



Lending Club Case Study

Submitted by:

Siddharth Sharma & Mani Kandan



Case Study Agenda

- Business Understanding & Objectives
- Data Sourcing
- Data Cleaning
- Statistics & Visualization
- Conclusion

Business Understanding & Objectives

- To Perform EDA for a consumer finance company which specializes in lending various types of loans.
- 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
- Our objective is to perform the detailed Analysis on the available data and try to understand the Trend, characteristics & identify the defaulters pattern which could help to decide if lending loans would be the right decision or not.

Data Sourcing

- The datasets are provided in the assignment were downloaded as below:
 - **Loan Data Set** : Dataset used for Case study having the data for loans issues between 2007 -2011
 - **Data Dictionary** : Refer to the variables meaning used for Analysis

Data Sourcing

```
[ ] #Downloading the data from the below file  
  
dataset=pd.read_csv('/content/drive/MyDrive/Colab Notebooks/Exploratory_Data_Analysis/LendingCaseStudy/loan.csv')
```

Data Cleaning

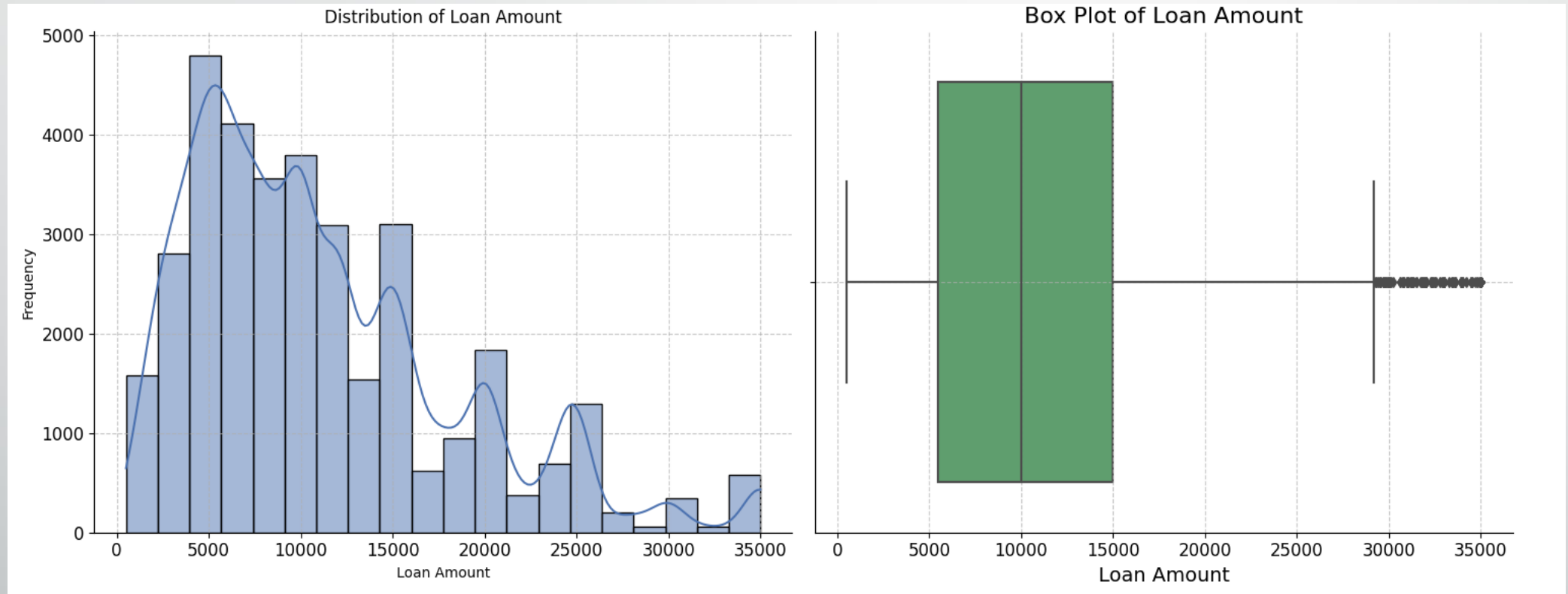
- Removed the Columns which had null values in the all records.
- Removed Columns which has more than 25% data is null values
- Removed Columns which were not useful for further analysis.
- Added Derived Variables as applicable

Univariate Analysis

Some of the key attributes identified for the analysis are listed as below:

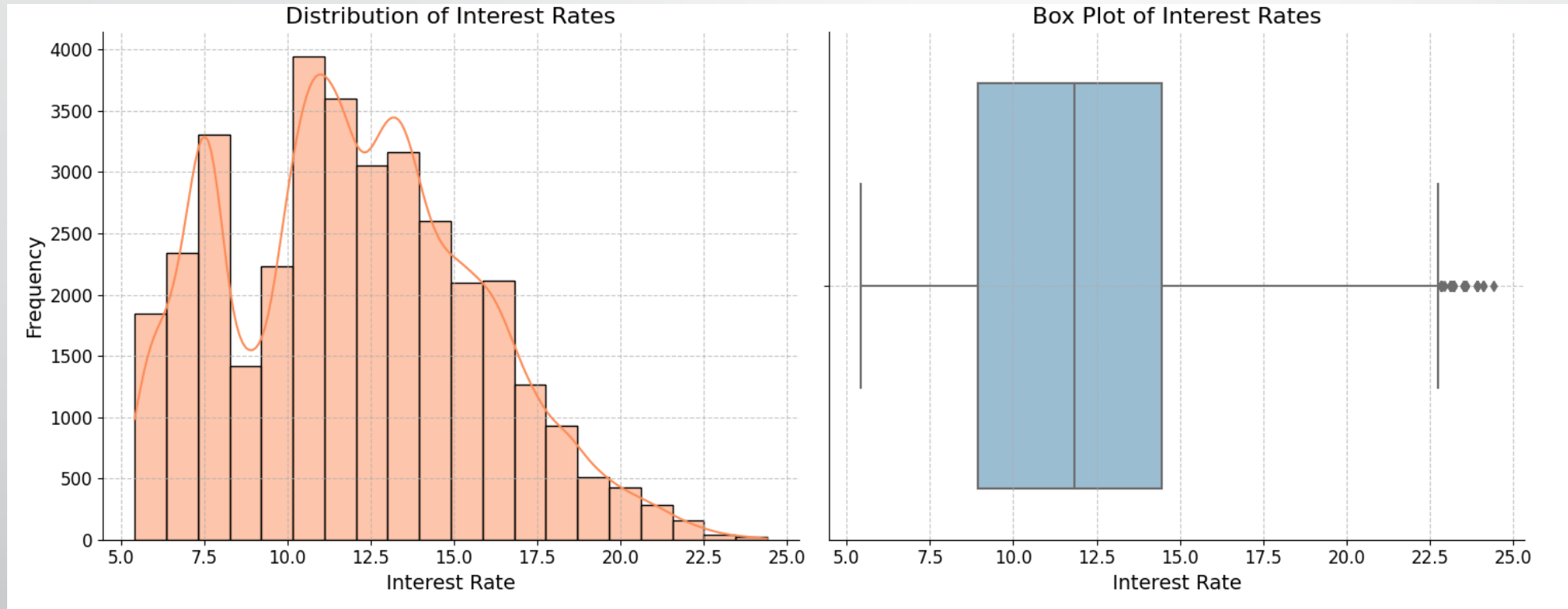
1. Loan Amount
2. Interest Rate
3. Annual income
4. Loan Purpose
5. Loan Status
6. Home Ownership
7. Grades

Loan Amount



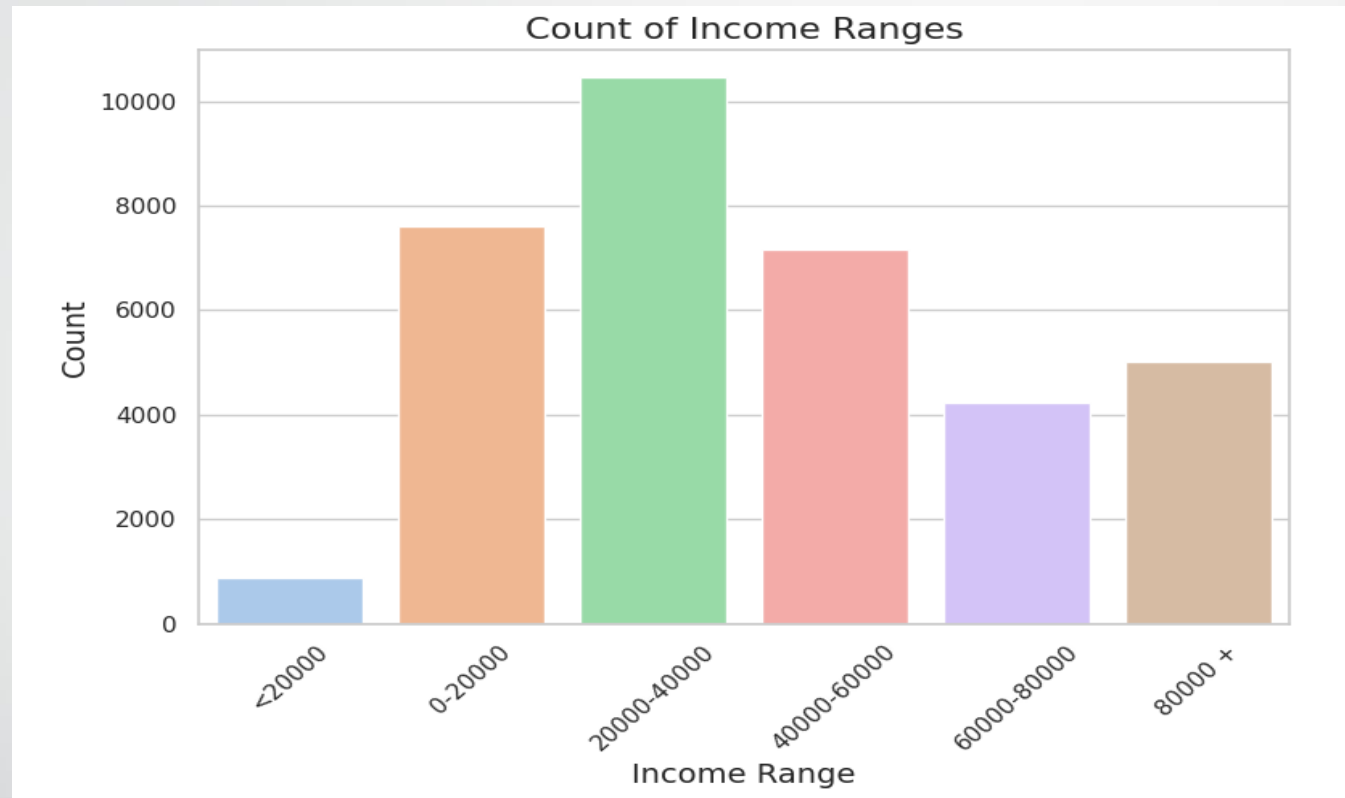
- Borrowers applied loan between 500 – 35000
- As depicted in the graphs, a substantial majority of individuals fall within the range of 'Loan Amount' between 5,500 and 15,000.

Interest Rate



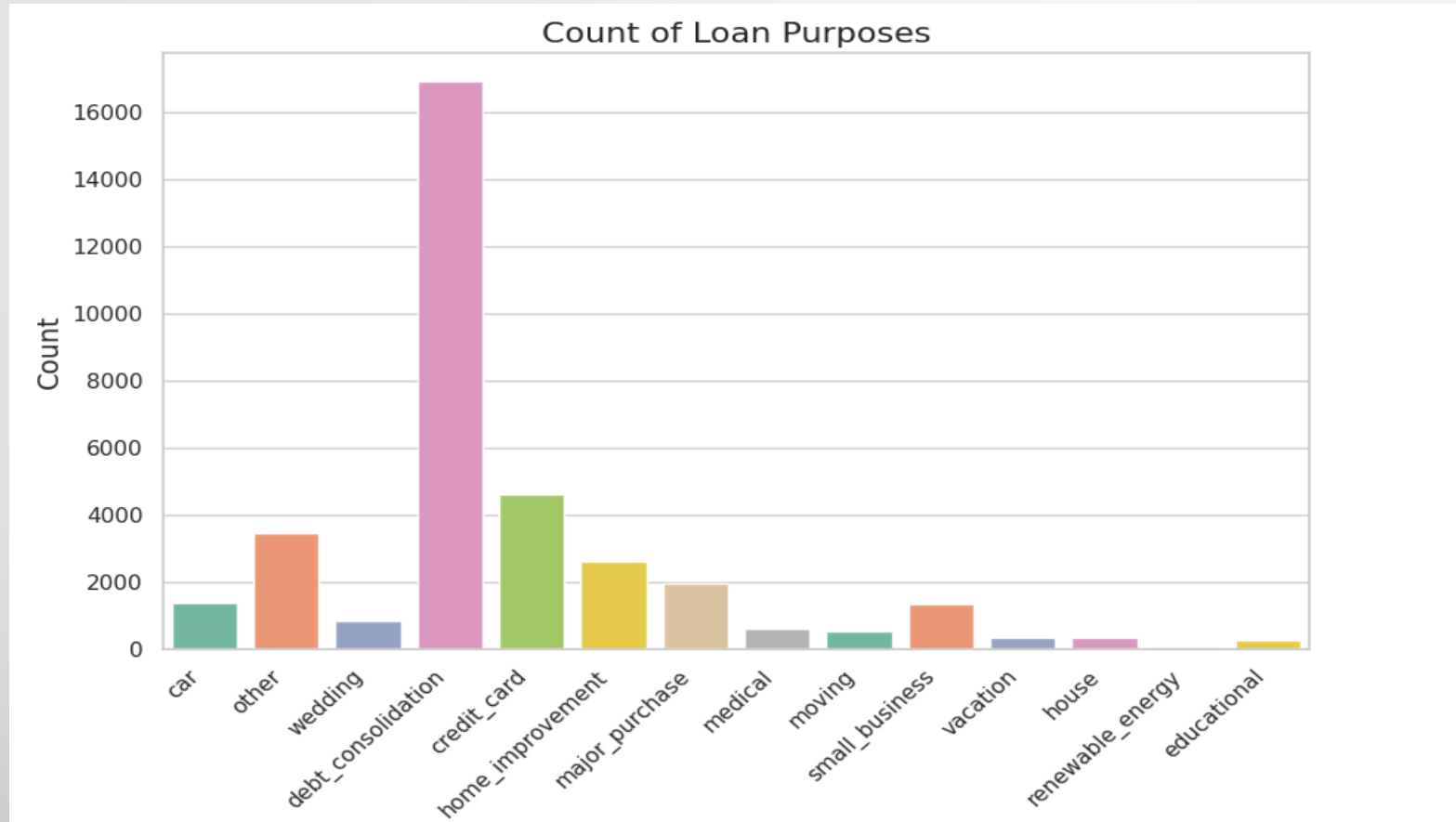
- Interest Rates vary between 5.5% to 24 %
- As depicted in the Box plot graphs, the 'Interest rates' commonly ranges from 10% to 14%; nevertheless, there are instances where the interest rates for loans can reach as high as 24%.

Annual Income



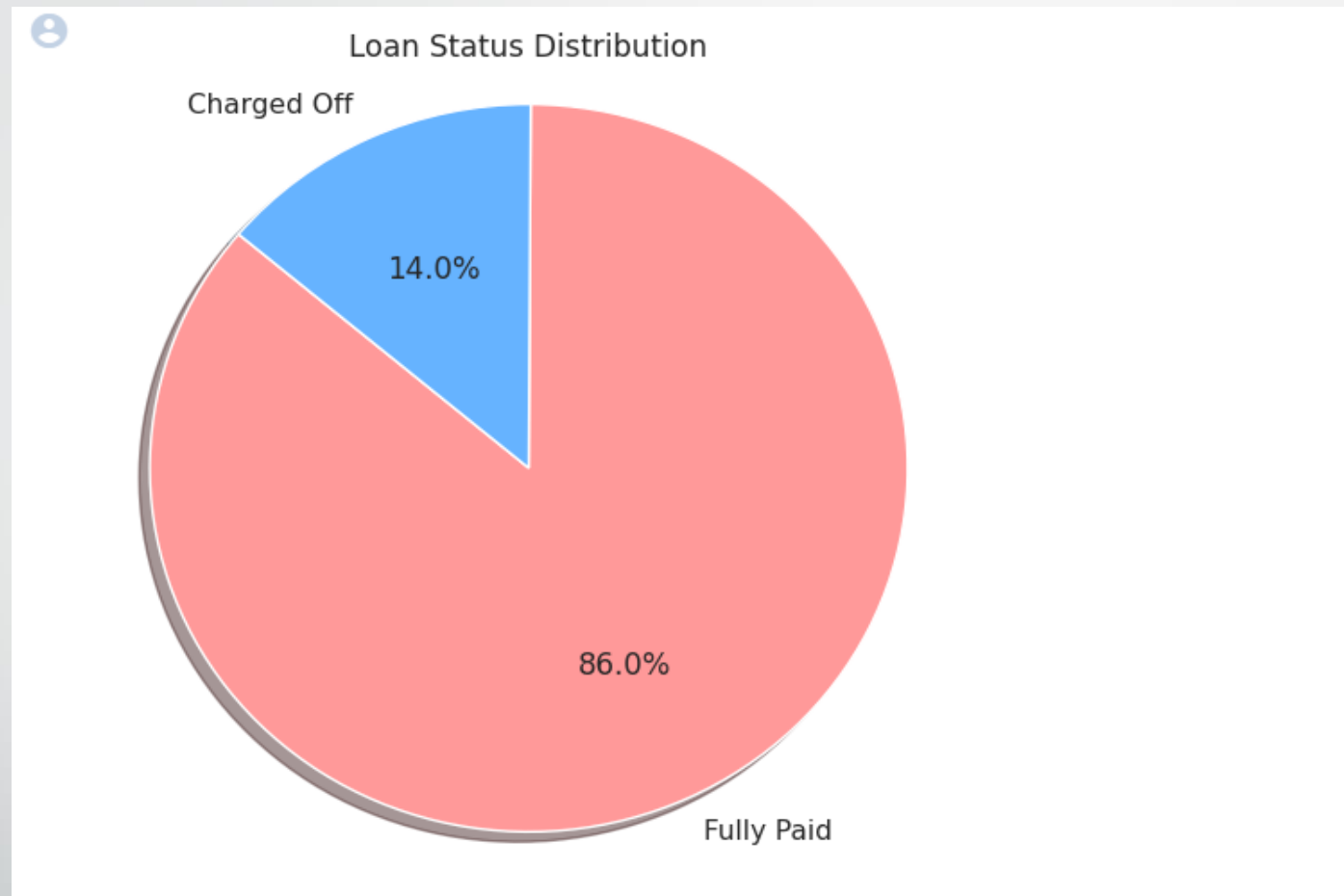
- The majority of loan applicants have reported an income falling within the range of 20,000 to 40,000.

Loan Purpose



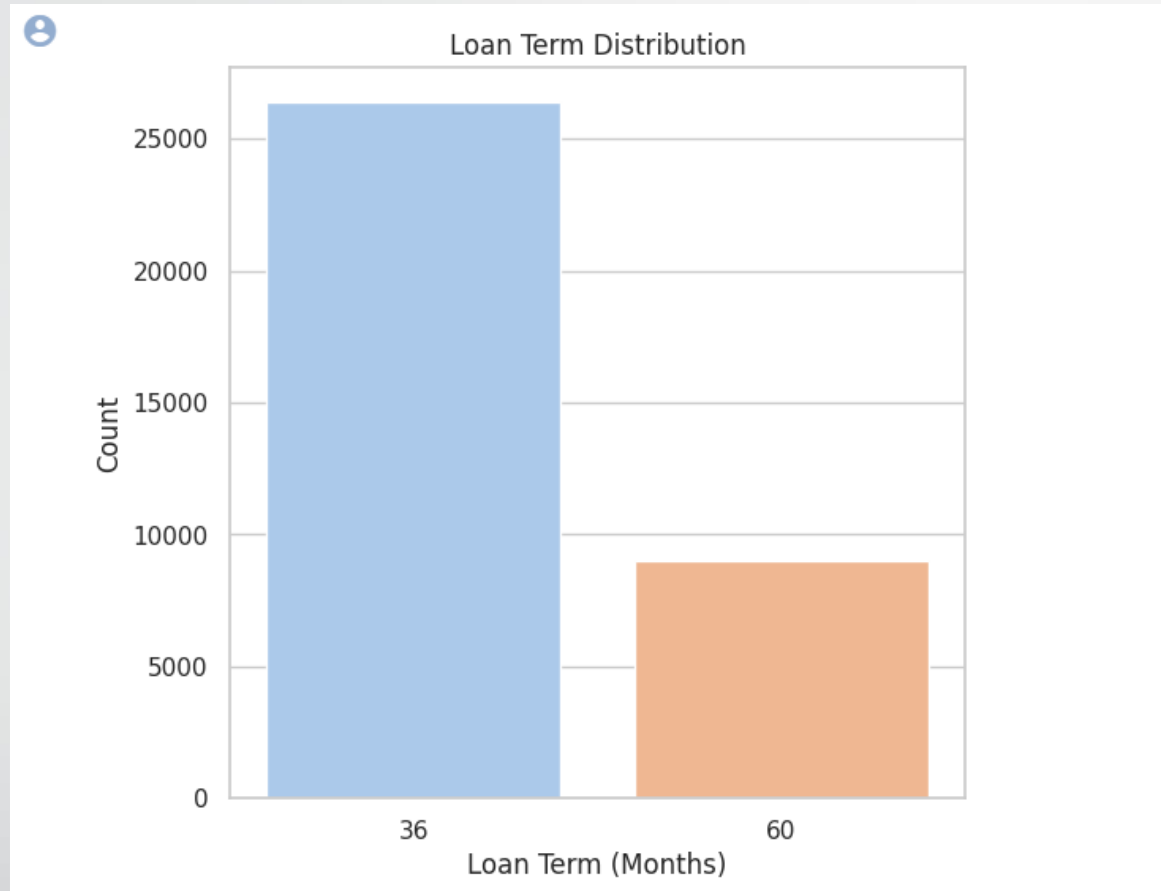
- The primary purpose for obtaining loans is predominantly debt consolidation, followed closely by credit card debt repayment.

Loan Status



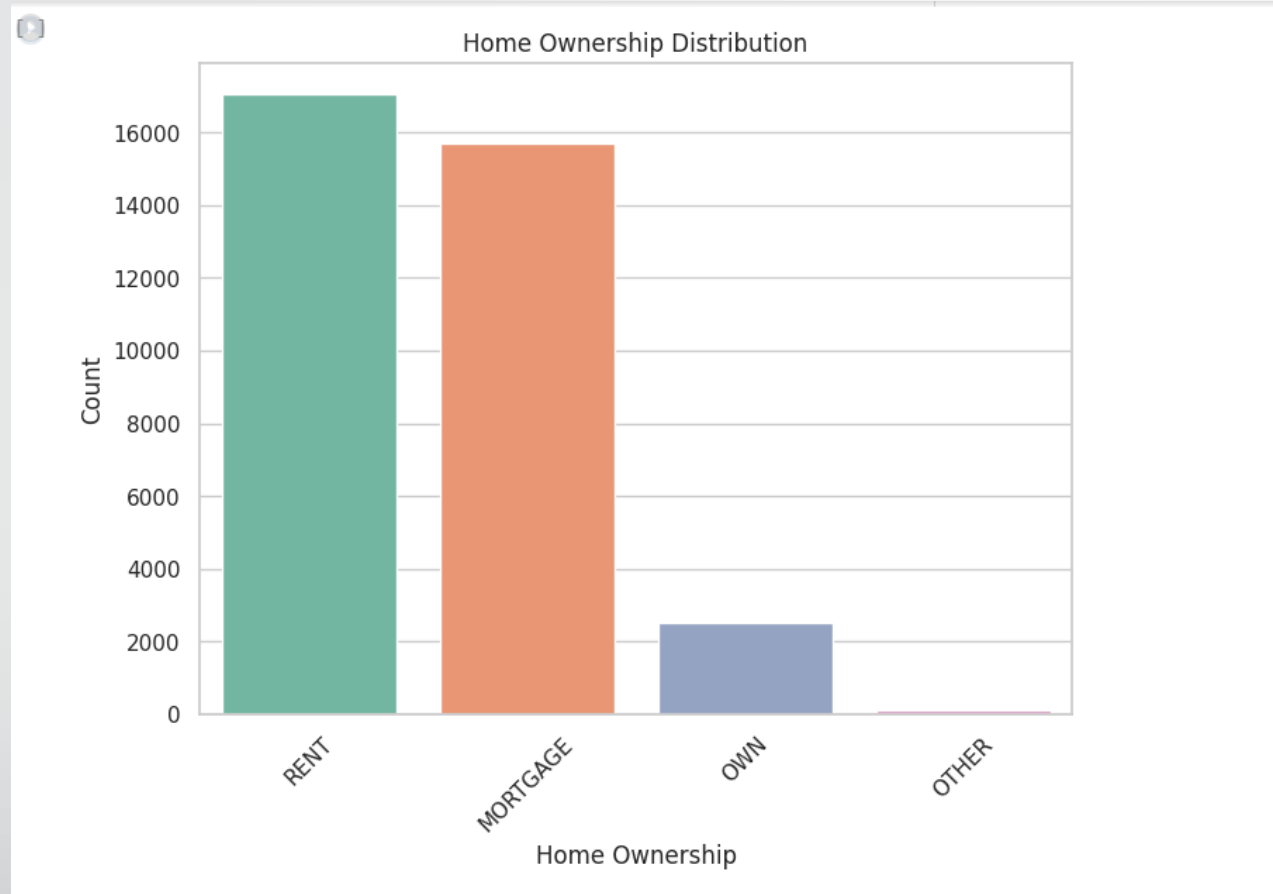
- It can be inferred that, out of a total of 50 loans, approximately 7 are projected to be charged off.

Loan Term



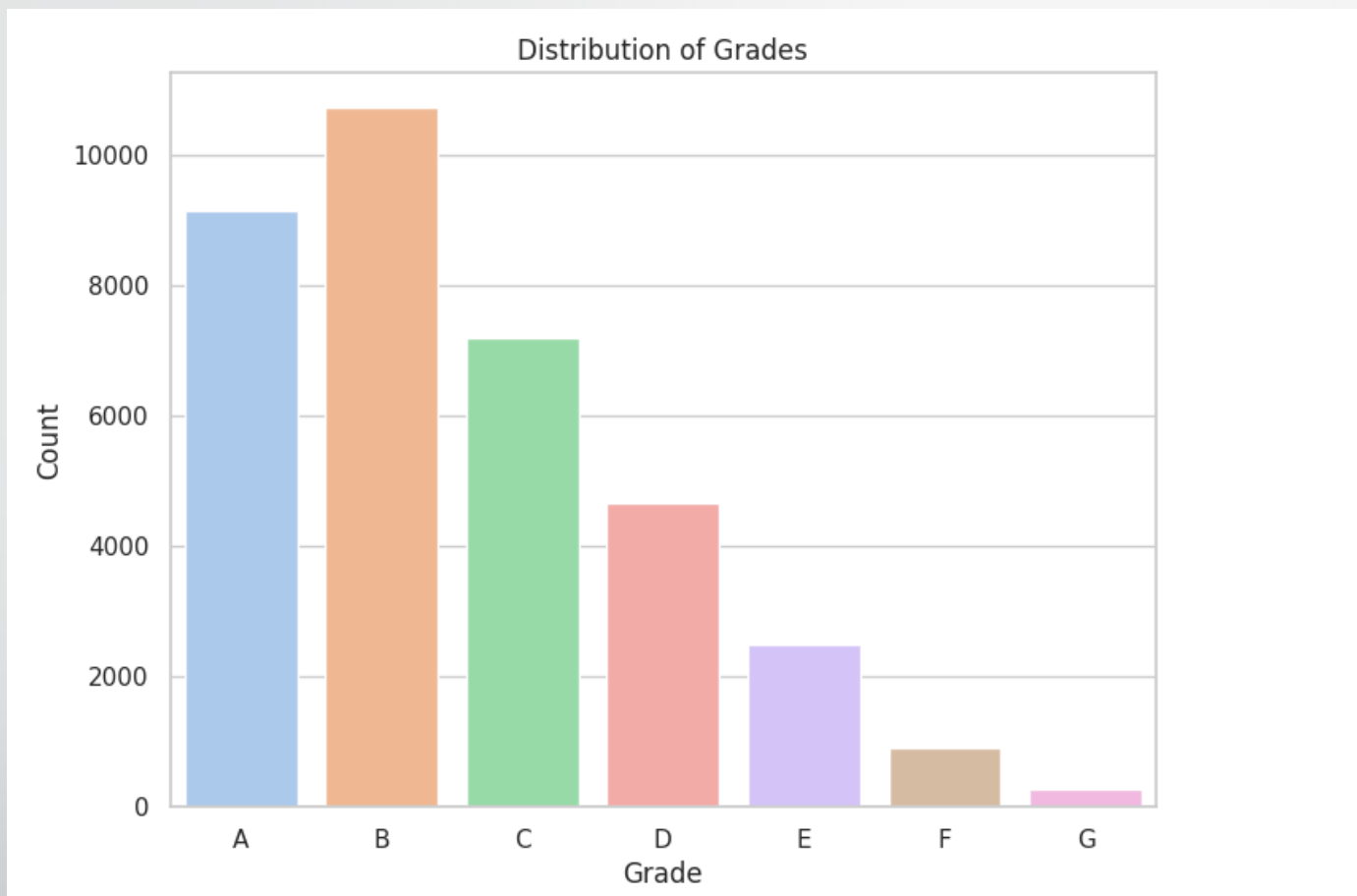
- The data clearly indicates that a higher proportion of borrowers exhibit a preference for shorter-term loans, opting for 30 months over 60 months

Home Ownership



- Individuals who are either renting or have a mortgage tend to take out loans more frequently compared to those who own their own houses.

Grades



- The B Grade category has demonstrated a higher propensity for loan acquisition compared to the other grade categories.

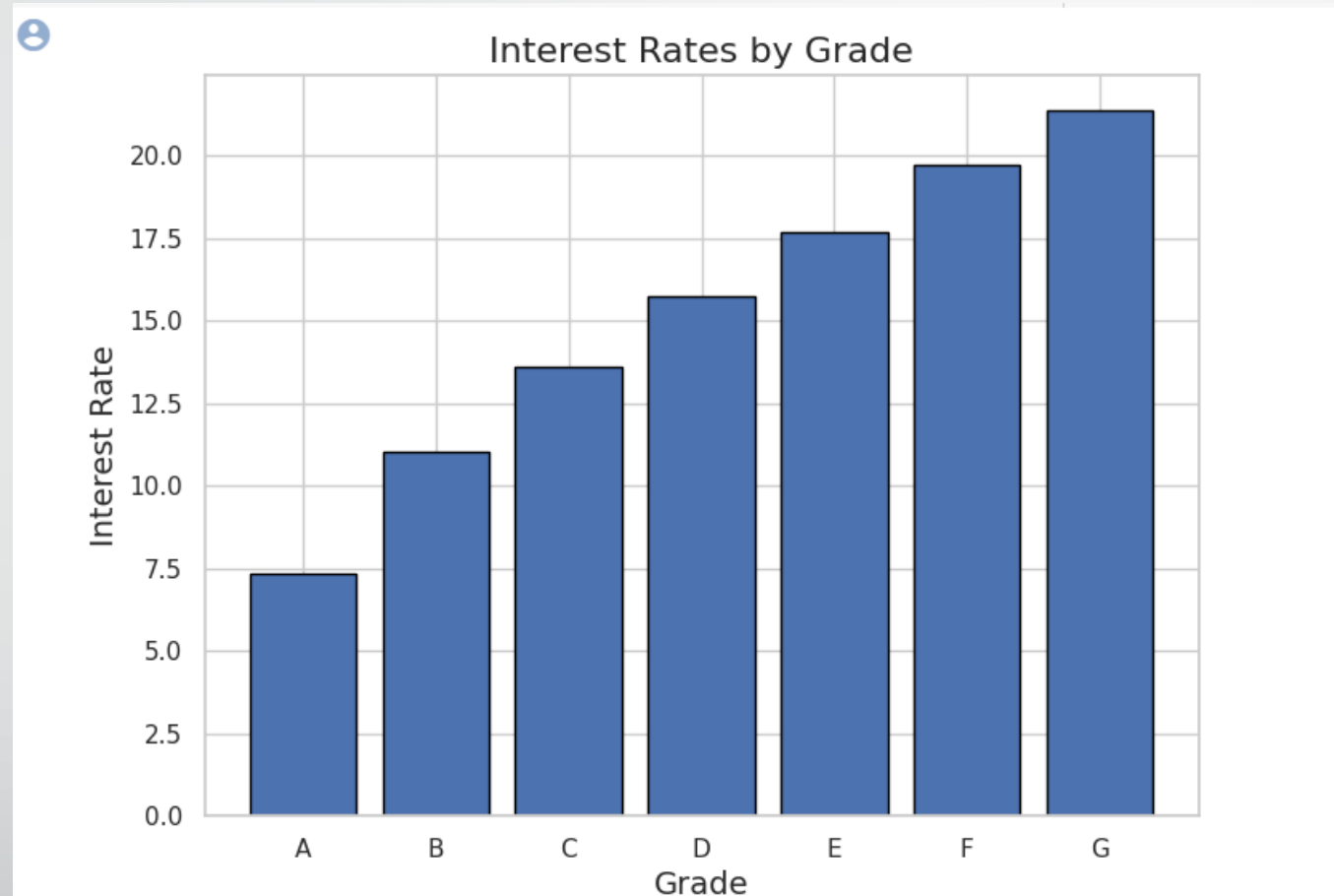
Conclusion:

Univariate Analysis

Some of the key attributes identified for the analysis are listed as below:

- The applicants primarily seek loans in the range of 5,500 to 15,000, with the maximum loan amount requested being 35,000.
- While the majority of applicants are charged interest rates ranging from 10% to 14%, there are instances where interest rates can go as high as 24%.
- Among loan purposes, debt consolidation ranks highest, with over 17,000 applicants selecting it, followed by 'Credit card' which has slightly less than 5,000 applications.
- Regarding loan status, approximately 85.7% of borrowers have successfully paid off their loans, while 14.3% have experienced a 'Charged off' status.
- Loans with a term of "36" months are preferred by over 25,000 borrowers, whereas those with a term of "60" months are requested by fewer than 10,000 applicants.
- The majority of borrowers indicate 'Rent' as their primary form of home ownership, followed closely by 'Mortgage'.

Segmented Analysis



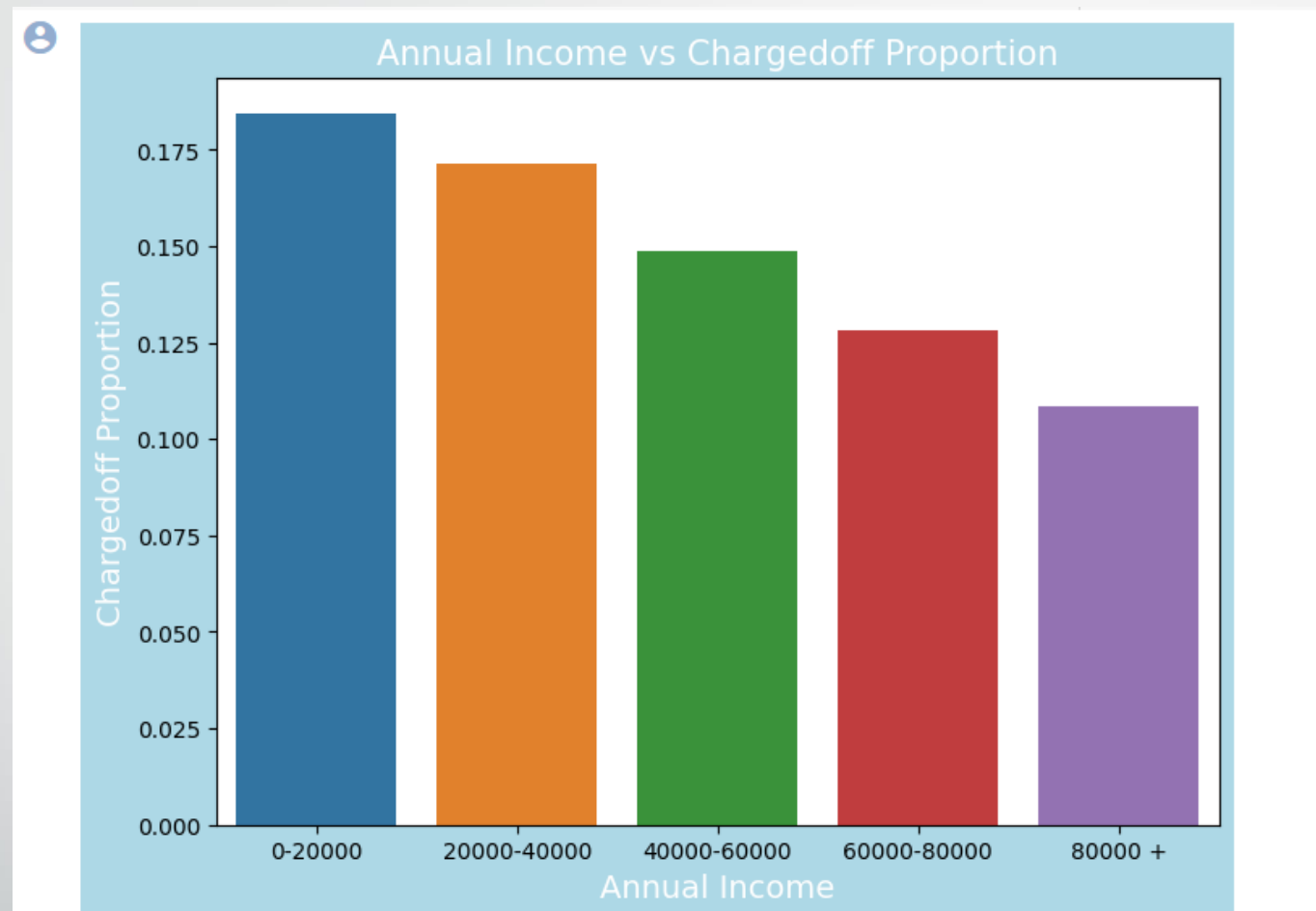
- It is evident that the mean 'Interest Rate' is highest for the 'G' grade, reaching nearly 24%. Conversely, for 'A' grade borrowers, the maximum 'Interest Rate' is below 7.5%.

Bivariate Analysis

Some of the key attributes identified for the analysis are listed as below:

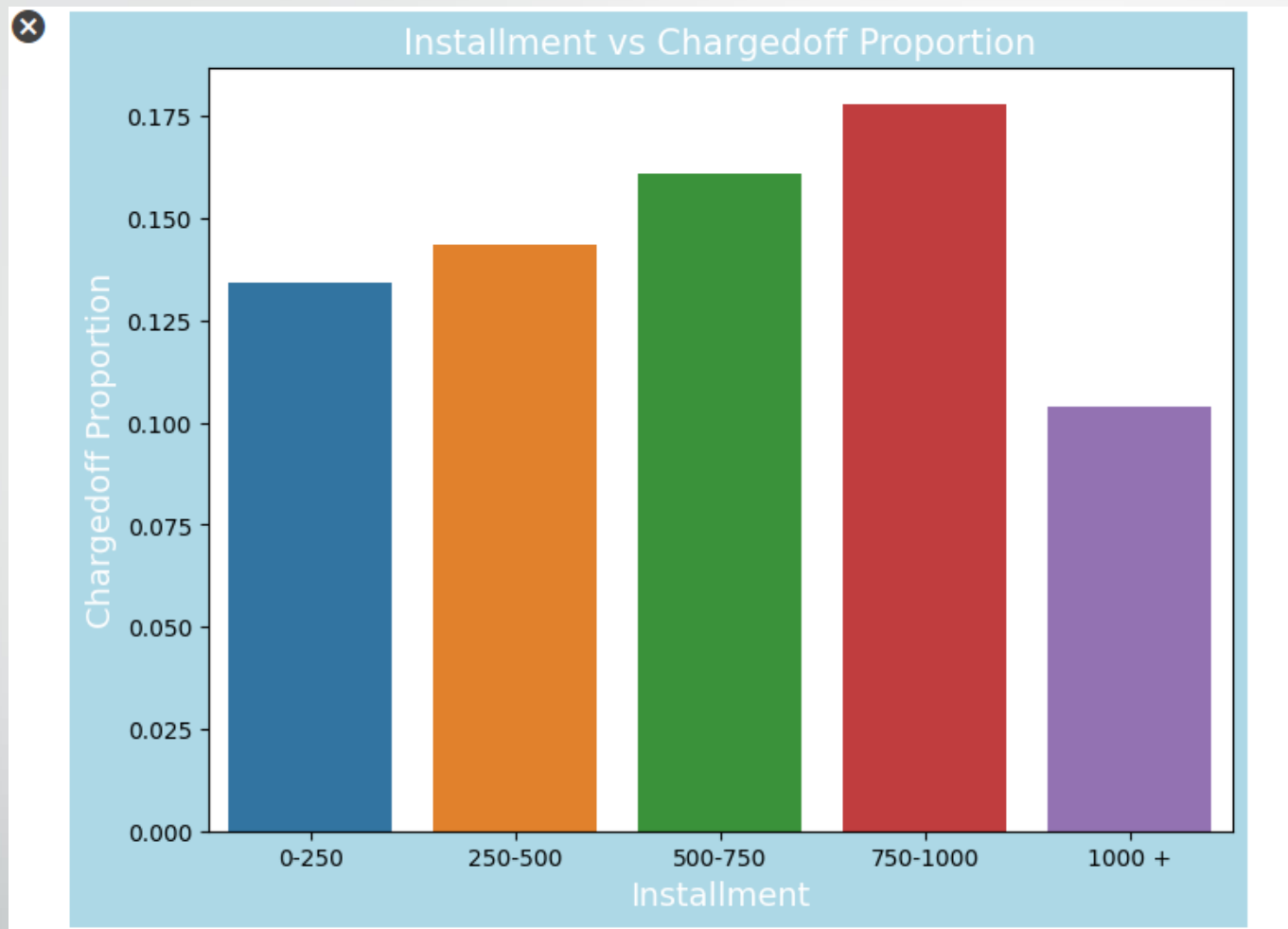
1. Annual Income vs Chargedoff Proportion
2. Installment vs Chargedoff Proportion
3. Purpose of loan vs Chargedoff Proportion
4. Grades vs Chargedoff Proportion
5. Interest Rates vs Chargedoff Proportion
6. Address State vs Chargedoff Proportion
7. Verification vs Chargedoff Proportion
8. Term of Loan vs Interest Rate
9. Loan Amount vs Interest Rate

Annual Income vs Chargedoff Proportion



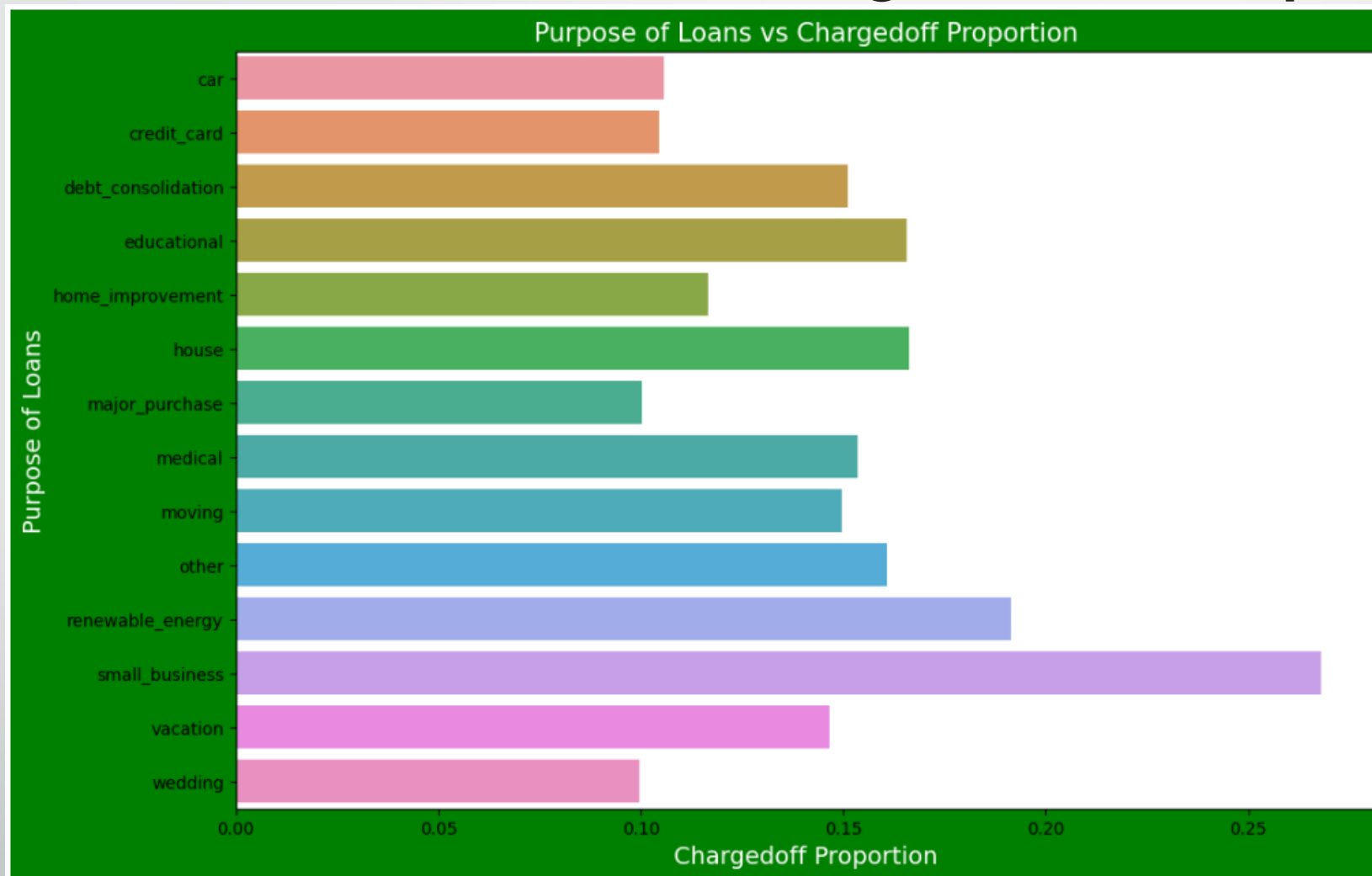
- As annual income increases, the likelihood of experiencing a charge-off decreases.

Installment vs Chargedoff Proportion



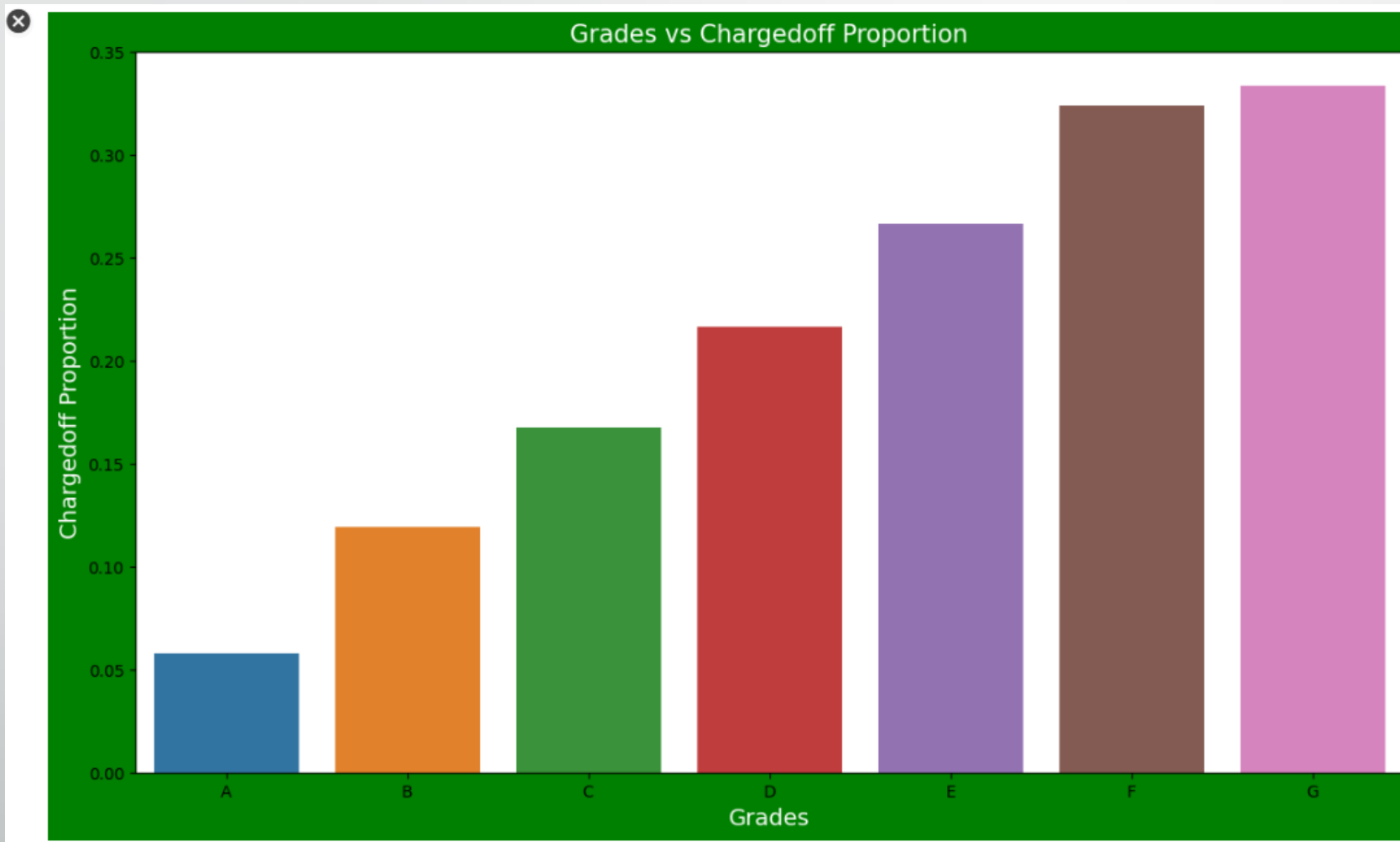
- Installment amounts exceeding 1000 are associated with a lower likelihood of experiencing a charge-off.

Purpose of Loans vs Chargedoff Proportion



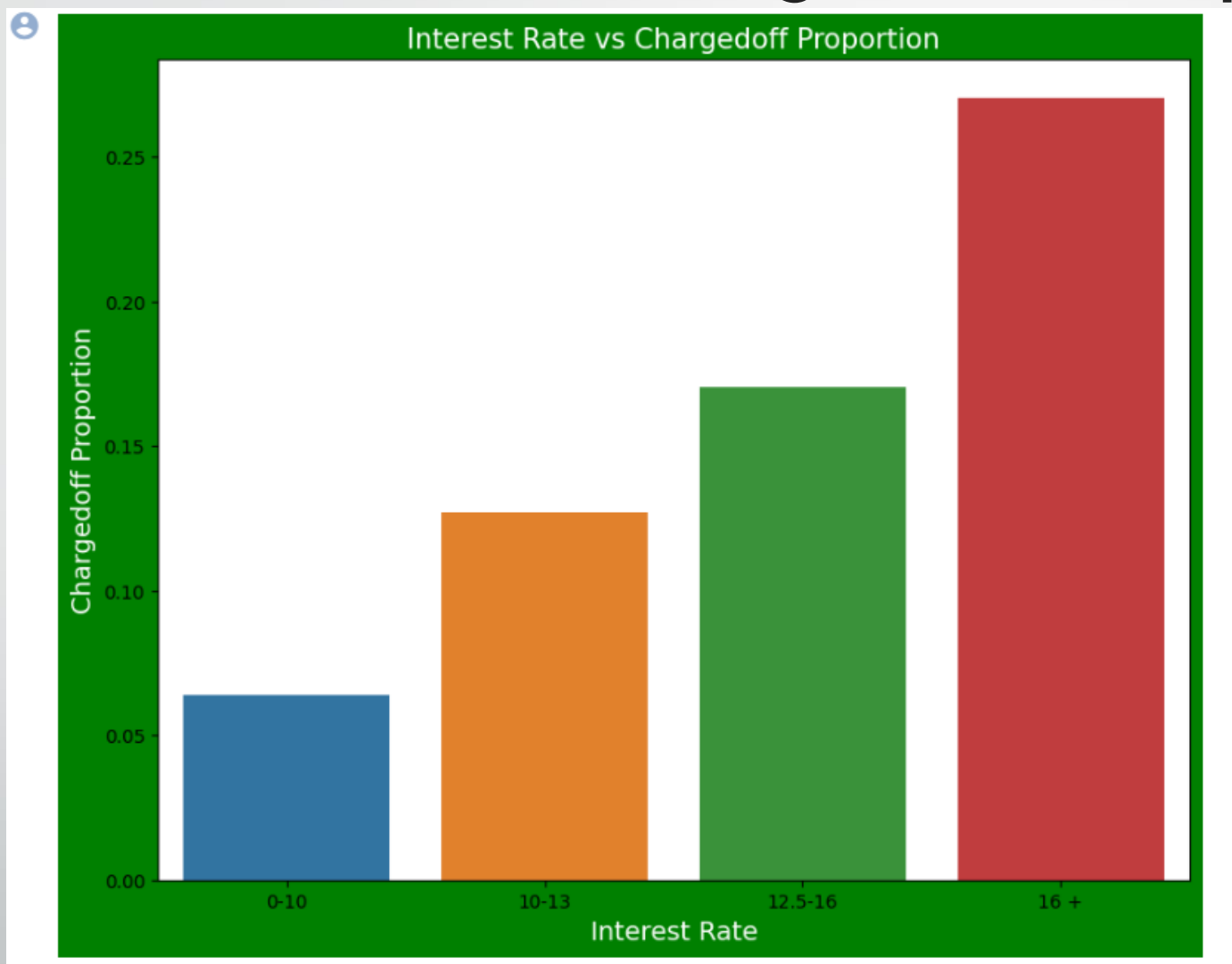
- Loans designated for the purpose of "Small Business" are more likely to result in a charge-off.

Grades vs Chargedoff Proportion



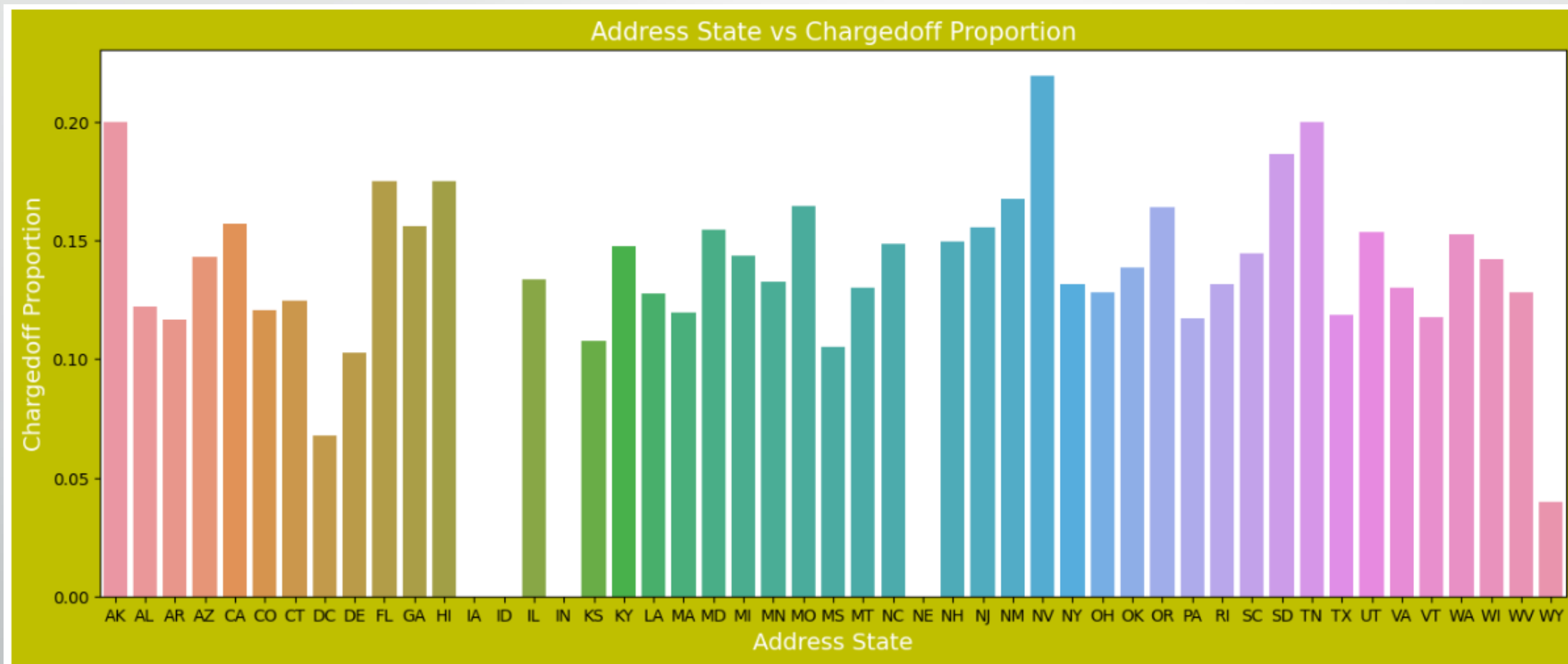
- The likelihood of a charge-off is significantly elevated for borrowers in G and F grades, while it is notably lower for those in the A grade.

Interest Rate vs Chargedoff Proportion



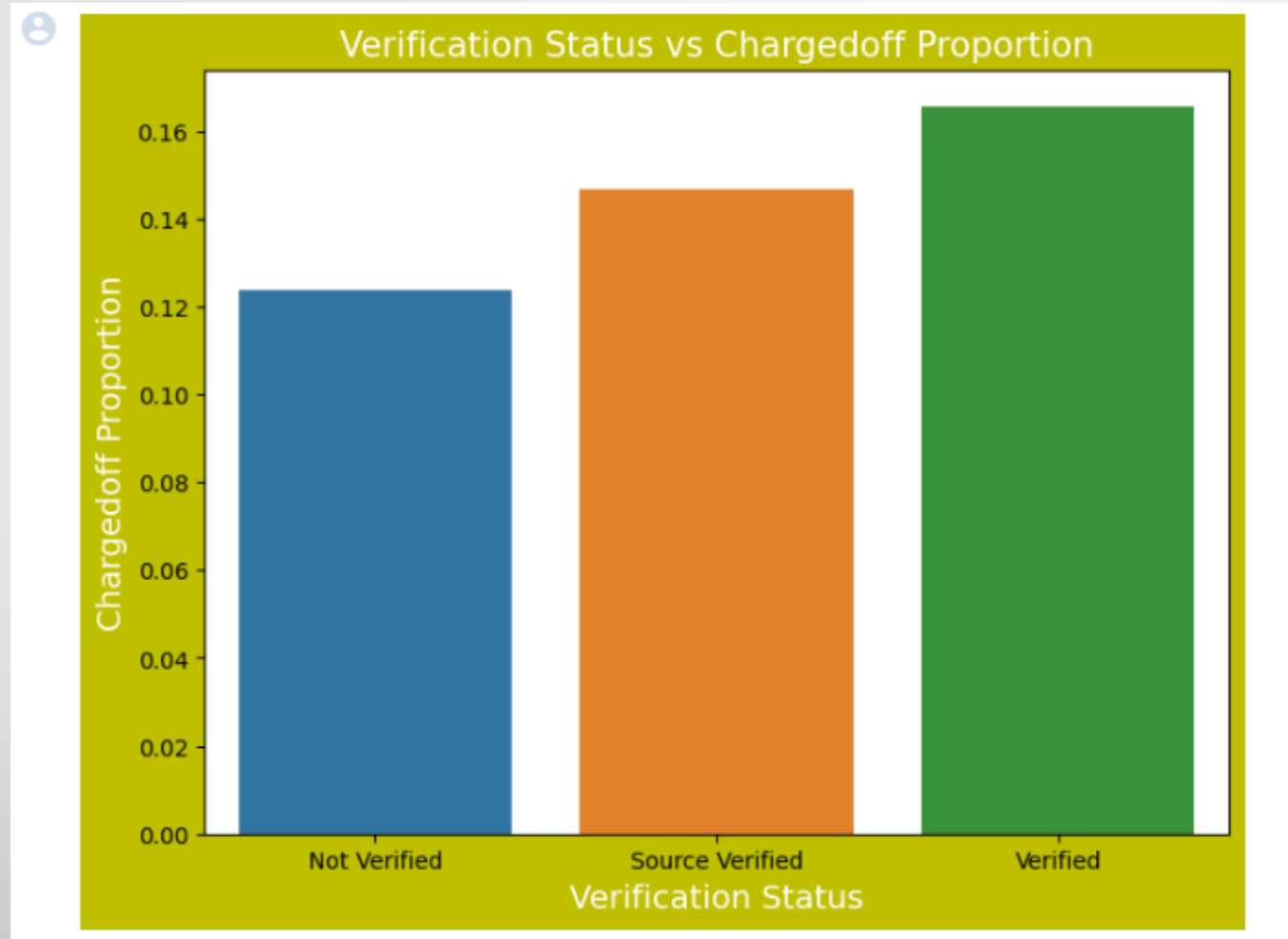
- As anticipated, an increase in the interest rate is strongly correlated with a higher likelihood of experiencing a charge-off.

Address State vs Chargedoff Proportion



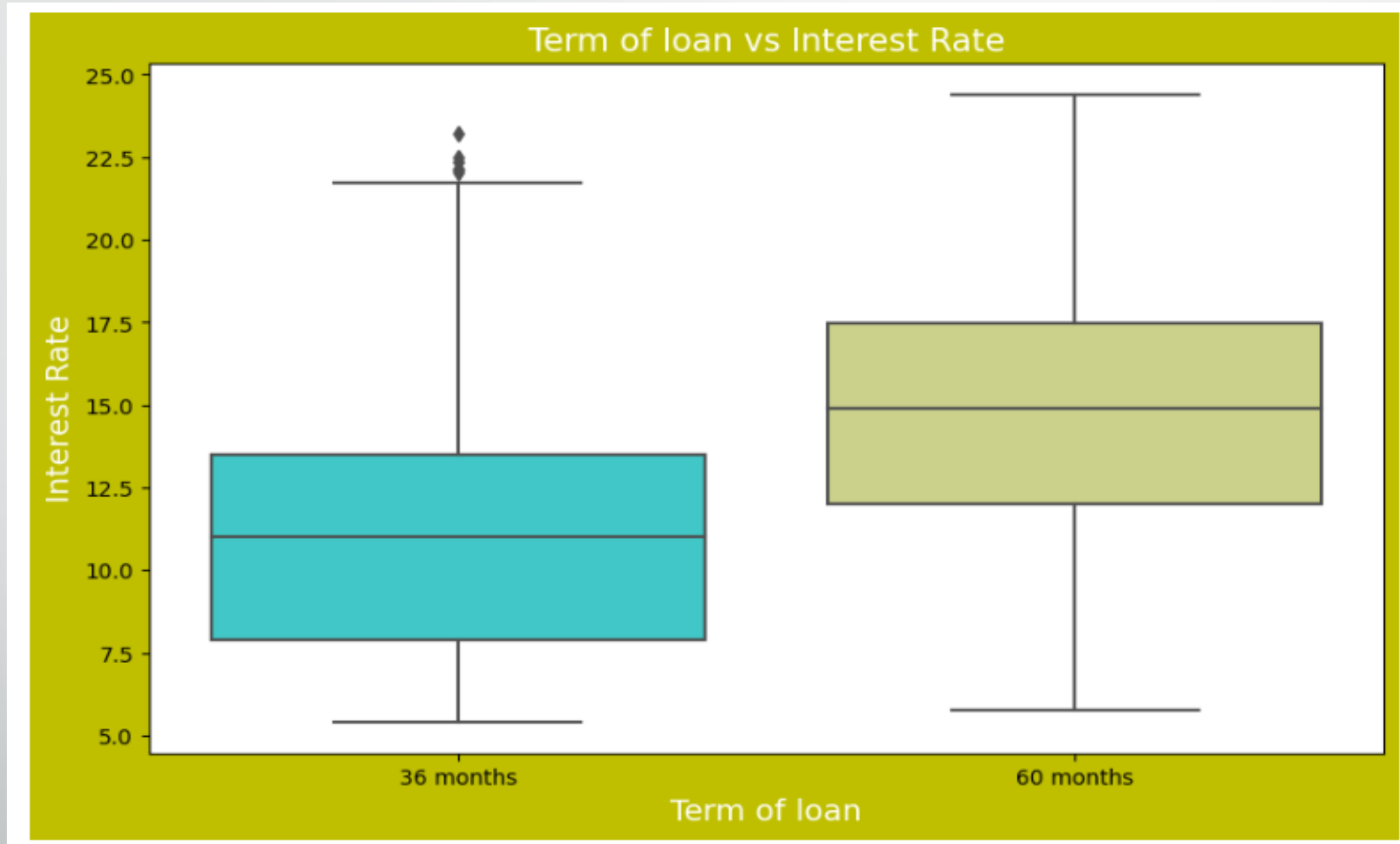
- The states of NV, TN, and SD exhibit a notable number of charged-off cases in relation to the overall volume of applications.

Verification Status vs Chargedoff Proportion



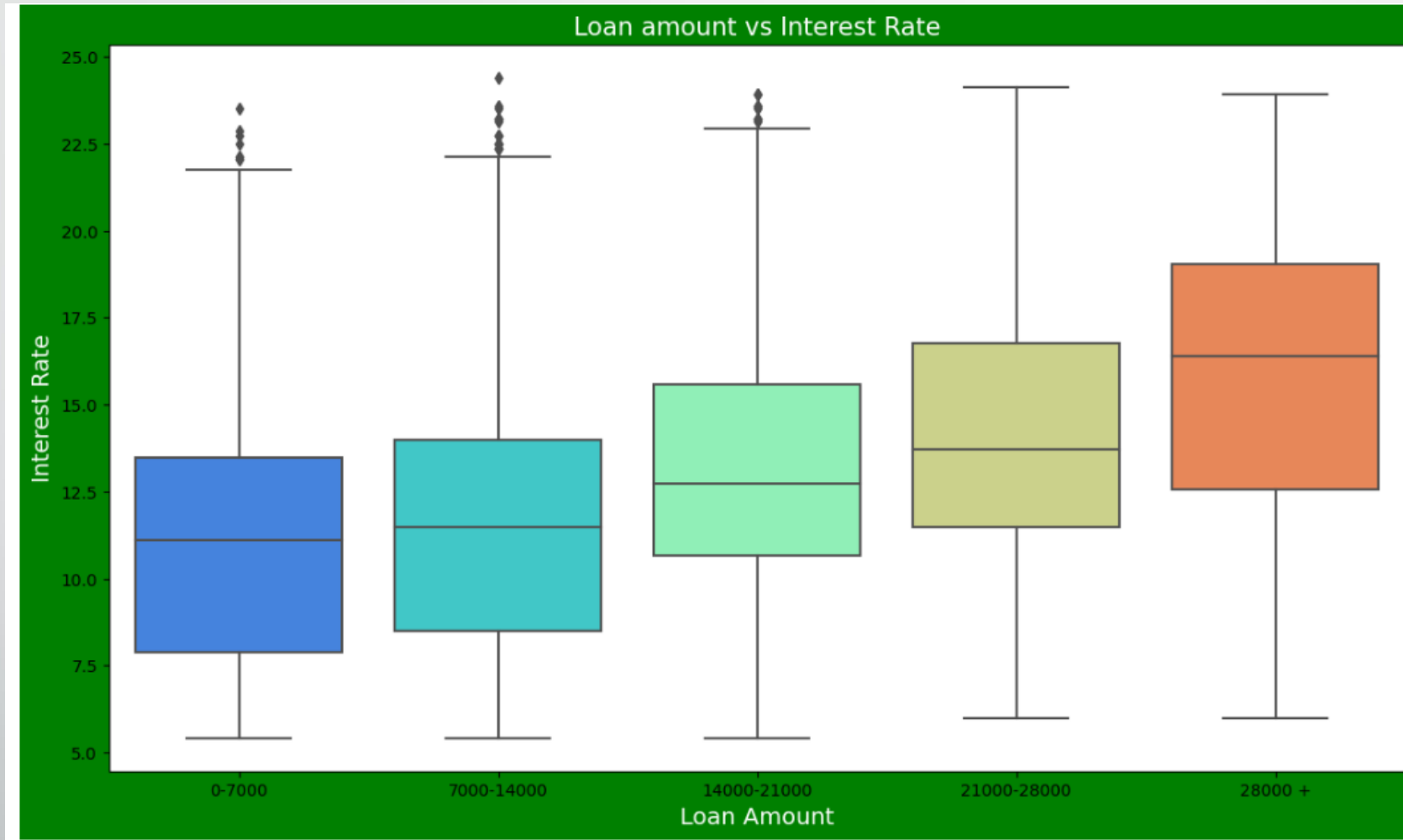
- The available data doesn't provide a significant enough distinction to draw a conclusive decision.

Term vs Interest Rate



- The interest rates for loans with a 60-month duration are considerably higher compared to those with shorter terms.

Loan Amount vs Interest Rate



- The interest rate demonstrates a positive coefficient with the loan amount. This means that as the loan amount increases, the interest rate also tends to rise.

Conclusion: Bivariate Analysis

Some of the key attributes identified for the analysis are listed as below:

- Borrowers in grades G and F are at higher risk of charge-offs, while those in grade A have a lower likelihood of charge-offs.
- Loans exceeding 1000 demonstrate a lower likelihood of charge-offs.
- 'Small Business' loans are associated with a higher probability of charge-offs.
- There is a positive correlation between interest rates and loan amounts, indicating that as loan amounts increase, interest rates tend to rise.
- States like NV, TN, and SD exhibit a notable number of charged-offs relative to the total number of applications.
- For loans with a 60-month duration, interest rates tend to be significantly higher compared to shorter-term loans.



Thank You