



Lending Club Case Study

Upendra Kumar

A large orange circle is positioned on the left side of the slide, partially cut off by the edge. The word "Content" is written in white text inside this circle.

Content

Business Objectives

Approach

Data Cleansing, Manipulation and Feature Engineering

Univariate Analysis

Segmented Univariate Analysis

Bivariate Analysis

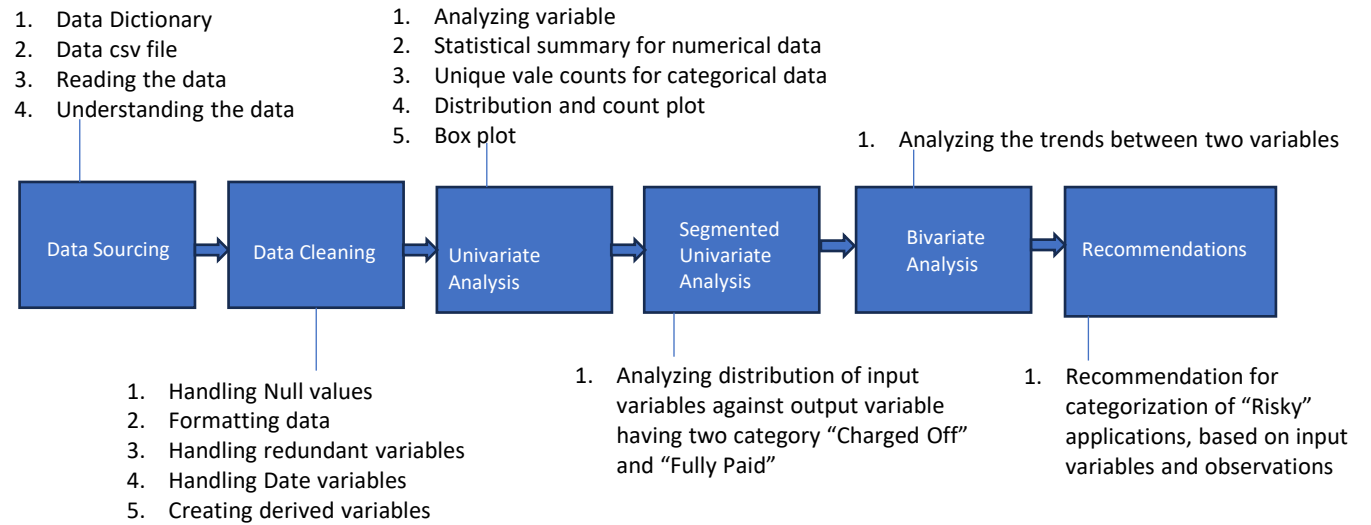
Recommendations

Business Objectives

The Consumer Finance company is the largest online loan marketplace, facilitating personal loans, business loans, and financing of medical procedures. Borrowers can easily access lower interest rate loans through a fast-online interface.

- Lending loans to 'risky' applicants is the largest source of financial loss (called credit loss).
- If one can identify these risky loan applicants, then such loans can be reduced thereby cutting down the amount of credit loss.
- The company wants to understand the **driving factors (or driver variables)** behind loan default, i.e., the variables which are strong indicators of default.
- Identification of such applicants' using EDA is the aim of this case study.

Approach



Data Cleansing, Manipulation and Feature Engineering

- Dropping the columns having null values greater than 50%.
- Drop “desc” column since same information is present in purpose column.
- Dropped records where null value is present for few columns.
- After analyzing the data, dropped the features which were not contributing to analysis.
- Change data type format for date columns
- Derived month and year features from datetime columns.
- Drop existing date columns after deriving month and year from those columns.
- Changing the format of few features as per requirement
- Removed all records having loan status ‘Current’ since the applicant is in the process of paying the instalments.

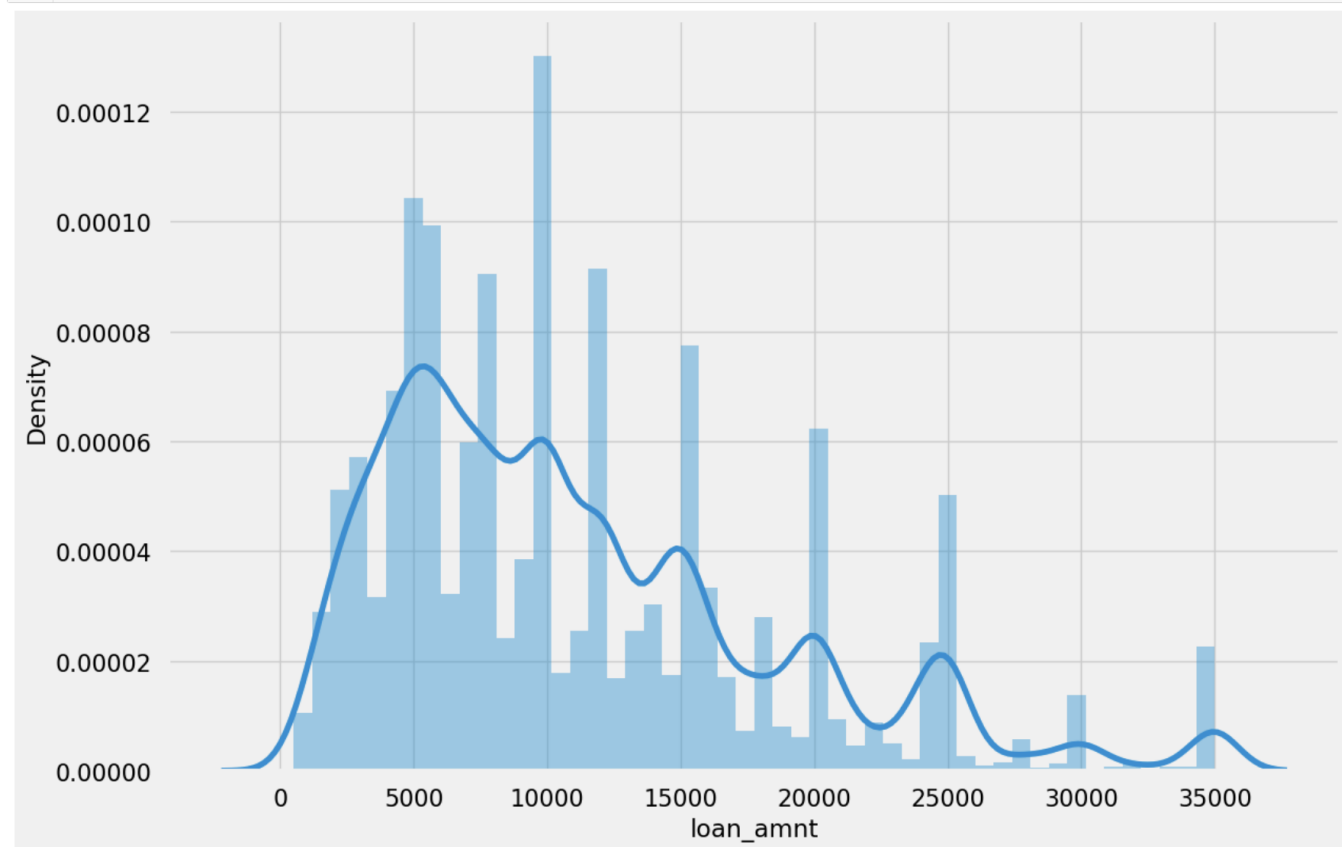
The slide features a solid orange background. A large white circle is centered on the page. The text "Univariate Analysis" is written in a black, sans-serif font across the middle of the white circle. On the left side of the white circle, there is a dashed yellow arc. On the bottom right edge of the white circle, there is a small solid blue circle.

Univariate Analysis

Loan Amount

Observation:

- Loan amount distribution is slightly right-skewed with mean greater than the median.
- Most of the loans granted are below 15000.
- A spike is observed around 5000.



Insight:

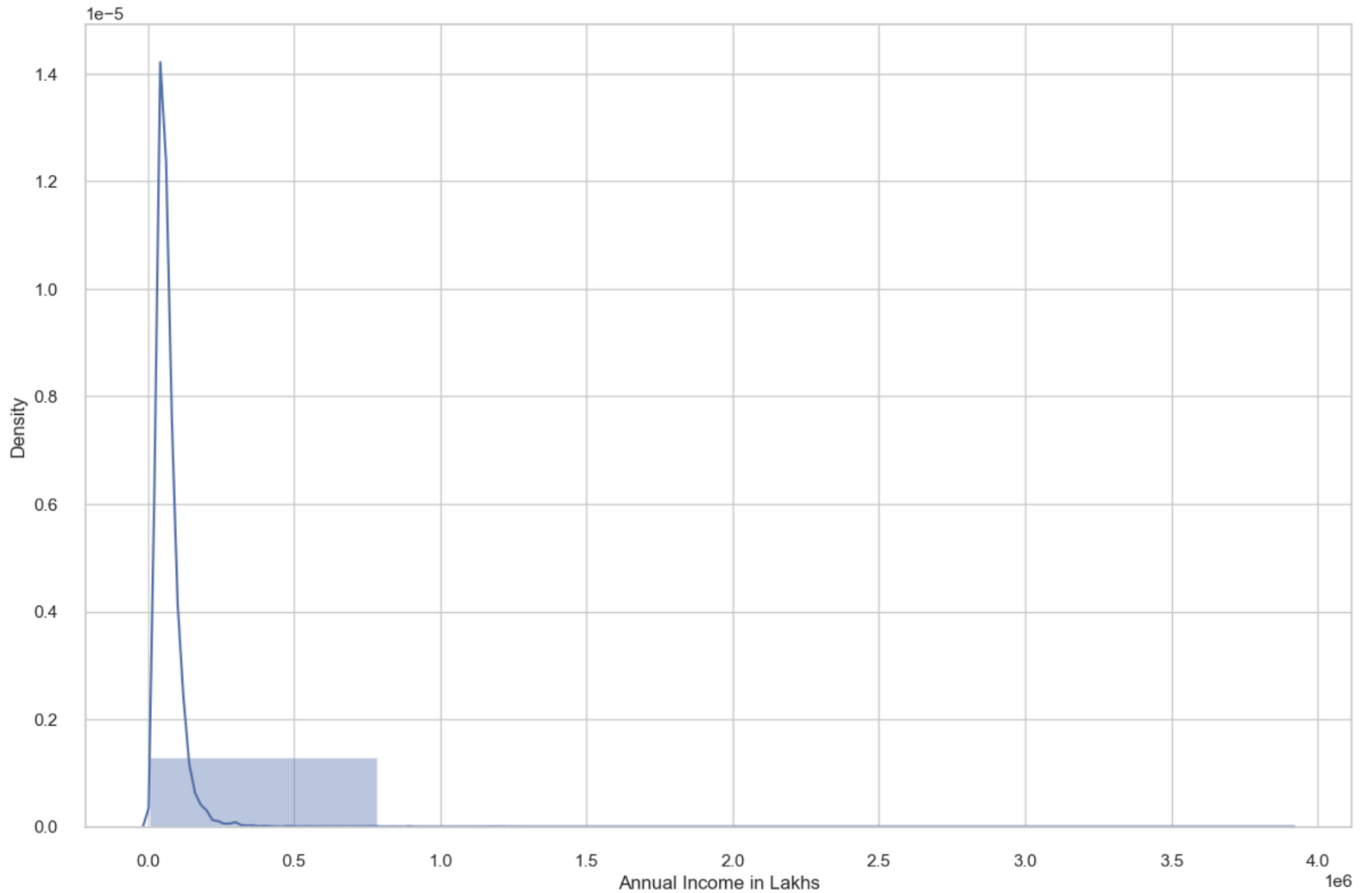
Loan Amount statistical summary:

- **Central Tendency:** On an average, loans tend to be around 11,133.
- **Variability:** The standard deviation is 7326.23. A higher standard deviation suggests greater variability in loan amounts, indicating a wider range of loan sizes within the dataset.
- **Distribution:** The median (50th percentile) is 10000 represents the middle value of the dataset and divides it into two equal parts. This indicates that 50% of the loan amounts are below 10,000, while the other 50% are above.
- **Range:** Min value is 500 and max value is 35000 which indicates that loans range from 500 to 35,000.
- **Common Loan Sizes:** The quartiles (25%, 50%, 75%) also provide insight into common loan sizes. 25% of loans are less than or equal to 5,500, 50% are less than or equal to 10,000, and 75% are less than or equal to 15,000.

Annual Income

Observation:

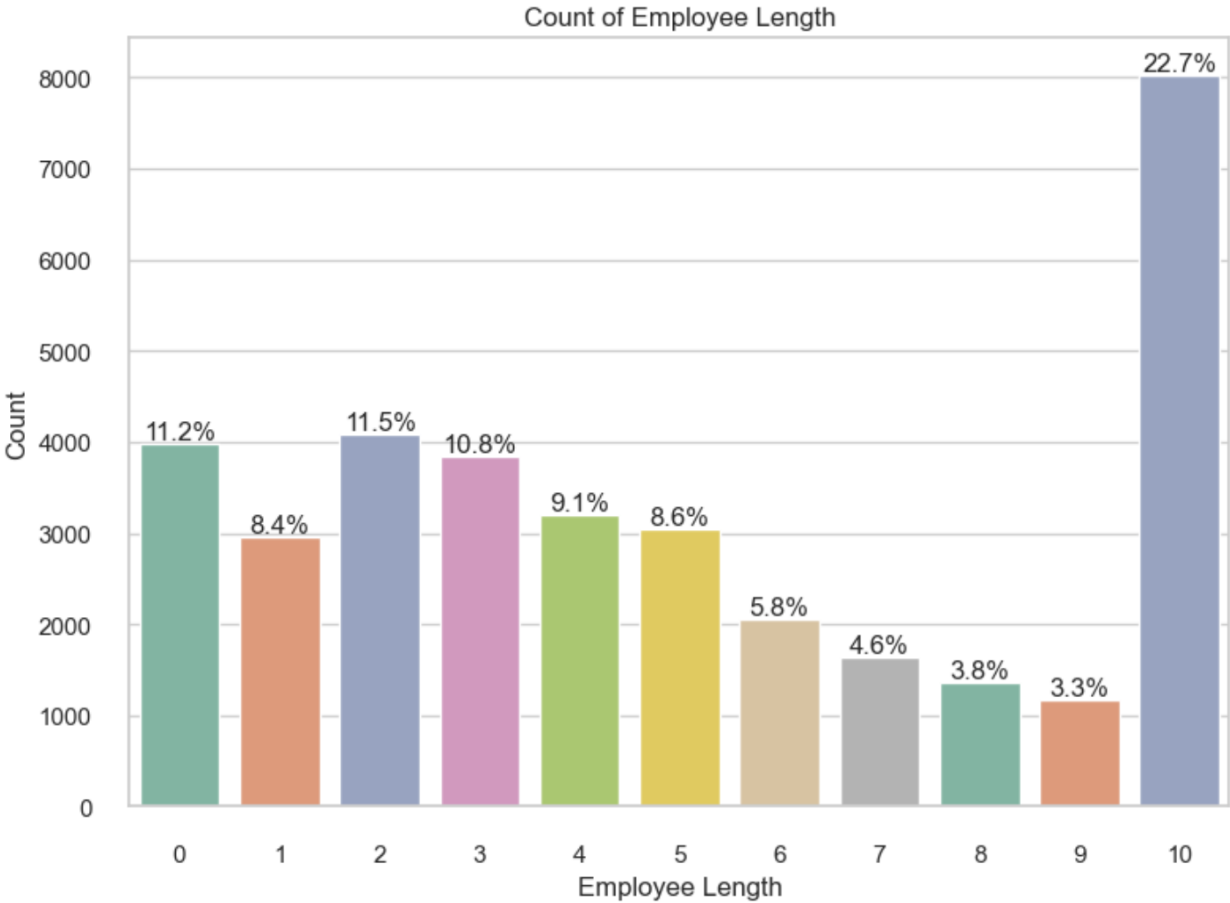
- Majority of loan applicants have income less than 5 lakhs



Employee Work Experience

Observation:

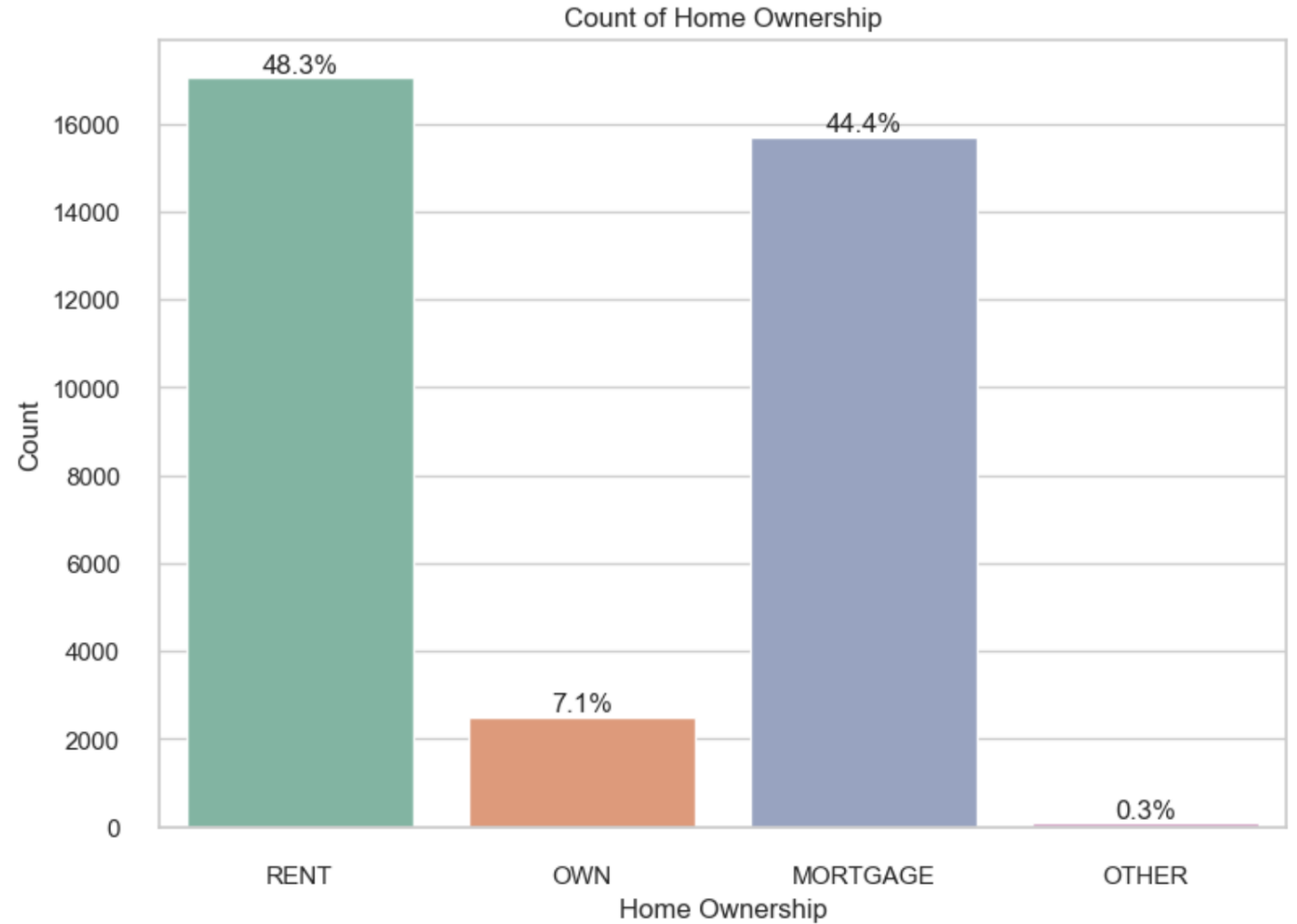
- 22.7% of the employees applying for the loan have more than 10 years of experience.
- 11.2% of the employee applying for the loan have zero years of experience.



Home Ownership

Observation:

- Applicant with RENT and MORTGAGE home ownership are more common.
- However, those with OWN home ownership are significantly fewer.
- This suggests that individuals who own their homes apply for loans less frequently..





Segmented Univariate Analysis

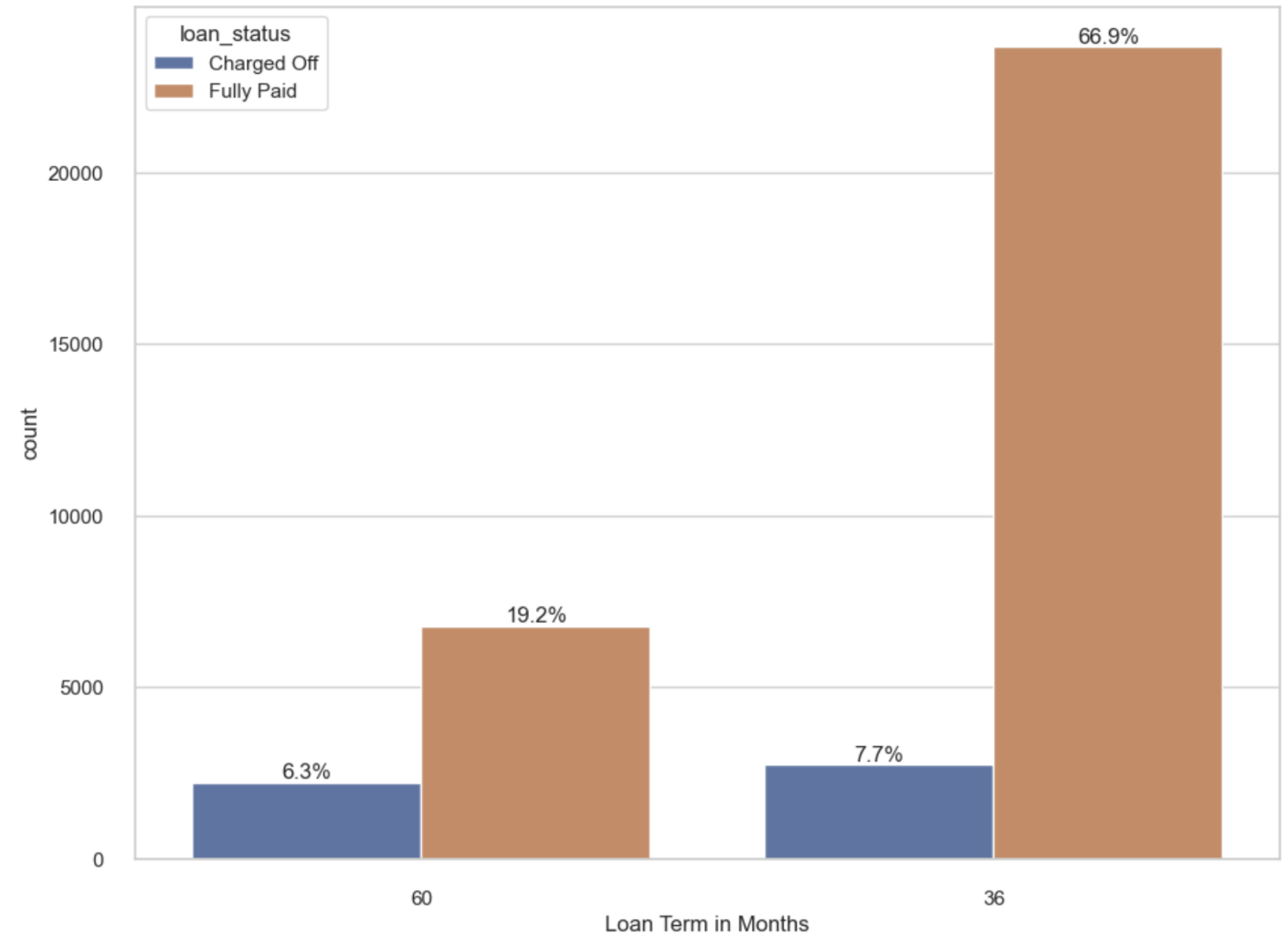
Loan Term and Loan Status

Observation:

- Term of 36 months have high chances to be fully paid, whereas defaulters for both terms are almost equal

Recommendation:

- Term of 36 months can be more profitable than that of 60 months



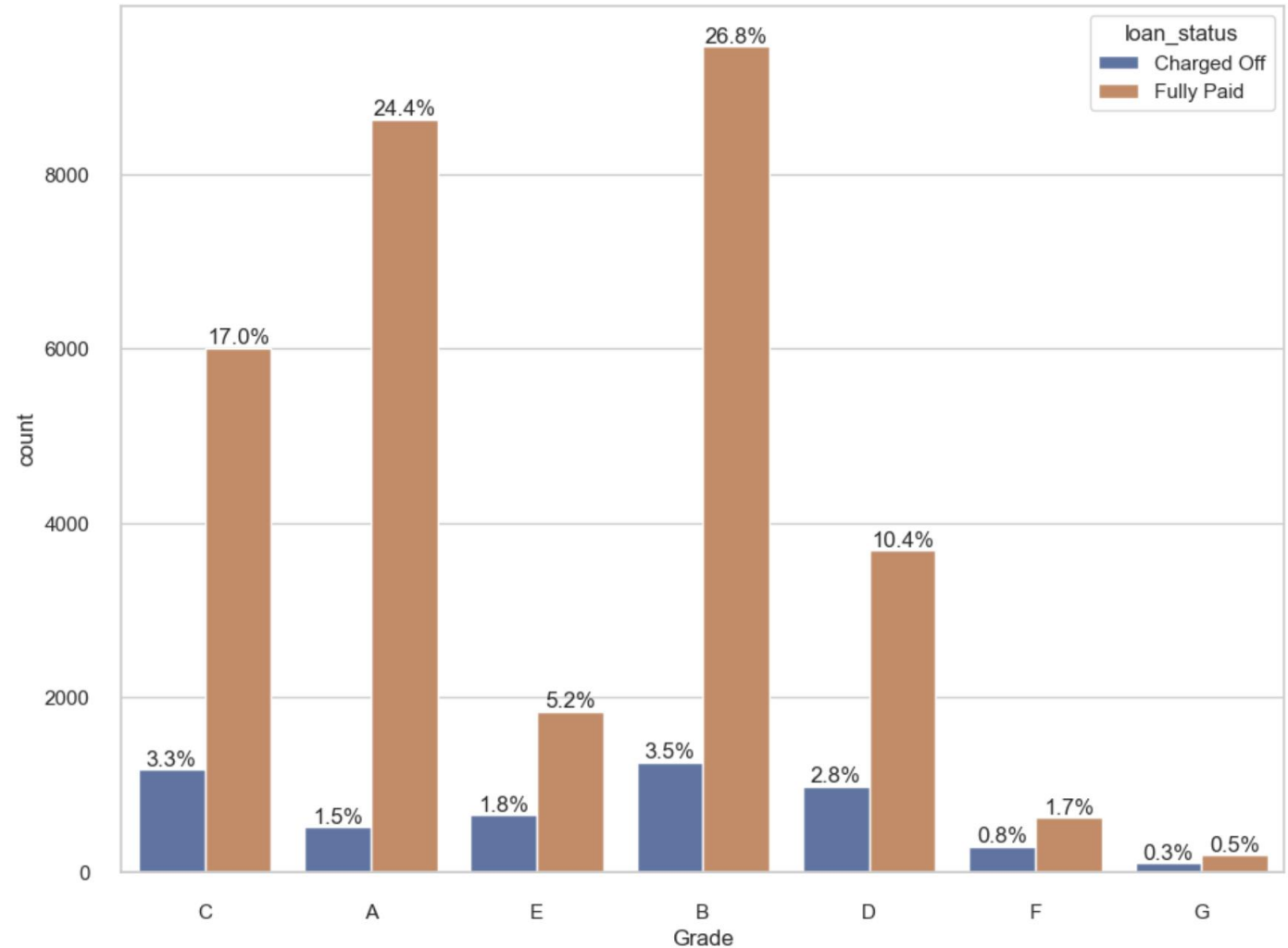
Grade and Loan Status

Observation:

- Default rates are high for B, C and D grades that means medium graded loan have high risk of defaulters.

Recommendation:

- Grades B, C and D are not recommended for loan approval.



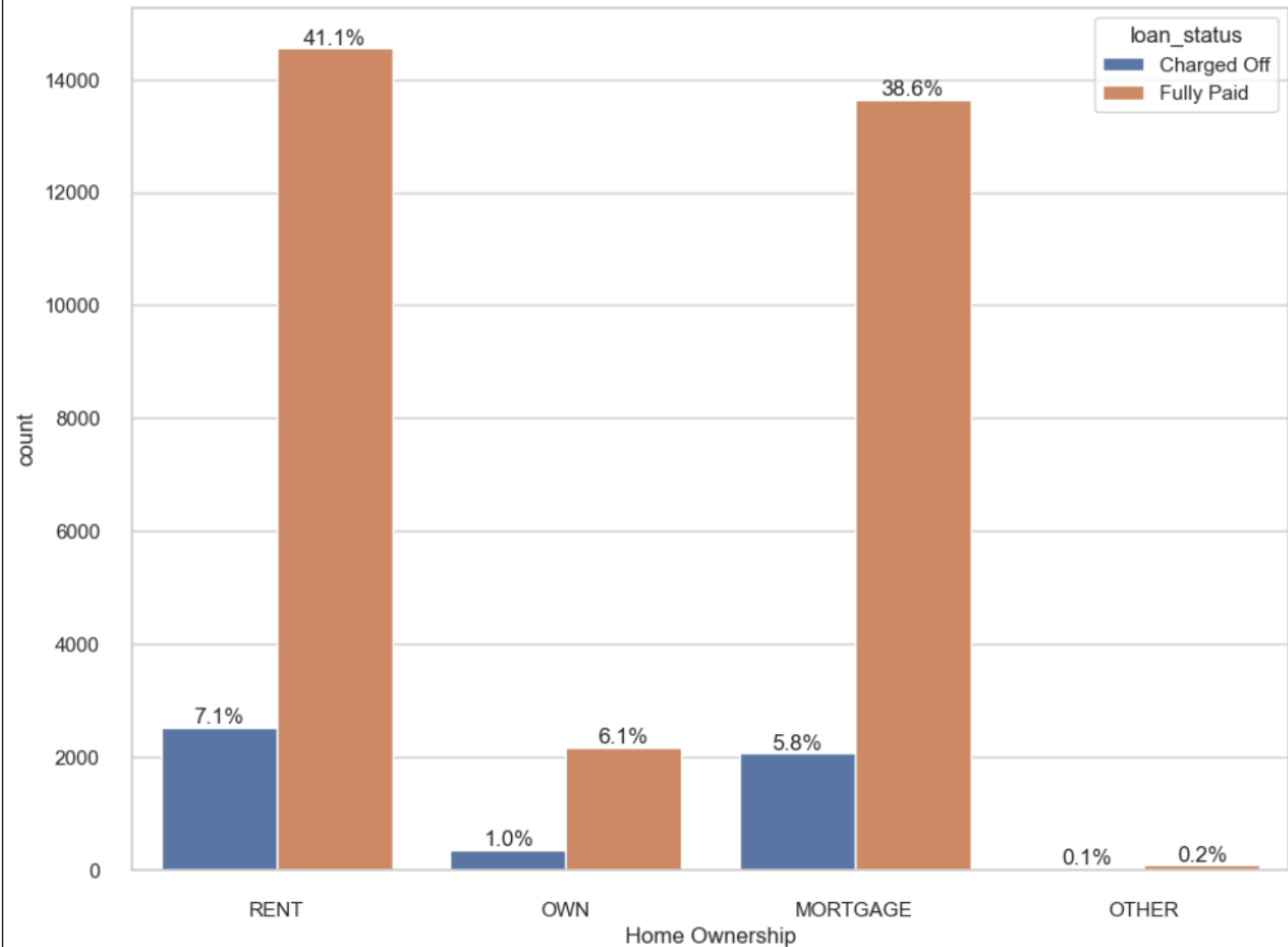
Home Ownership and Loan Status

Observation:

- The percentage of defaults is higher for individuals with RENT and MORTGAGE home ownership compared to others.
- People who stay in Rented house or Mortgage their house are high risk of being defaulters.

Recommendation:

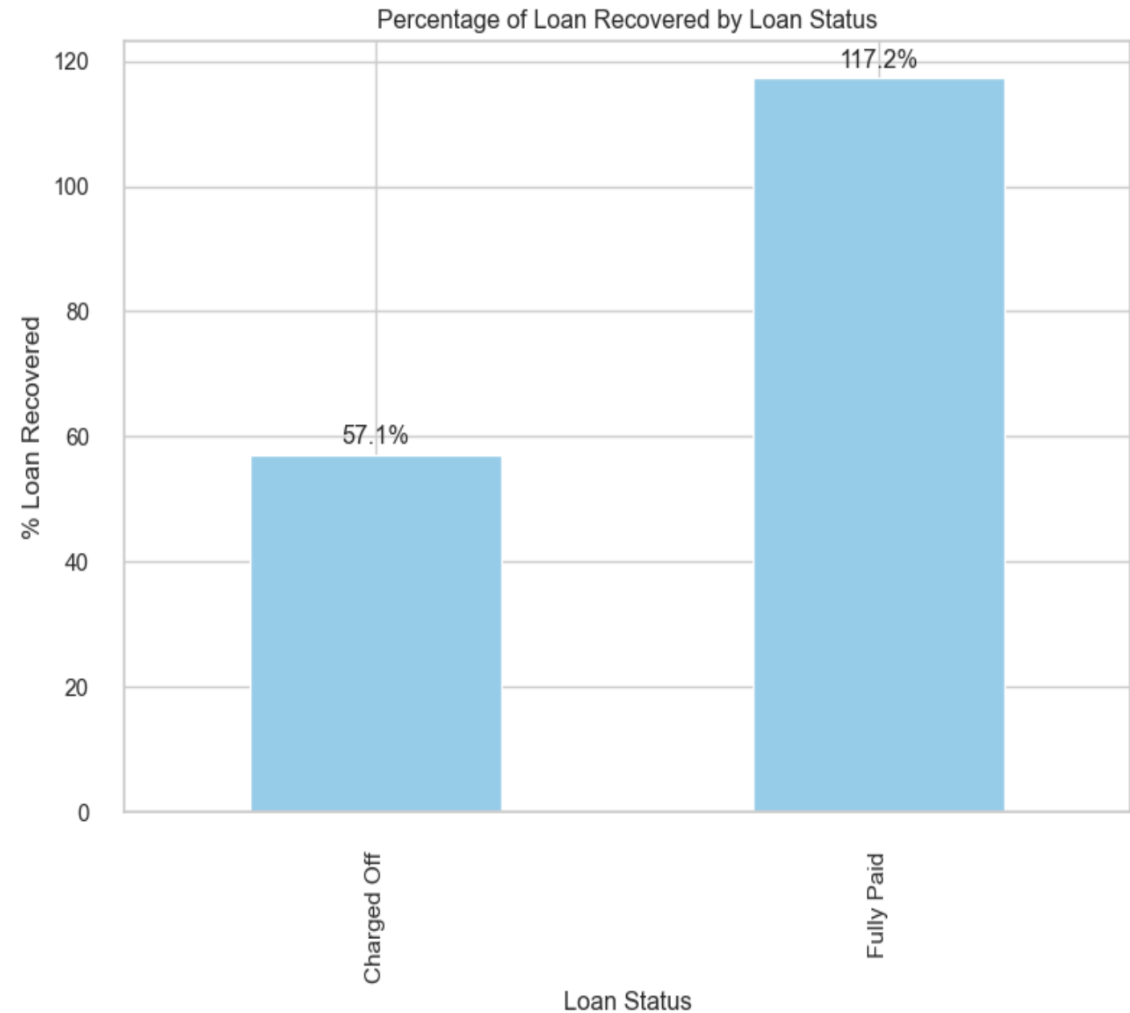
- Double verification for the credibility for applicant having RENT and MORTGAGE home ownership.



Loan Recovery Status

Observation:

- Lending Club only recovers 57% of the loan amount when loans are defaulted and on fully-paid up loans, the company makes 17% profit.



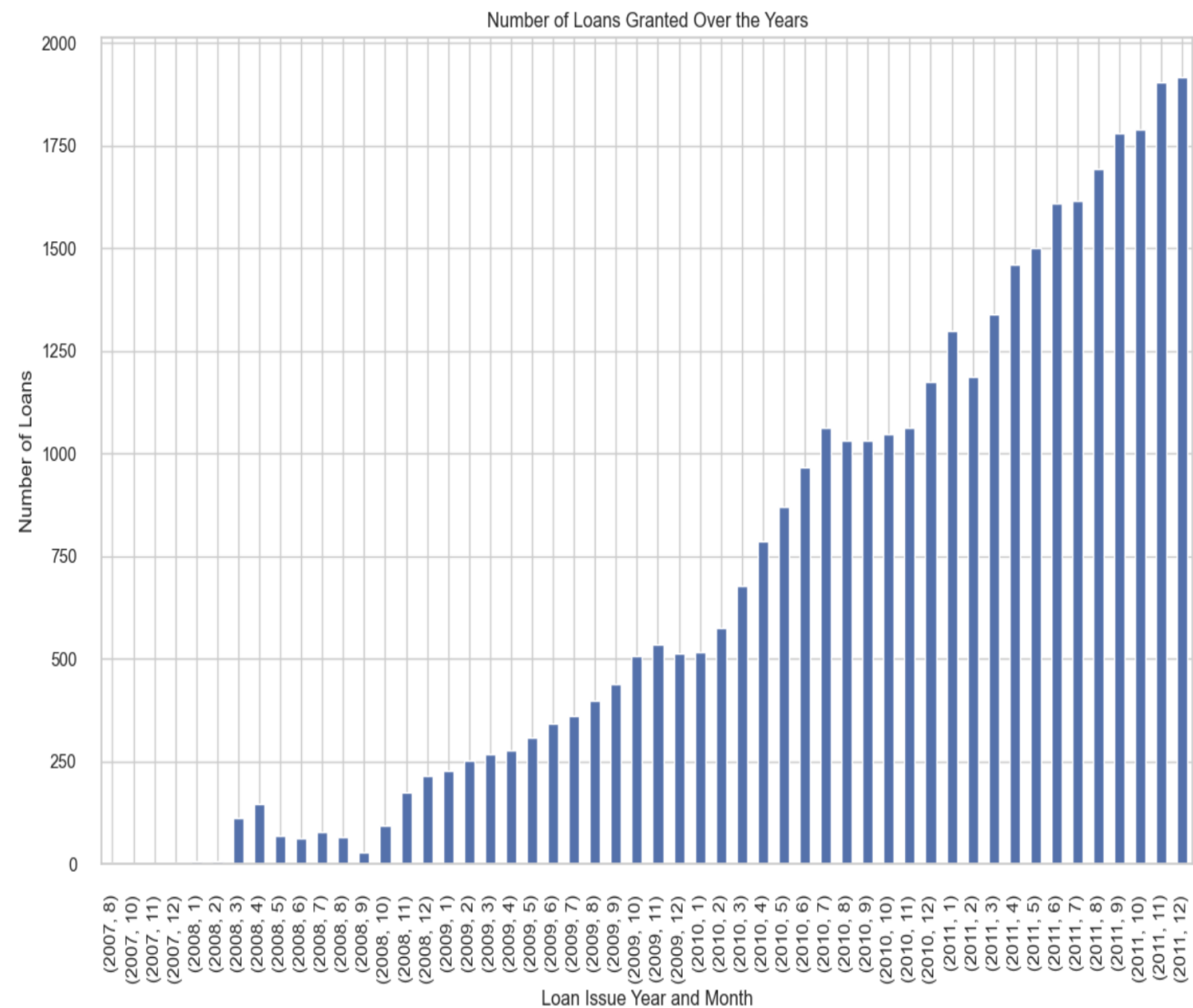
The slide features a solid orange background. A large white circle is centered on the page. The text "Bivariate Analysis" is written in a black, sans-serif font across the middle of the white circle. On the left side of the white circle, there is a dashed yellow arc. On the bottom right edge of the white circle, there is a small solid blue circle.

Bivariate Analysis

Number of loan applications over the years

Observation:

- Maximum applicants are observed in year 2011 and it is observed that loan applicants are exponentially increasing from 2007 to 2011.



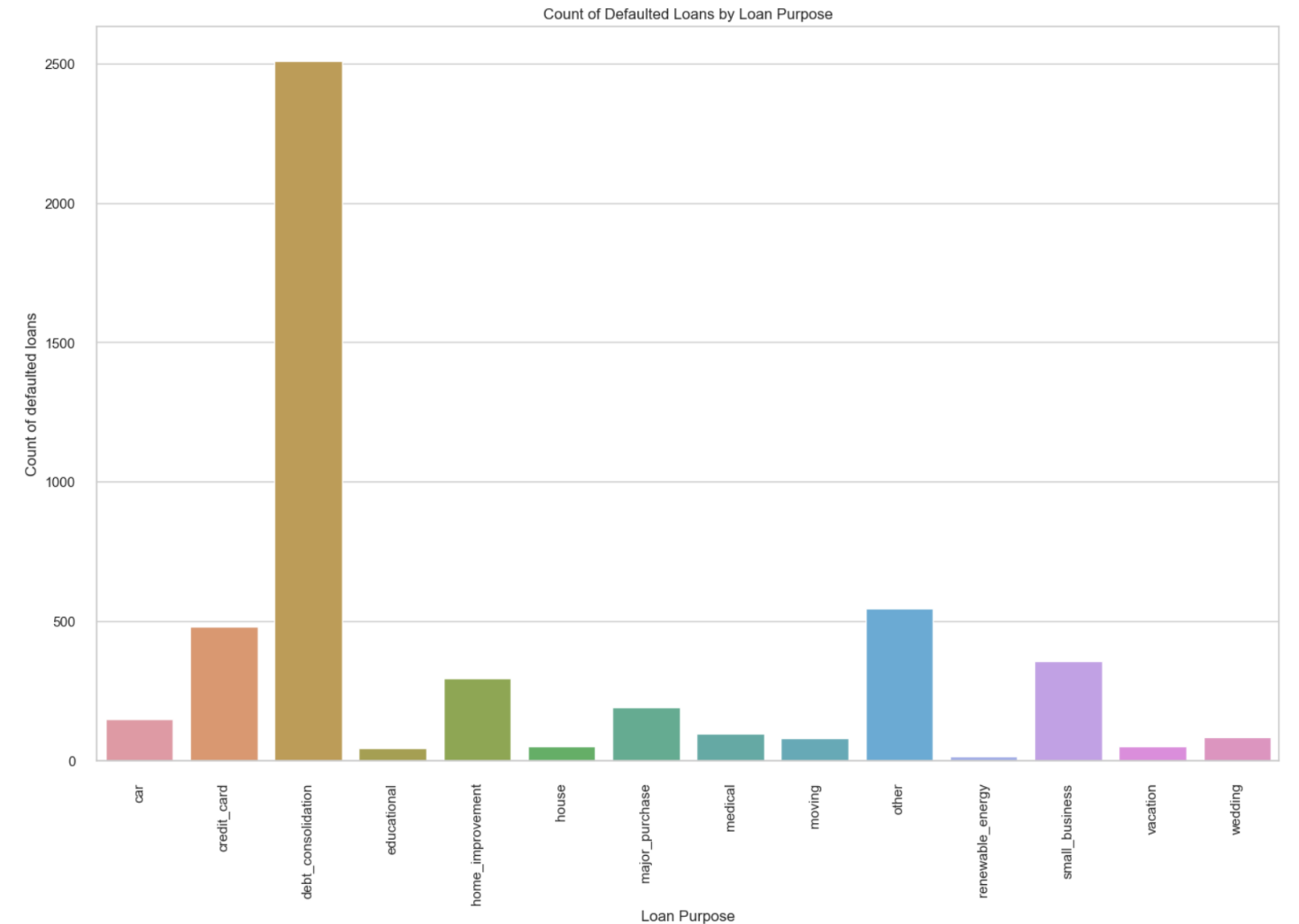
Defaulters based on Loan purpose

Observation:

- Majority of defaulters have been for the debt consolidation purpose.

Recommendation:

- Loan should not be granted for purpose of debt consolidation.



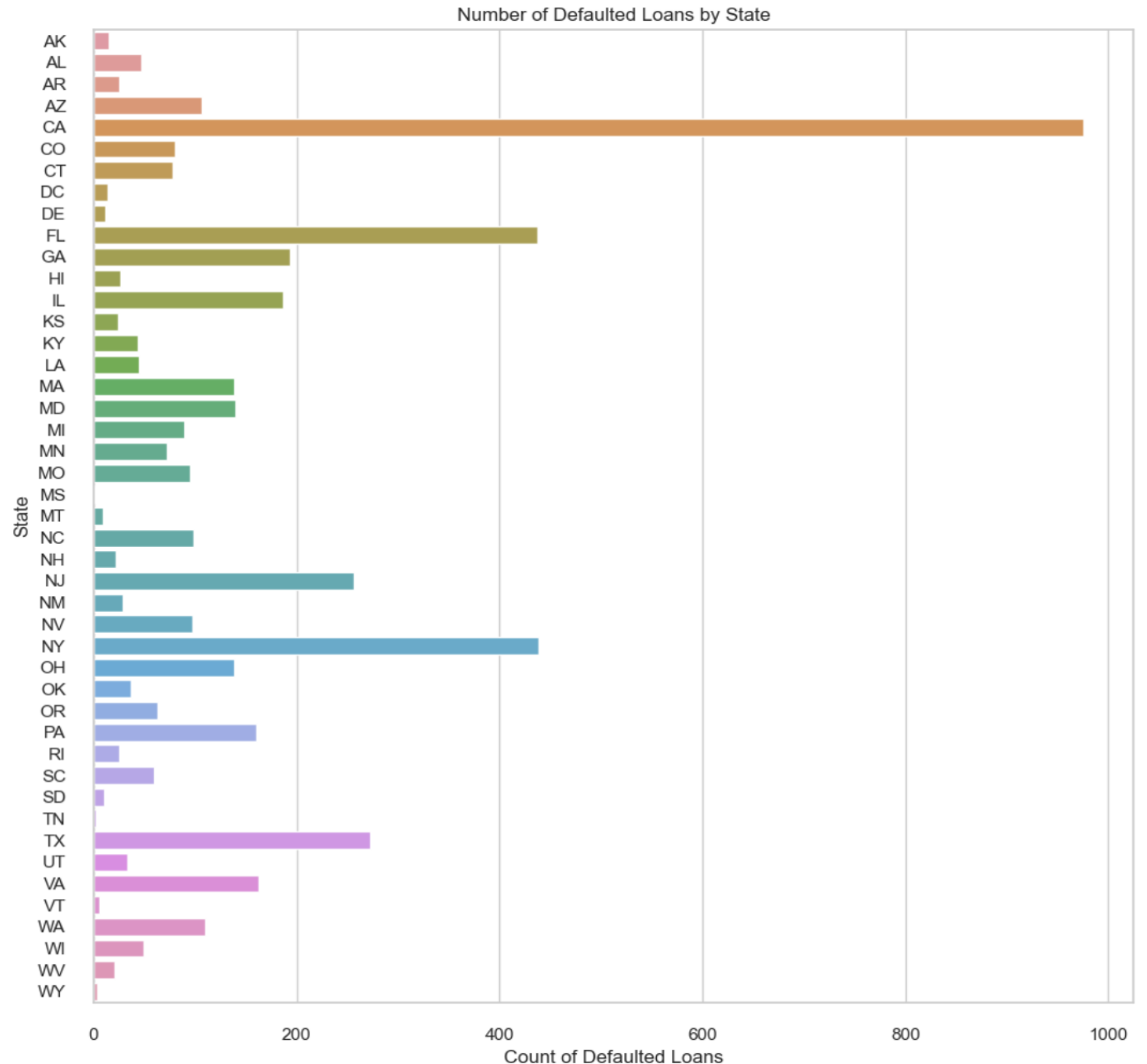
Defaulters based on State

Observation:

- Applicants from the state CA have a high probability of default.

Recommendation:

- Loan can be granted to people from states like TN, MS, WY and avoid CA, FL, NY.



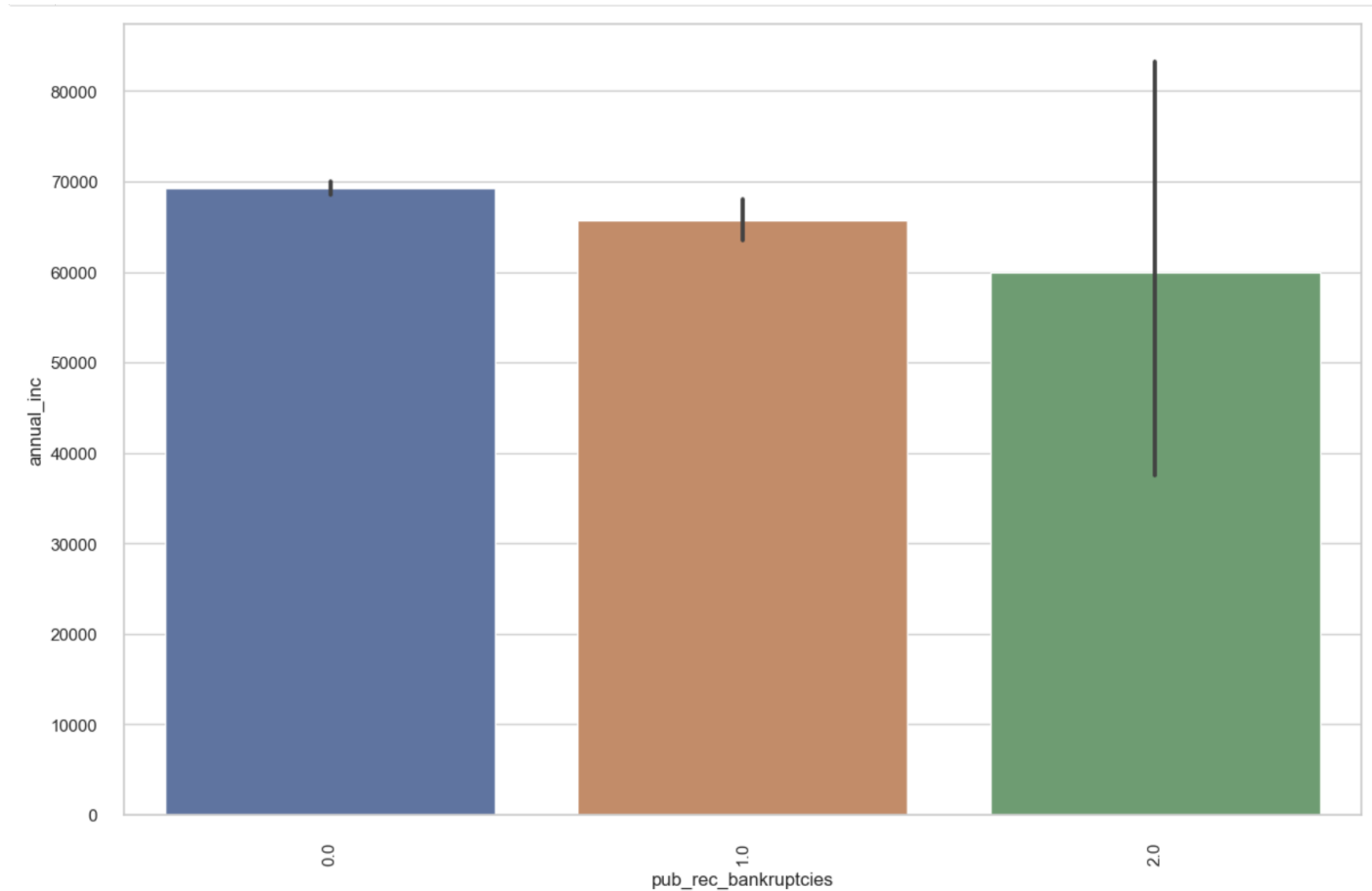
Bankruptcy Records based on Income

Observation:

- Employees with low income have maximum bankruptcies

Recommendation:

- Loan can be granted to people with 0 or 1 public bankruptcy record with high annual income.



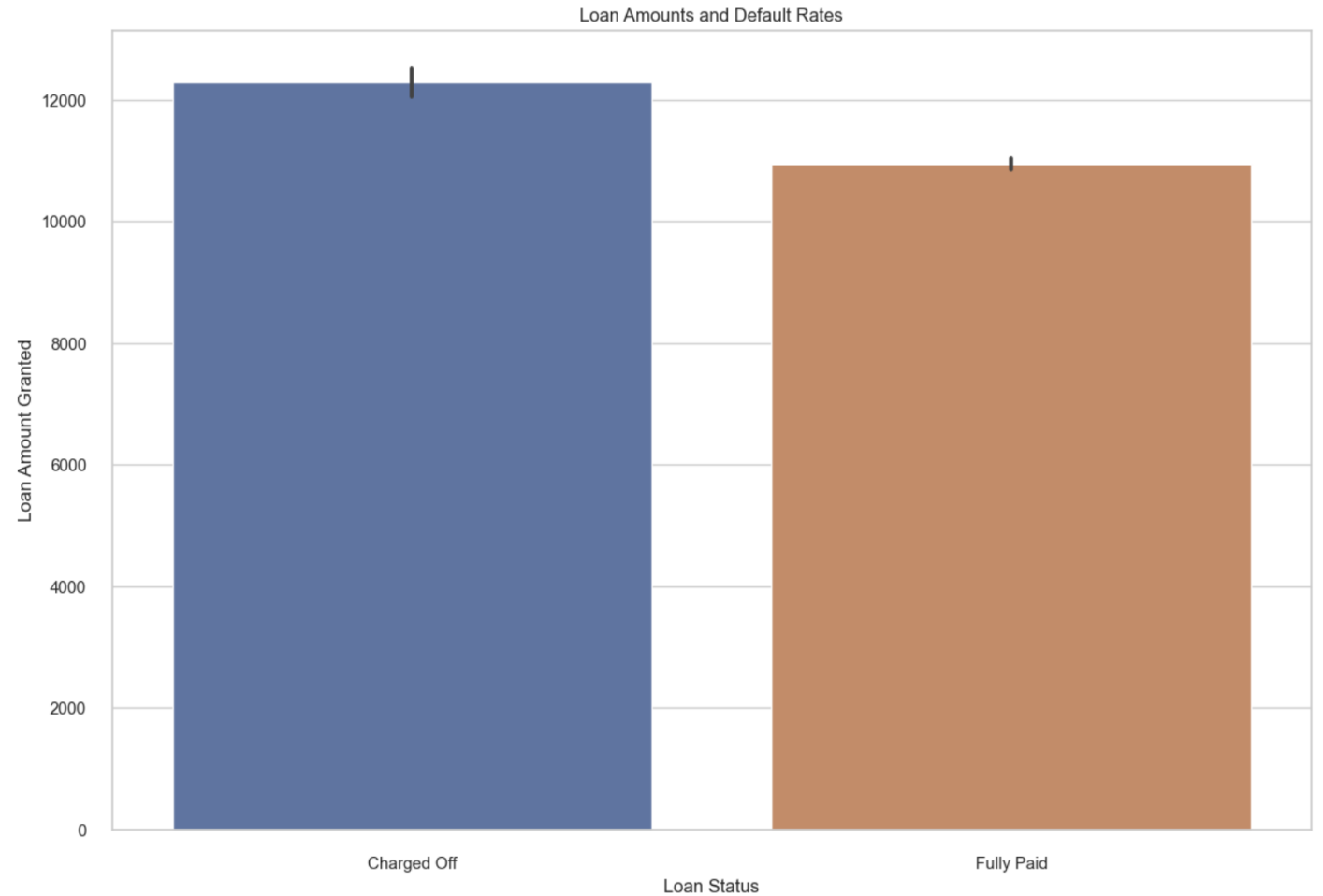
Defaulters based on loan amount

Observation:

- It is observed that defaulters are the one whose loan amount are greater than 12000.

Recommendation:

- Loan should not be granted for high amount.



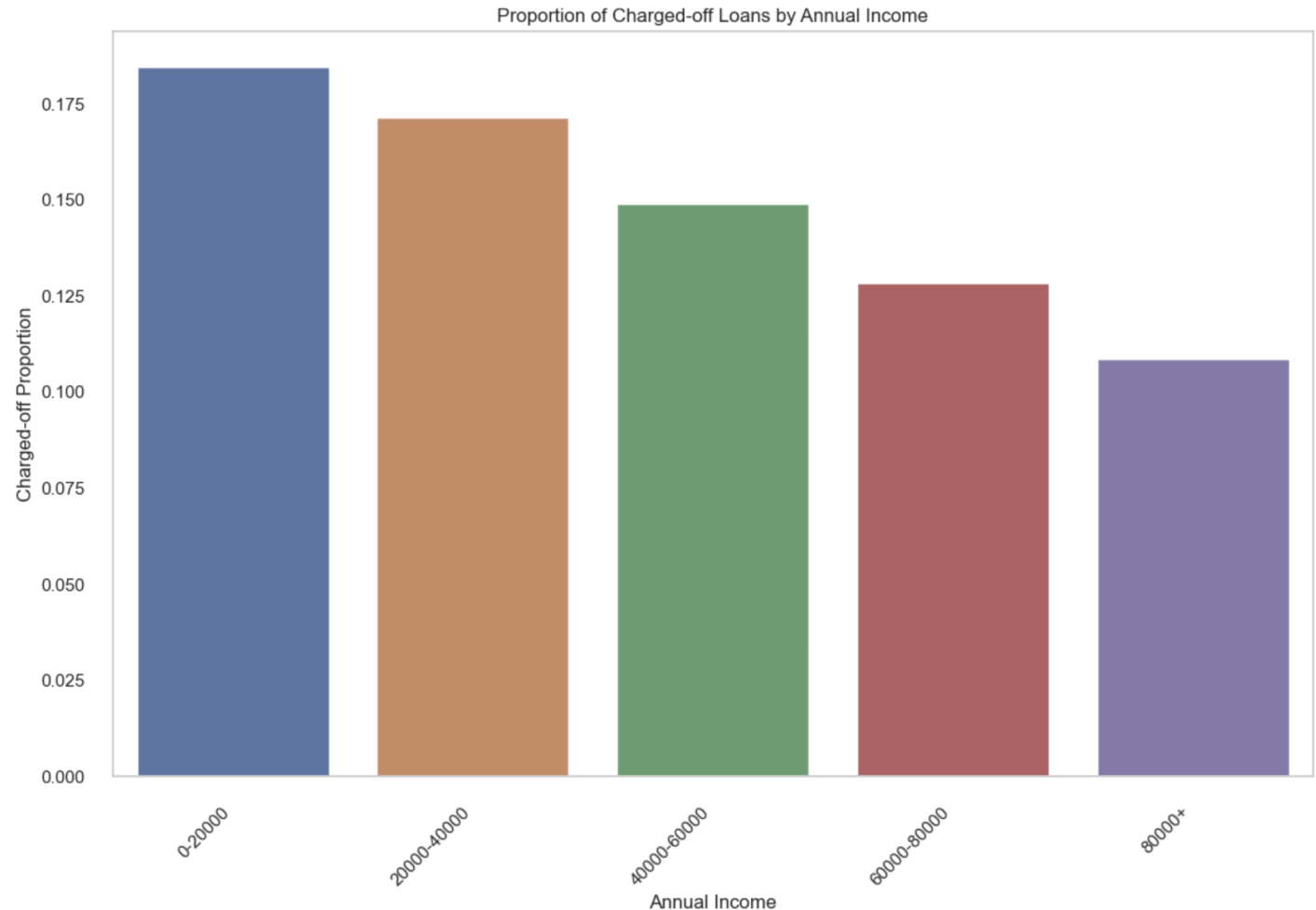
Defaulters based on Annual Income

Observation:

- Income range 80000+ has less chances of charged off.
- Income range 0-20000 has high chances of charged off.
- Notice that with increase in annual income charged off proportion got decreased.

Recommendation:

- Loan should be granted to people with high annual income.



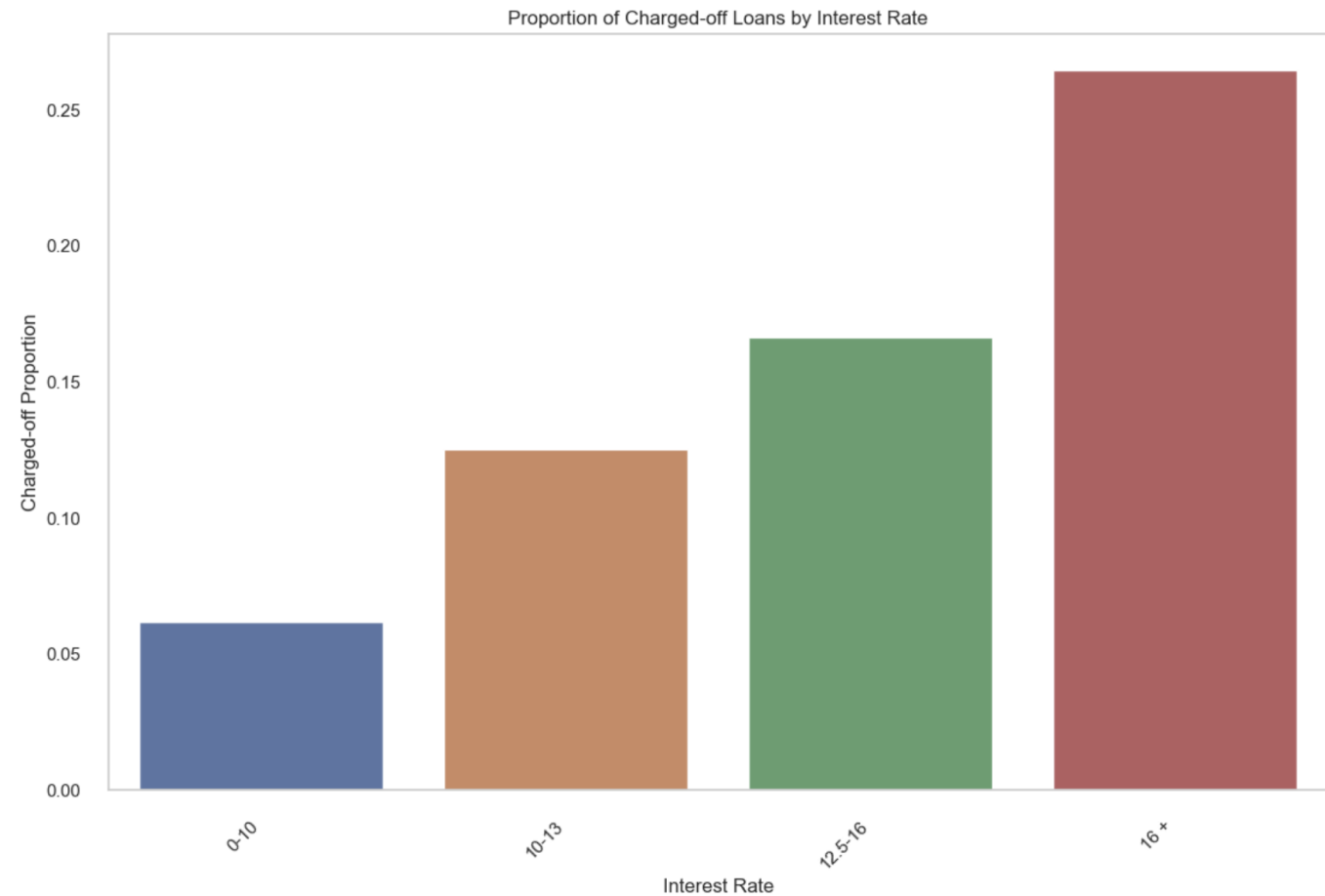
Defaulters based on Interest Rate

Observation:

- Interest rate less than 10% has very less chances of charged off.
- Interest rates are starting from minimum 5 %. Interest rate more than 16% has good chances of charged off as compared to other category interest rates.
- Charged off proportion is increasing with higher interest rates.

Recommendation:

- Loan should be granted to people with low interest rate.



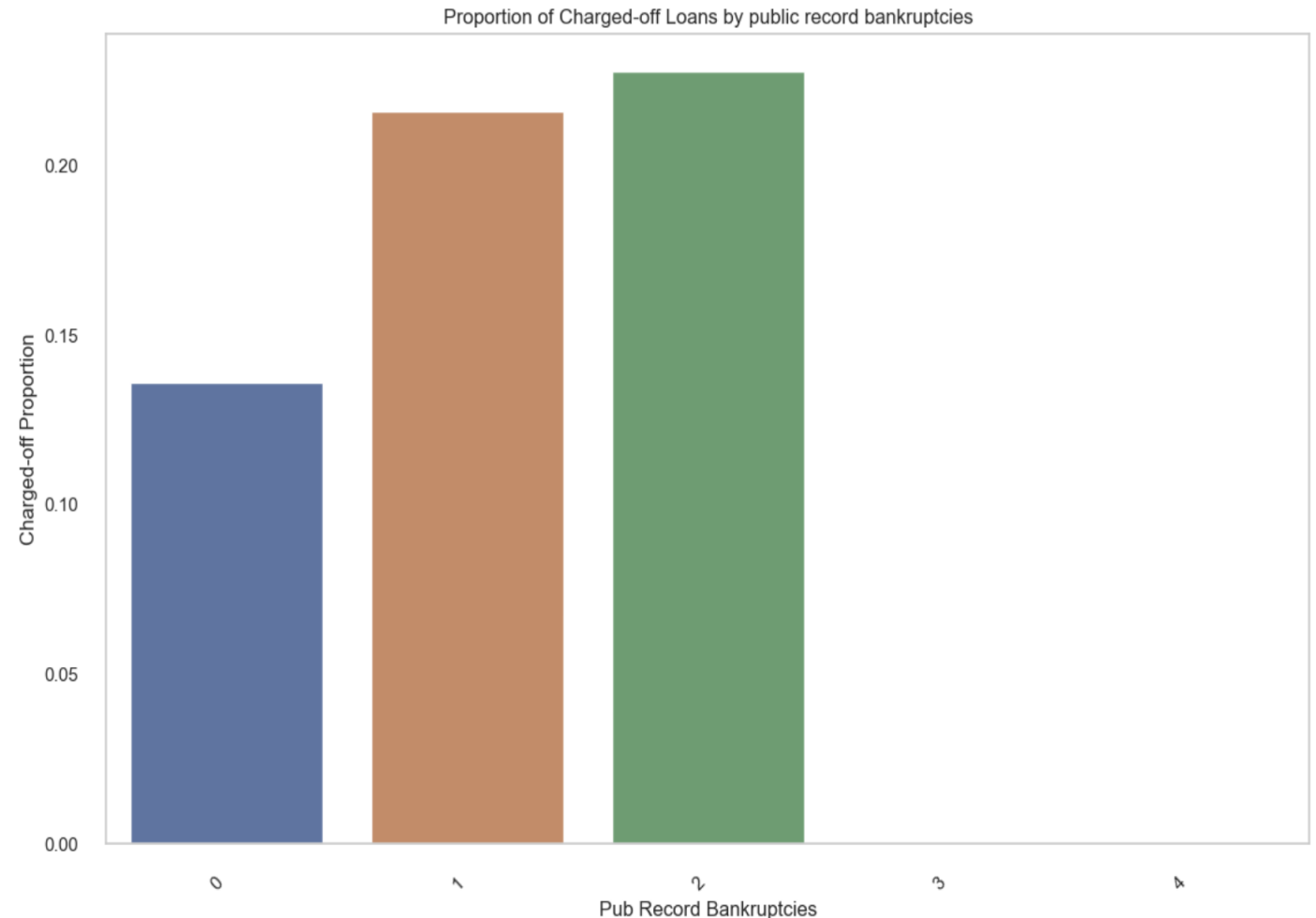
Defaulters based on Derogatory Public Records

Observation:

- Those who already have derogatory public record value 1 or 2 have charged off chances higher than those who have no Derogatory Public Record.
- Don't have enough data for count 3-4 so cannot reach on any conclusions.

Recommendation:

- Loan should be granted to people with 0 derogatory public record..



Recommendations

Sr. No.	Significant Attributes	Recommendation
1	Term	60 months term
2	Number of Public Record Bankruptcies	0 or 1 record
3	Derogatory Public Records	0 record
4	Funded amount	Applied for loan amount more than 15000
5	Employee Length	High
6	Annual Income	High
7	Grade	Other than B, C, D
8	Purpose	Vacation, Renewable Energy, Educational
9	State	Tennessee, Mississippi, Wyoming
10	Home Ownership	Other
11	Interest rate	Low
12	Loan amount	Low



Thank You
