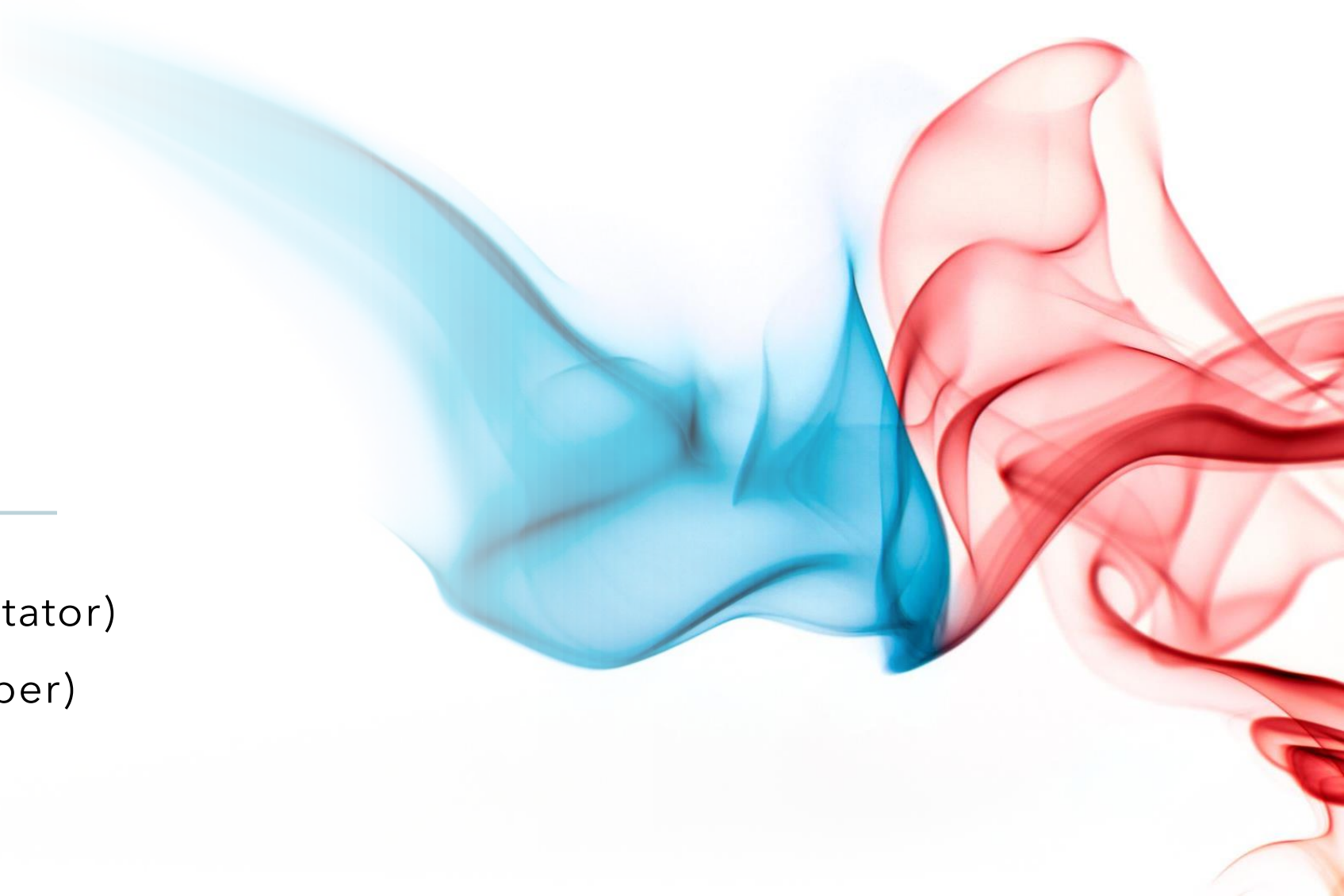




Loan application case study

Sumanta Basak (Group facilitator)

Pradeep Kumar (Team member)



Business Objective & Background

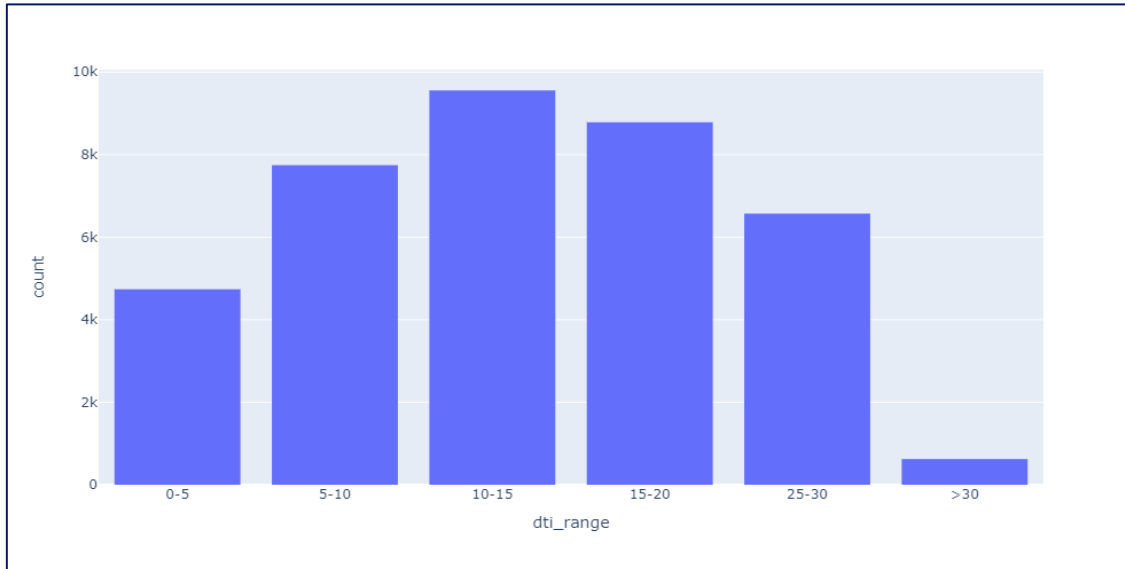
Objective - The objective is to identify predictors of default so that at the time of loan application, we can use those variables for approval/rejection of the loan

Background - Lending club is a marketplace where investors are willing to lend money and borrowers are willing to take loan. Now, identifying the borrowers is key to business and the drivers of them as otherwise business might face a loss, in case the borrower doesn't pay the loan.

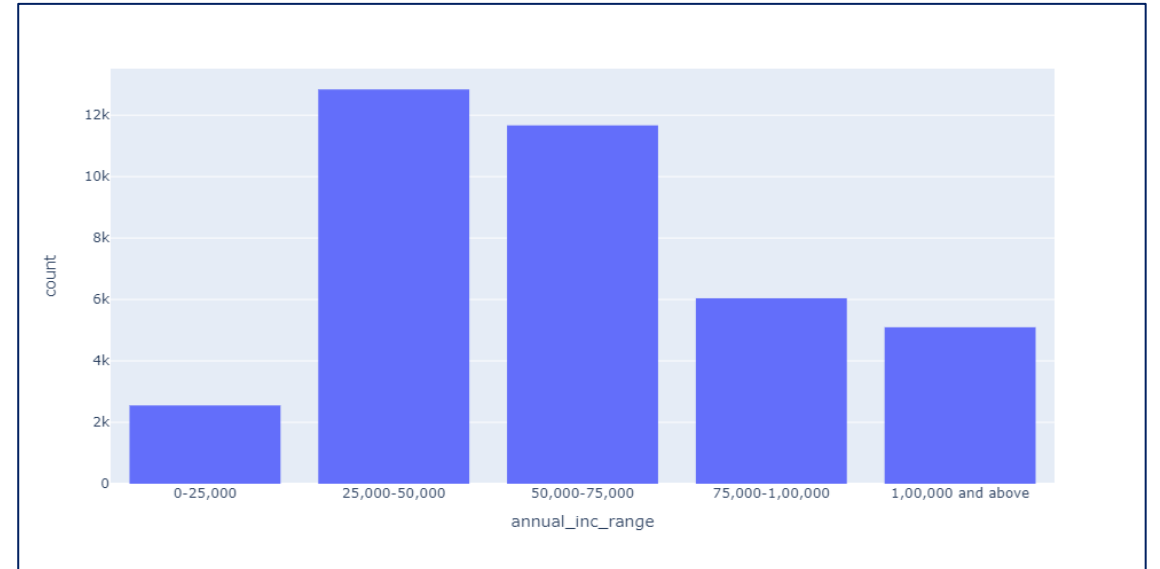
Approach of analysis

- Data understanding - The data comprises of 39,717 observations and 111 features. We're trying to understand the feature 'Loan Status' and drivers of this variable and interaction with other features, mainly to understand what may cause 'charge-off'. Loan Status is either of 'Fully paid' or 'Charge-off'.
- Data cleaning - The features with 25% or more missing values were removed. Out of 111 features, 58 were removed keeping 53 in data. Further few more, which may not be of any use (As they are mainly not available at customer level when they are borrowing). Now the features is further reduced to 28.
- Approach - With the help of univariate, bi-variate & multi-variate analysis, we tried to see what can help in explaining the driver of paying off loans.

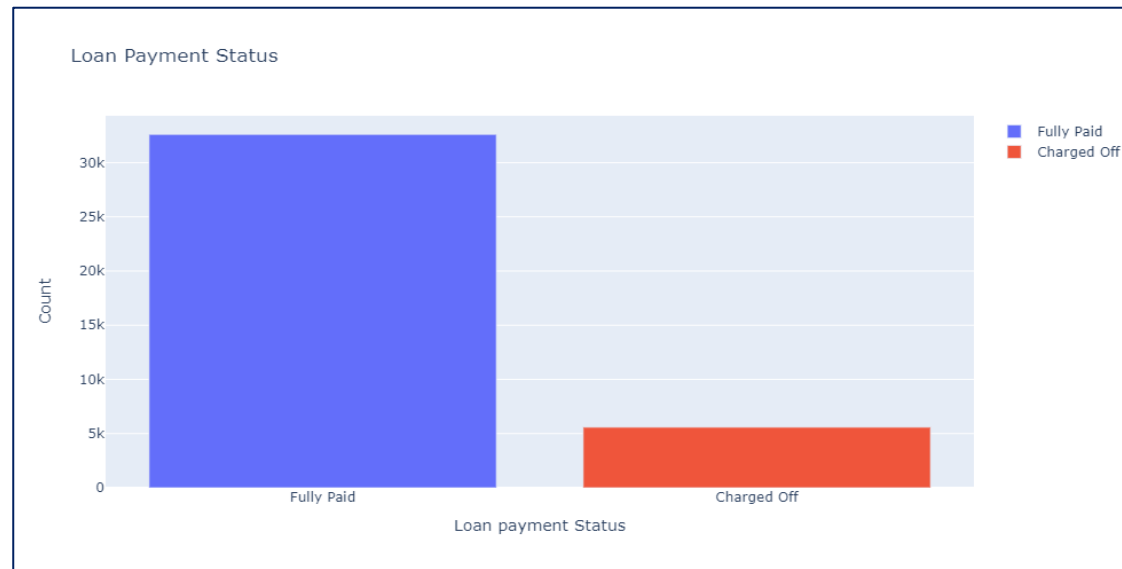
Distribution of features



Most loans are disbursed who has Debt-to-Income ratio in the range 10-15

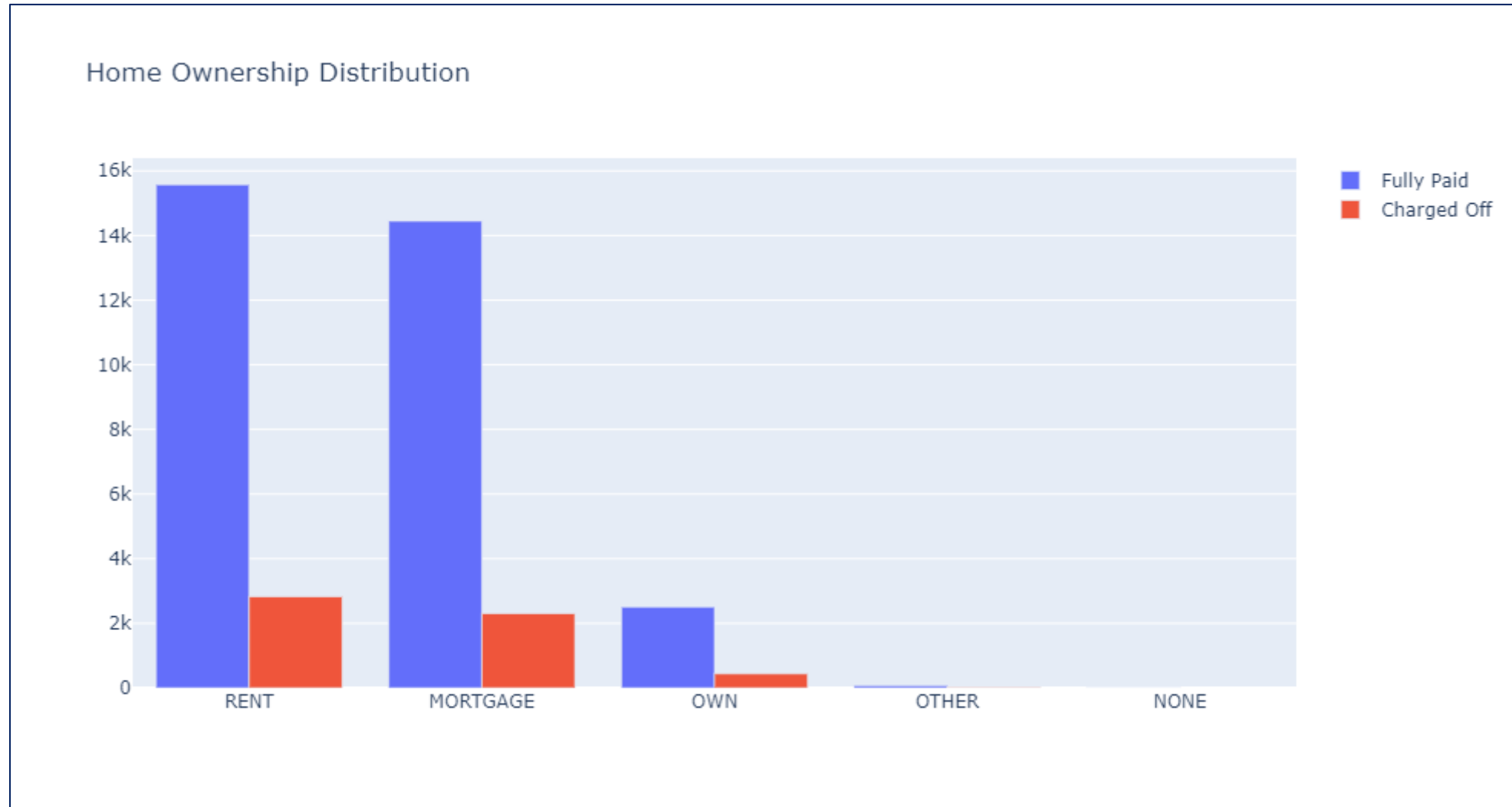


With annual income in range 25K-50K, has highest loans disbursed



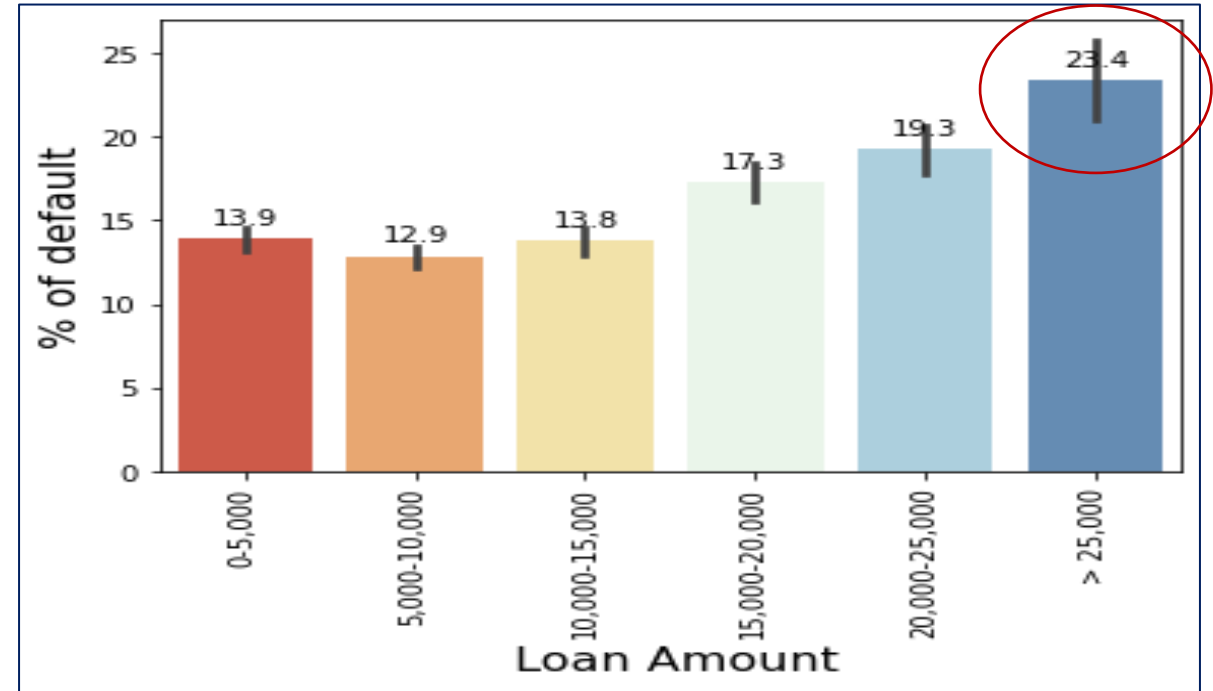
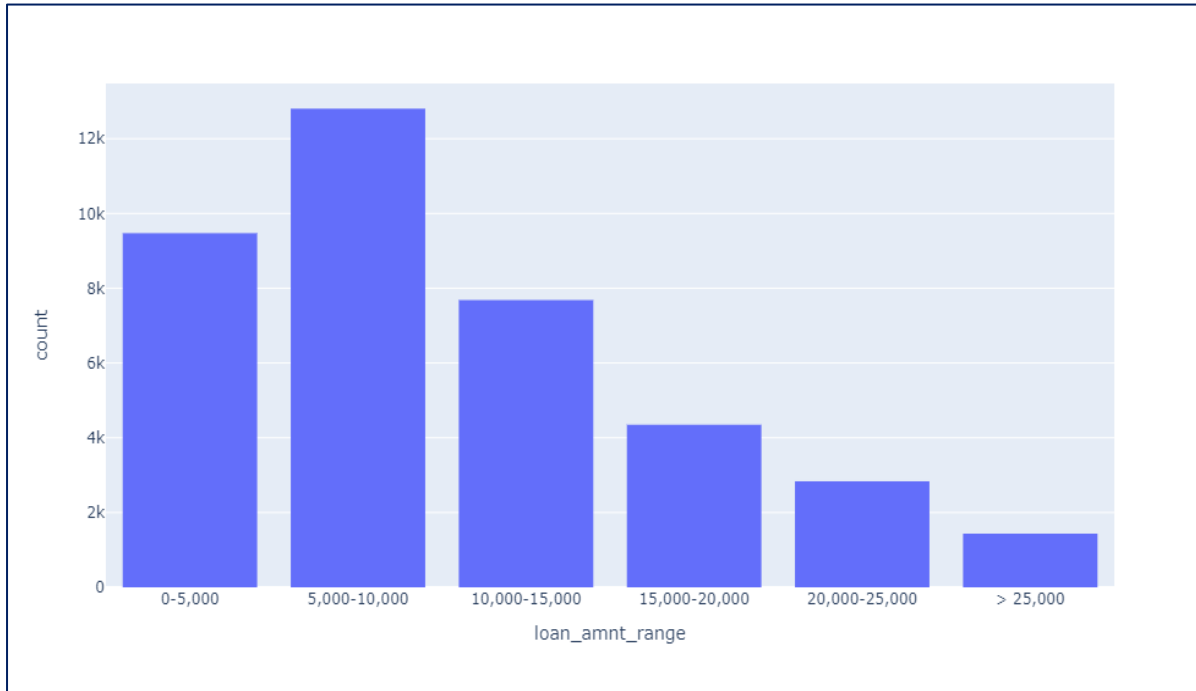
Around 15% are loans charge-off

Distribution of loan applicants by home ownership



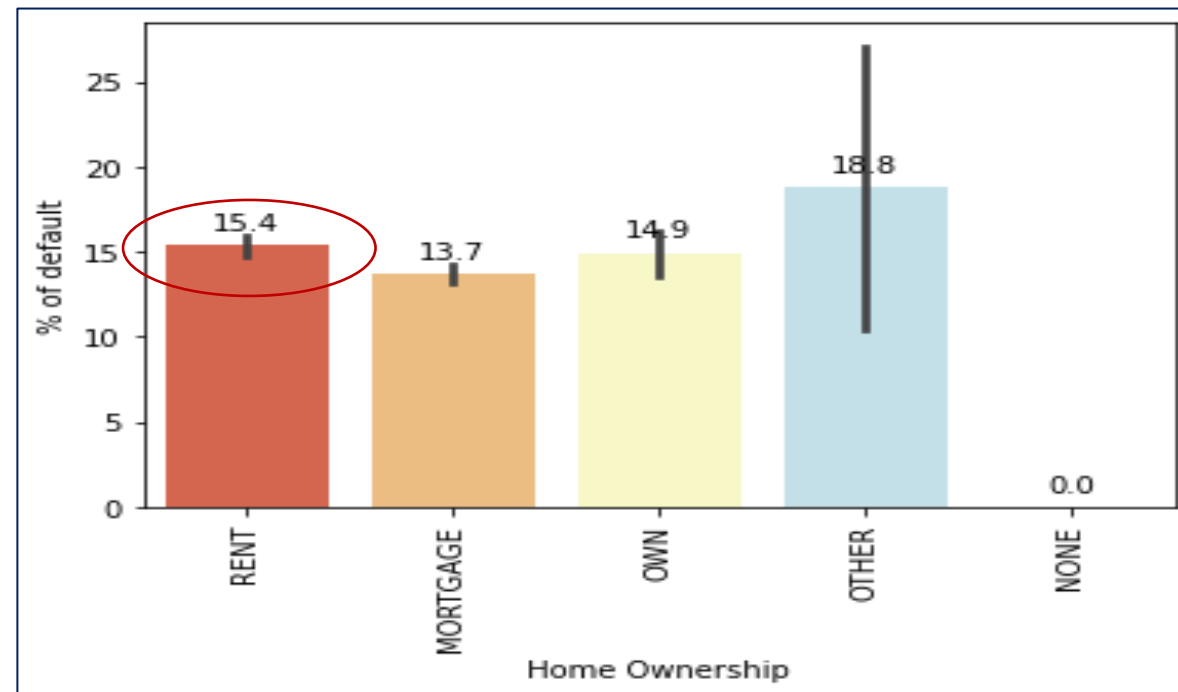
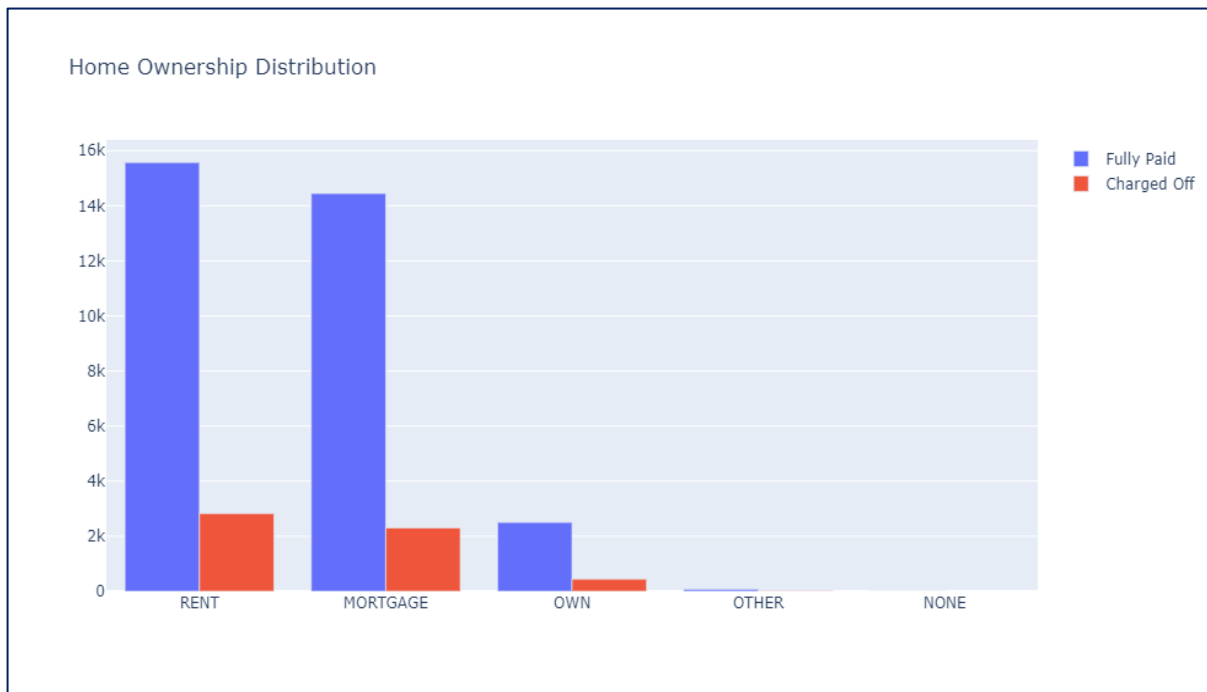
Those who are in rented home, has the highest chance of getting charged off, the owners have the least chance

Loan amount distribution & effect on charge off



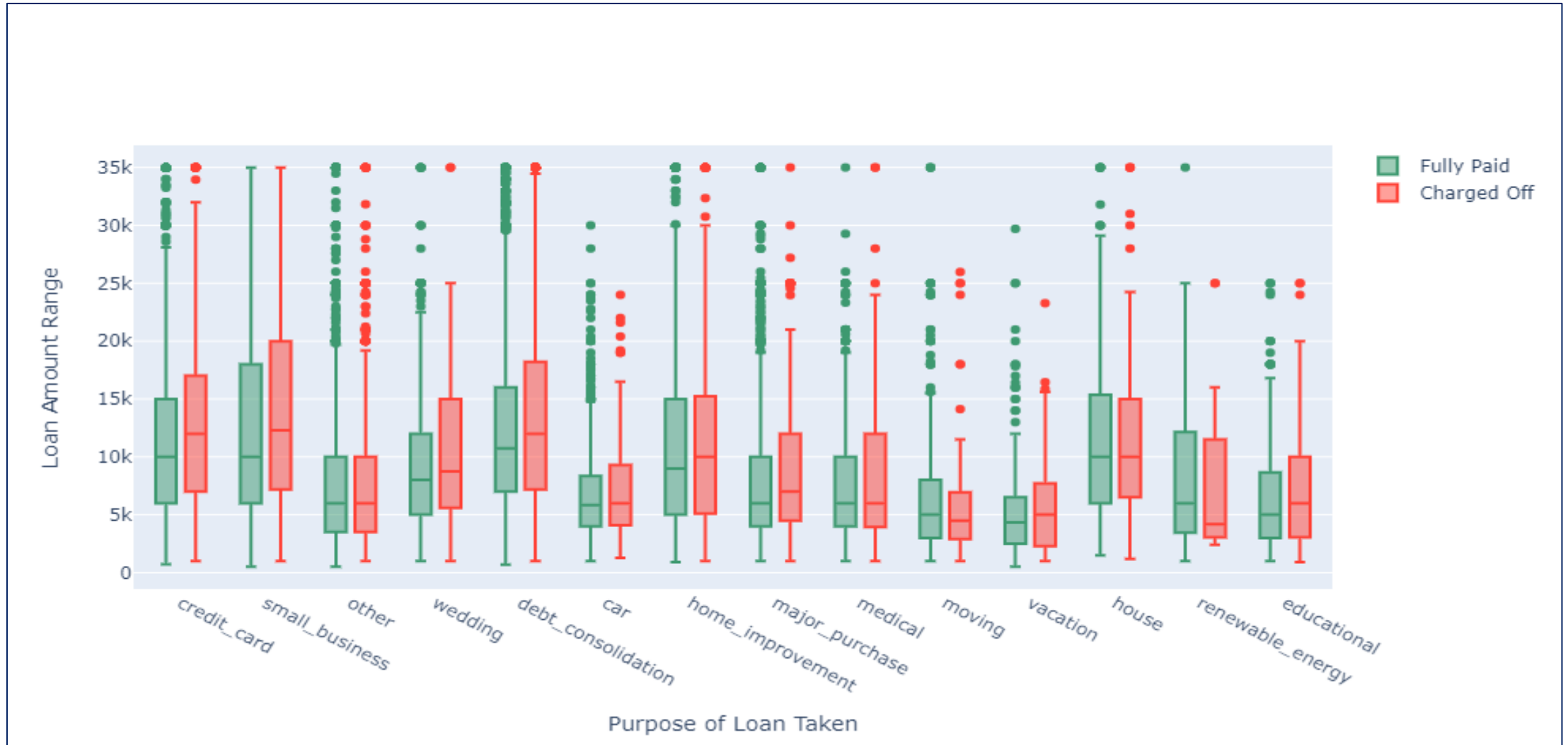
Most of the loans are disbursed in 5K – 10K range, >25K being the highest category, but big-ticket size loans prove to be quite risky

Home Ownership distribution & effect on charge off



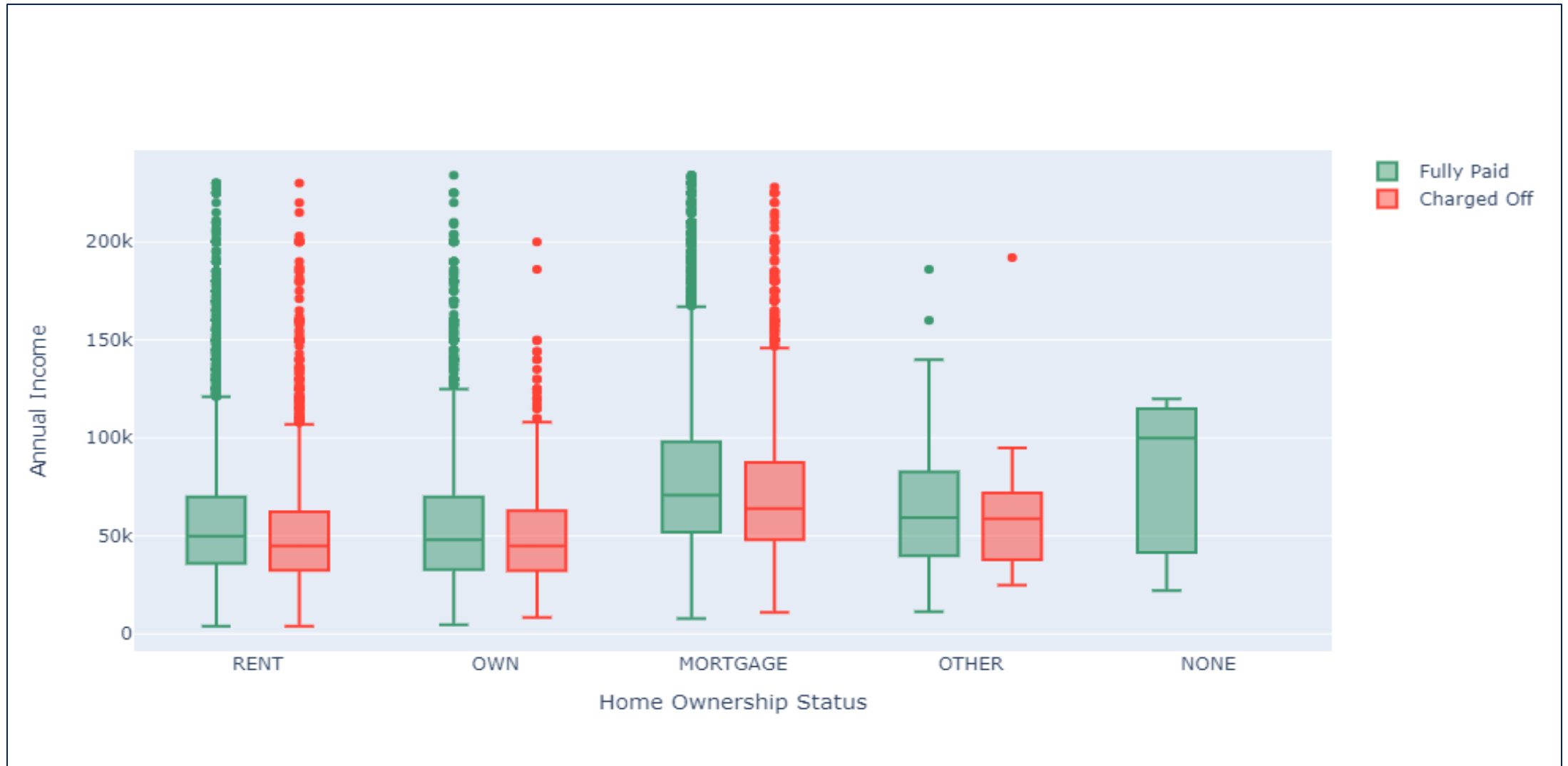
Most of the loans are disbursed in for Rent & Mortgage home ownership, but borrowers who take rent has highest chance of getting charged off

Relation between Loan amount taken and Loan purpose & effect on charge-off



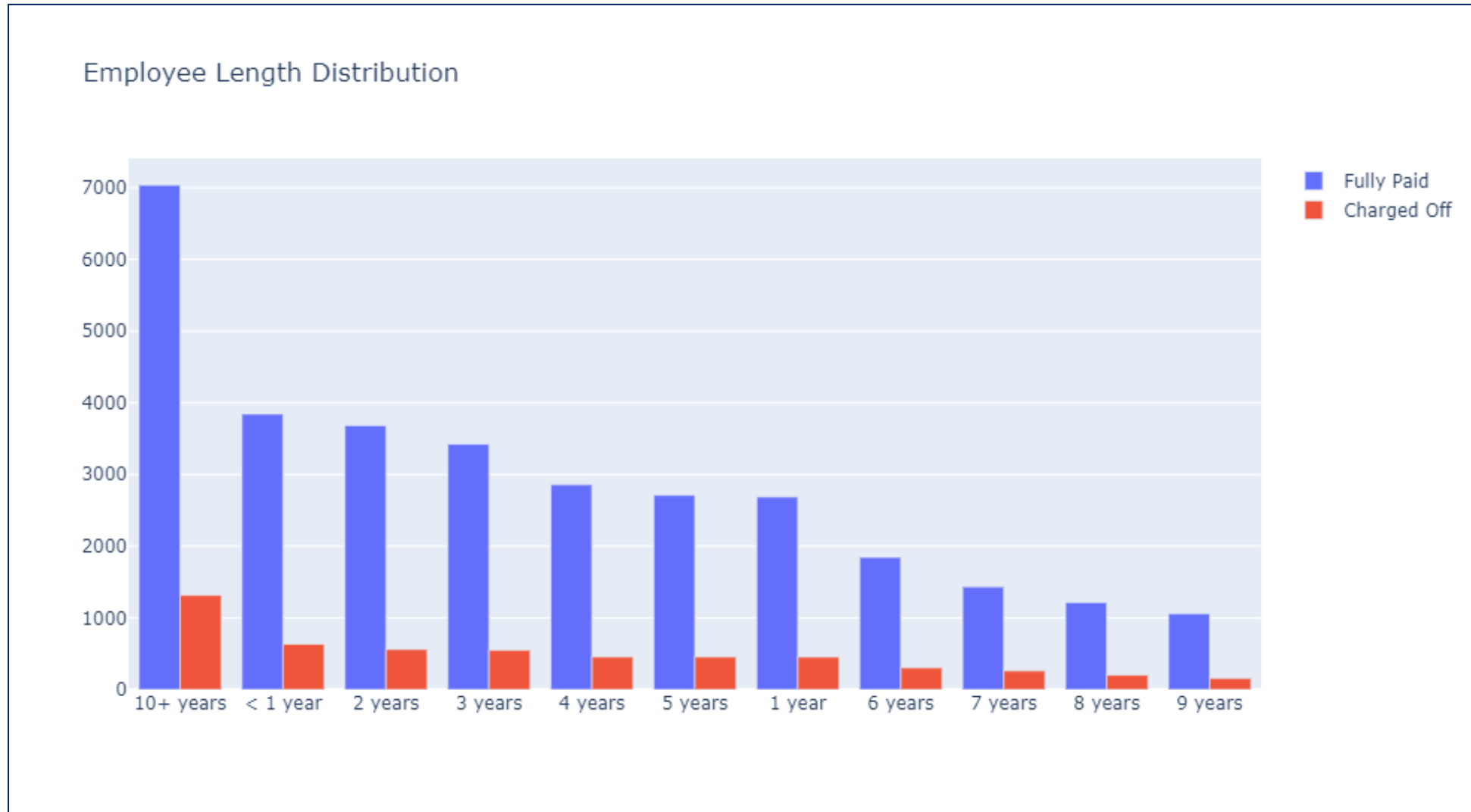
Loans taken for small business (as high as 20K) and debt consolidation (as high as 18K) are highly likely to get charge off

Relation between Annual Income and Home Ownership & effect on charge-off



Home with Mortgage with high median income as 70K are most likely to get charged off

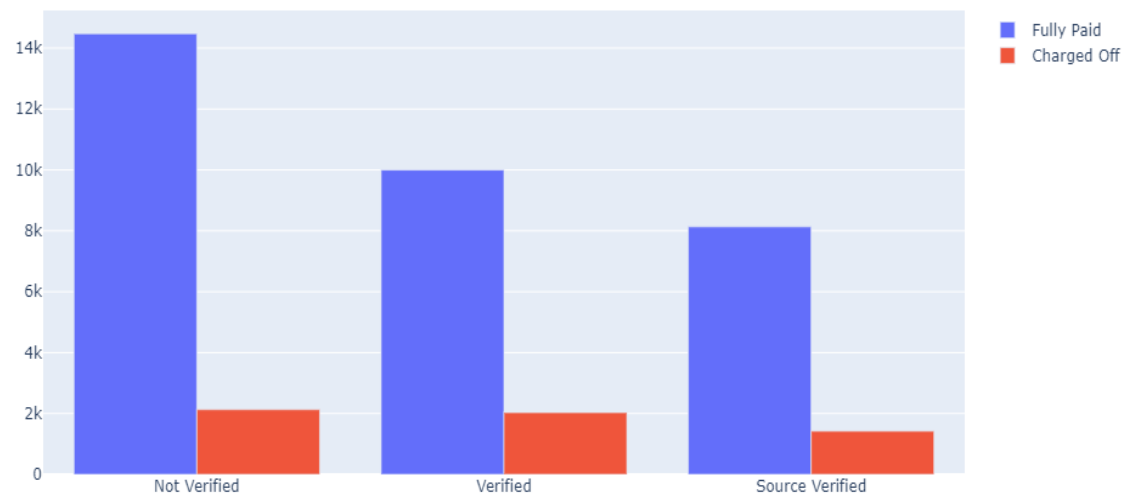
Relation between Employment length and effect on charge-off



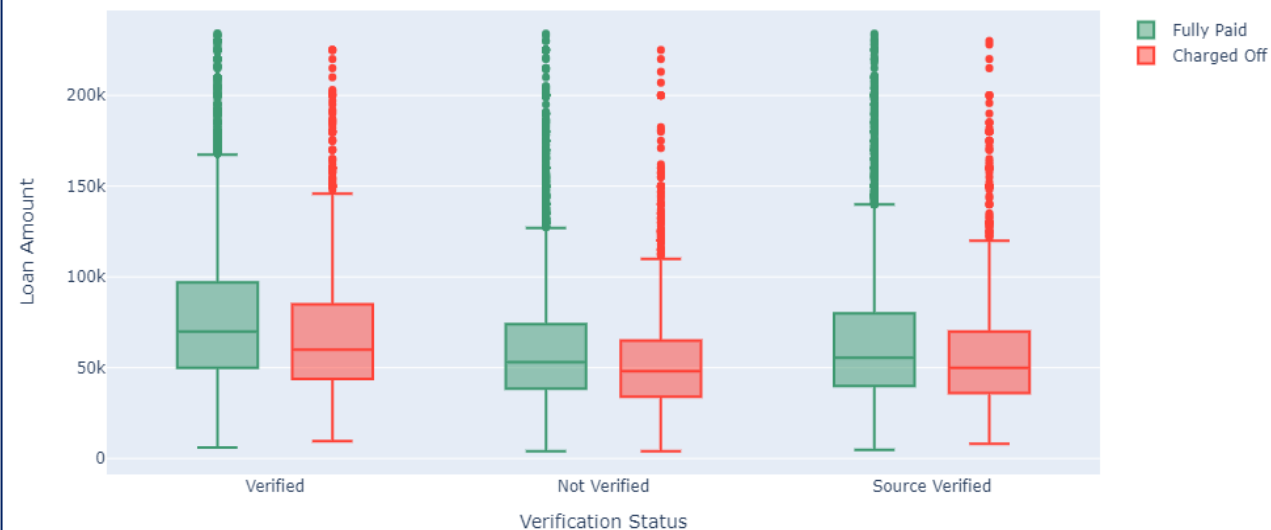
Employees with more than 10 years of employment most likely to get charged off

Relation between source verification and effect on charge-off

Income Verification Status



In terms of charge-off, source verified or not, does not make a big difference



Those who has income source verified, generally get a better loan amount disbursed. Loan amount in 'verified' category is highest to be charged off

THANK YOU