Lending Club Case Study

- BY
  - Dipak Sah
  - Nitesh Kumar

- Date : 7th Feb 2024

Business Understanding

- **Lending Club is a consumer finance company.**
- company receives multiple 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

Problem Statement

- Lending Club is the largest available loan marketplace for different sorts of loans.

- Minimize the risk of lending money by using EDA technique to find out driving factors for an applicant/borrower which can default and create loss of finance.

Analysis Approach

Data Sourcing

Data Loading

Data Cleaning

- Missing Value Treatment
- Outlier Treatment

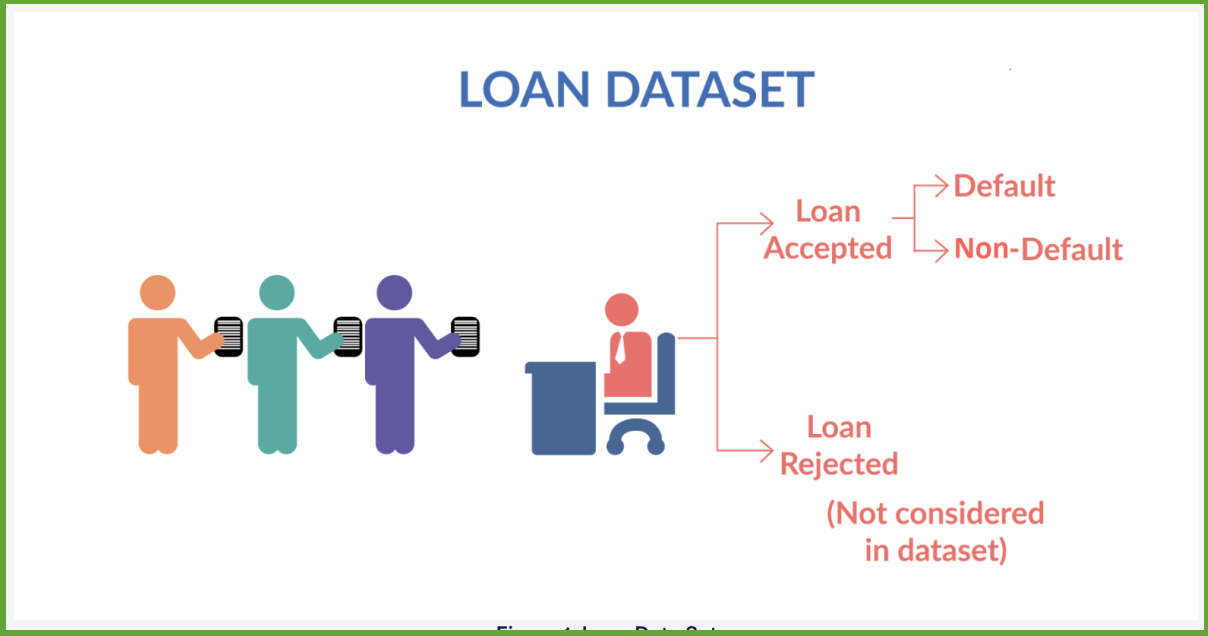
Data Analysis

- Univariate Analysis
- Bivariate Analysis
- Derived Metrics
- Multivariate Analysis

Business Recommendations

Dataset Understanding

Dataset File : loan.csv
Data Dictionary: Data_Dictionary.xlsx



Data Cleaning

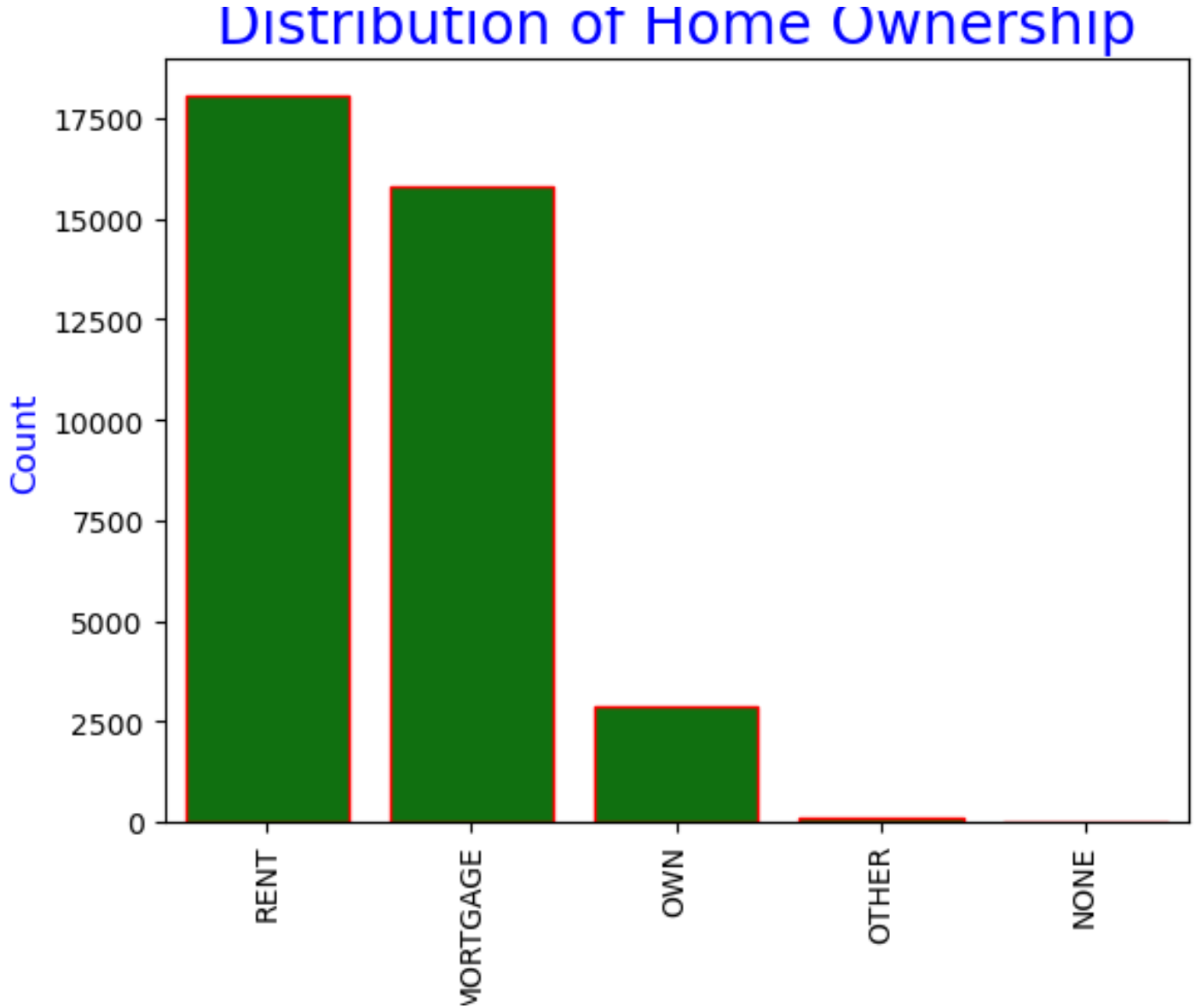
- Missing Value Treatment
- Imputing/dropping columns and rows.
- Outlier Treatment

Univariate Analysis

Home ownership vs Frequency

Observation :

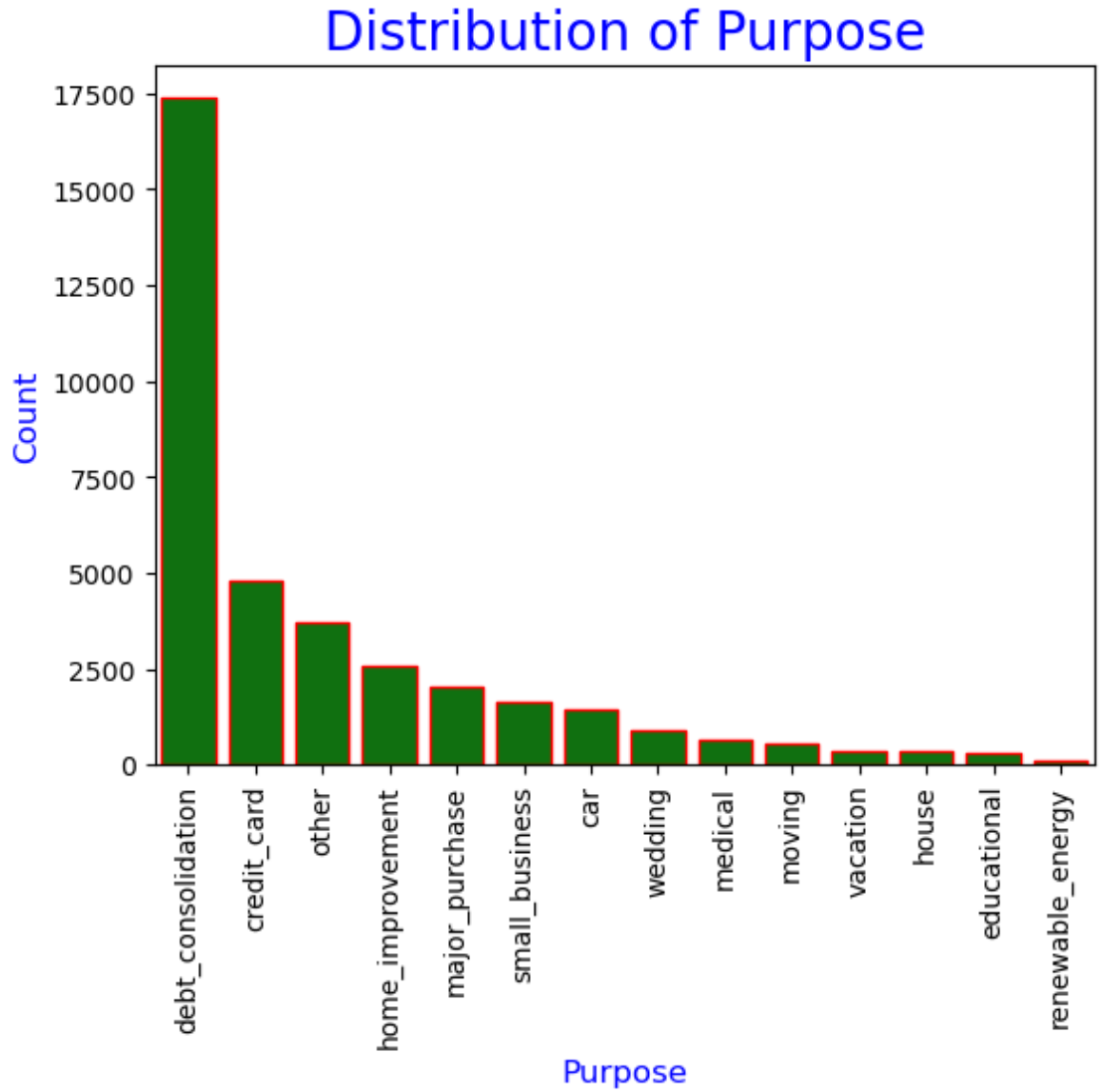
- Most of the customer (close to 92%) are either on rent or having mortgaged home.
- Very less applicant have their own home.



Purpose of Loan vs Frequency

Observation :

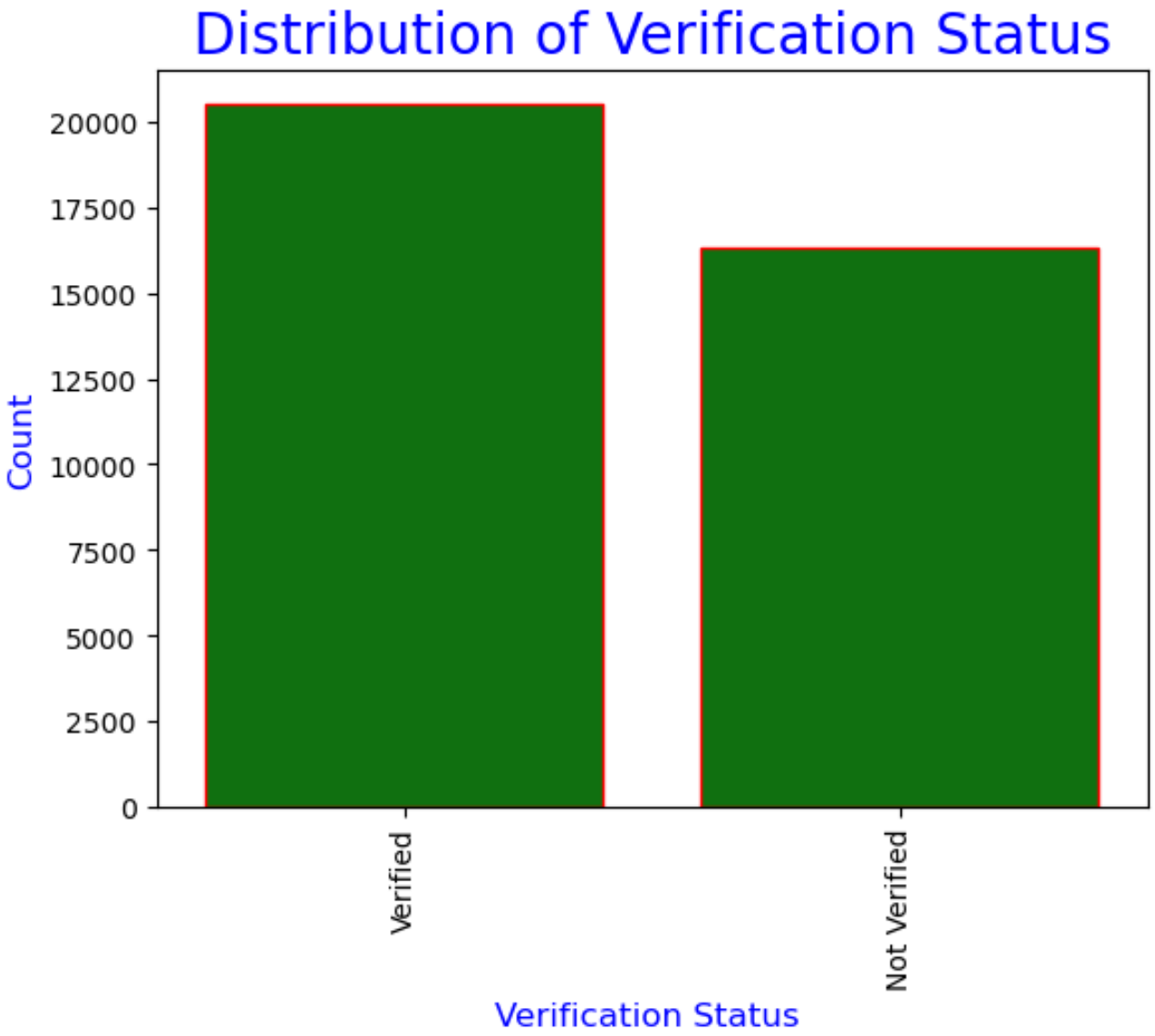
- Almost half of the applicants have applied loan with purpose of debt consolidation.
- Around 13 % of the applicants have applied for loan with purpose of paying credit card bills.
- Conclusion is that more than 60% of the applicants/customers have applied for loan with purpose of paying some other existing debt or credit bills.



Verification Status vs Frequency

Observation :

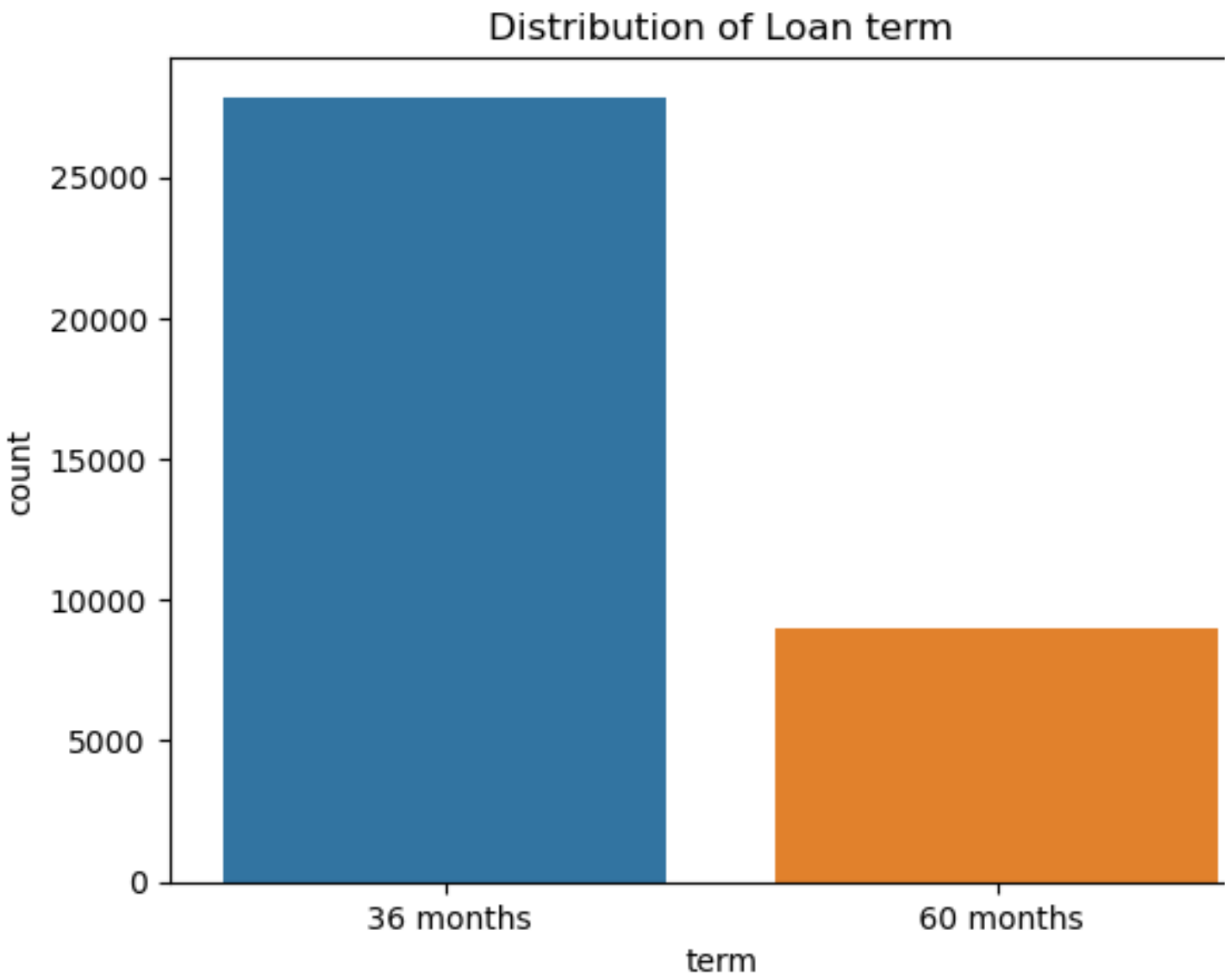
- Every 2 out of 5 applicant's verification is not completed which is very alarming in the *lending business.*



Term vs Frequency

Observation:

- 3 out 4 applicants go for 3 year loan term i.e., low tenure



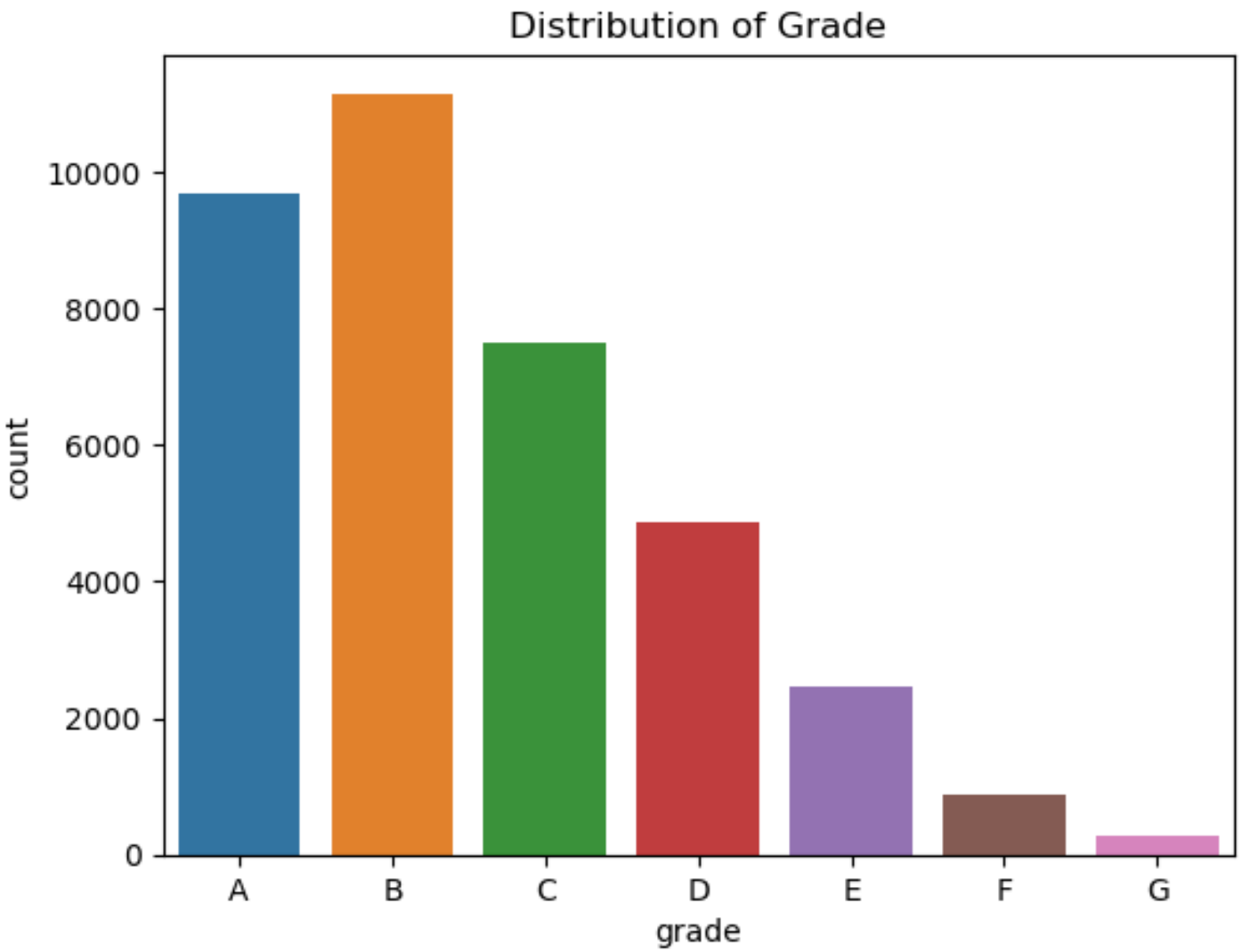
Grade vs Frequency

Observation:

- 3 out 4 applicants belong to grade A, B & C.
- Highest applicant count belongs to applicant B.
- Lowest applicant belongs to G.

Assumption:

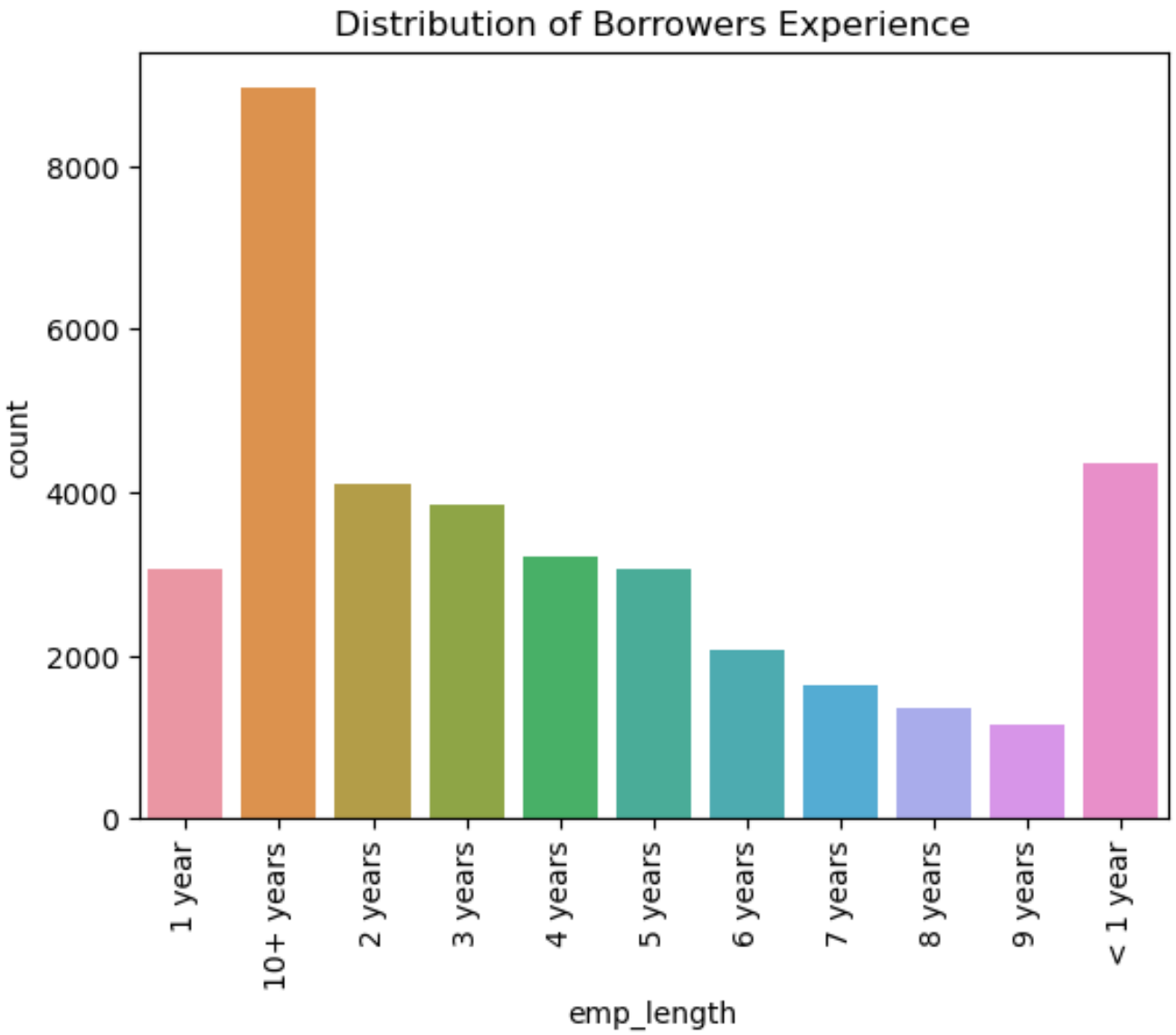
Assumption is A is higher grade and G is lower grade



Borrowers Work Experience vs Frequency

Observation:

- 25% of customers are having 10+ years of experience.
- 33% of customers are having less than or equal to 3 years of experience.

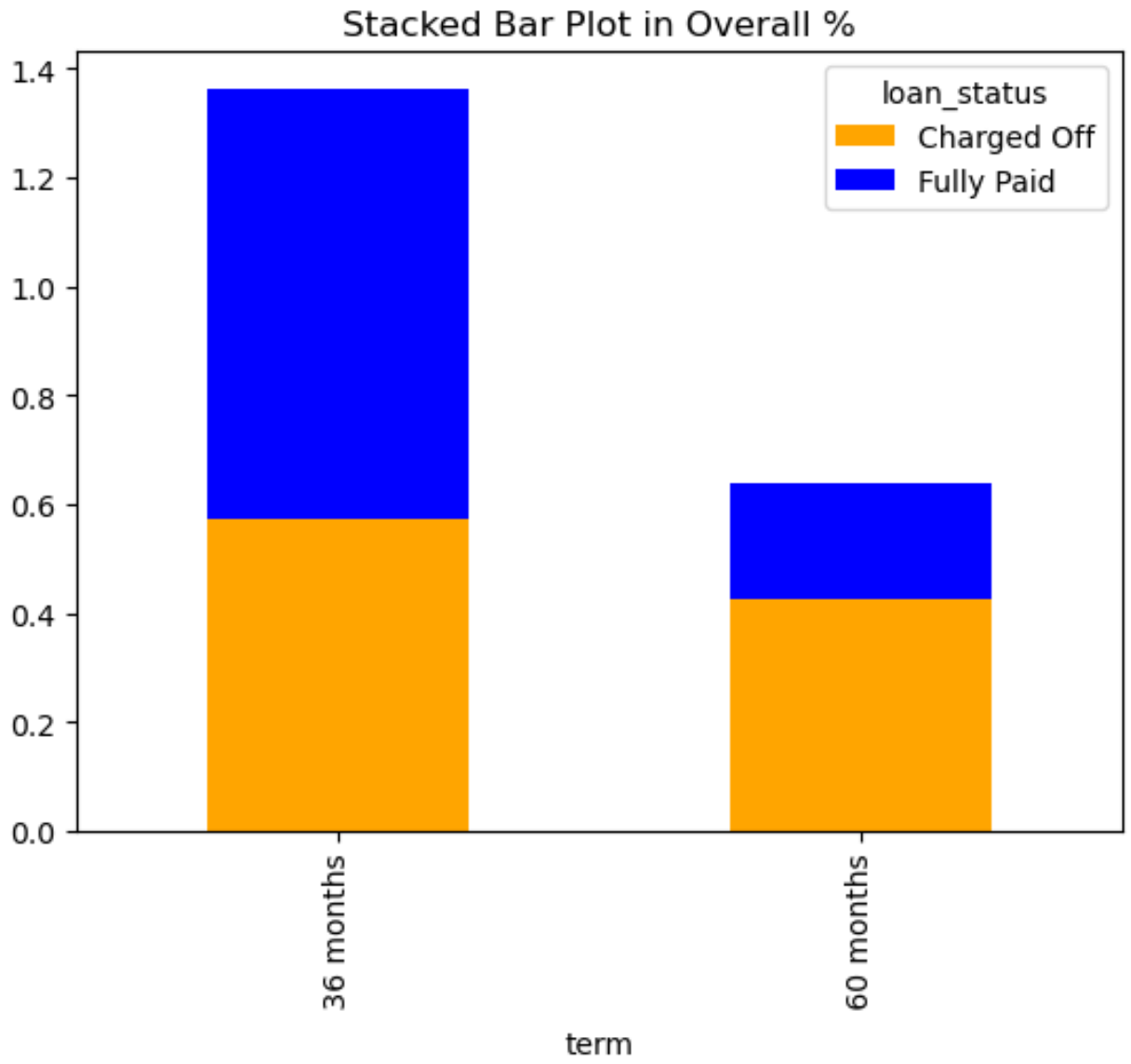


Bivariate Analysis

Term vs Loan Status

Observation:

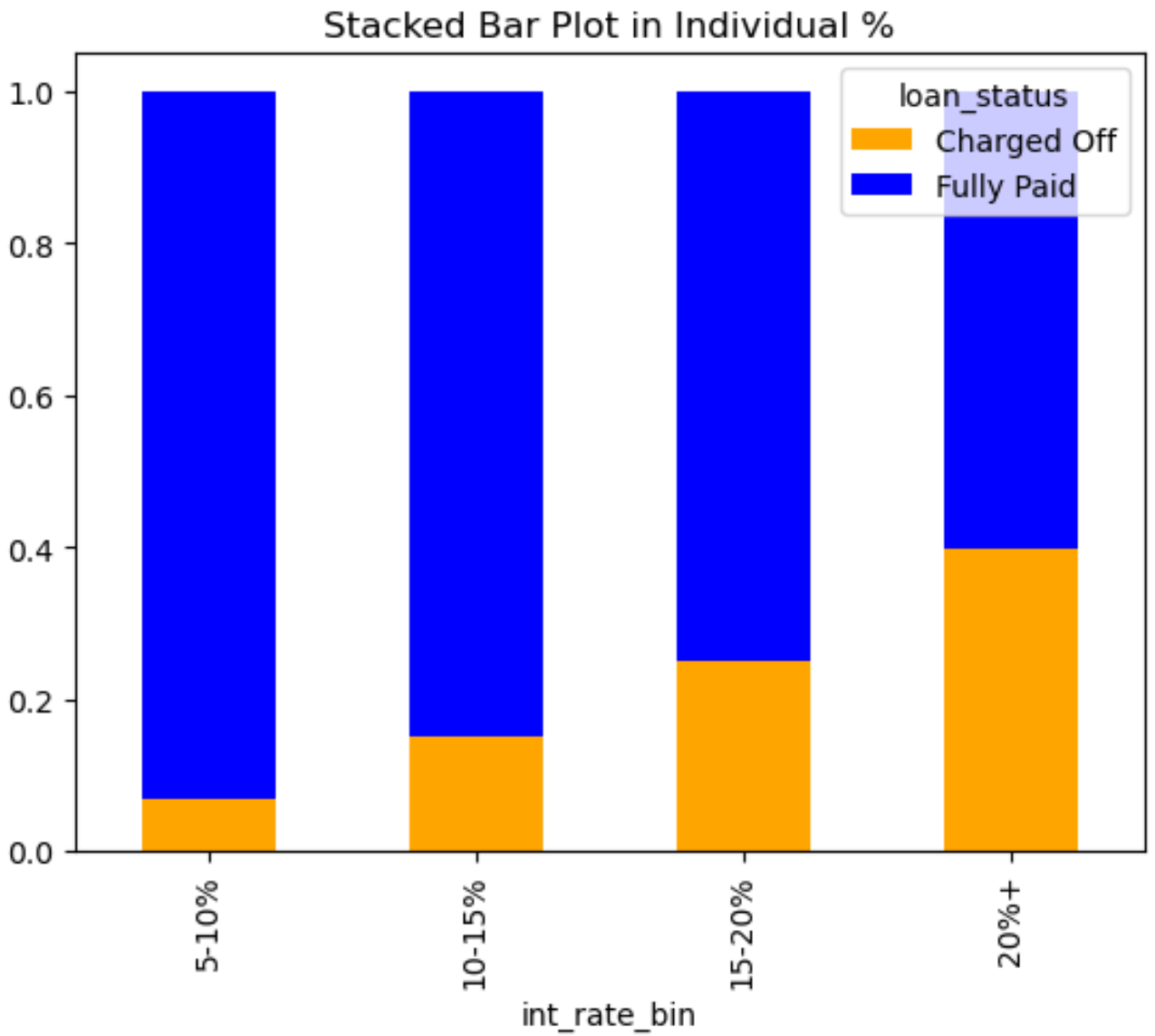
- Borrowers with 60 months term loan are more likely to default.



Interest rate vs Loan Status

Observation:

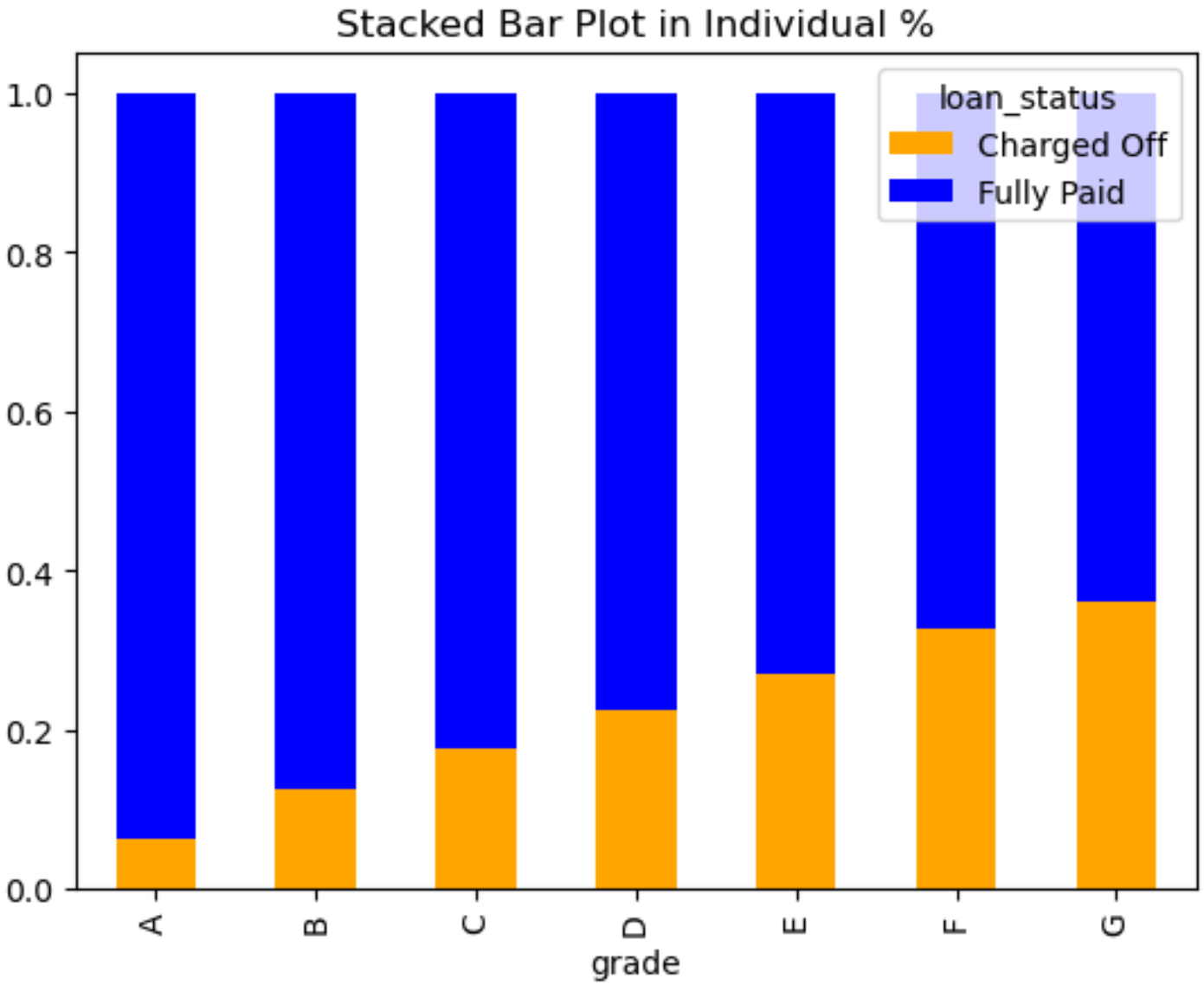
- Borrowers with 20%+ interest rate are more likely to default.



Grade vs Loan Status

Observation:

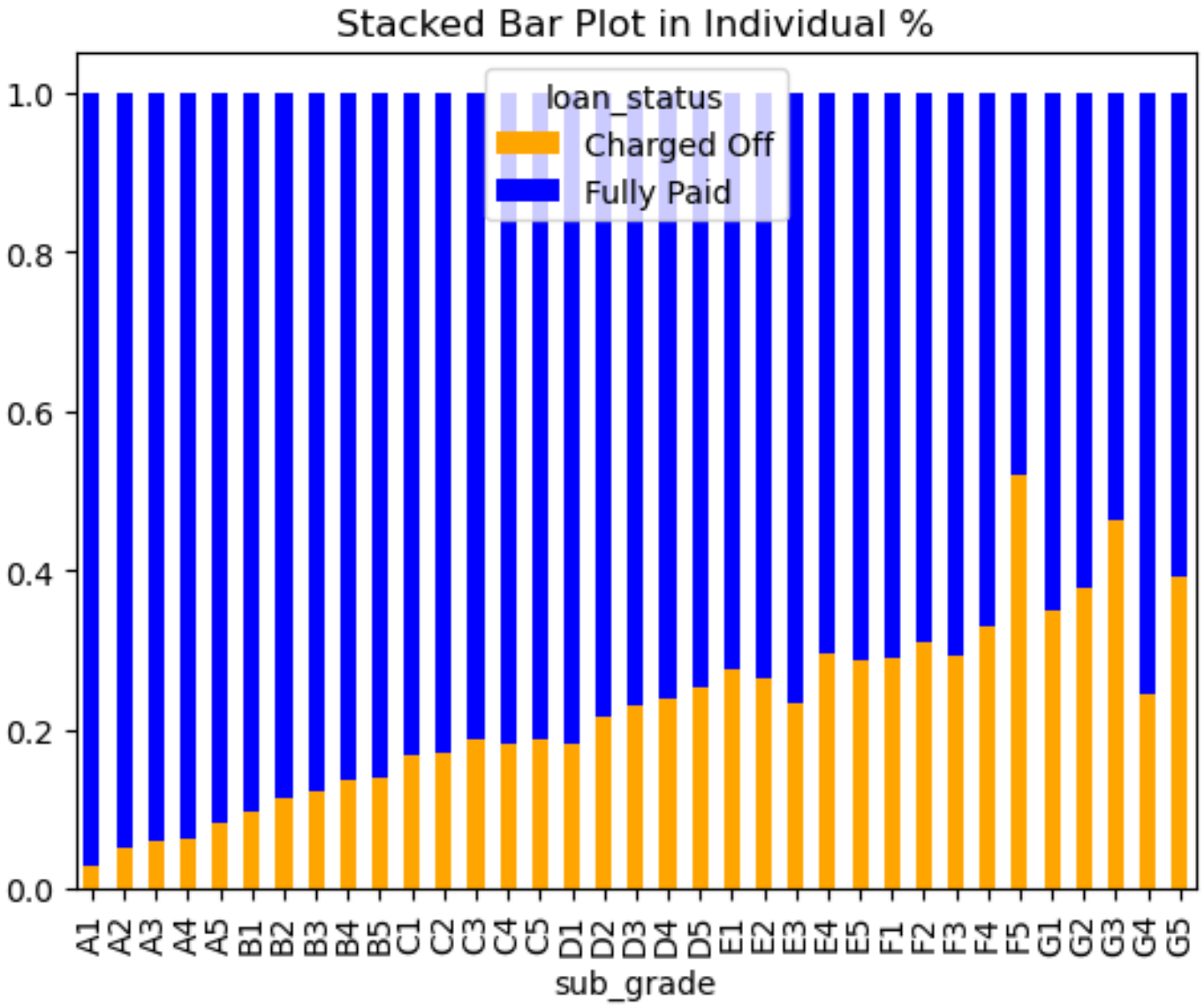
- Borrowers with lower grade are more likely to default.



Sub Grade vs Loan Status

Observation:

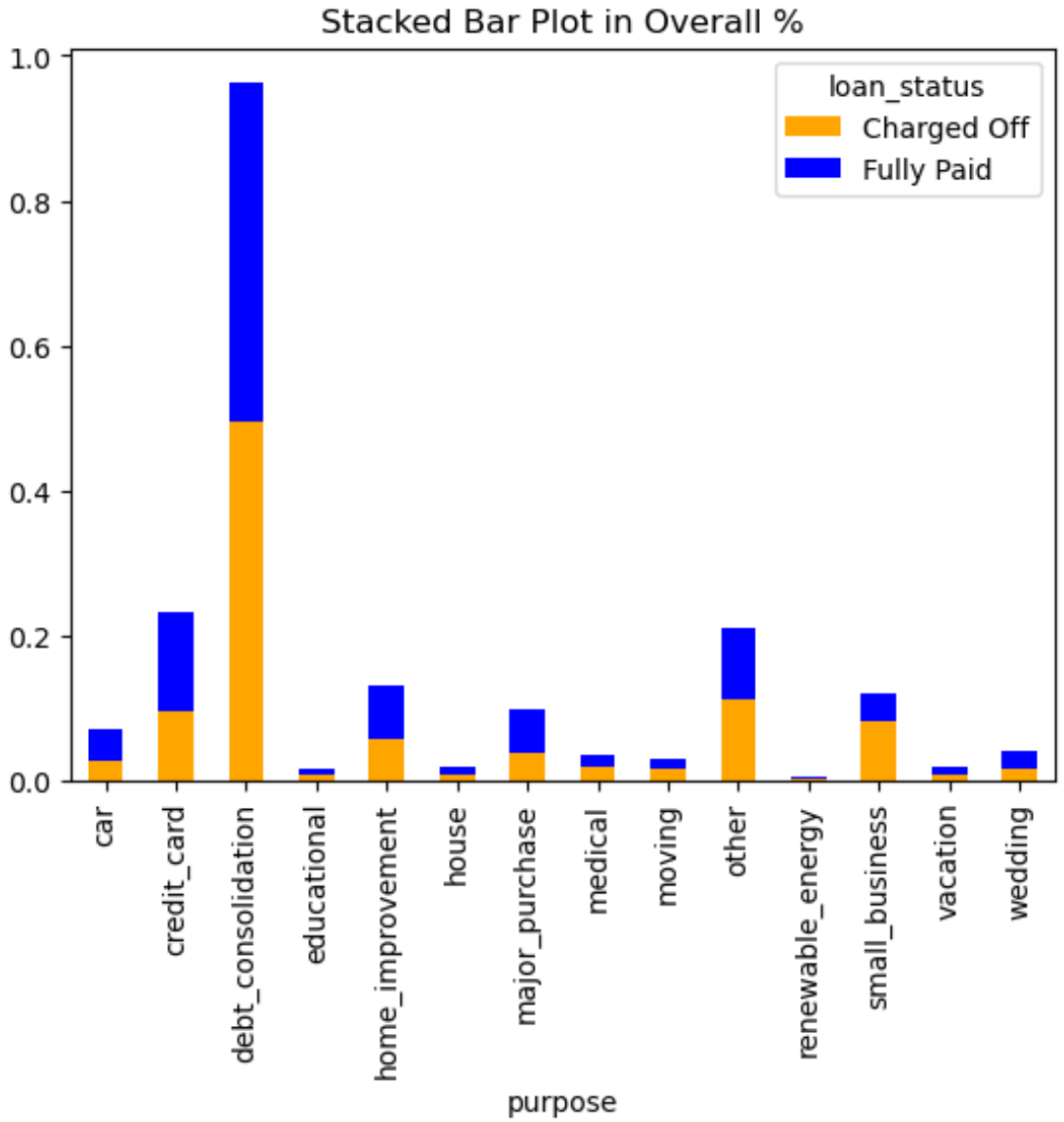
- Borrowers with sub grades of F5, G3 & G5 are more likely to default.



Purpose vs Loan Status

Observation:

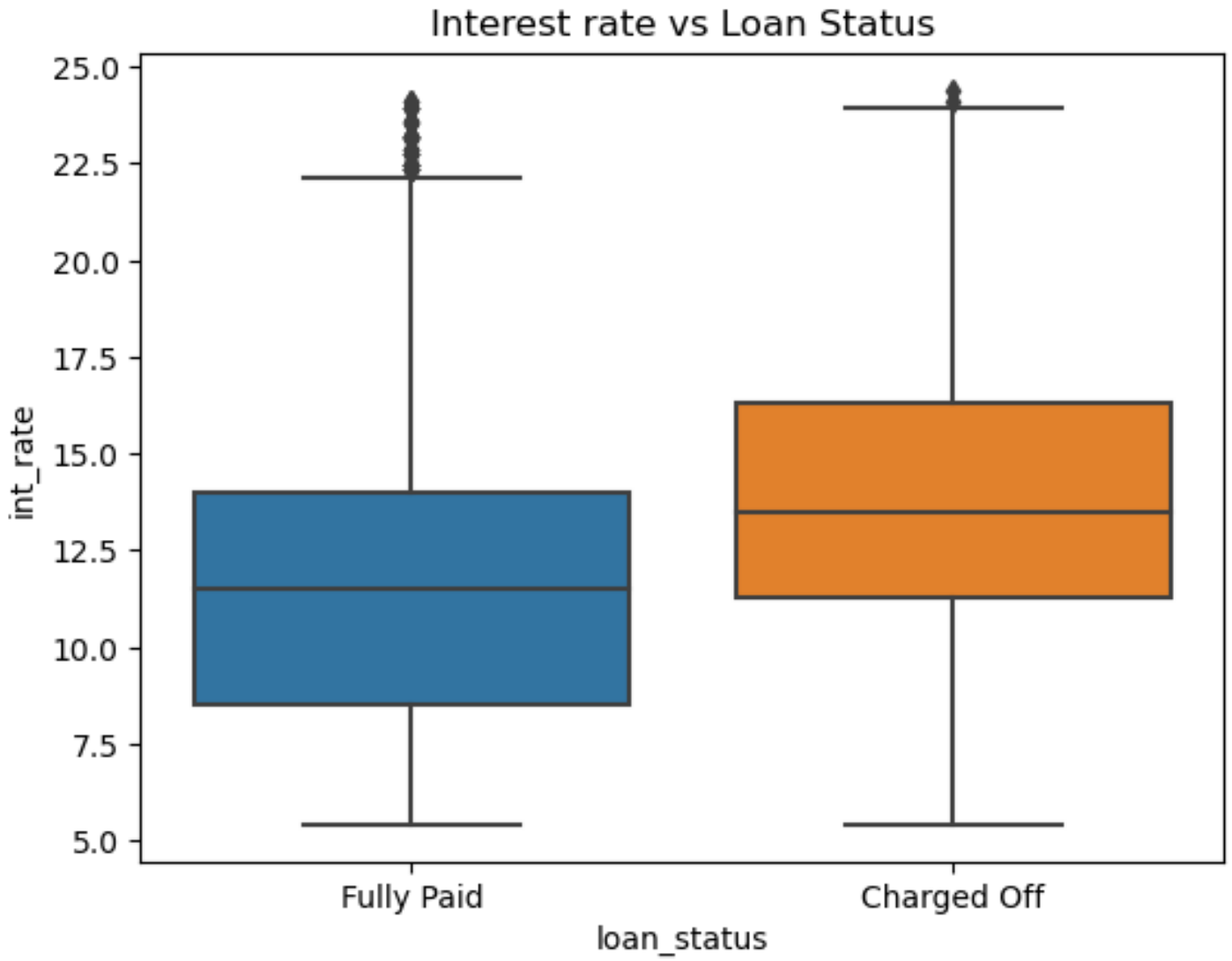
- Borrowers with purpose of below items are more likely to default.
  - Debt Consolidation
  - Credit Card
  - Small Business
  - Others



Interest Rate vs Loan Status

Observation:

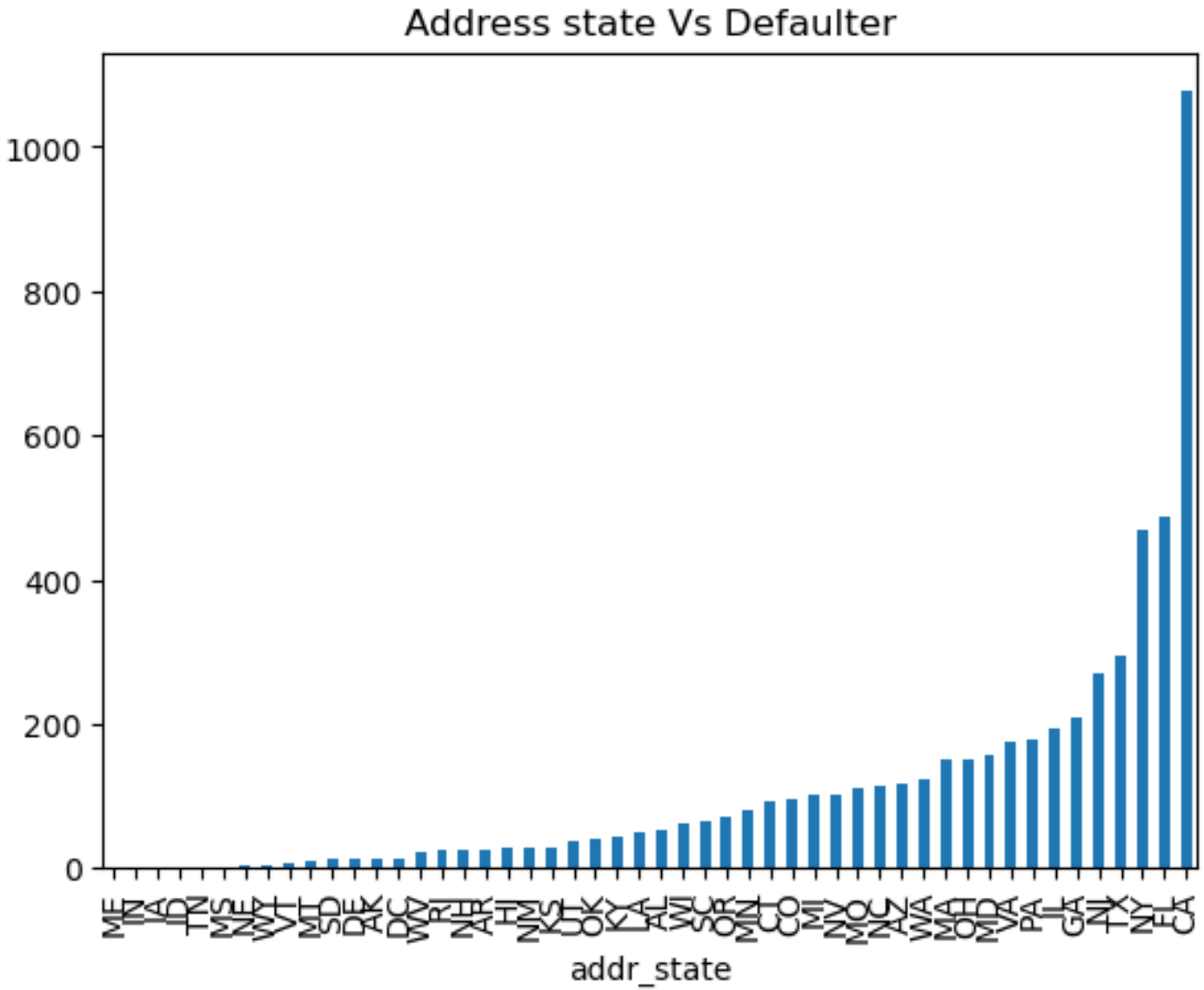
- Borrowers with higher interest rate are more likely to default.



Address State vs Defaulter

Observation:

- Borrowers from California are more likely to default.

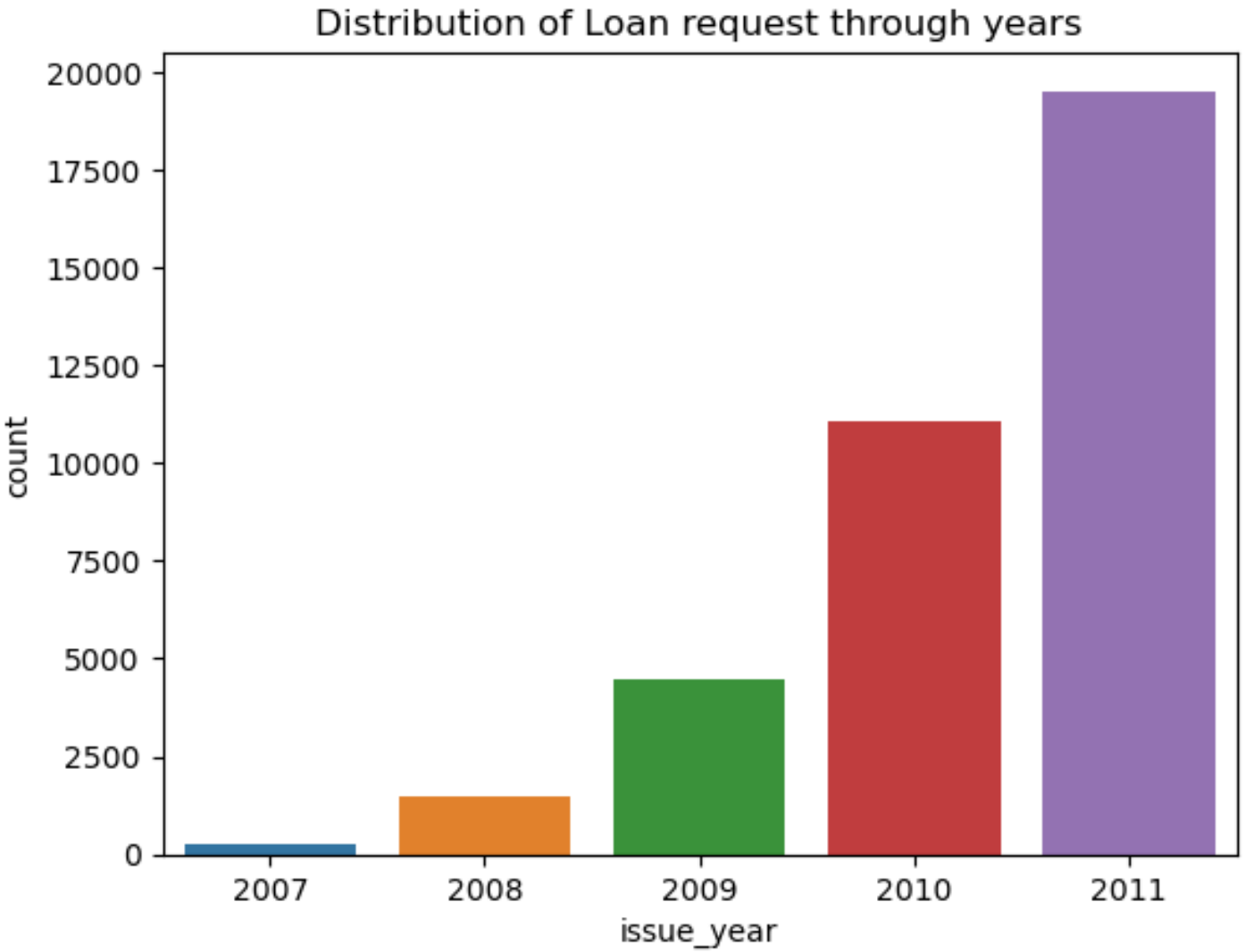


Derived Metrics

Loan request in year vs frequency

Observation:

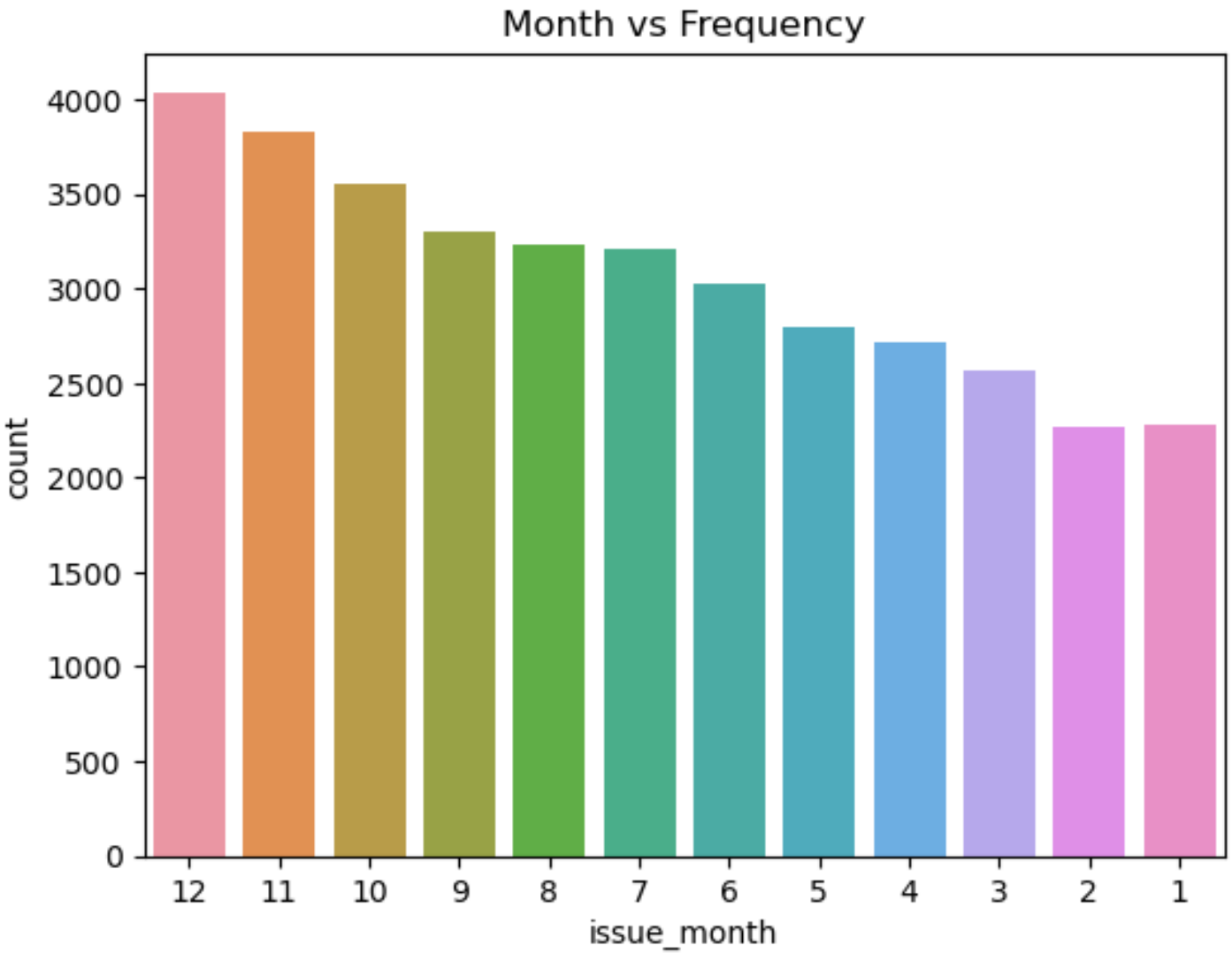
- With every passing year number of customer increased
- Lending club's business is in bullish trend.



Months vs Frequency

Observation:

- With each passing months in a year number of loans approved are increasing.

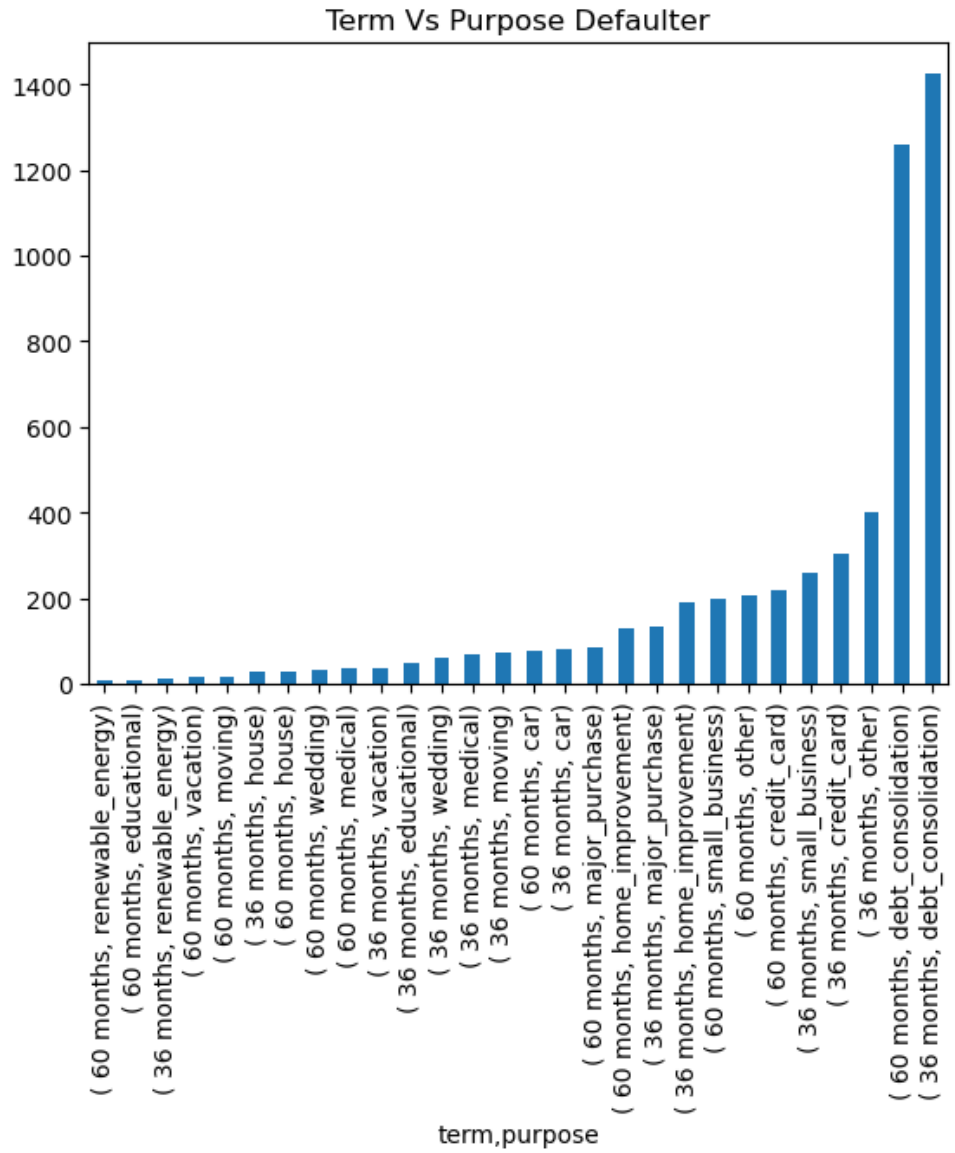


Multivariate Analysis

Term vs Purpose Defaulter

Observation:

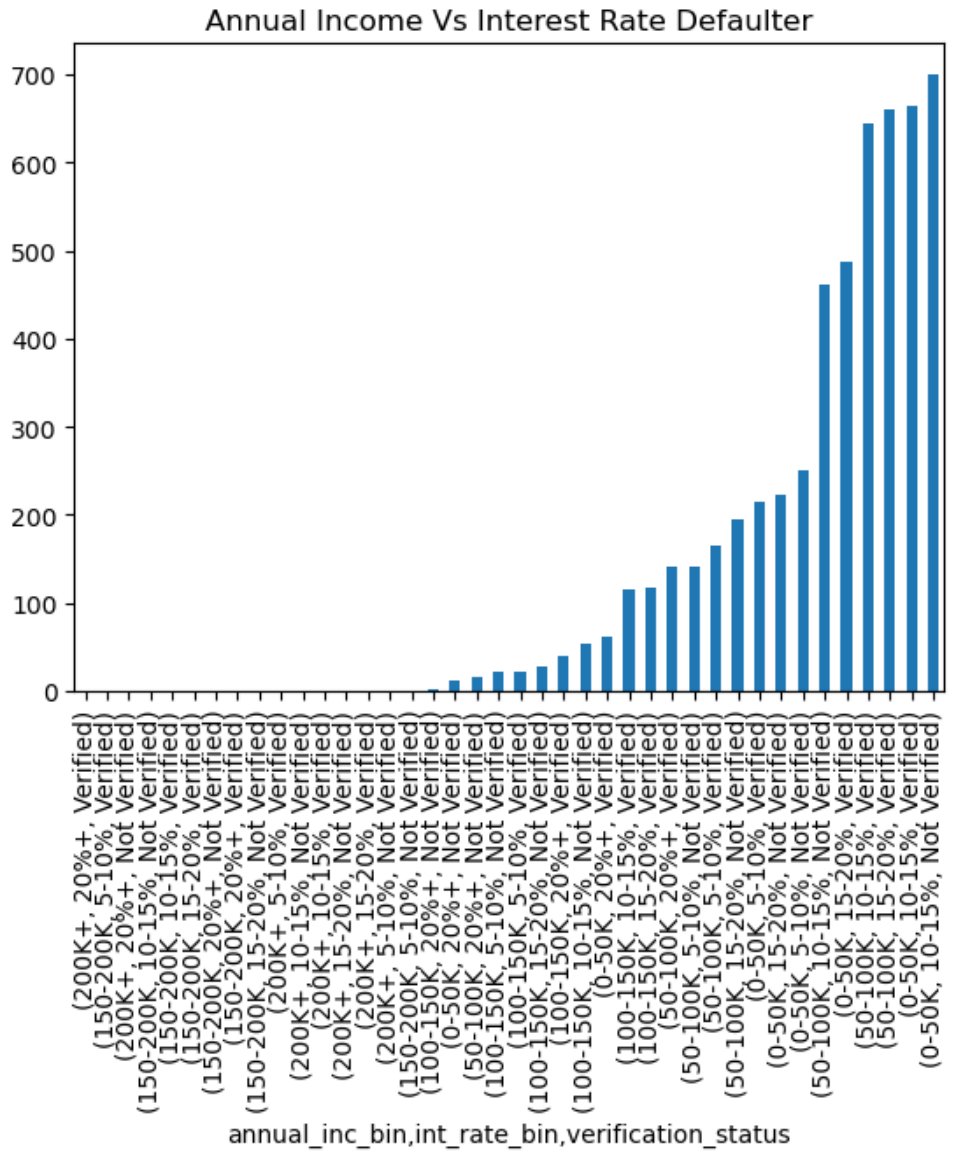
- Borrower applying with purpose of debt consolidation and 36 month low tenure are highly defaulting.



Interest rate vs Annual Income vs Verification Status

Observation:

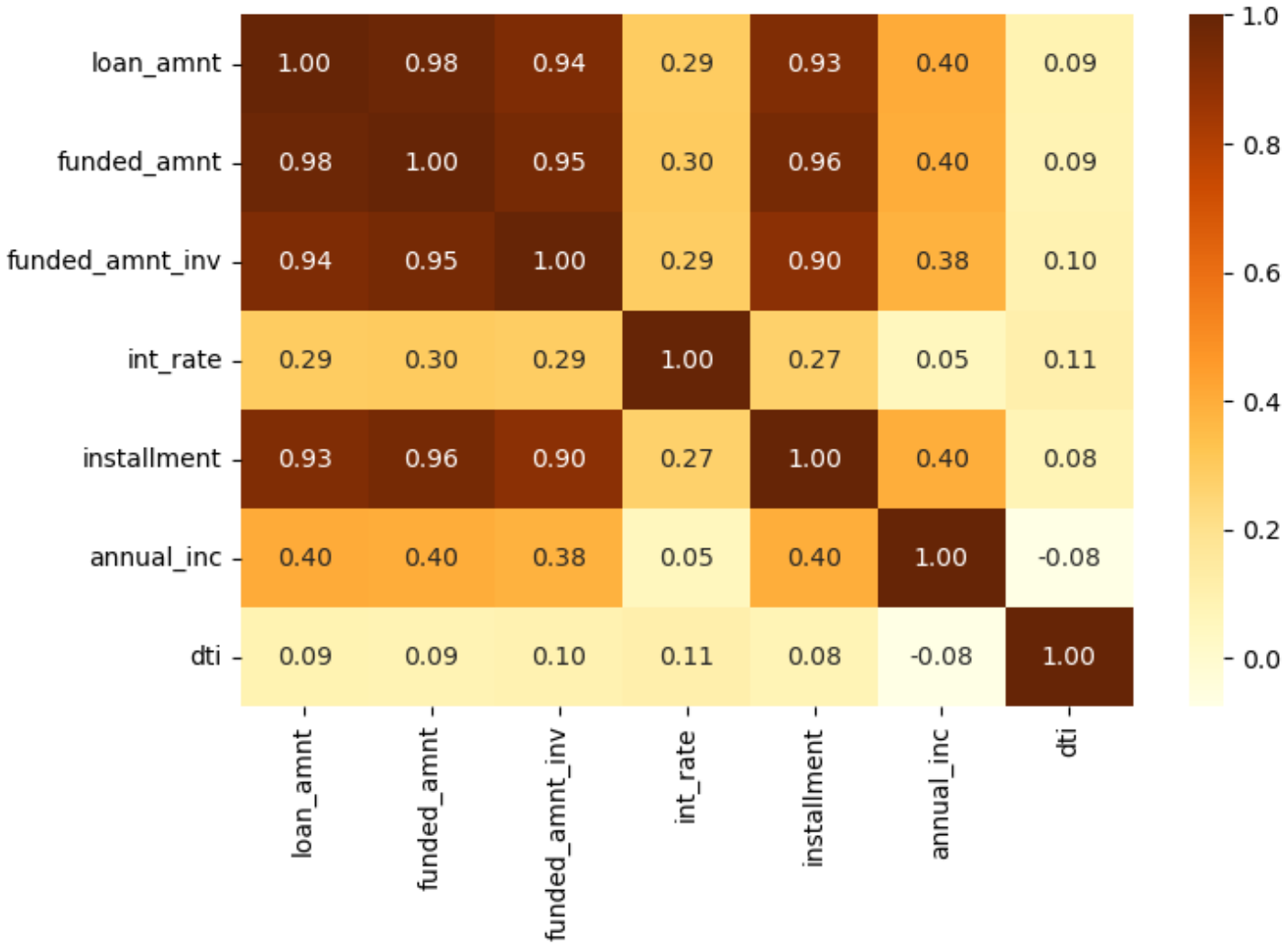
- Borrower with income less than 50K and applying for loan with interest rate 10% to 15% are more likely to default if not verified.



Correlation Matrix

Observation:

- Annual Income is inversely correlated with DTI (Debt to Income Ratio)
- Loan Amount, Funded Amount & Funded Amount by Investor are highly correlated.



Consolidated Observations & Conclusion

Statistically, borrowers with 36 months tenure loans are more likely to default than 60 months

Our assumption is that A is highest grade and G is the lowest. Borrowers with lower grade are more likely to default.

Loans with purpose of debt consolidation, credit card , small business and others are more likely to default.

Low Income borrowers (<50K) are more likely to default.

Every year with each passing month number of borrowers increases.

Indicators for defaulting

Borrowers with 36 **months Tenure**

Borrowers with **annual income less than 50K**

Borrowers with **interest rate more than 10%**

Borrowers with **purpose of debt consolidation, credit card, small business and others.**

Borrowers with **10+ years of experience**

Borrowers with **<3 years of experience**

Borrowers **with lower grade**

Recommendations

- Lending Club must not approve loan with **purpose of others.**
- Borrowers with higher interest rate calculated must be verified completely.
- Lending club should reduce the interest rate for borrowers with income less than 50K.
- Lending Club should avoid borrowers with higher DTI (Debt to Income Ratio).
- Minimize number of loans approval towards requests with debt consolidation purpose.
- Lending Club can promote 60 Months loan over 36 Month to enable borrowers with low installment amounts.