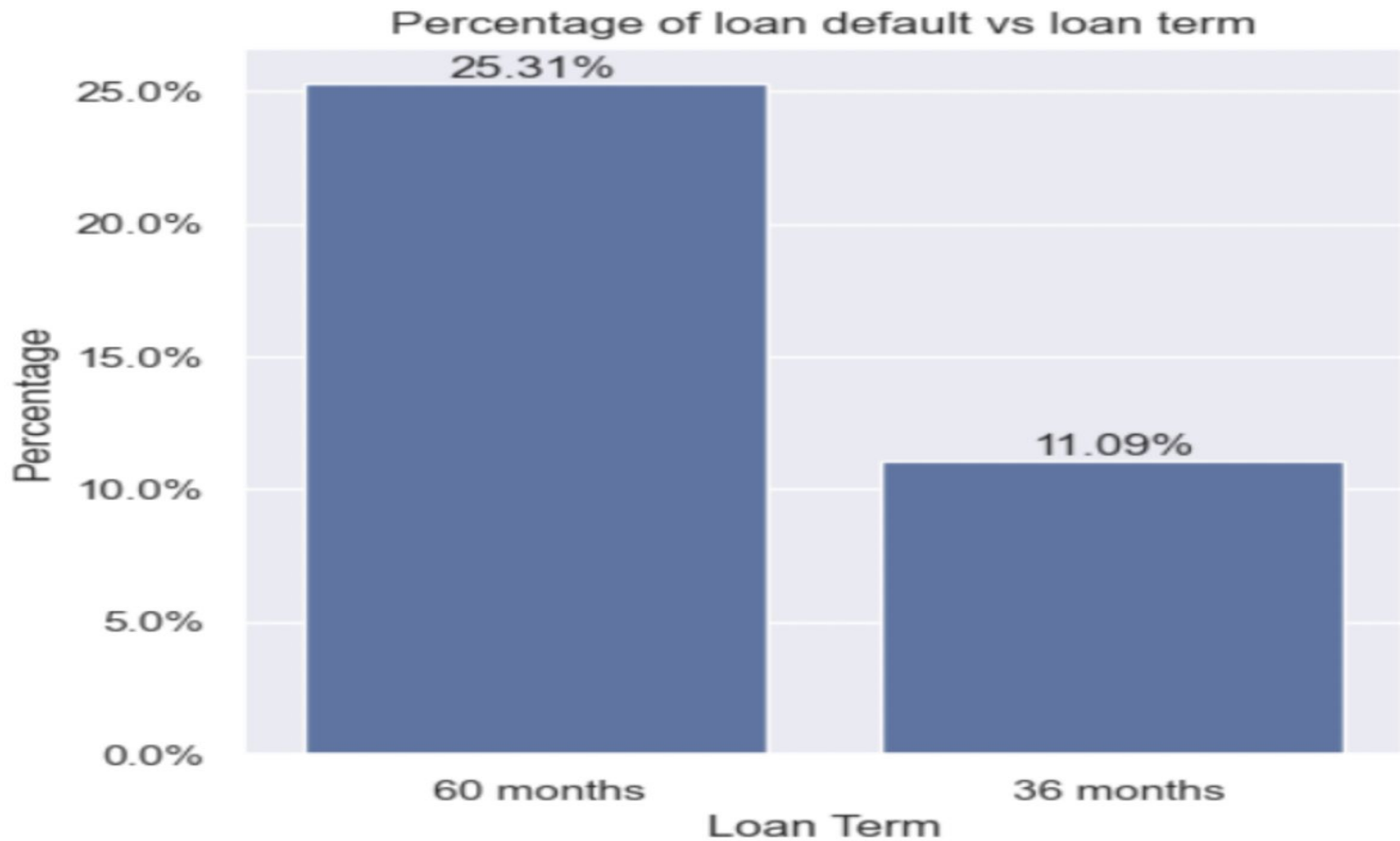
# Loan Count vs Loan Terms



- Nearly 2/3rd of the loans have 36 months tenure.
- But the ratio of default rate looks higher for 60 months tenure.

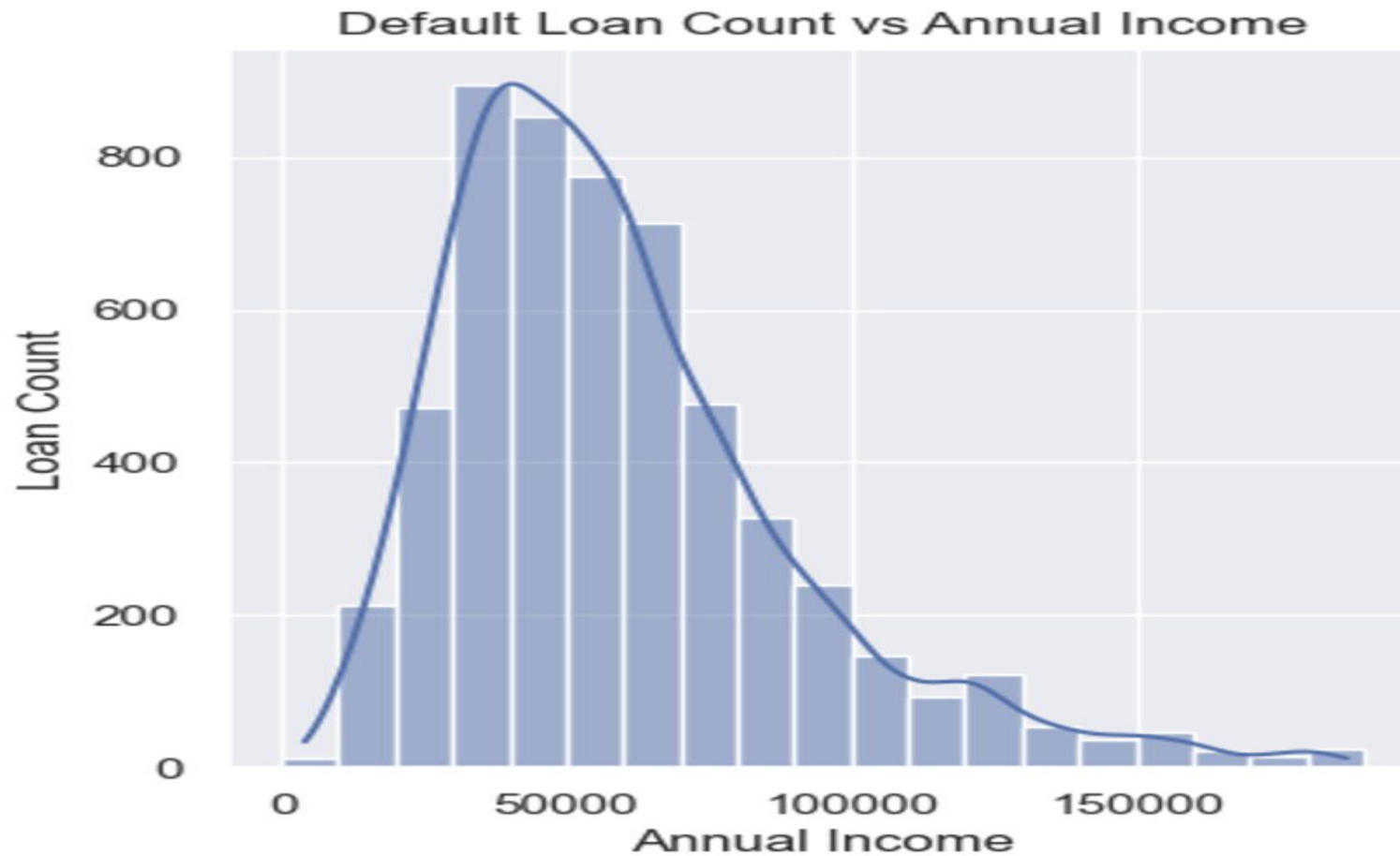


# Default Rate vs Loan Terms



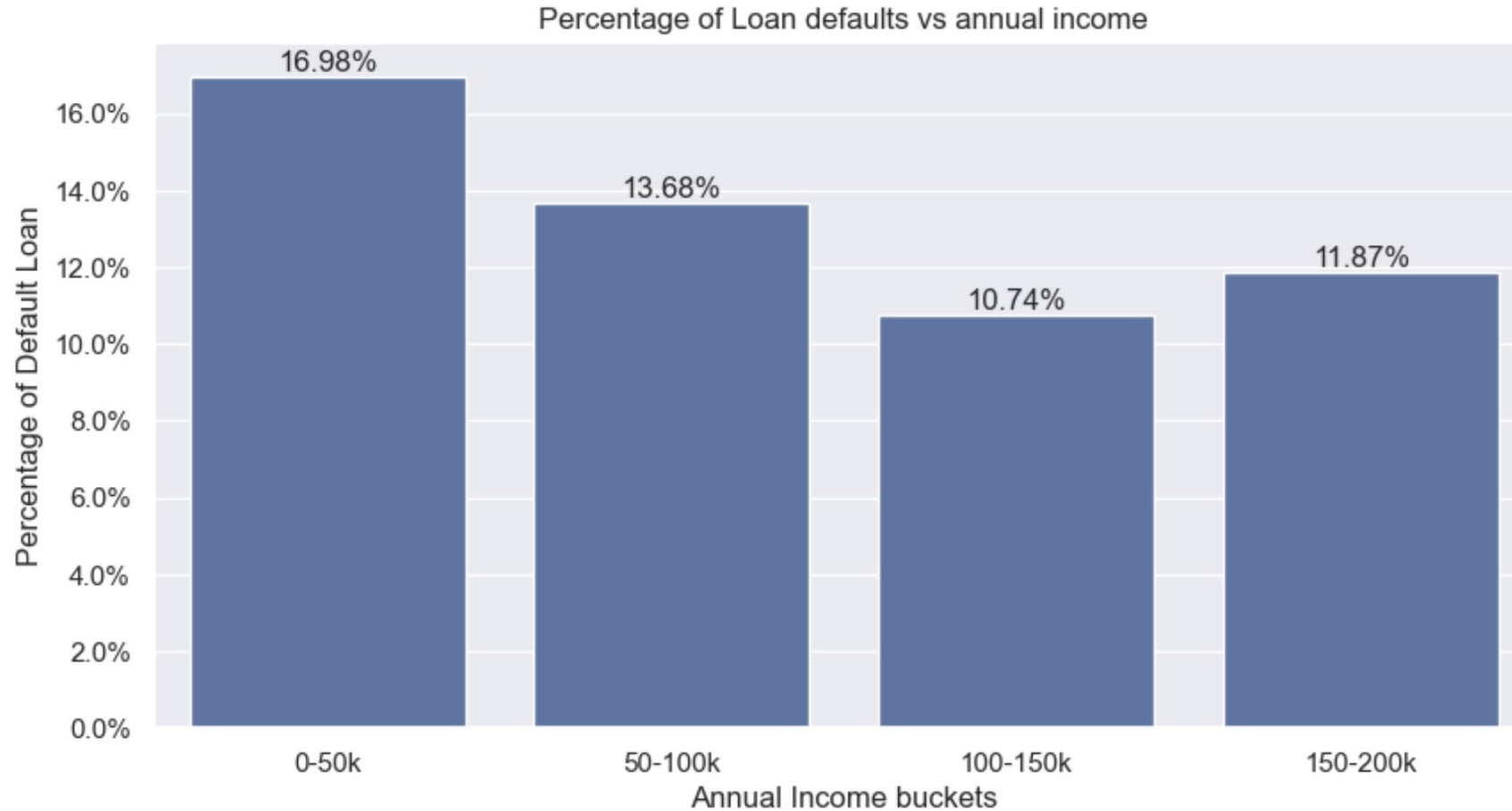
- '60 months' loans are twice likely to being defaulted compared to '36 months' loans.

# Loan Count vs Annual Income



- Most borrower's income are in the range 40000 to 80000

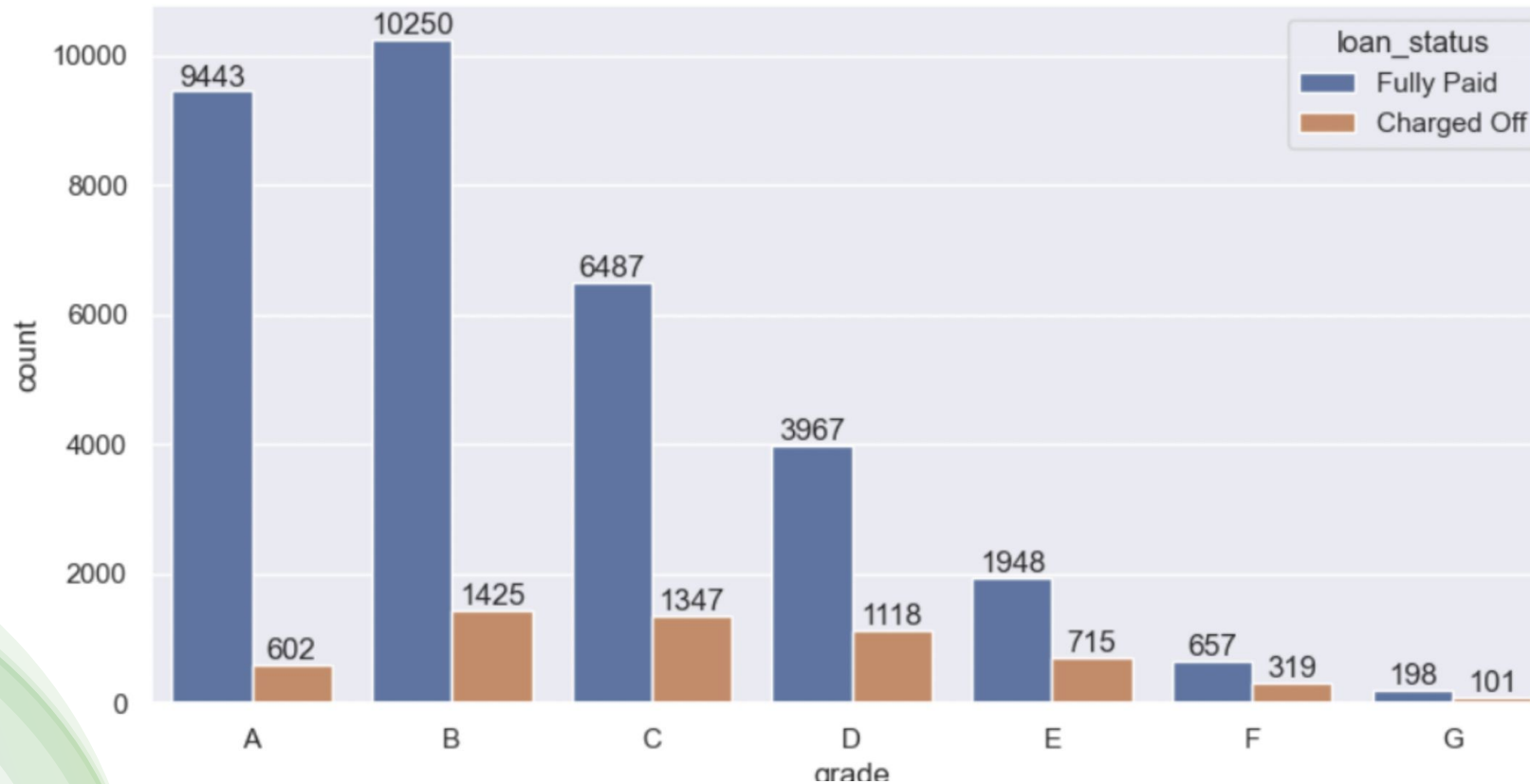
# Default Rate vs Annual Income



- Chances of default rate is higher for low income group (specifically for 0-50K income bracket)

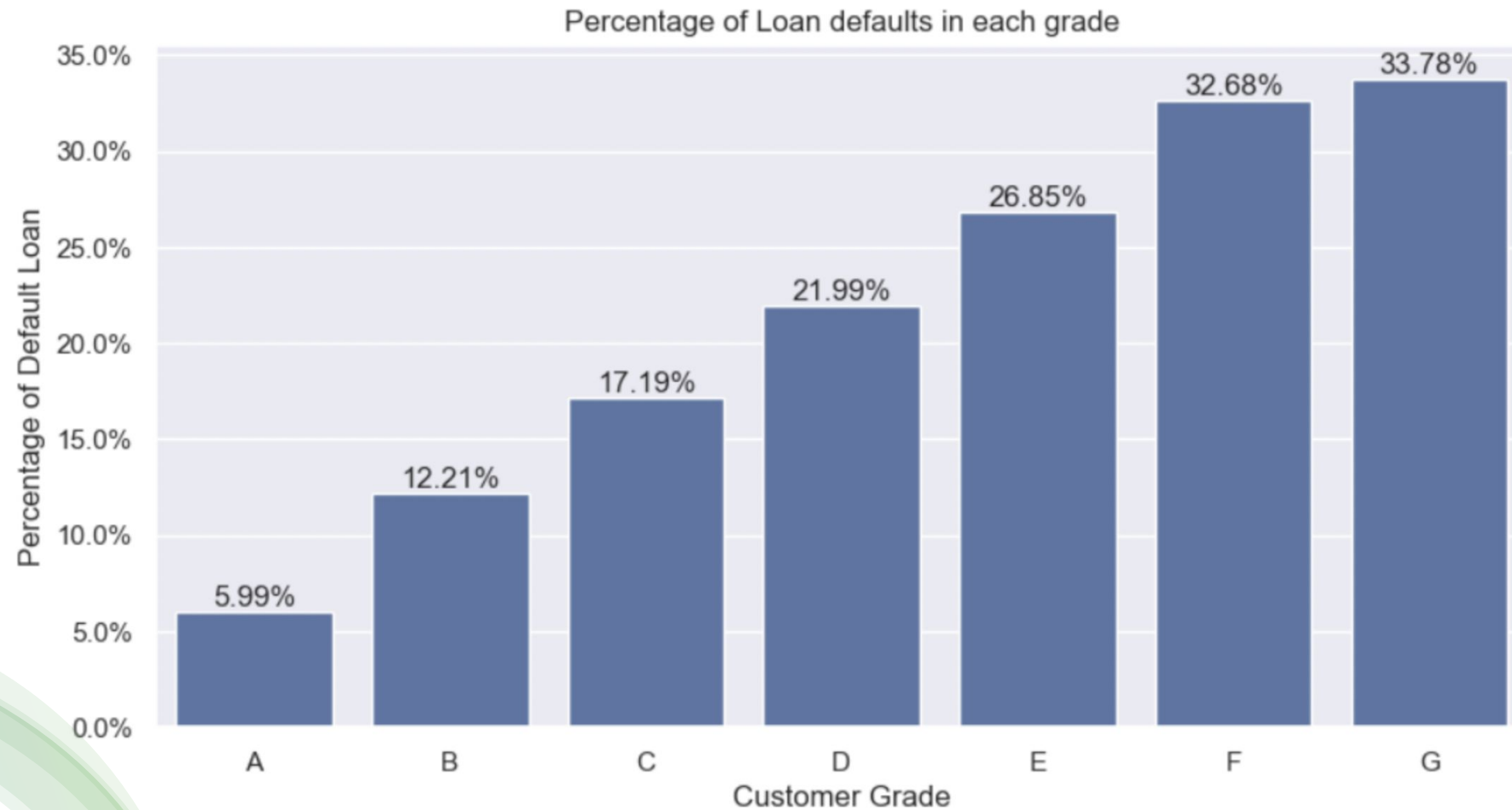
# Loan Status vs Grades

Count of Grades vs Loan Status



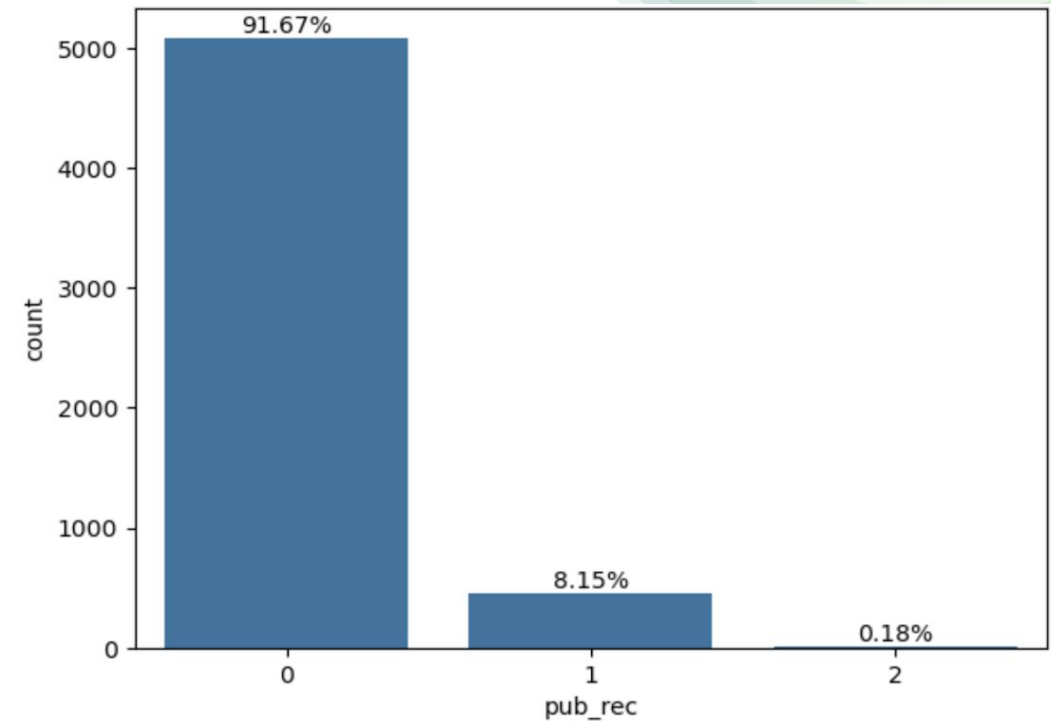
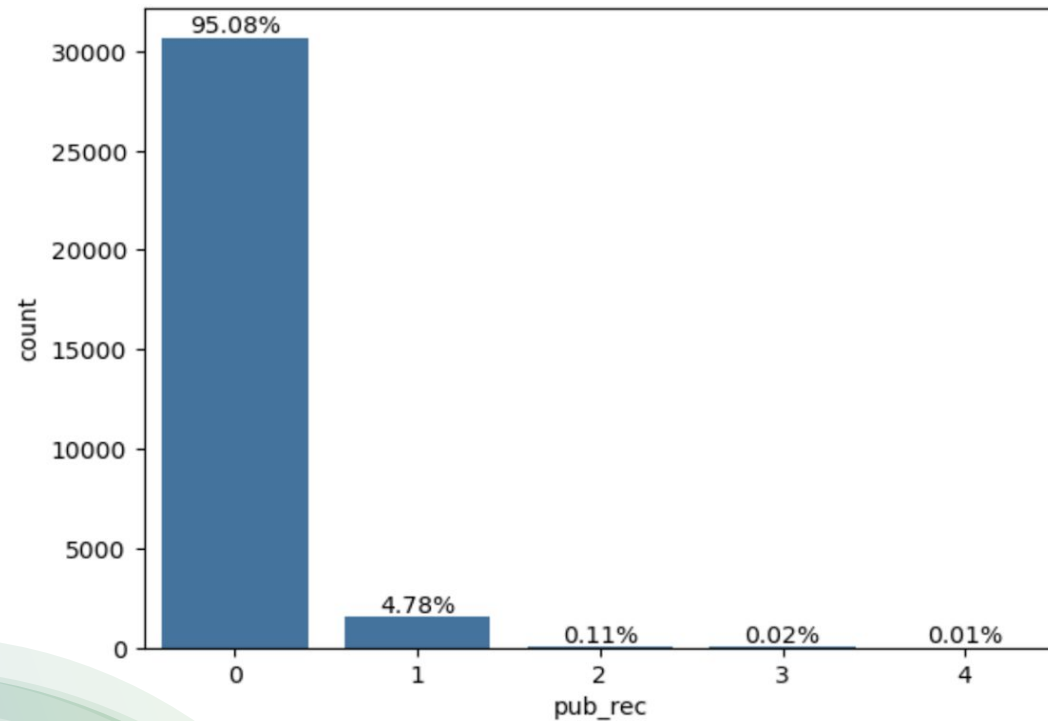
- More loans given to category A, B and C.
- Grade A has very less default rate compared to number of loans process in that bucket.

# Risk Analysis of defaulters acc to Grades



- Customers with Grades of E,F,G has more chances of defaulting the loan

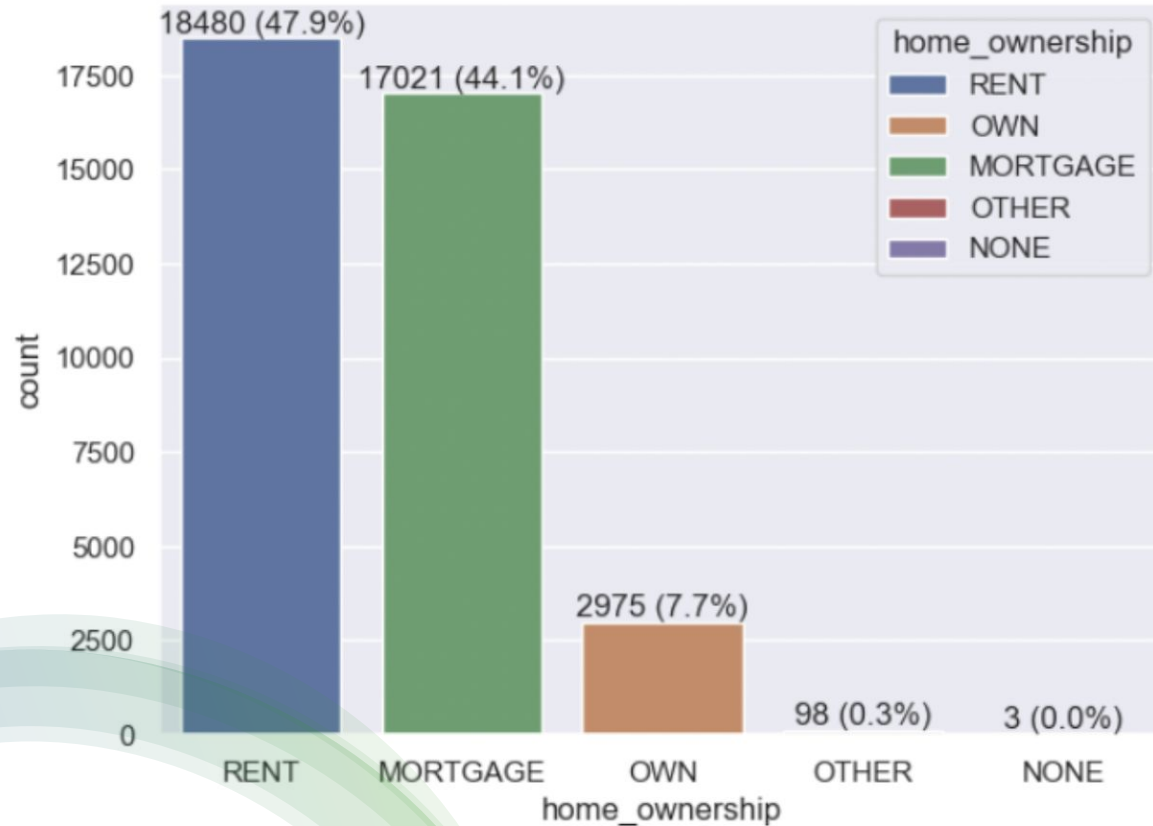
# Loan Default vs Public Record



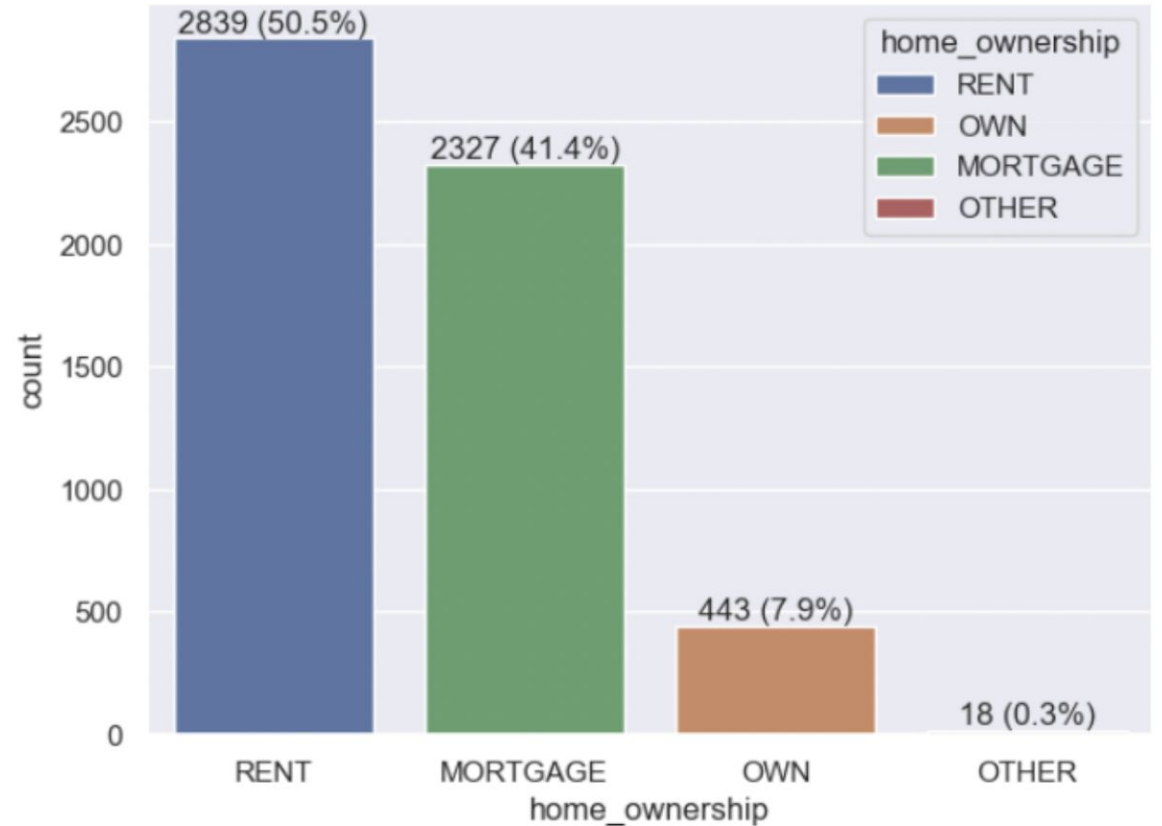
- The number of loans having any public record is very less to make any inference. But applicants with 1 public record have higher chances of default.

# Loan Count vs Home Ownership

Fully Paid Loans

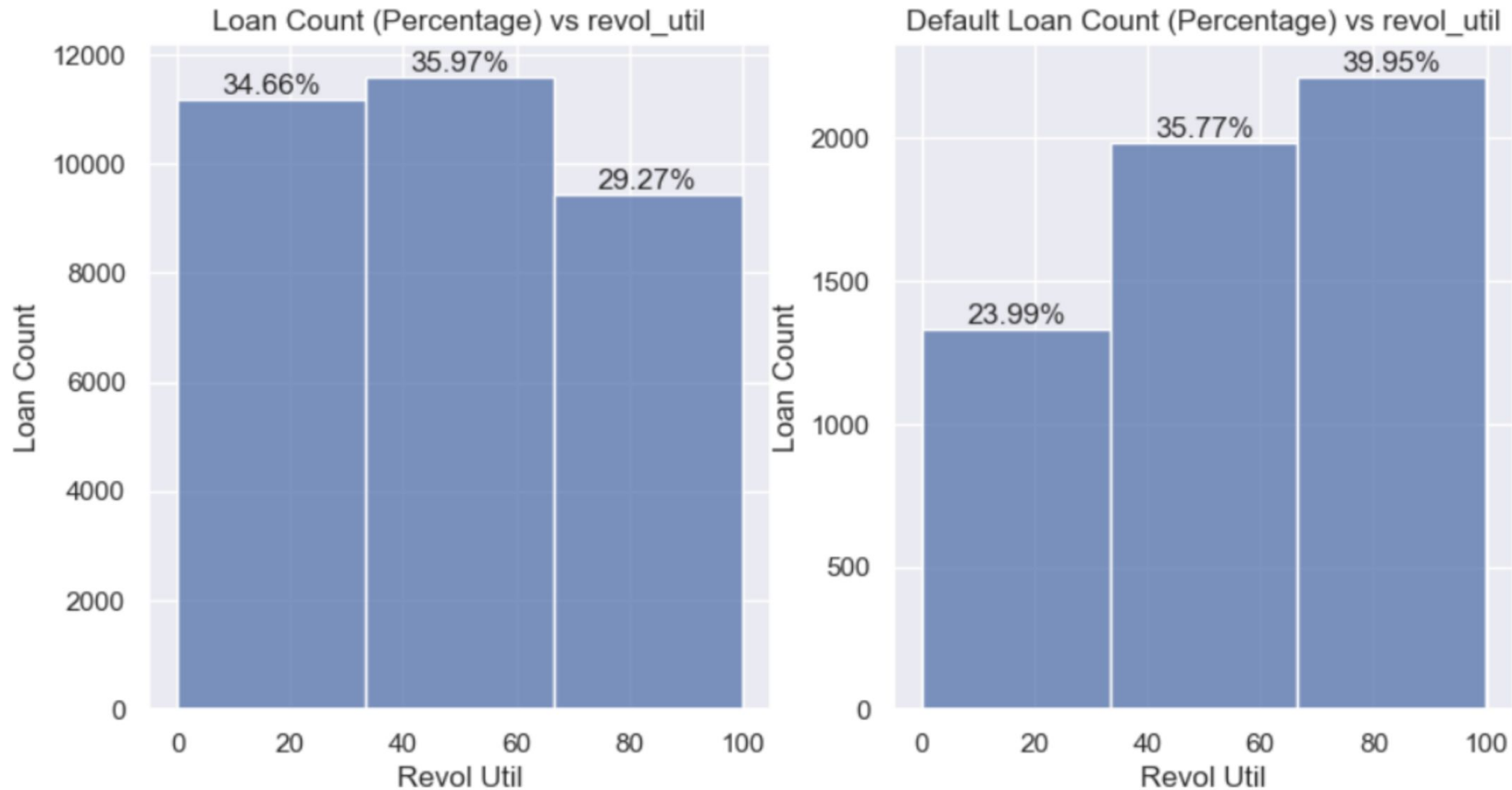


Charged Off Loans



- Maximum loans granted to Rent and Mortgage Category.
- Also Rent and Mortgage constitutes higher proportion in default cases as well.

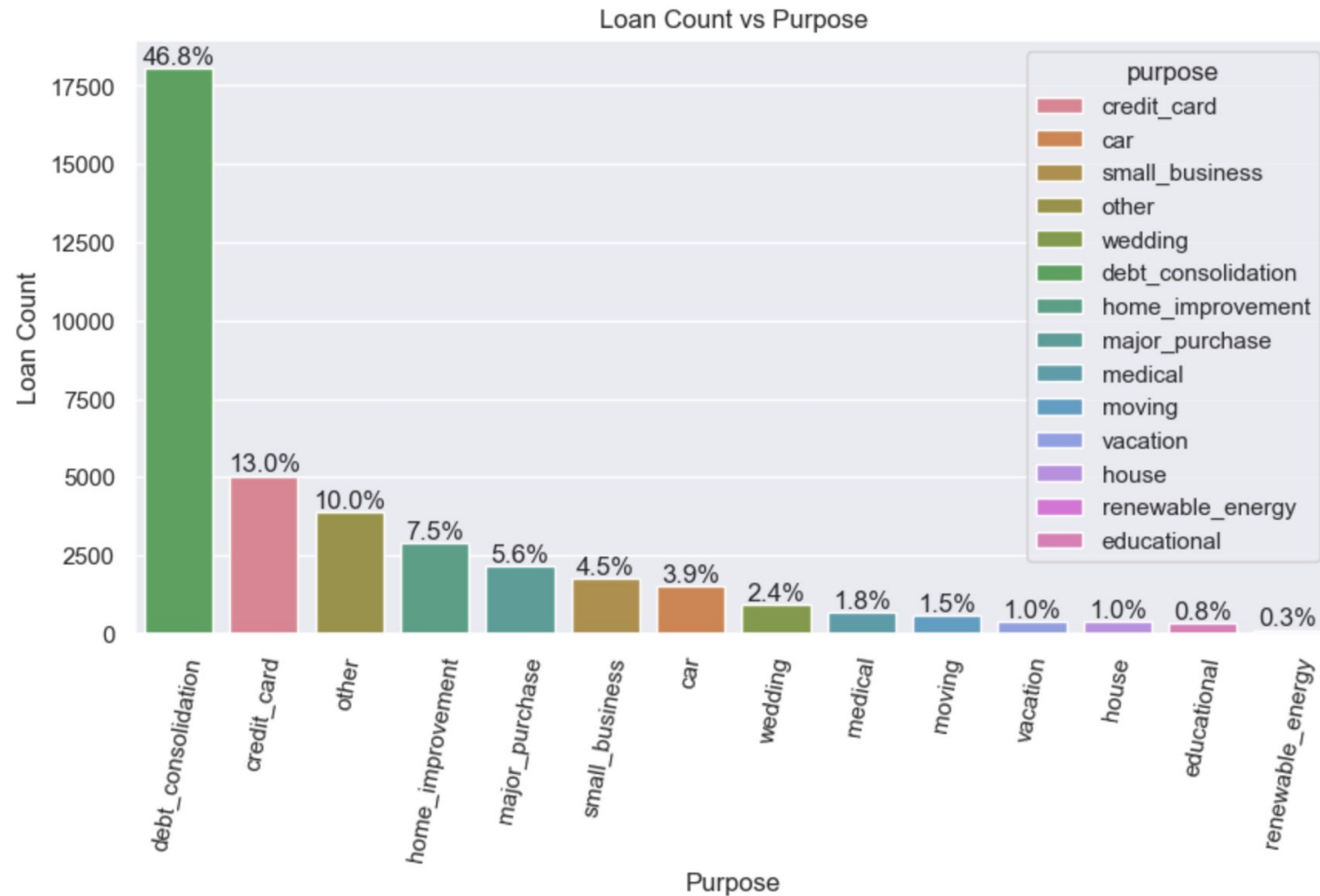
# Default Rate vs Revolving Credit utilisation



- In the defaulted loans, the revolving credit utilisation is greater than 70% for one third of the cases. Even 50% of the default loans have revolving credit utilisation greater than 60%.

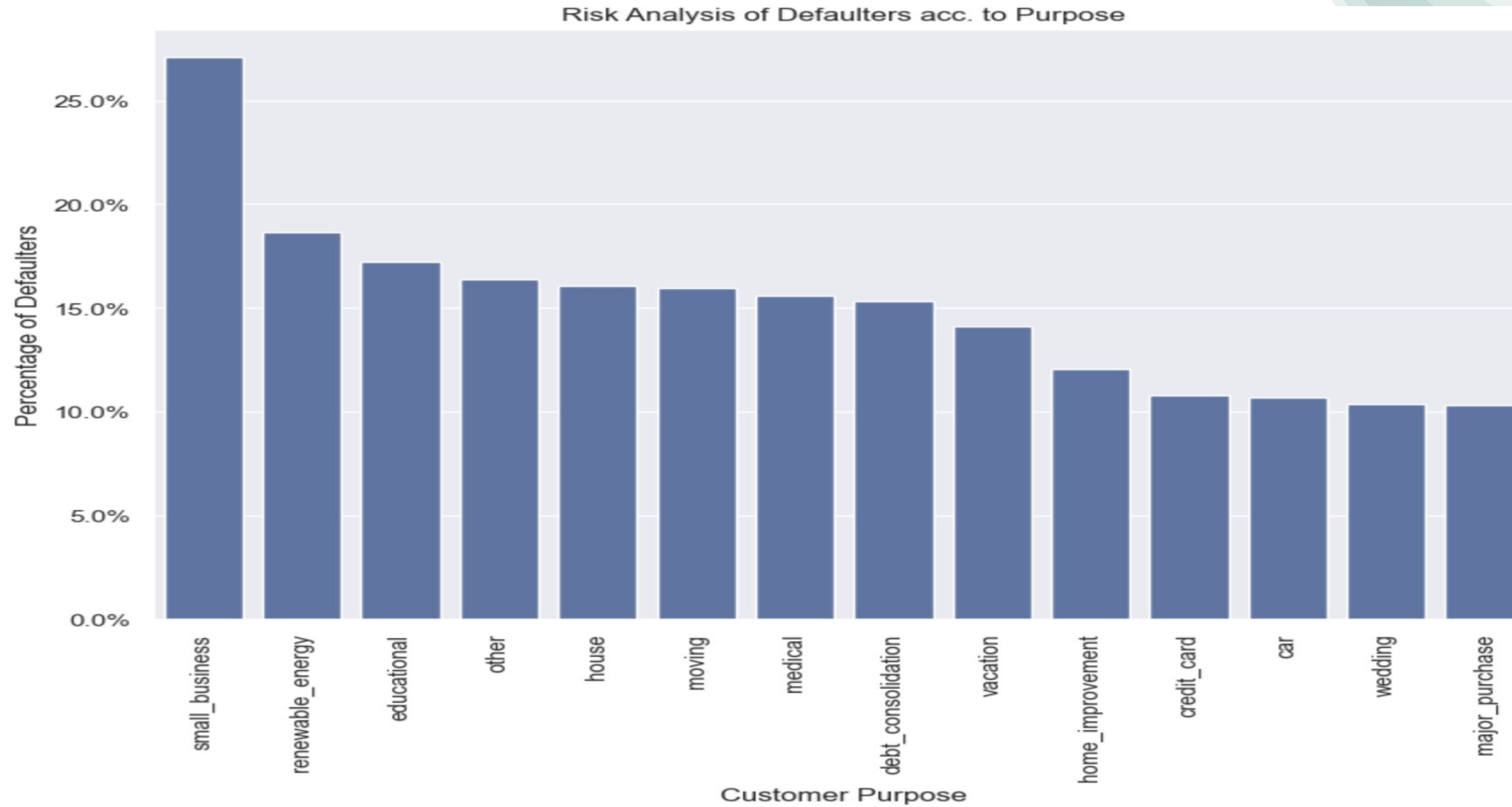


# Loan Count vs Purpose



- Debt Consolidation is the highly stated reason for taking loans. They constitute ~47% of overall loans.
- In the defaulted loans as well, it follows the similar trend.

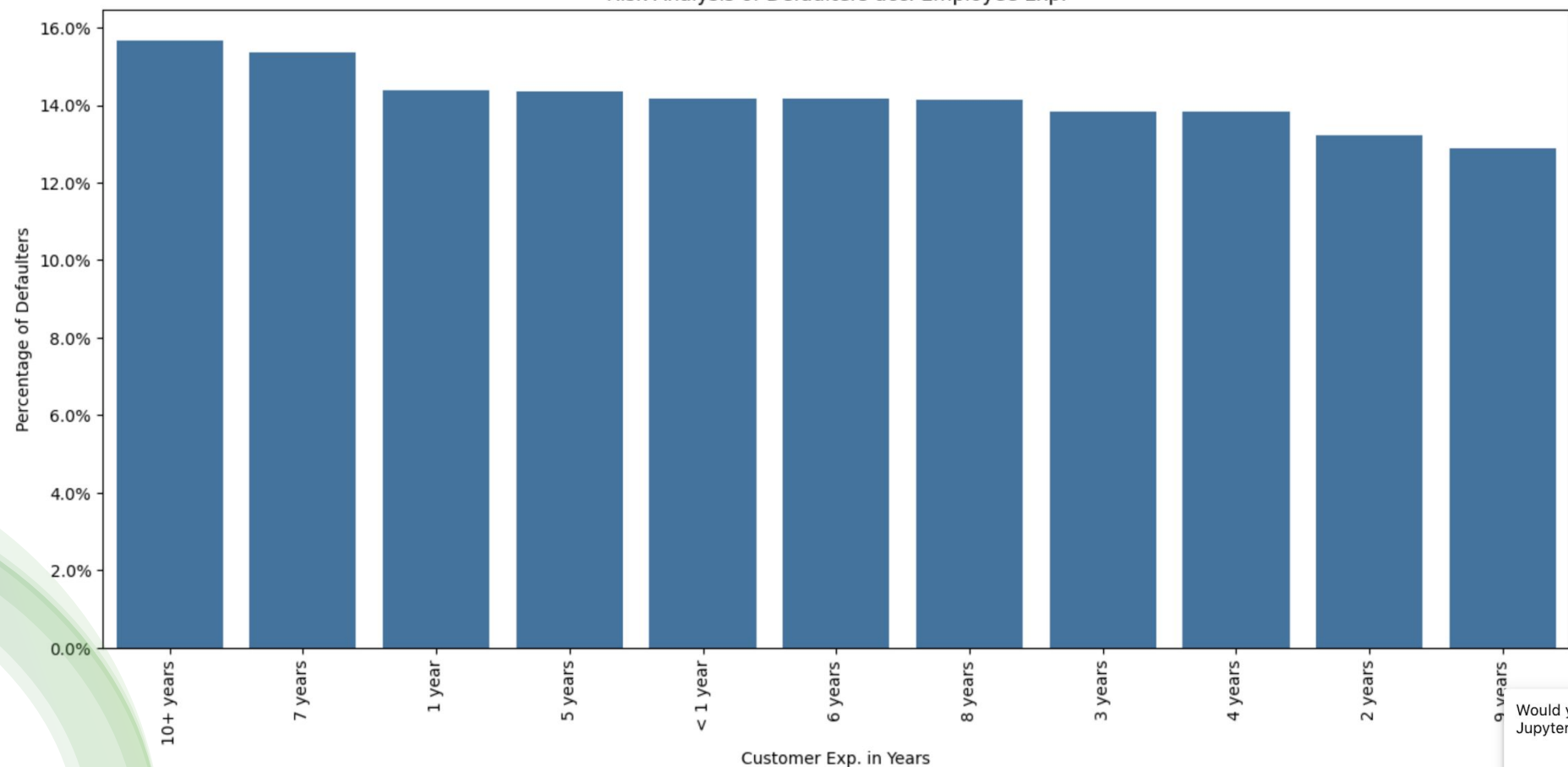
# Default Rate vs Purpose



- Loan taken for small business have higher default rate though. Debt consolidation is not even in top 5!

# Risk Analysis of defaulters acc to Employee Experience

Risk Analysis of Defaulters acc. Employee Exp.



Would you like to go  
Jupyter news?

# Observations

- **Loan Amount** - The default rate increases significantly as the loan amount increase from 15K
- **Loan Term** - The probability of default is more than twice for 60 months (25%) compared to that of 36 months (11%)
- **Interest rate** - Higher interest rates lead to higher rate of default.
- **Grades** - The default rate increases from Grade A to grade G. But Grade A has very less default rate.
- **Annual Income** - Default rate is higher in lower income bucket i.e between 0-50k
- **Revolving Credit Utilisation** - Default loans have higher revolving utilisation rates in general. Nearly 60% of default loans have more than 50% revolving credit utilisation.
- **Purpose** - Debt consolidation has higher loan applications and also default rate. But the loans provided for small business have higher chances of default.

# Closing Notes

- Tighten evaluation for all applications where loan amount is greater than 15k.
- If the credit score is arriving at higher interest rate, then the probability of default is also higher. Hence the company should have appropriate risk management measures.
- Avoid lending money for borrowers having very less annual income. Target people who have more than 40k in general.
- Grade A, B category borrowers are less risky. Hence target more of those.
- If `revol_util` is greater than 50%, then tighten evaluation process
- Preference should be given to 36months term instead of 60 months
- Tighten lending process for small business.



Thank you