

Lending Club Case Study

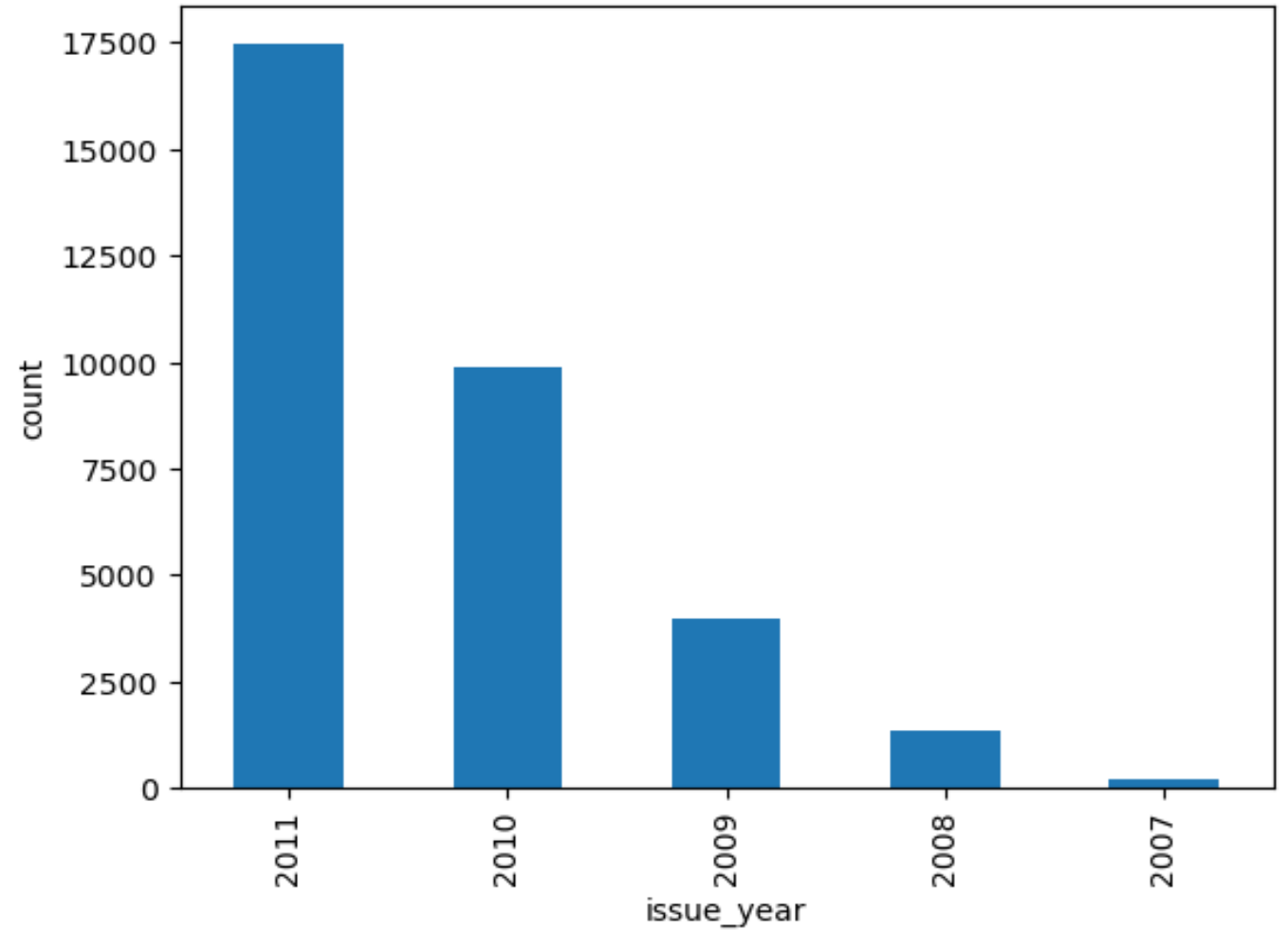
Data Cleaning

1. Drop all the columns with NULL values.
2. Drop all the rows, with all the NULL column values (None found in our case).
3. Remove the columns, which contains information unrelated to the approval of loan.
4. Find if any duplicates are present (None in our case).
5. Find the columns with a unique value in it and remove them.
6. Finally imputing the missing values, standardizing the dates and other columns.
7. Finally only filtering the rows, with loan is not current.
8. Outlier Treatment.

Derived Column Creation

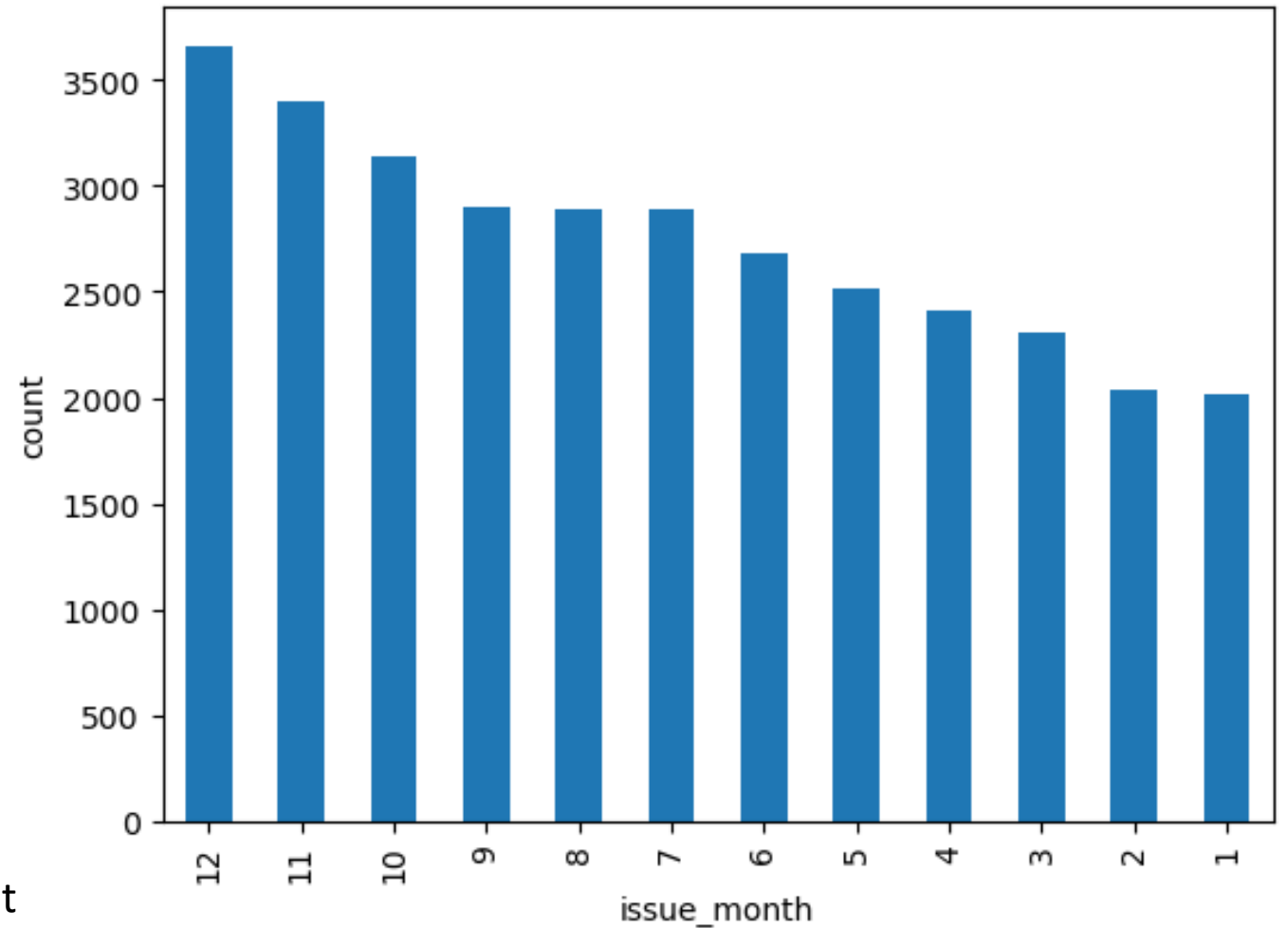
1. Based on the employment tenure we have created their categories
 1. Junior (0 - 3)
 2. Senior (3 - 7)
 3. Specialist (above 7)
2. Based on the annual income
 1. Low
 2. Moderate
 3. High
 4. Very High

Univariate Analysis



This clearly depicts that the borrowers are increasing over each year.

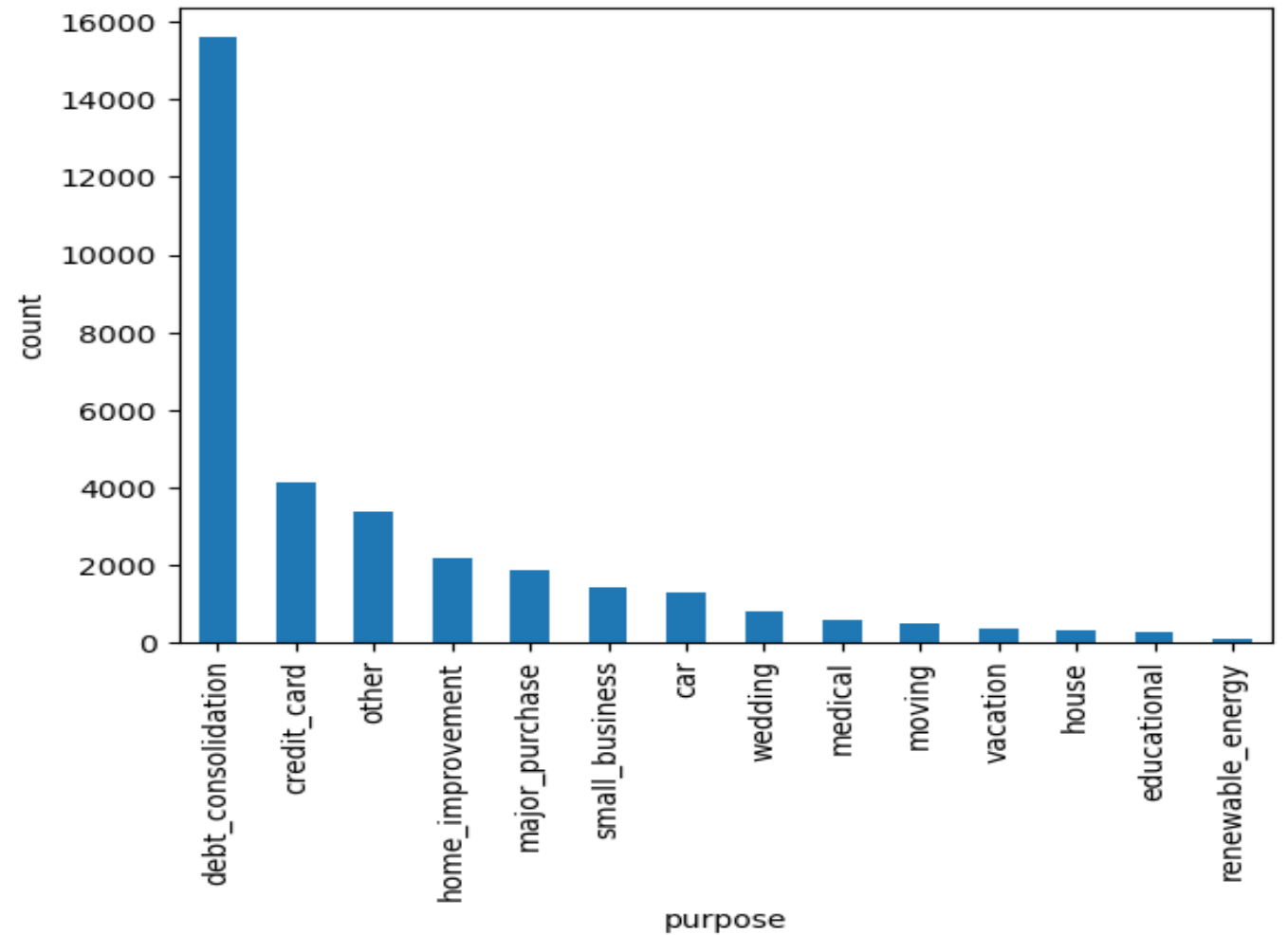
Univariate Analysis



The number of borrowers tend to take loan more at the end of the year, maybe due to

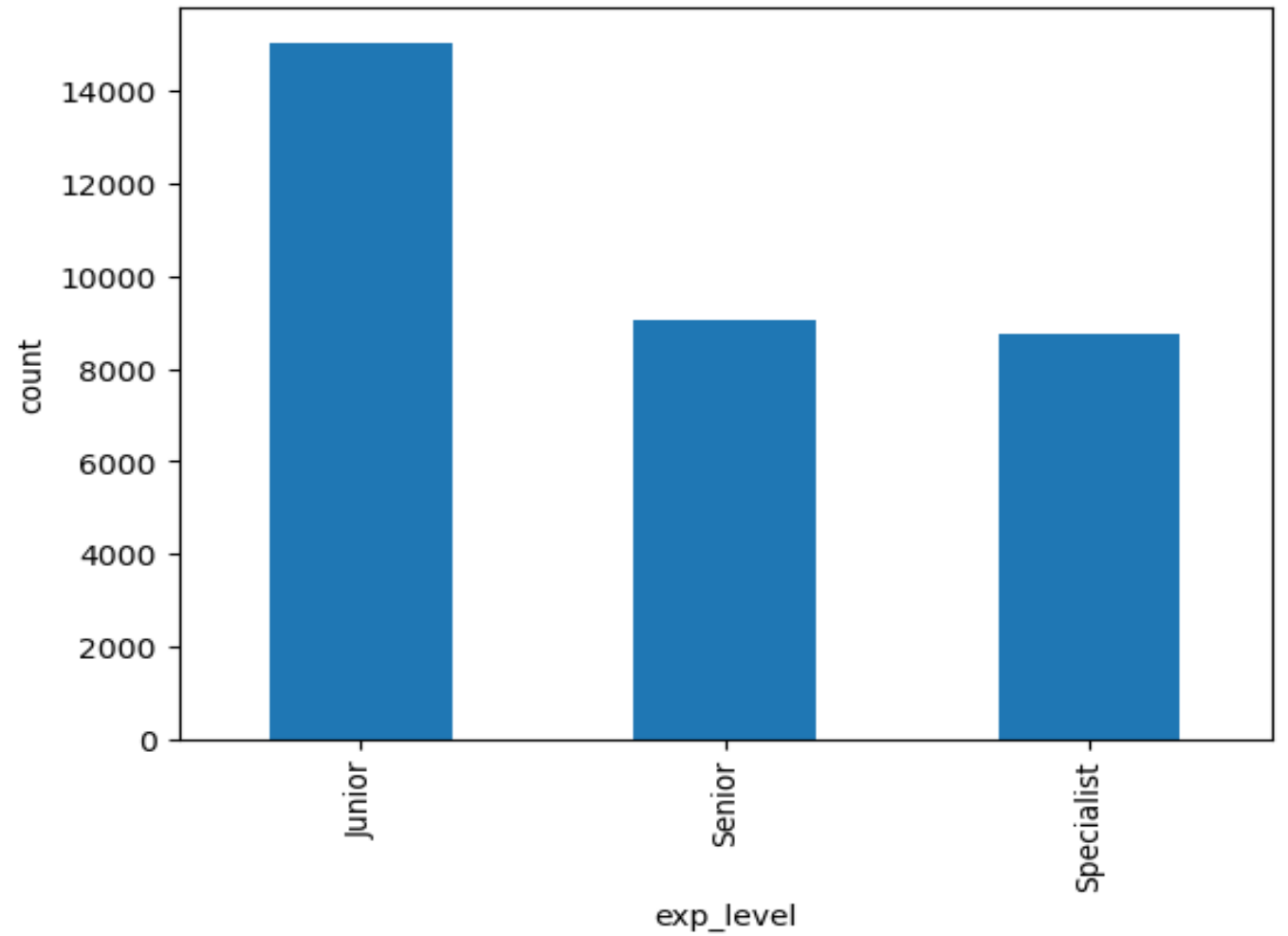
- 1.Number of festivals mainly occurs at the end of the year.
- 2.There are a lot of year end sales at the end of the year, hence people need a lot of money.

Univariate Analysis



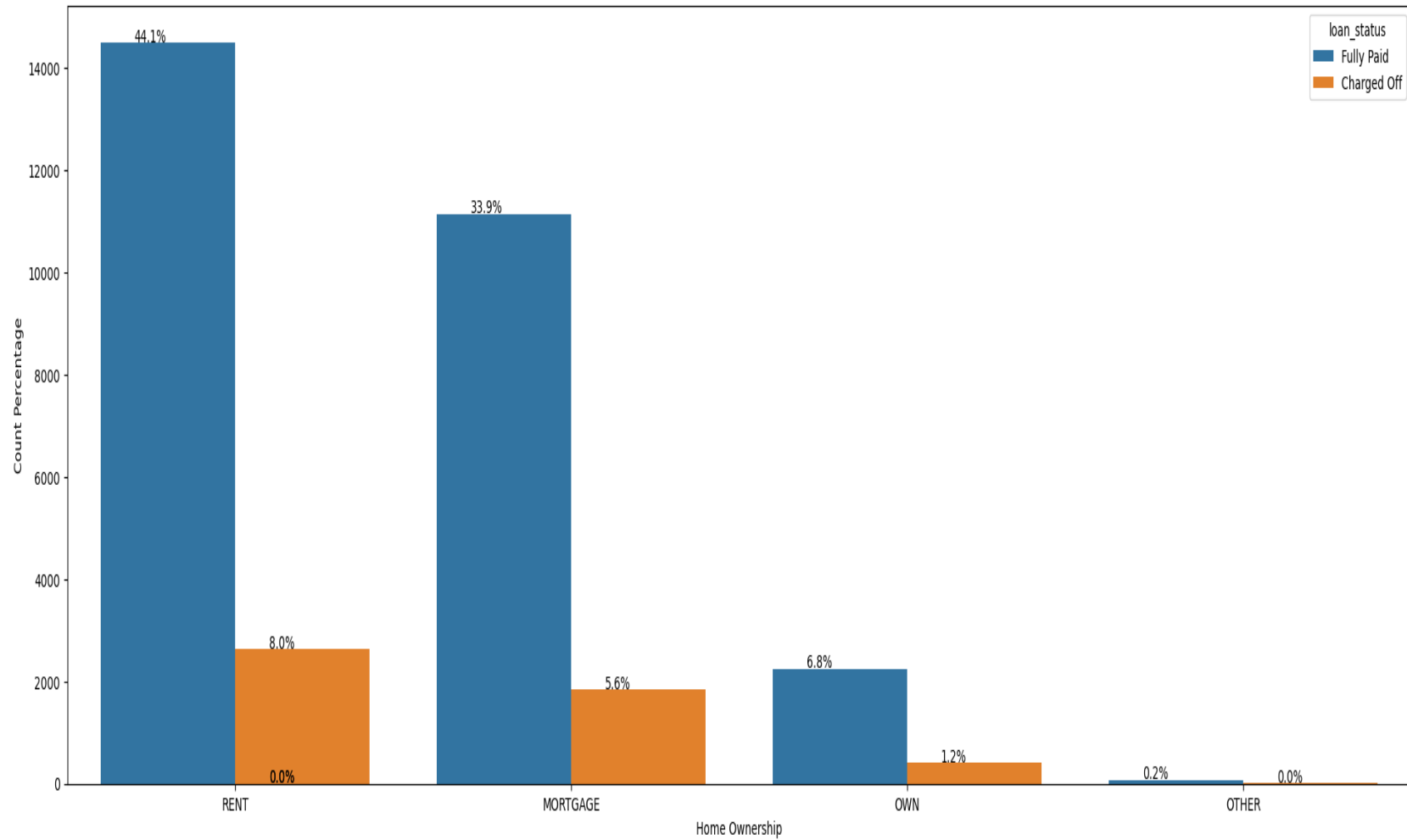
It seems that people who take new loans mainly to pay off the old loans is very high in volume

Univariate Analysis



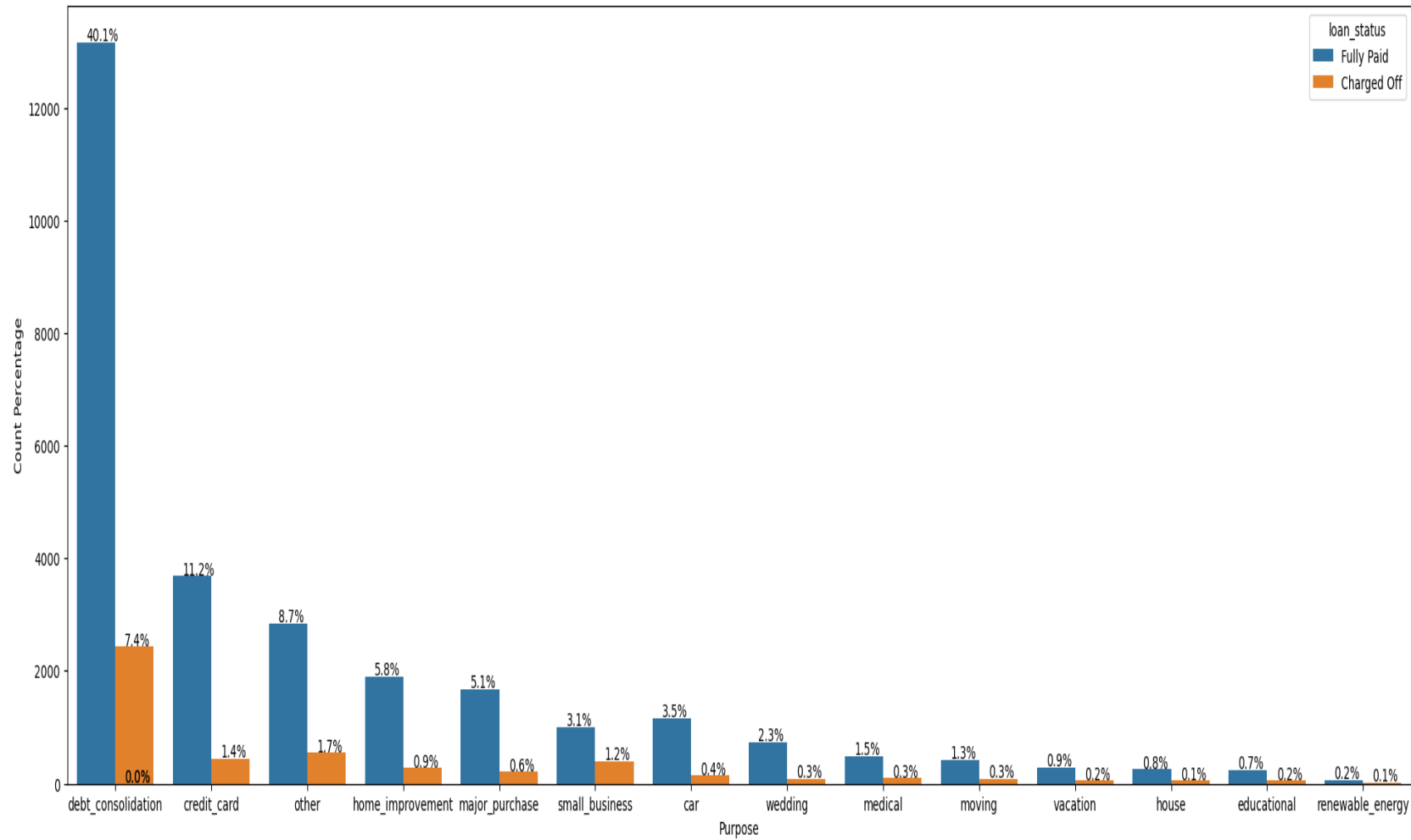
It seems Junior employees tend to take loan more.

Bivariate Analysis



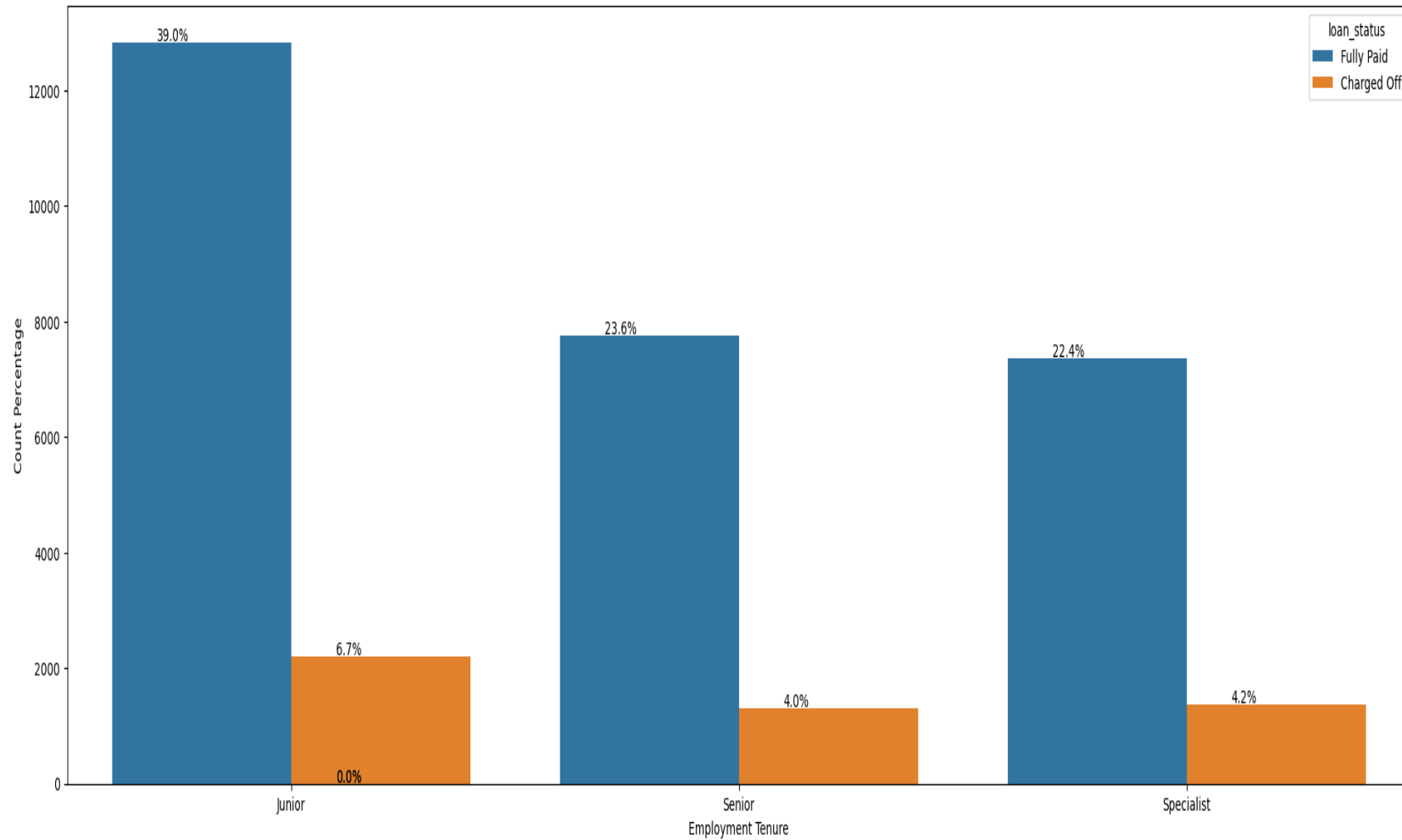
It seems, that Mortgage and rent housing type borrowers are riskiest.

Bivariate Analysis



It seems that debt consolidation is riskiest of all with very high volume of charged off borrowers.

Bivariate Analysis



It seems that the junior employees are the riskiest as their volume is much high for charged off accounts. On the other hand, Senior employees are comparatively better as their charged off percentage is less.

Heatmap Visualization

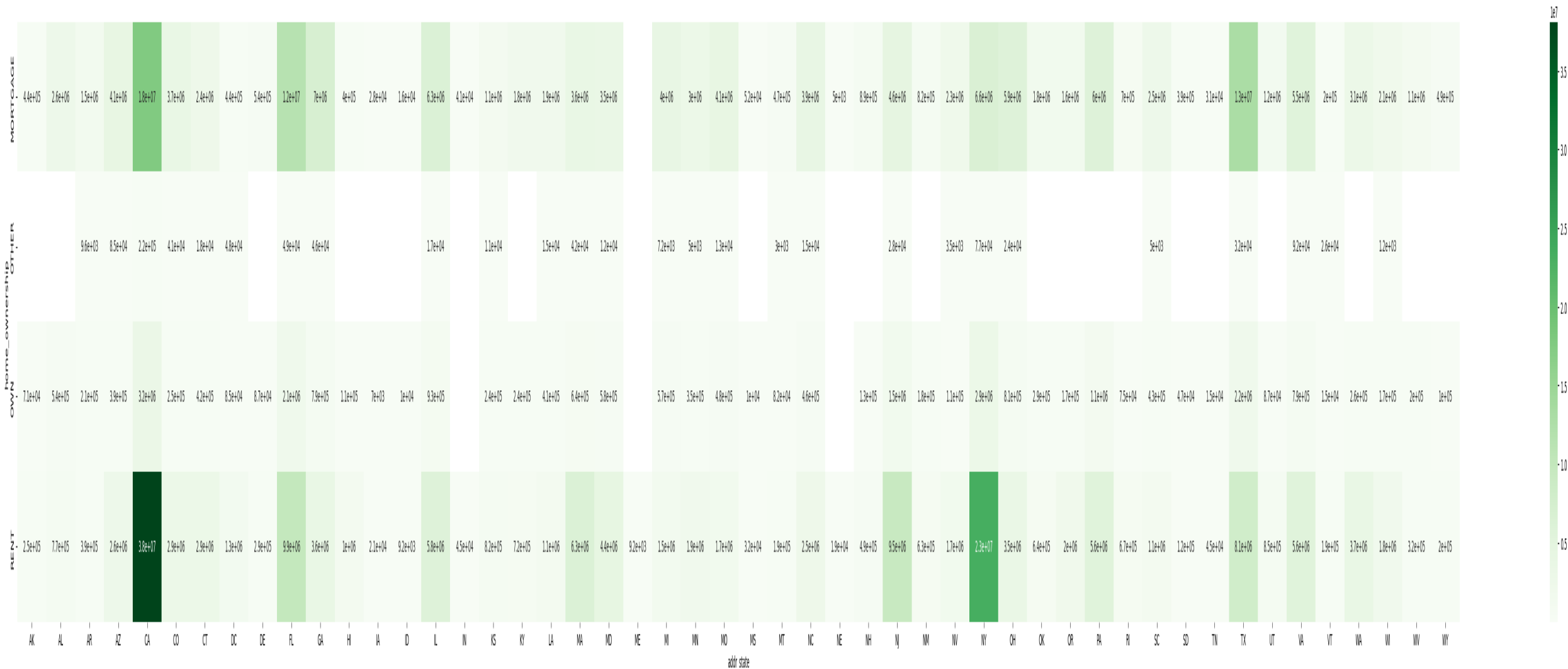


Fig 1

Heatmap Visualization

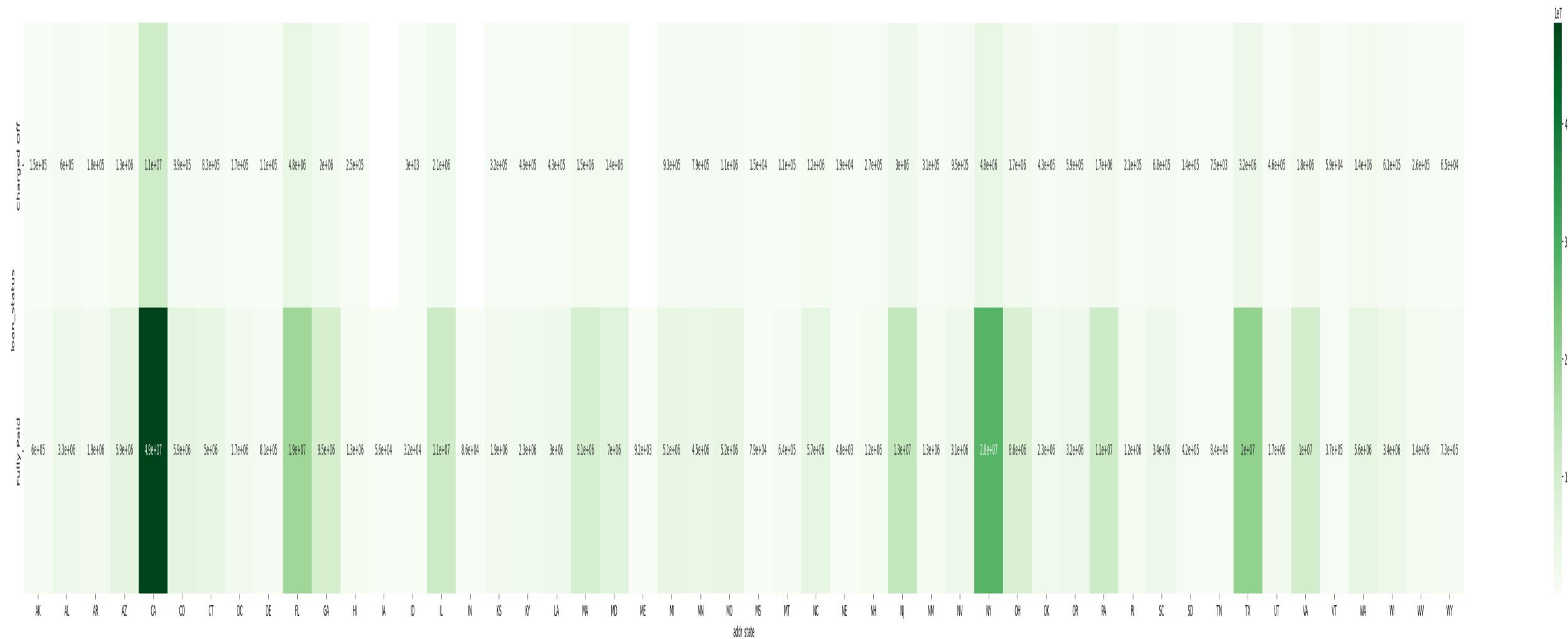


Fig 2

Analysis of the heatmap

- As seen from the previous Fig 1 and Fig 2,
- It seems that the borrowers living in RENT facility in CA state takes the highest amount of loan.
- They have highest fully paid accounts, and
- Also are having highest charged off accounts as well among any other state.