

Apparao Mulpuri and Challa Maruthi Ratna Kishore

# Lending Club Case Study

# Objective

- The objective of this exercise is to identify patterns that indicate a person is likely to default, which can then be used to take steps such as refusing the loan, lowering the loan amount, or lending (to riskier applicants) at a higher interest rate.

# Steps followed in this exercise

- Data Cleaning
  - a. Remove columns that contain no valuable data and have a high number of null values
  - Identify null rows and delete or replace them with pertinent data.
- Identify outliers and delete those rows.
- Create additional columns
- Univariate Data Analysis
- Bivariate Data Analysis
- Multivariate Data Analysis

# Data Cleaning

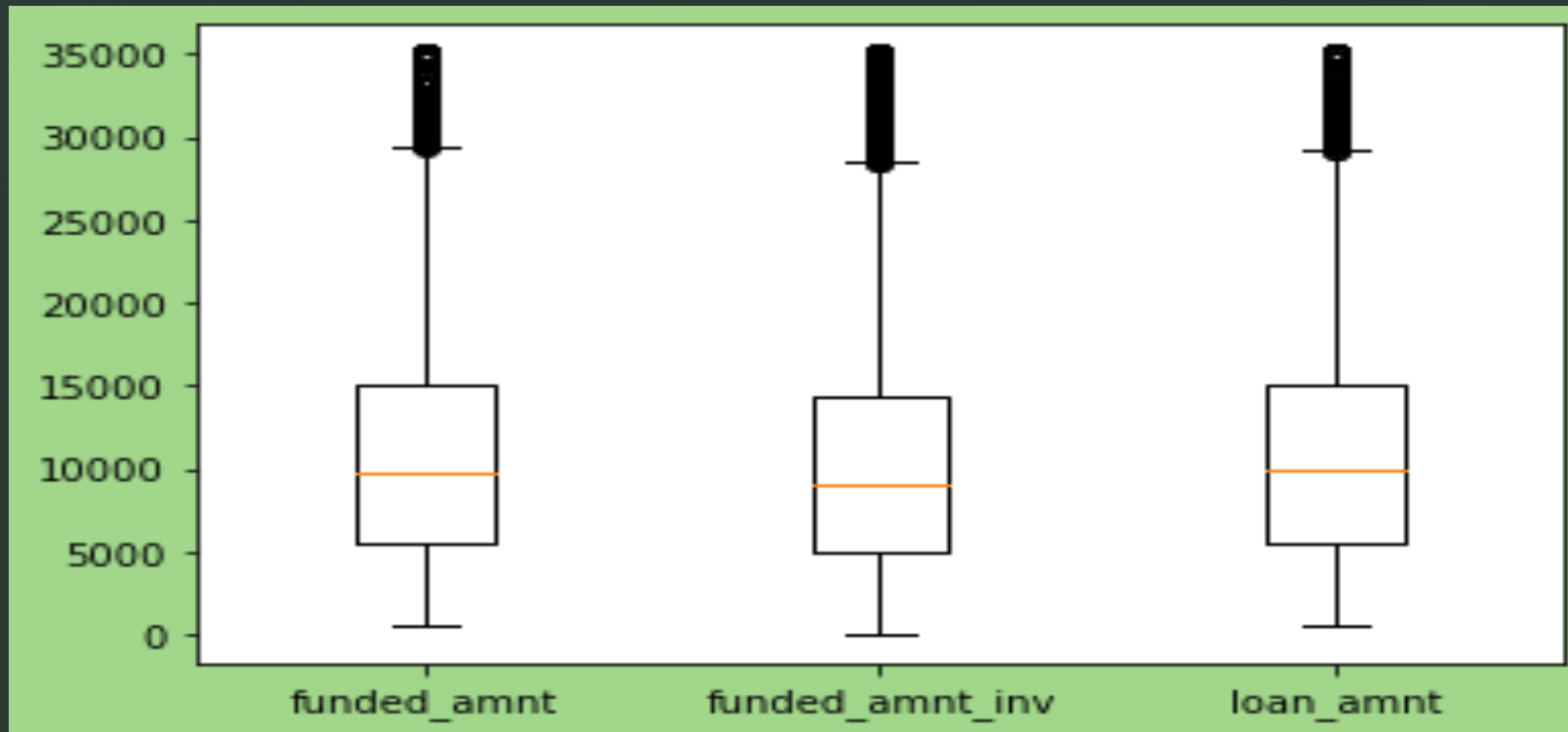
- Dropped multiple columns from the loan data set because of most of the values are NULL
  - tot\_hi\_cred\_lim, total\_bc\_limit, total\_bal\_ex\_mort, total\_il\_high\_credit\_limit, num\_tl\_op\_past\_12m, pct\_tl\_nvr\_dlq, percent\_bc\_gt\_75, 'num\_tl\_120dpd\_2m', 'num\_tl\_30dpd', 'num\_tl\_90g\_dpd\_24m', 'num\_rev\_accts', 'num\_rev\_tl\_bal\_gt\_0', 'num\_sats', 'num\_bc\_tl', 'num\_il\_tl', 'num\_op\_rev\_tl', 'num\_actv\_bc\_tl', 'num\_actv\_rev\_tl', 'num\_bc\_sats', 'mths\_since\_recent\_inq', 'mths\_since\_recent\_revol\_delinq', 'num\_accts\_ever\_120\_pd', 'mort\_acc', 'mths\_since\_recent\_bc', 'mths\_since\_recent\_bc\_dlq', 'mo\_sin\_old\_rev\_tl\_op', 'mo\_sin\_rcnt\_rev\_tl\_op', 'mo\_sin\_rcnt\_tl', 'mo\_sin\_old\_il\_acct', 'bc\_util', 'bc\_open\_to\_buy', 'avg\_cur\_bal', 'acc\_open\_past\_24mths', 'inq\_last\_12m', 'total\_cu\_tl', 'inq-fi', 'open\_rv\_12m', 'open\_rv\_24m', 'max\_bal\_bc', 'all\_util', 'total\_rev\_hi\_lim', 'il\_util', 'open\_acc\_6m', 'open\_il\_6m', 'open\_il\_12m', 'open\_il\_24m', 'mths\_since\_rcnt\_il', 'total\_bal\_il', 'annual\_inc\_joint', 'dti\_joint', 'verification\_status\_joint', 'tot\_coll\_amt', 'tot\_cur\_bal', 'next\_pymnt\_d', 'mths\_since\_last\_major\_derog', 'mths\_since\_last\_record', 'mths\_since\_last\_delinq', desc, emp\_title
- In **emp\_length**, removed years and < > and + symbols and fill 0 for the null values
- pub\_rec\_bankruptcies, 'title', 'total\_acc', 'last\_pymnt\_d', 'collections\_12\_mths\_ex\_med', 'chargeoff\_within\_12\_mths', 'tax\_liens' – filled NA for the null values

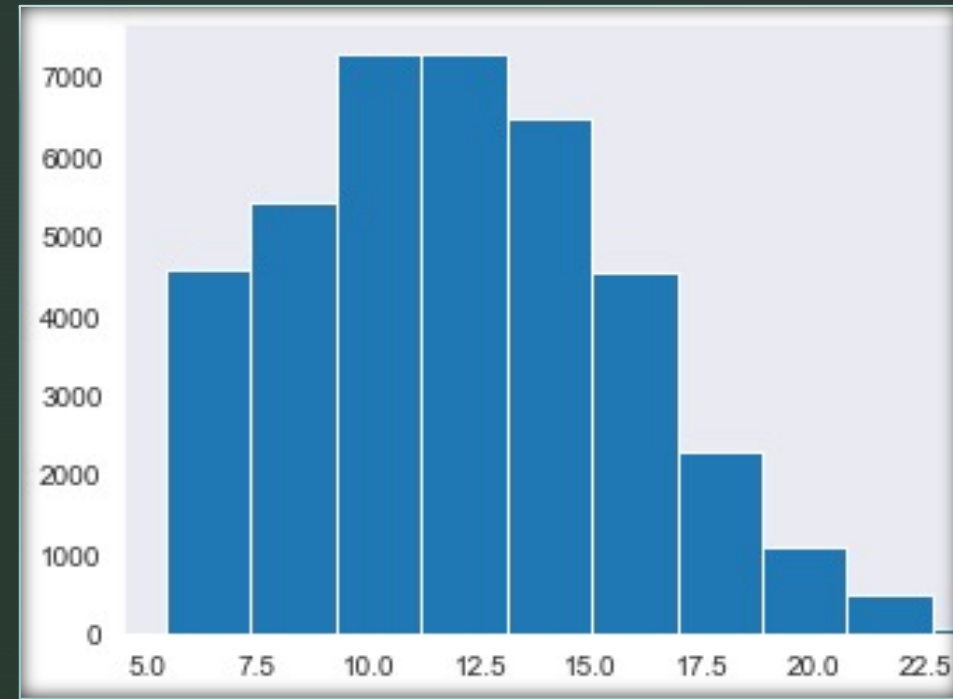
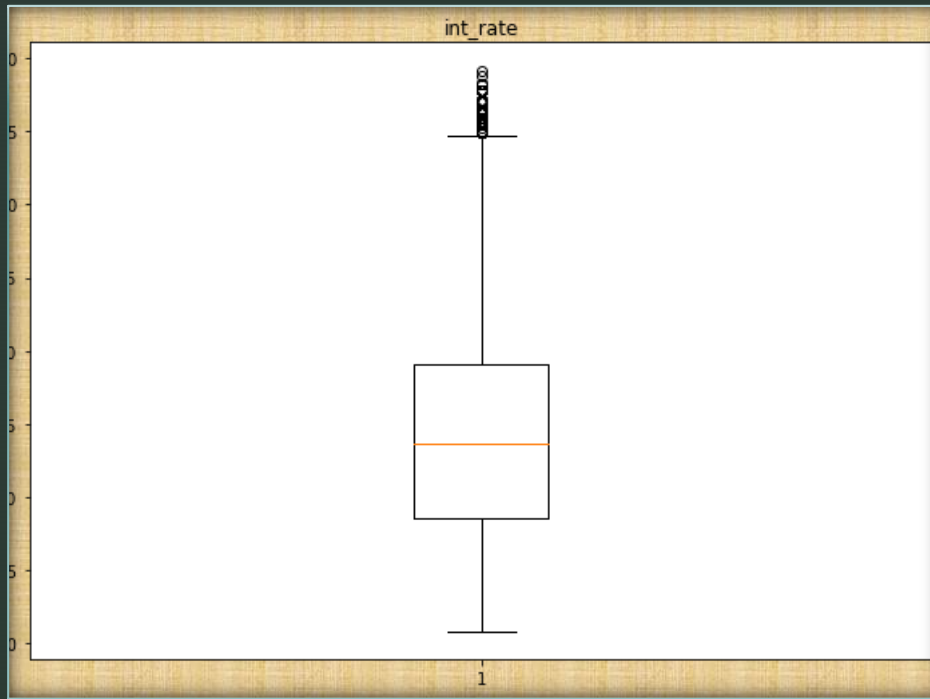
# Data Cleaning

- Divided the Date columns into multiple columns: Created issue\_year, issue\_month from issue\_d and last\_pymnt\_year, last\_pymnt\_month from last\_pymnt\_d
- Removed percentage symbols from revol\_util and int\_rate columnns
- Converted columns to numeric type for plots. 'loan\_amnt', 'funded\_amnt', 'funded\_amnt\_inv', 'int\_rate', 'installment', 'emp\_length', 'annual\_inc', 'dti', 'delinq\_2yrs', 'inq\_last\_6mths', 'open\_acc', 'pub\_rec', 'revol\_bal', 'revol\_util', 'total\_acc', 'out\_prncp\_inv', 'total\_pymnt', 'total\_pymnt\_inv', 'total\_rec\_prncp', 'total\_rec\_int', 'total\_rec\_late\_fee', 'recoveries', 'collection\_recovery\_fee', 'last\_pymnt\_amnt']



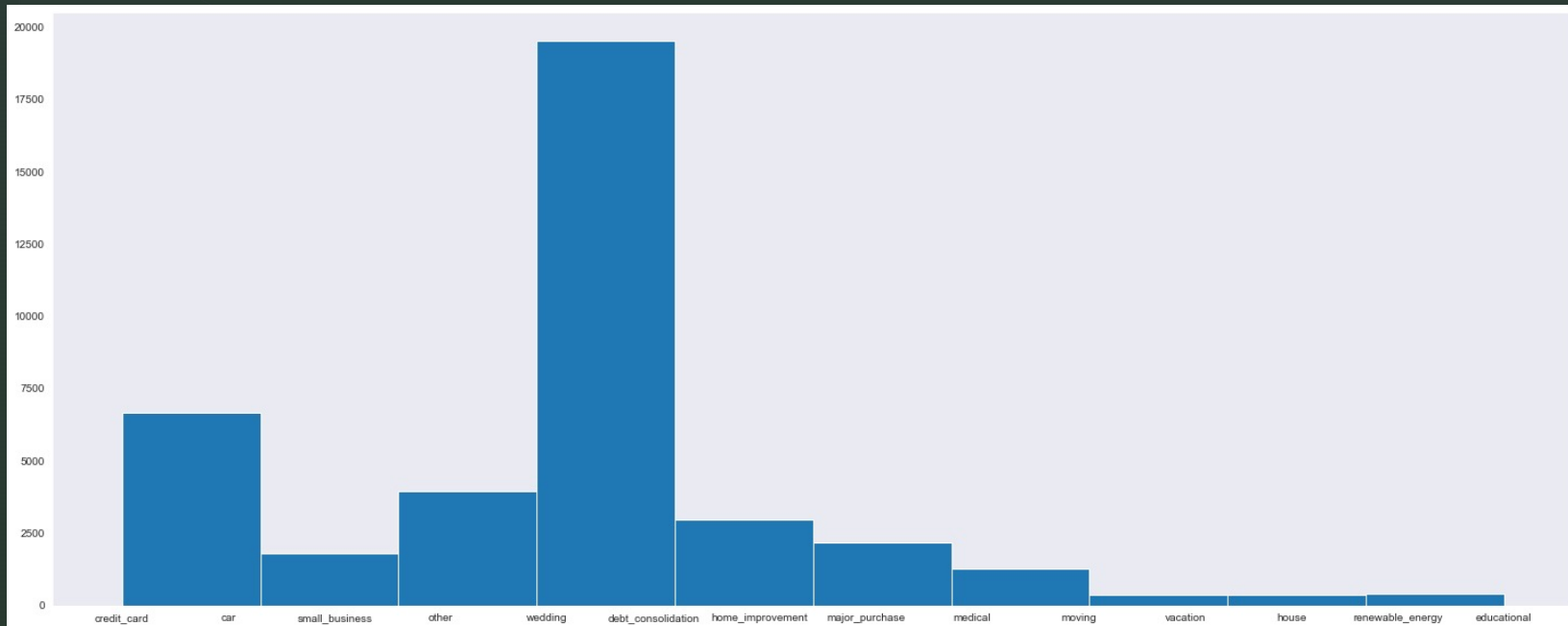
Observation: Distribution of loan amount and funded amounts are almost same





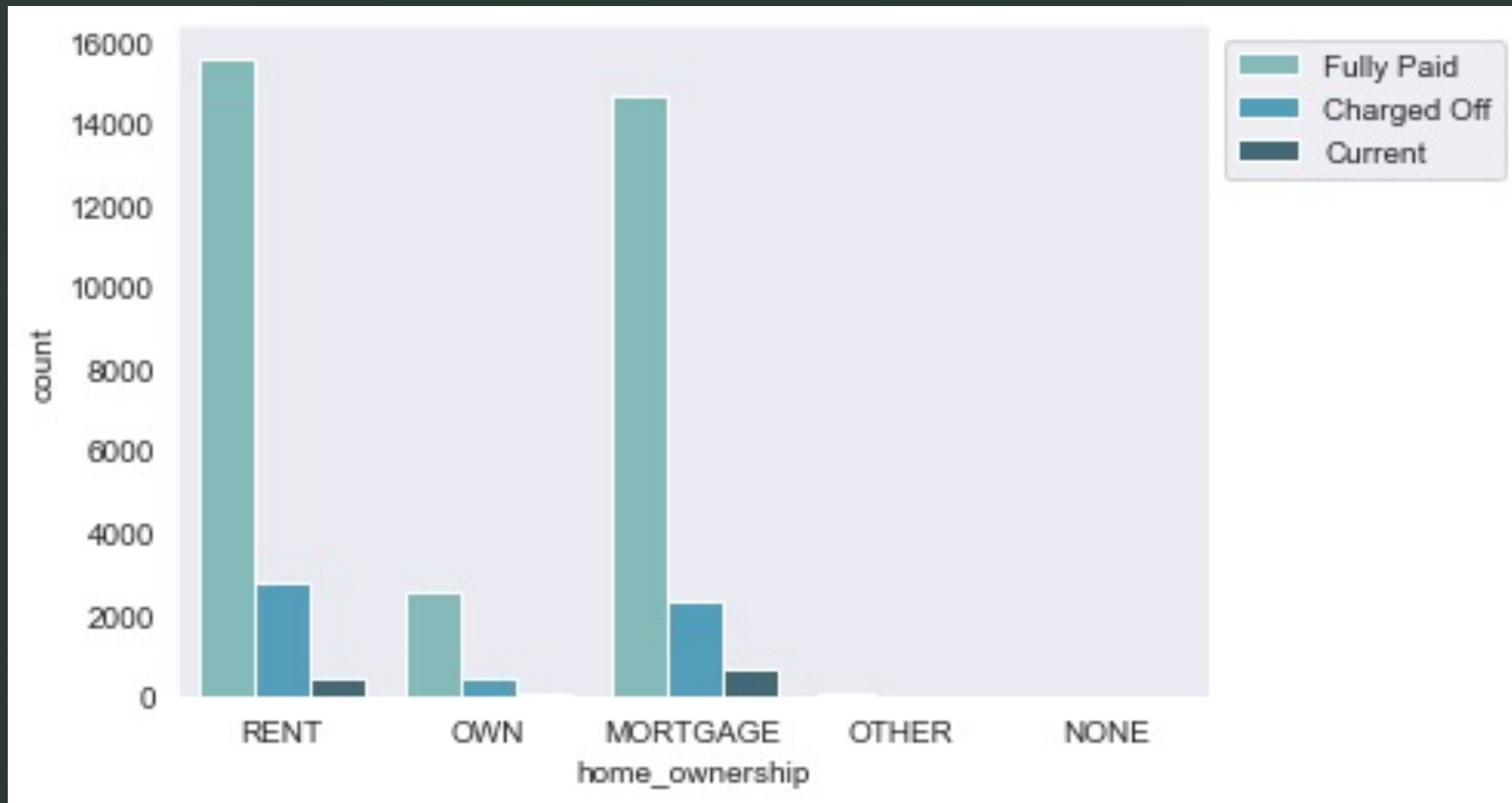
Observation: Most of the loans are taken with the interest rates between 10 -15%

Observation: most of the loans are taken for the debt consolidation purpose

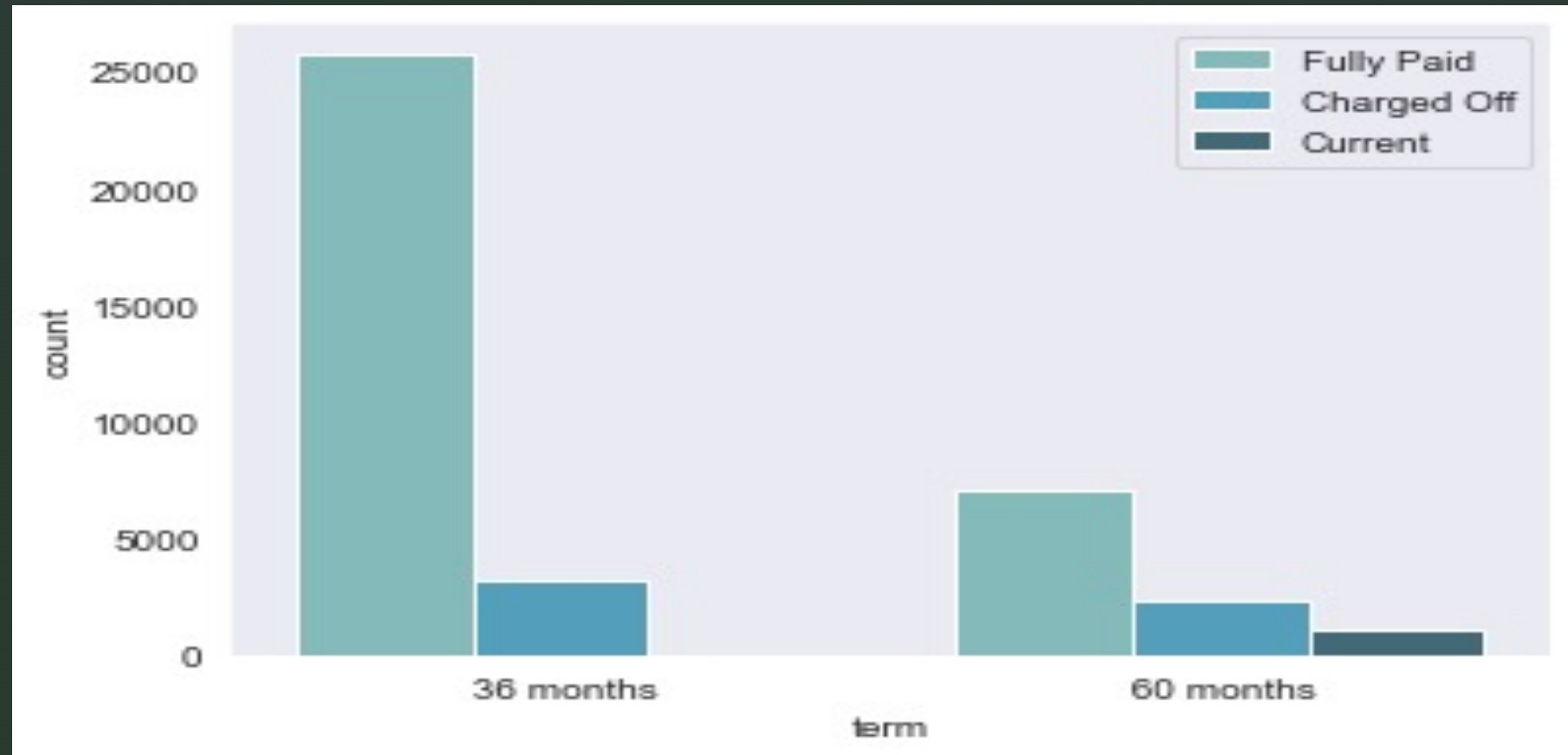




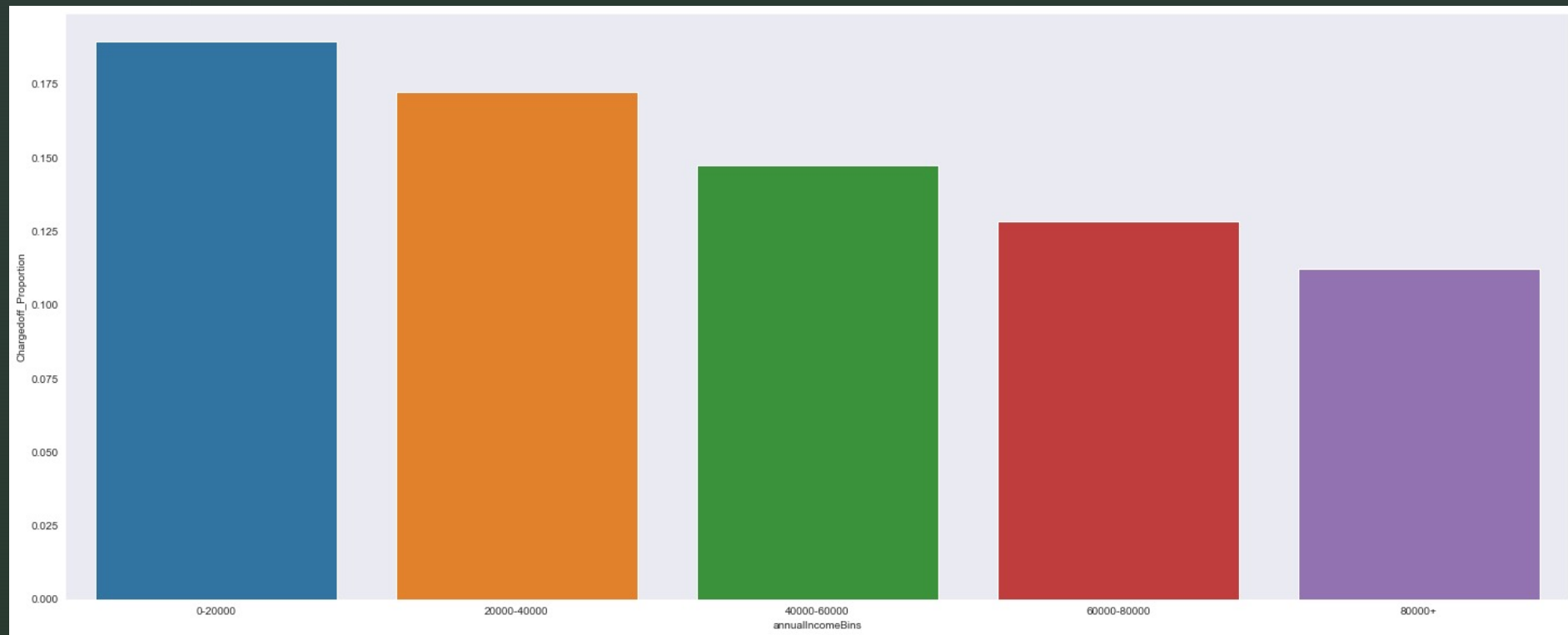
Observation: Most of the loans are taken by the people living in rent house or to mortgage their home. Charged Off is high from these categories



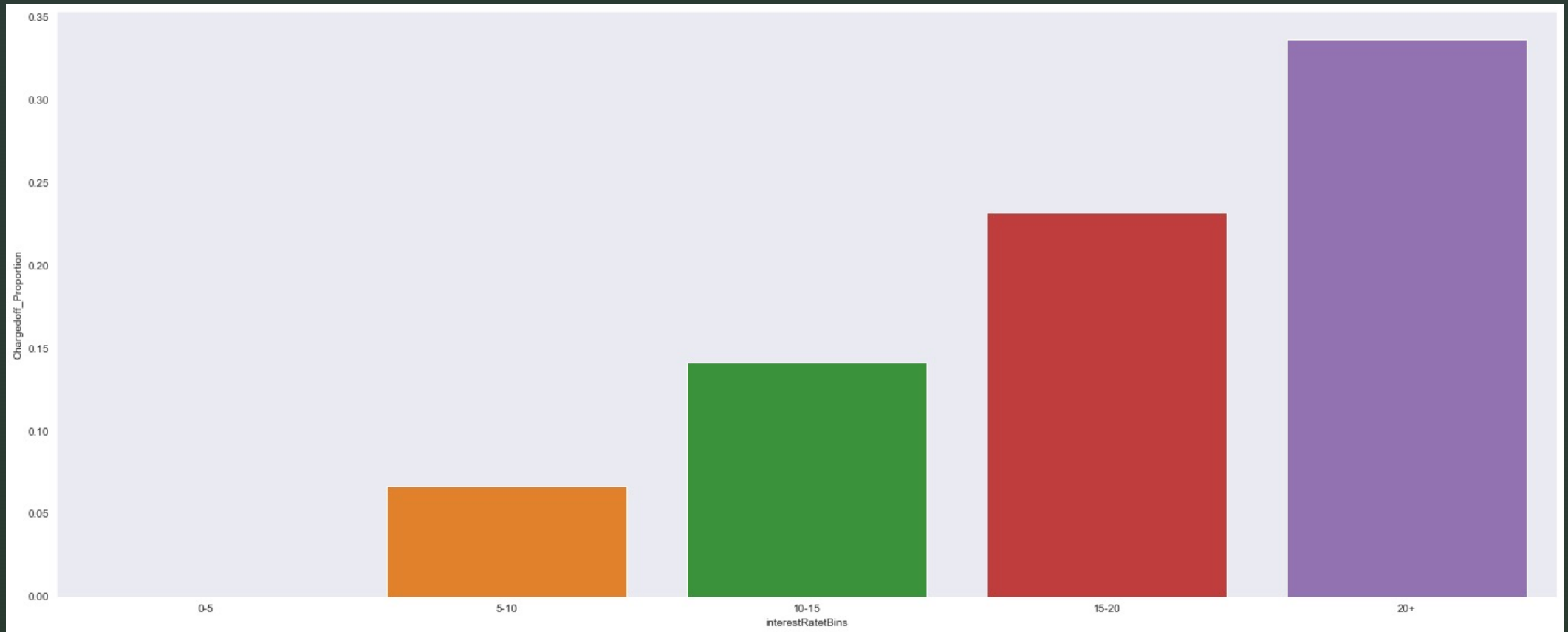
Observation: Most of the loans are taken with less duration and charged off portion is high with less duration loans



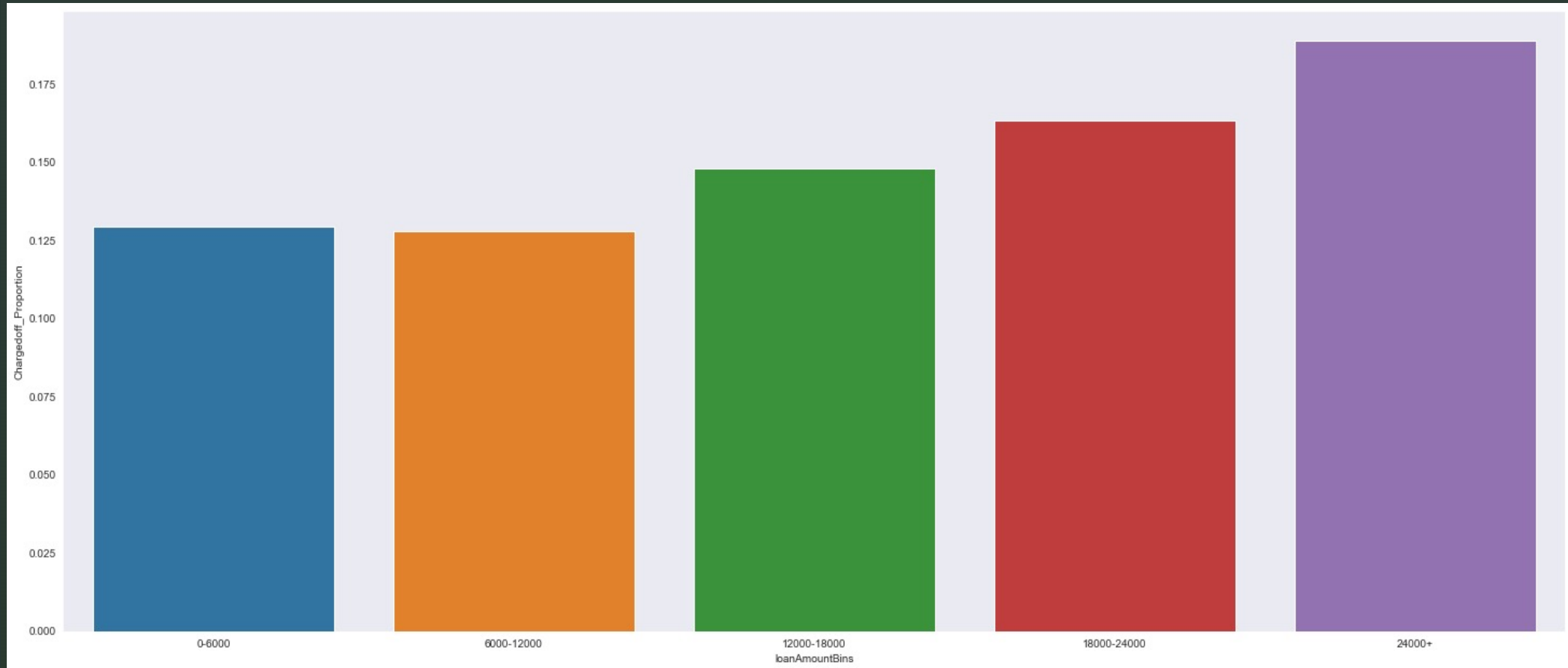
Observation: The charged-off part is significantly higher in low-income households. As revenue grows, the charged-off share decreases.



Observation: Charged off portion is increasing with high interest loans



Observation: Charged off portion is increasing as the loan amount increased.





## Recommendations

- Avoid high-interest loans, particularly those with a rate greater than 15%.
- Avoid lending to individuals who have previously defaulted.
- Provide less loans to persons who are unemployed or have been unemployed for less than a year.
- Avoid providing high-grade loans at a high interest rate to individuals seeking financing for small businesses.
- Provide less loans to persons who are staying in the rented house