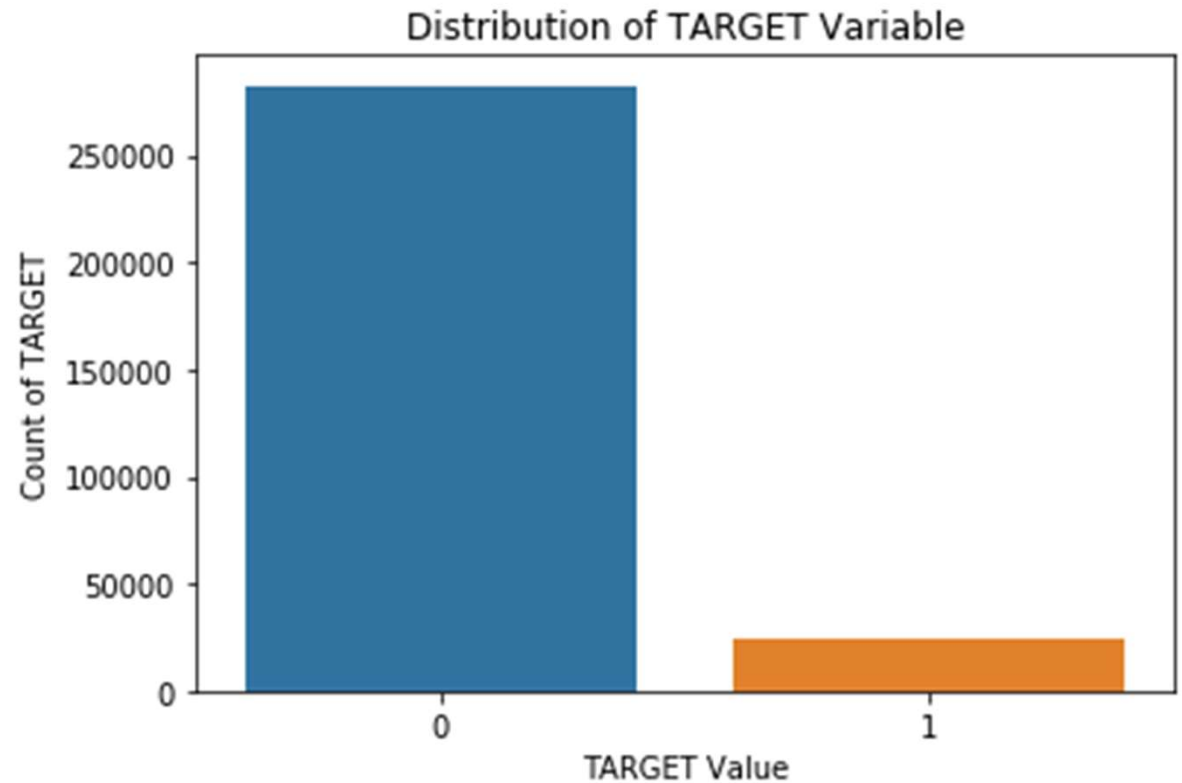
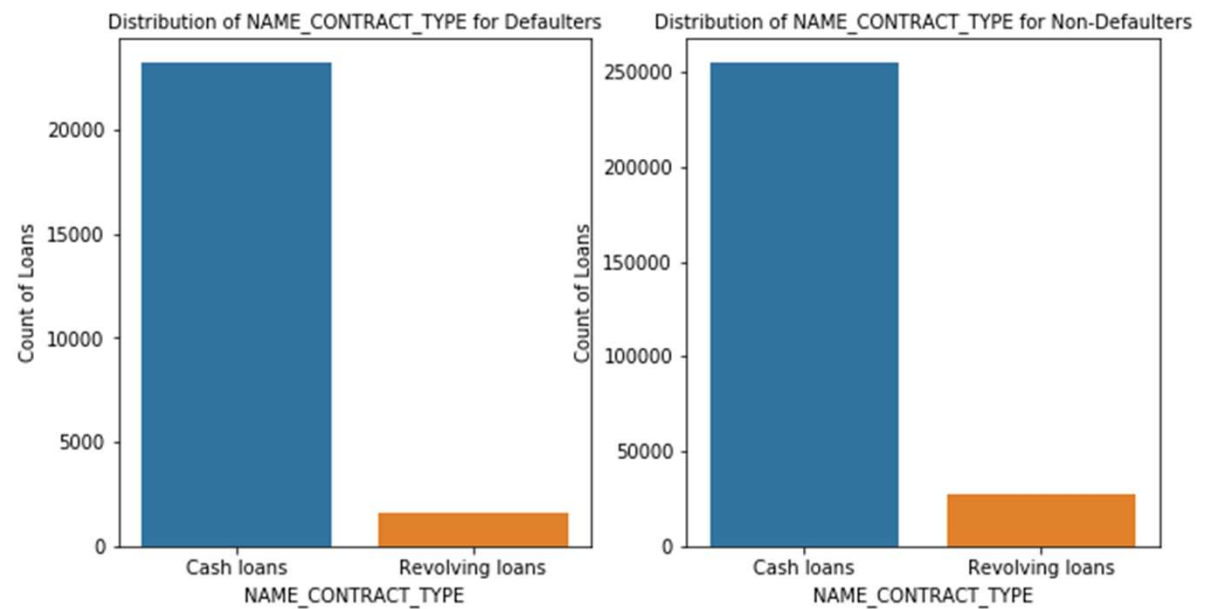


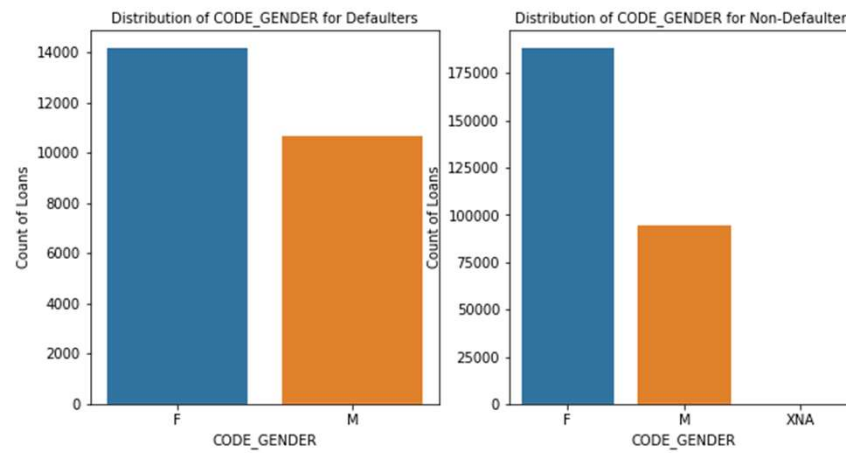
# Credit EDA - Case Study

- We see this is an imbalanced dataset.
- There are far more loans that were repaid on time than loans that were not repaid.
- More than 25000 loans were repaid, Less than 5000 loans were not repaid.

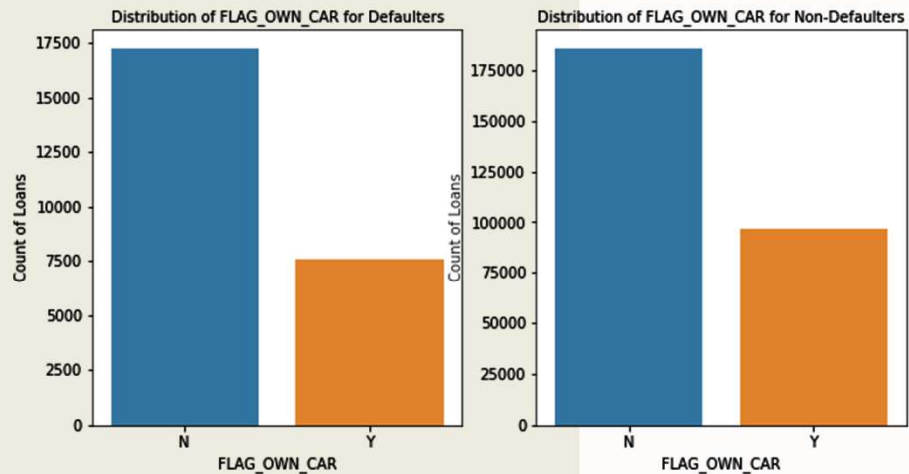


- 
- We observe that the number of **Cash loans** is much higher than the number of **Revolving loans** for both Target = 0 and Target = 1





- We observe that the number of **Females** taking loans is much higher than the number of **Males** for both Target = 0 and Target = 1

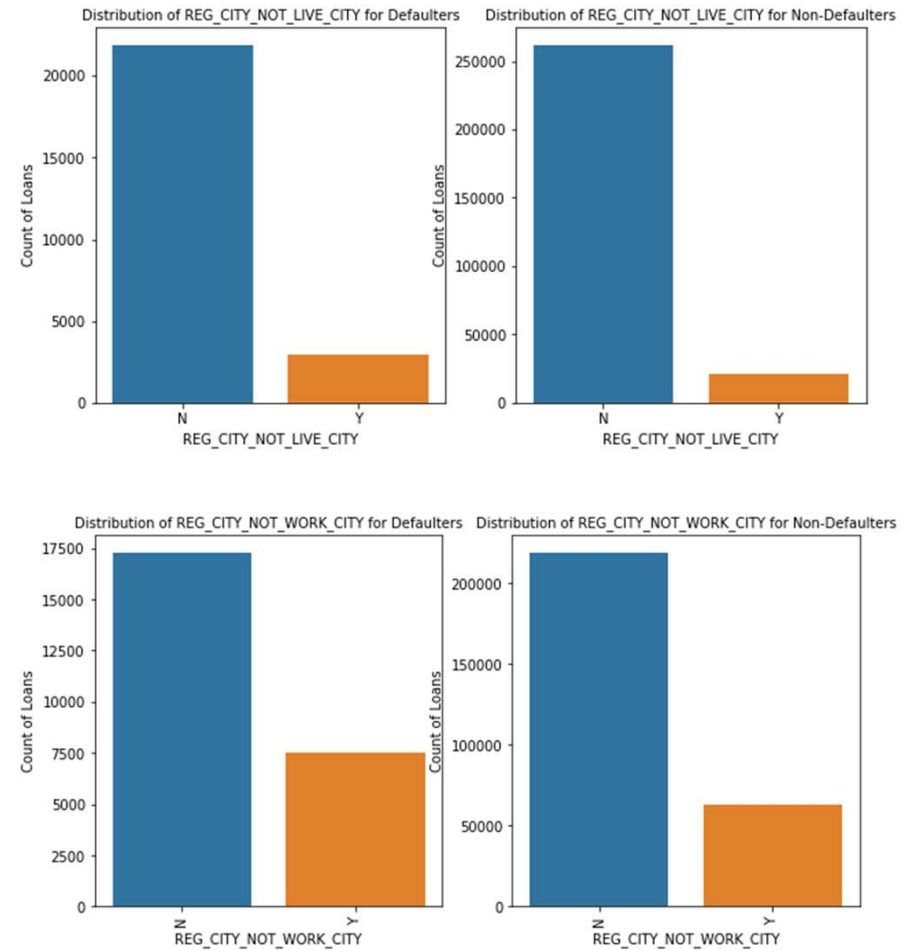


We observe that the number of most people applying for loan **do not own a car**.



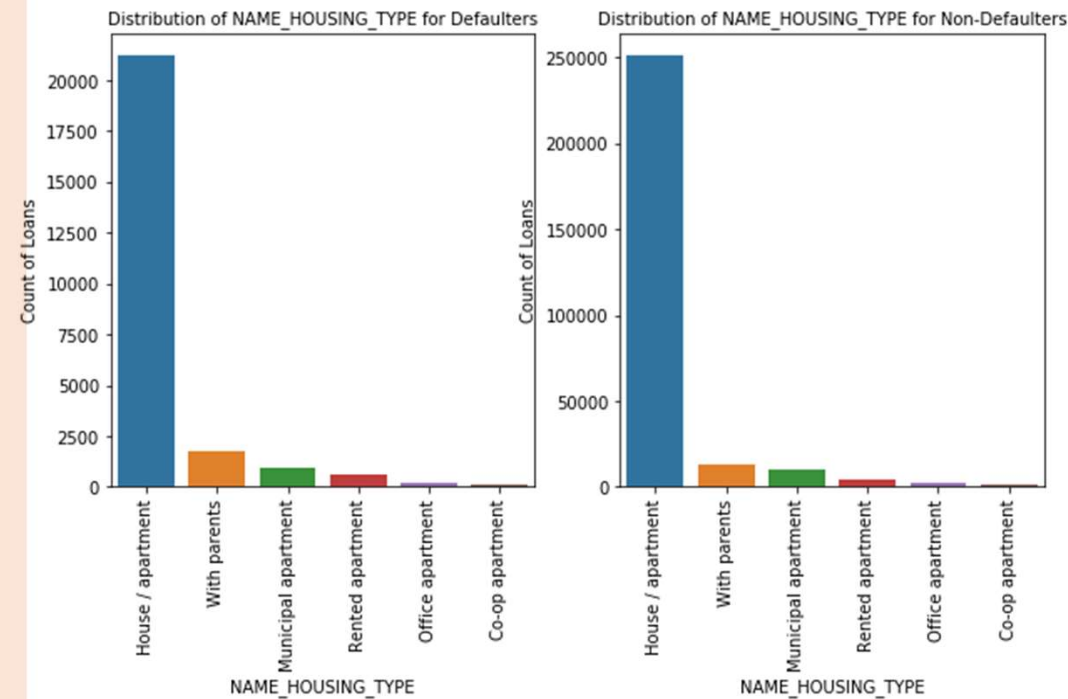
We also observe that the **ratio of people who own a car is higher for non-defaulters**

- We observe that the Ratio of people whose Registration City is not the same as live city or work city is higher in case of defaulters are compared to defaulters.
- It tells us that people who live or work in a city different than the registration city are more likely to have payment difficulties.

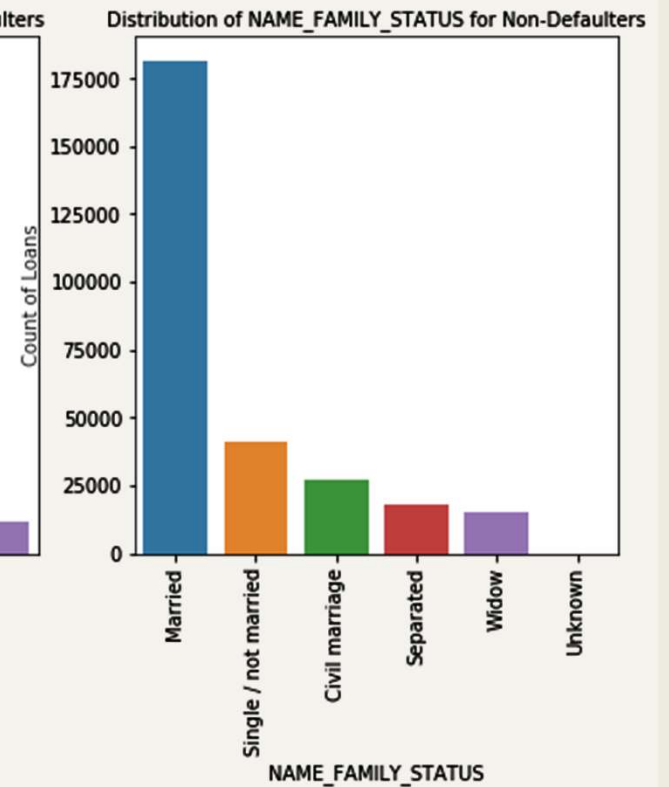
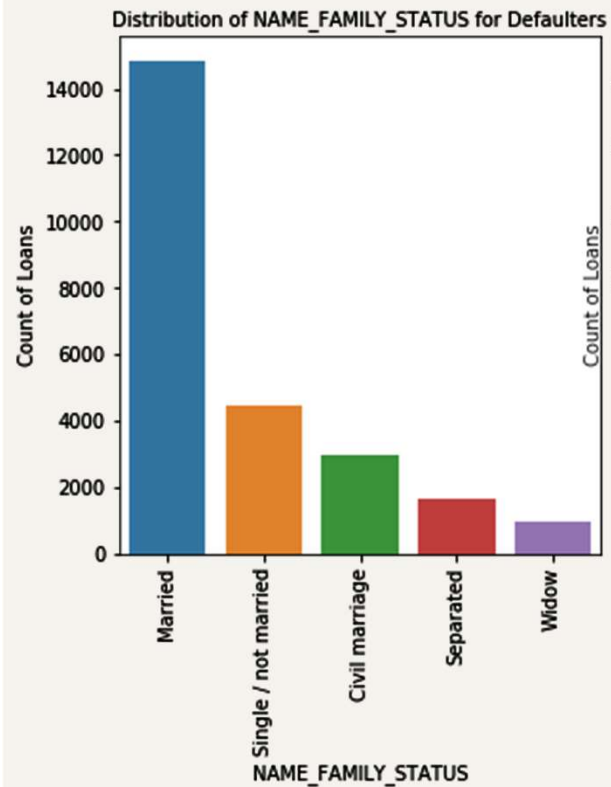


## Observation:

- Most people live in a House/Apartment
- Ratio of People who live **With Parents** is more for defaulter than non-defaulters. It tells us that applicant who live with parents have a higher chance of having payment difficulties.

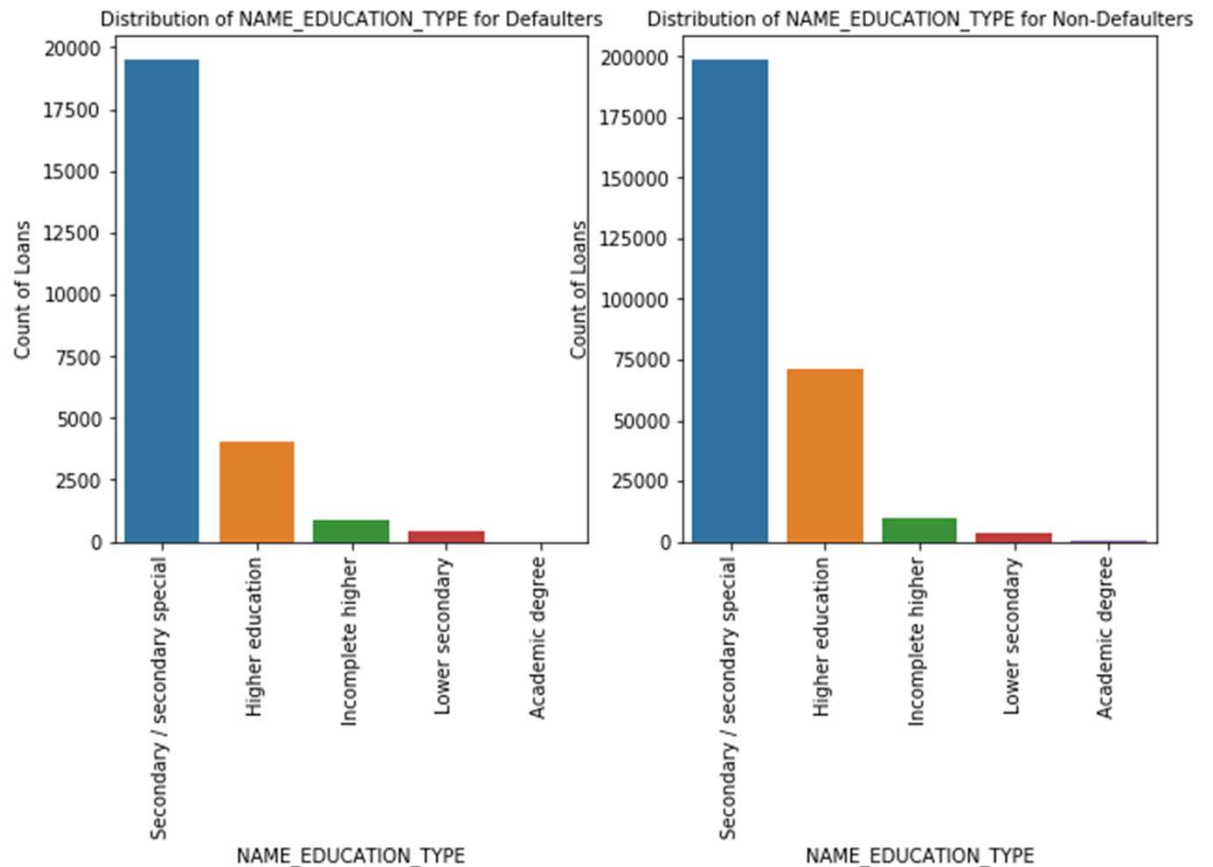


- Ratio of **Single/Unmarried** people is more in the left graph.
- Single/Unmarried people are more likely to have payment difficulties

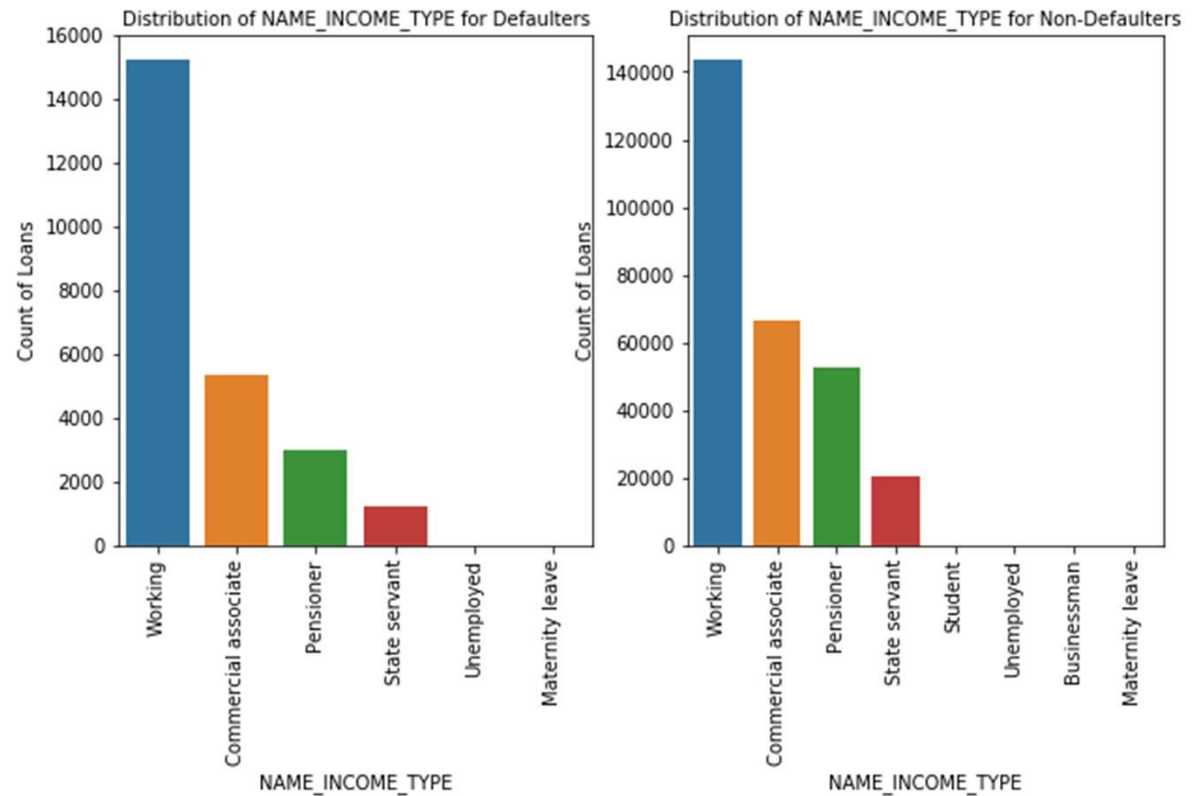




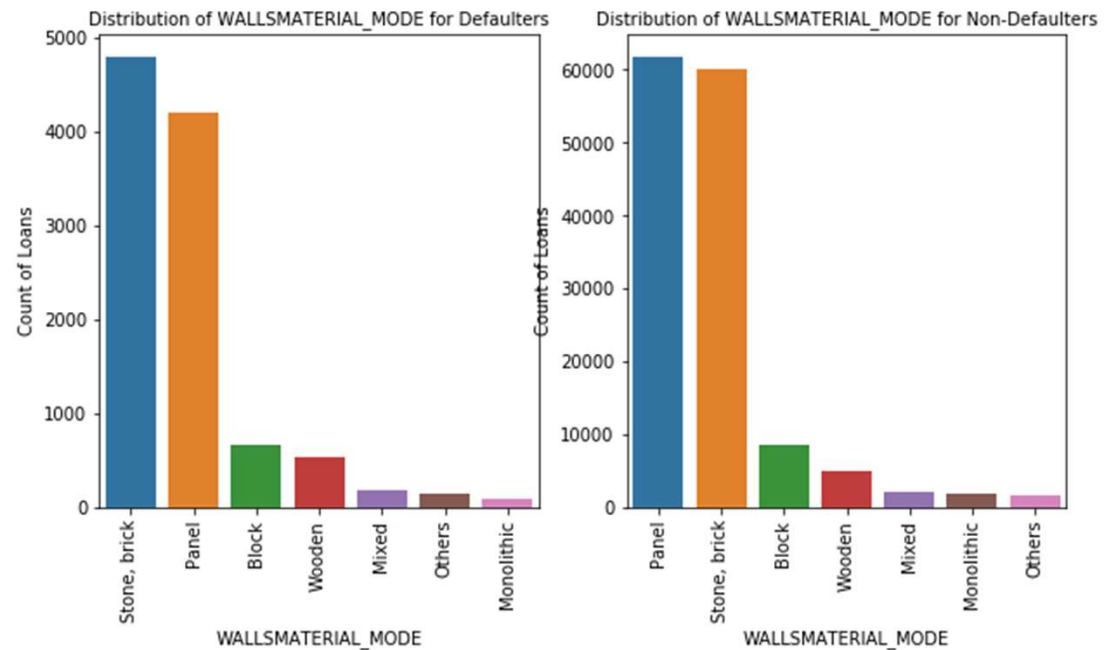
- While the category with highest count remains same.
- This chart tells us that people with Academic Degree rarely take loans and are rarely defaulters. So they are potentially good customers.
- People with higher education are less likely to have payment difficulties. The Ratio is higher for non-defaulters than defaulters.

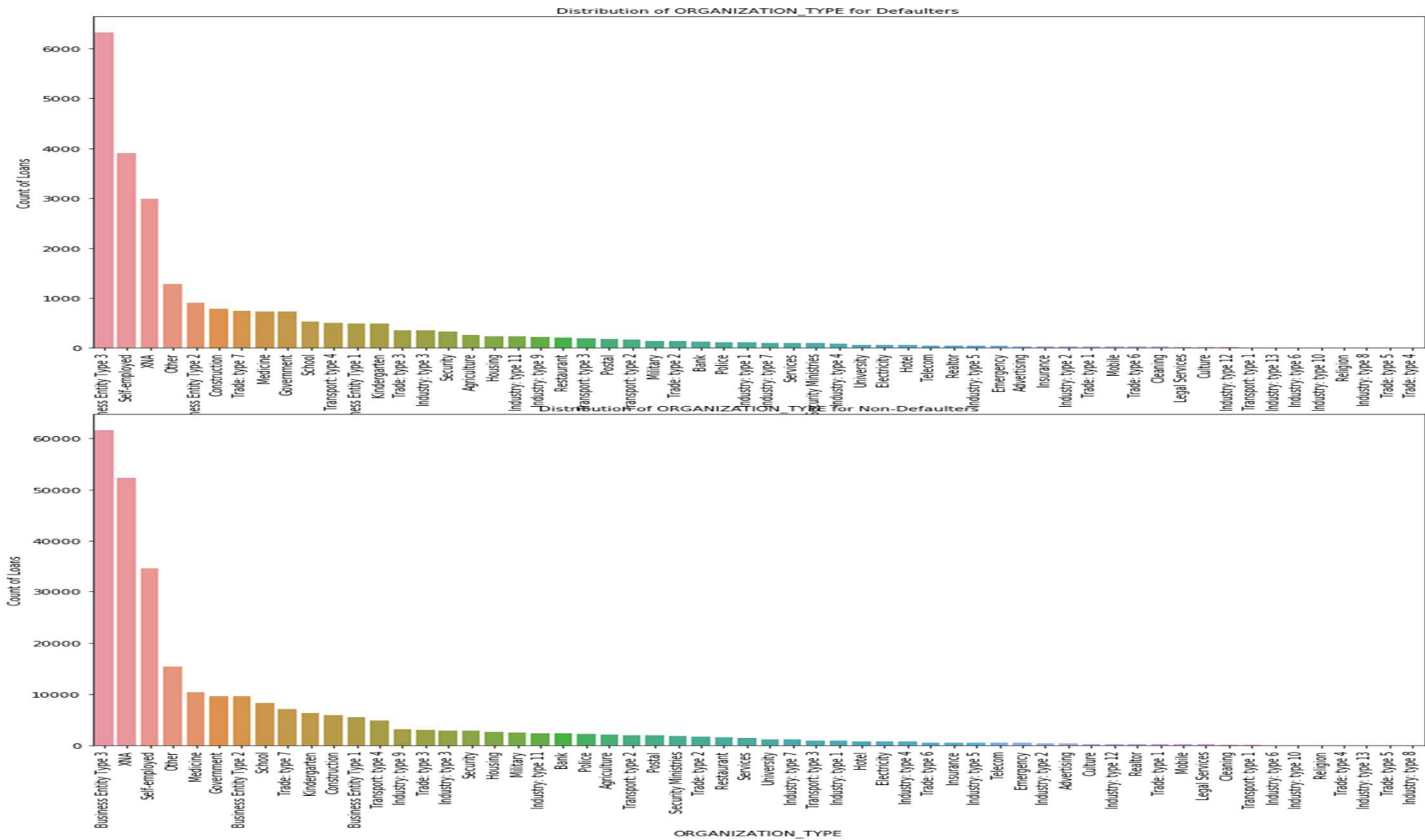


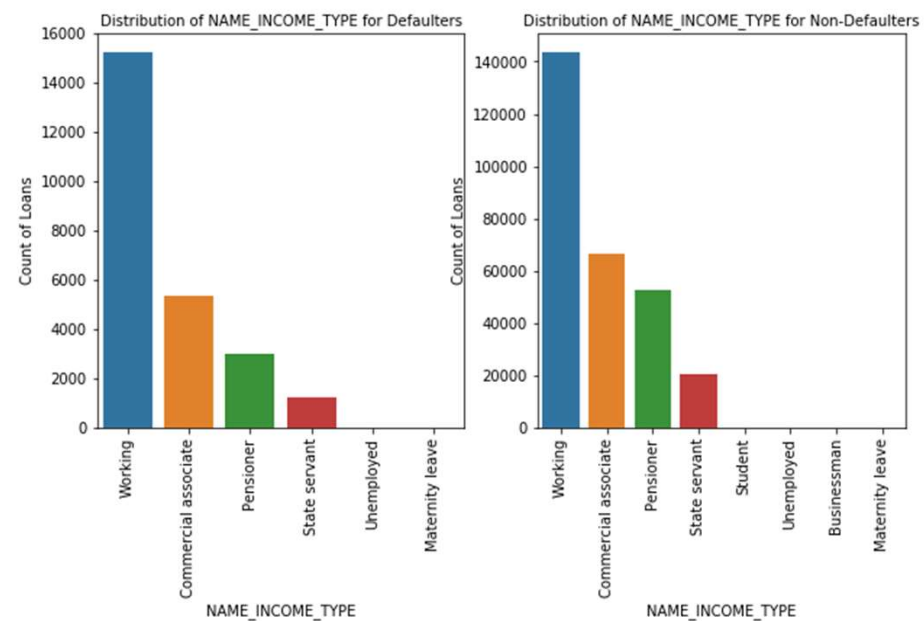
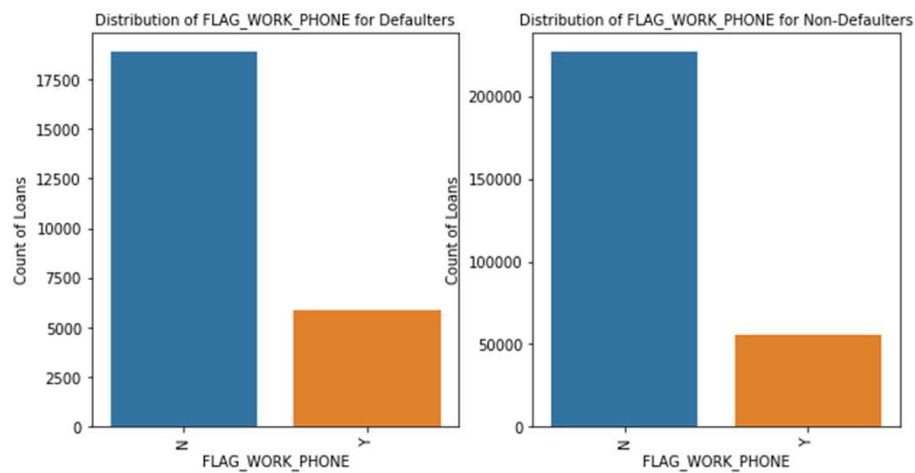
Commercial  
associates,  
Pensioner, State  
Servants have a  
higher ratio to total  
in non-defaulter

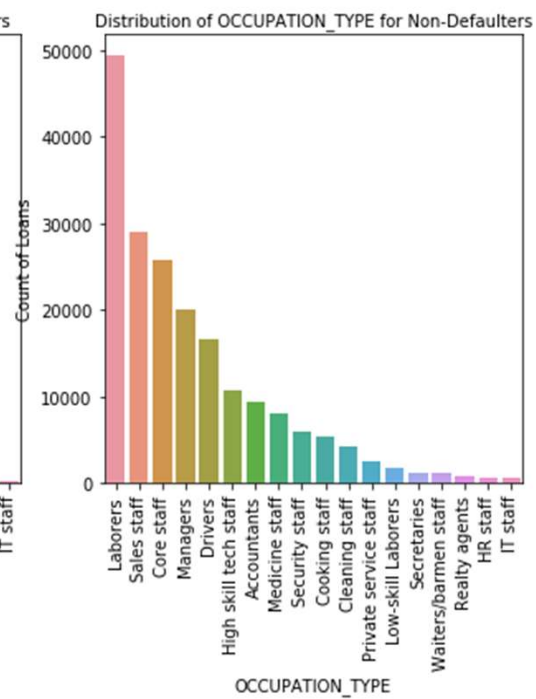
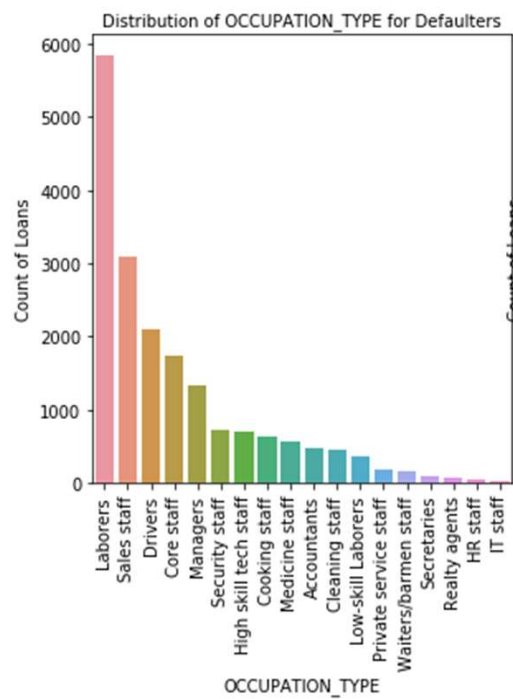


- This interesting chart tells us that most defaulters have houses made of stone and brick while most non-defaulters have houses made of Panel

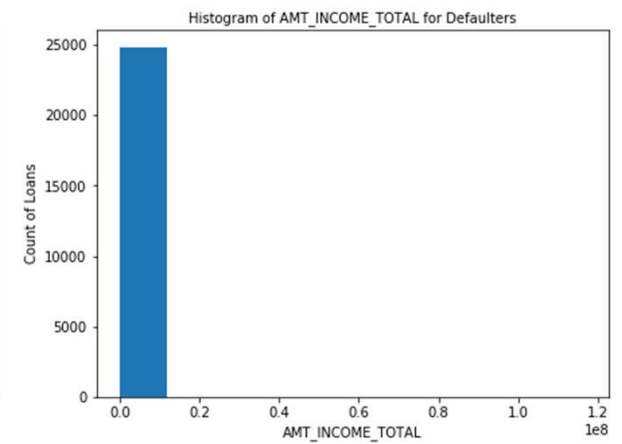
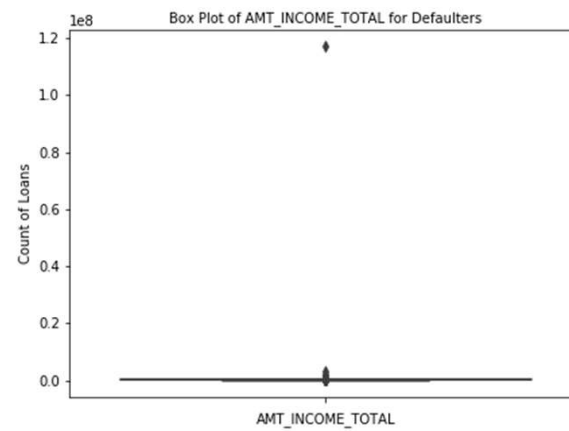




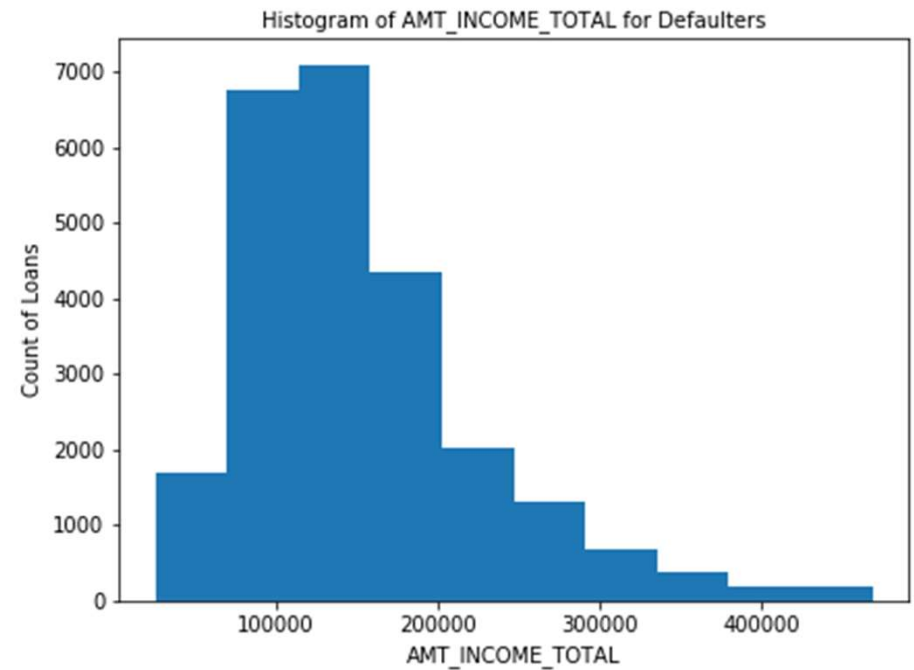
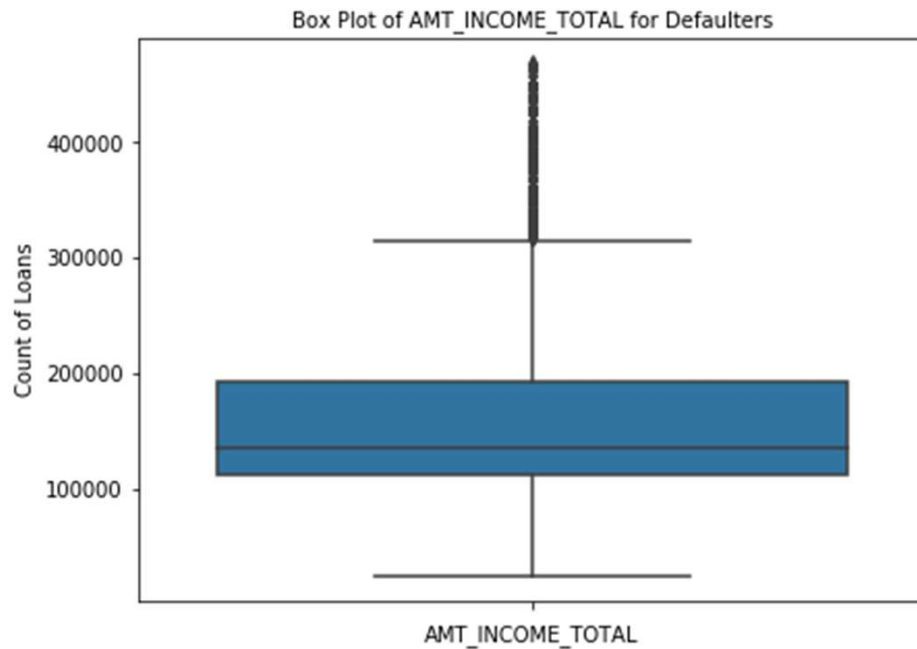




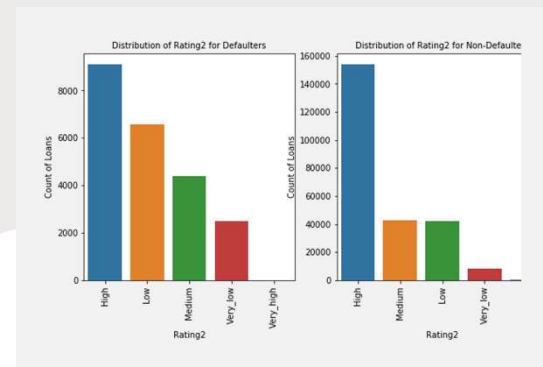
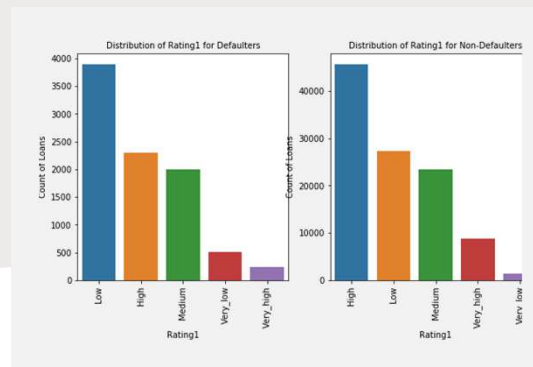
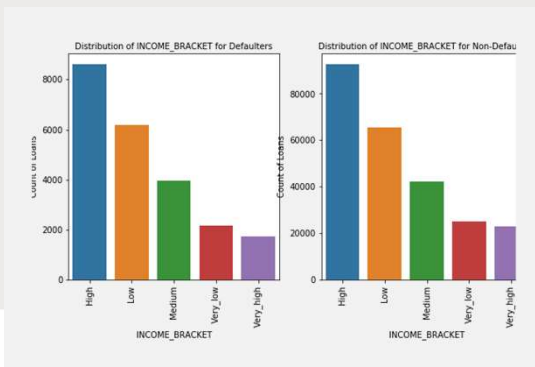
- We can see that there are some outliers and the graph looks like this to accomodate those outli



- This tell us that most people with payment have incomes in the lower range between 100000 to 200000 which some on the higher end some on the lower

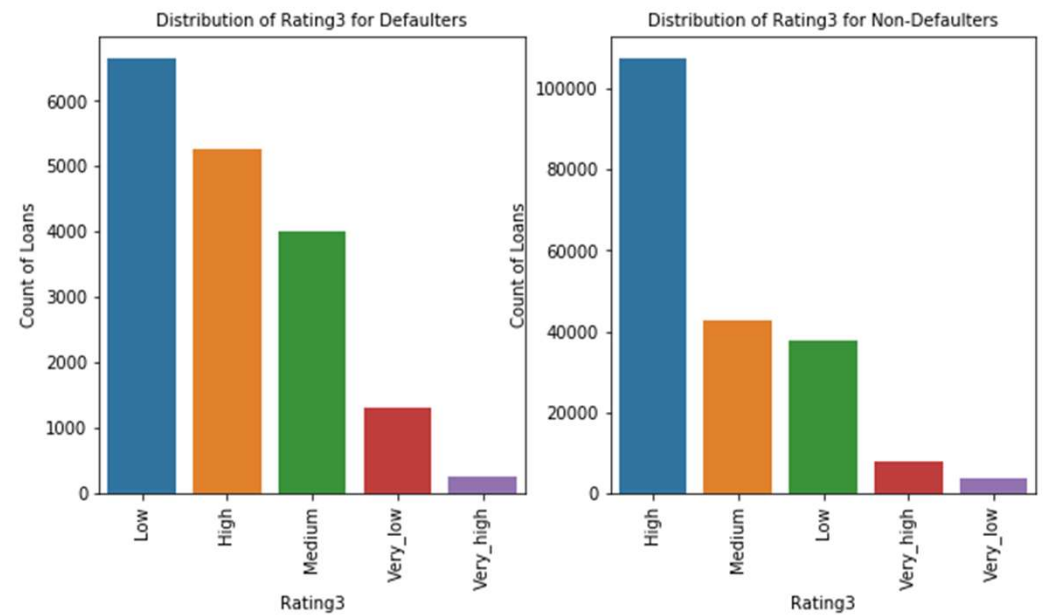




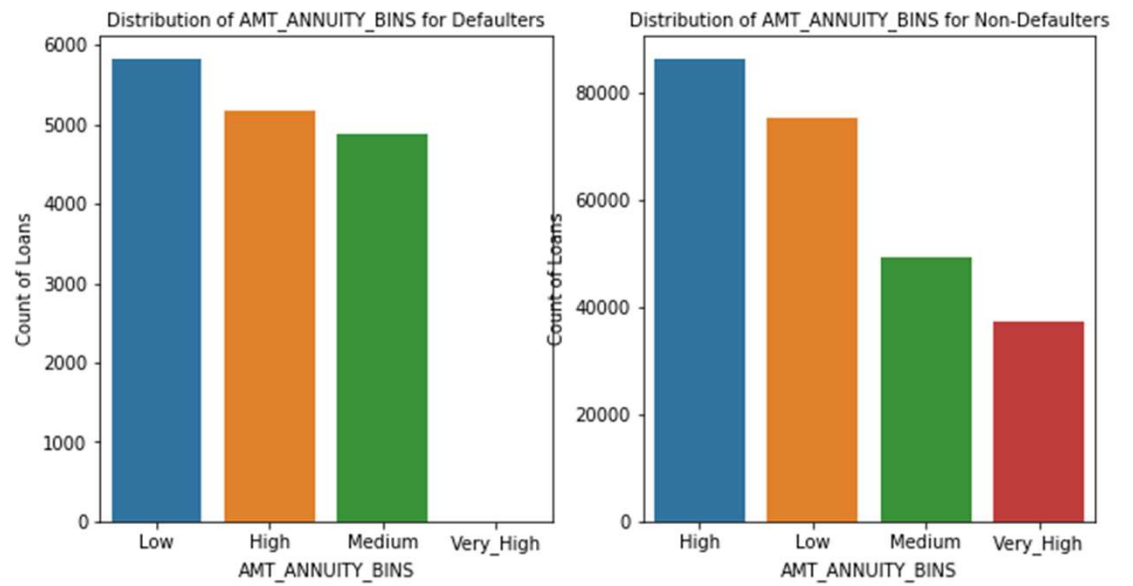


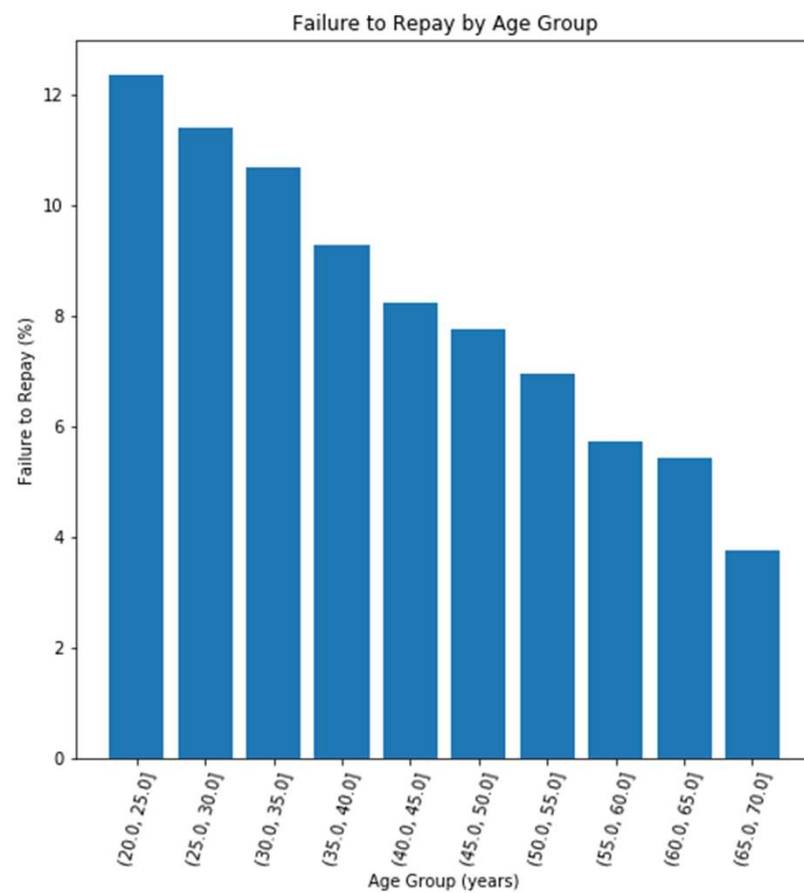
- A large number of defaulters have Low rating, while a large number of non-defaulters have a high rating

- A large number of defaulters have very Low rating, while a large number of non-defaulters have a high rating.

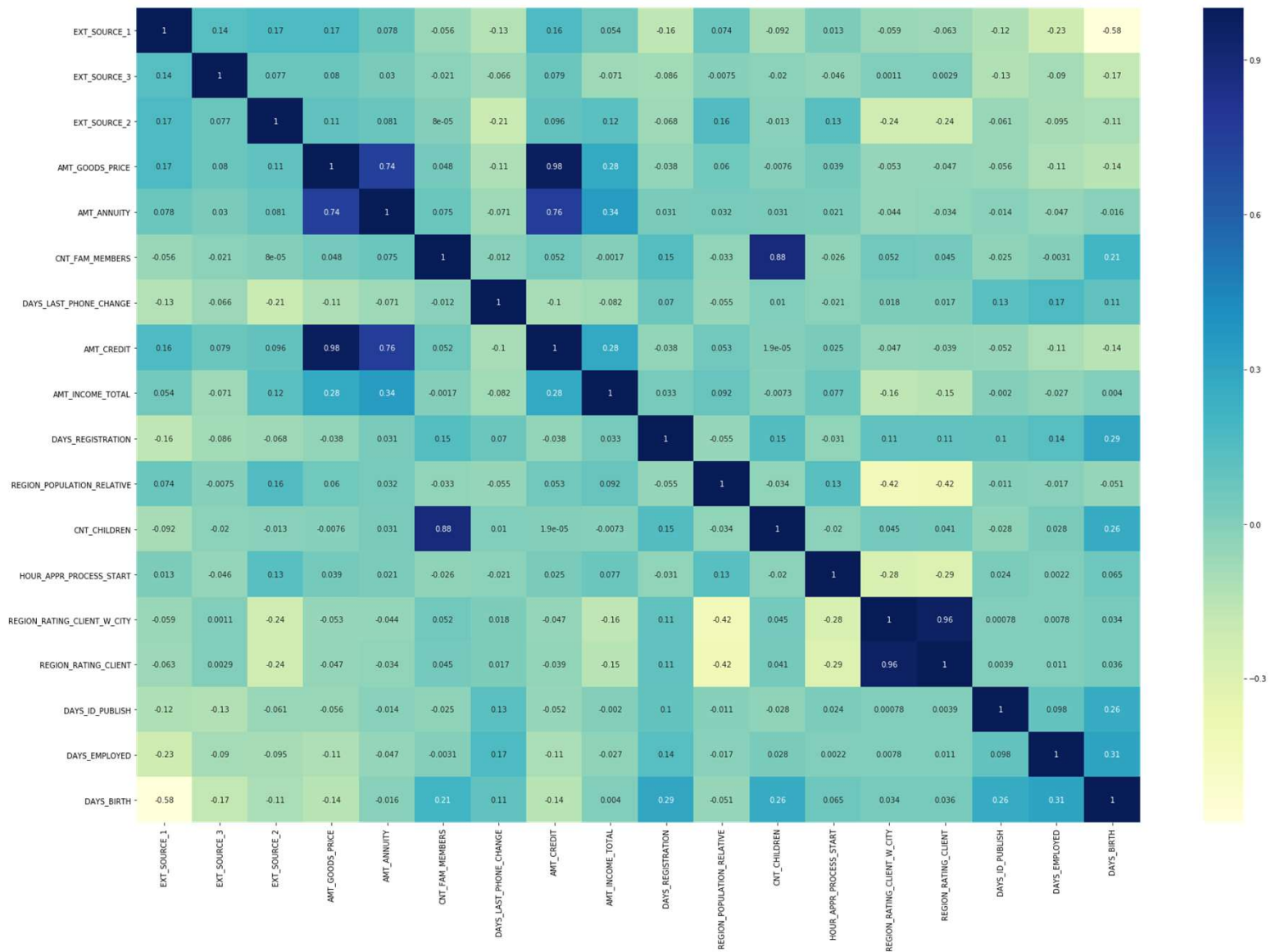


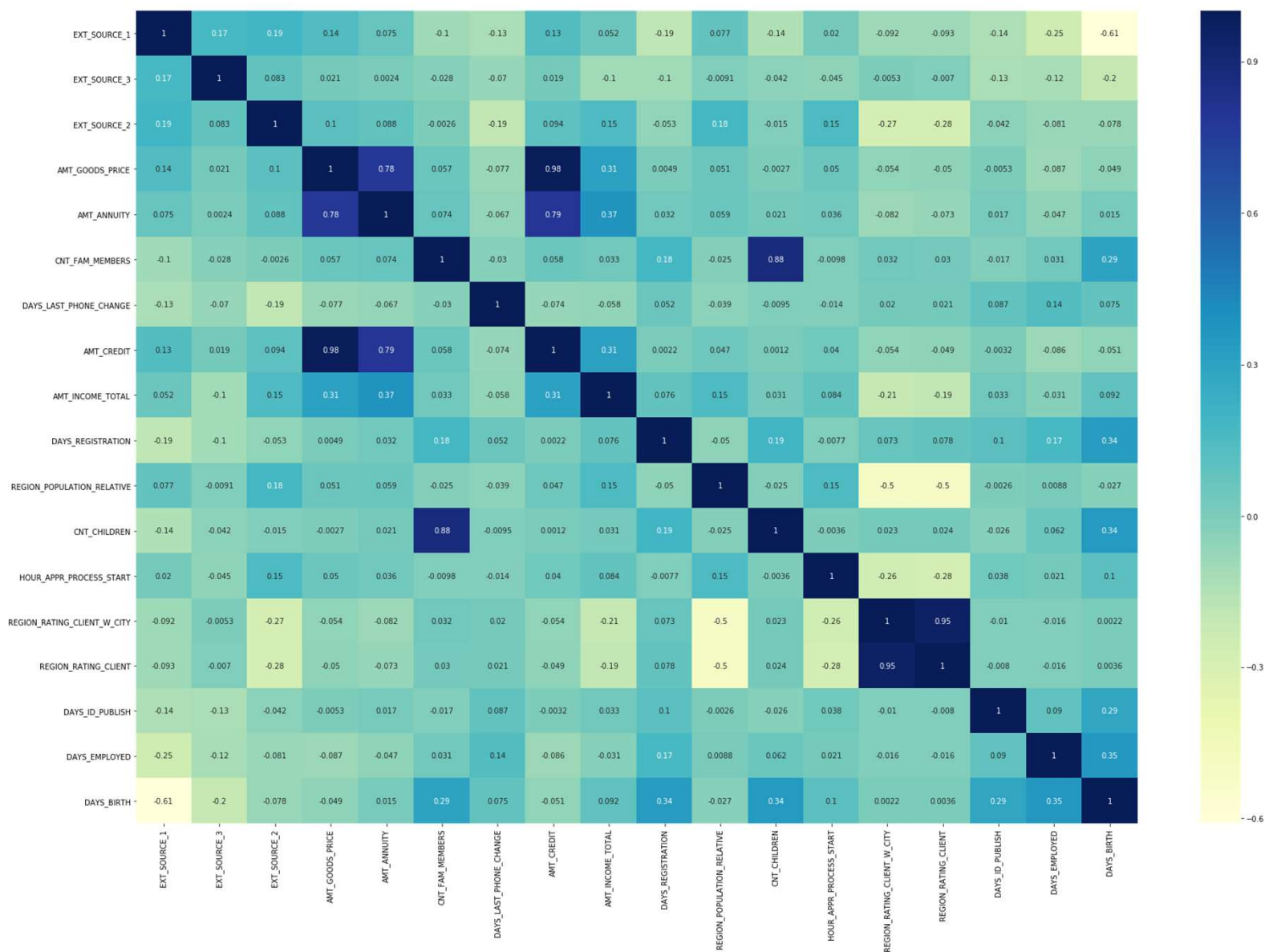
- 
- maximum number of defaulters have Low\_annuity Values, while maximum number of non-defaulters have high annuity



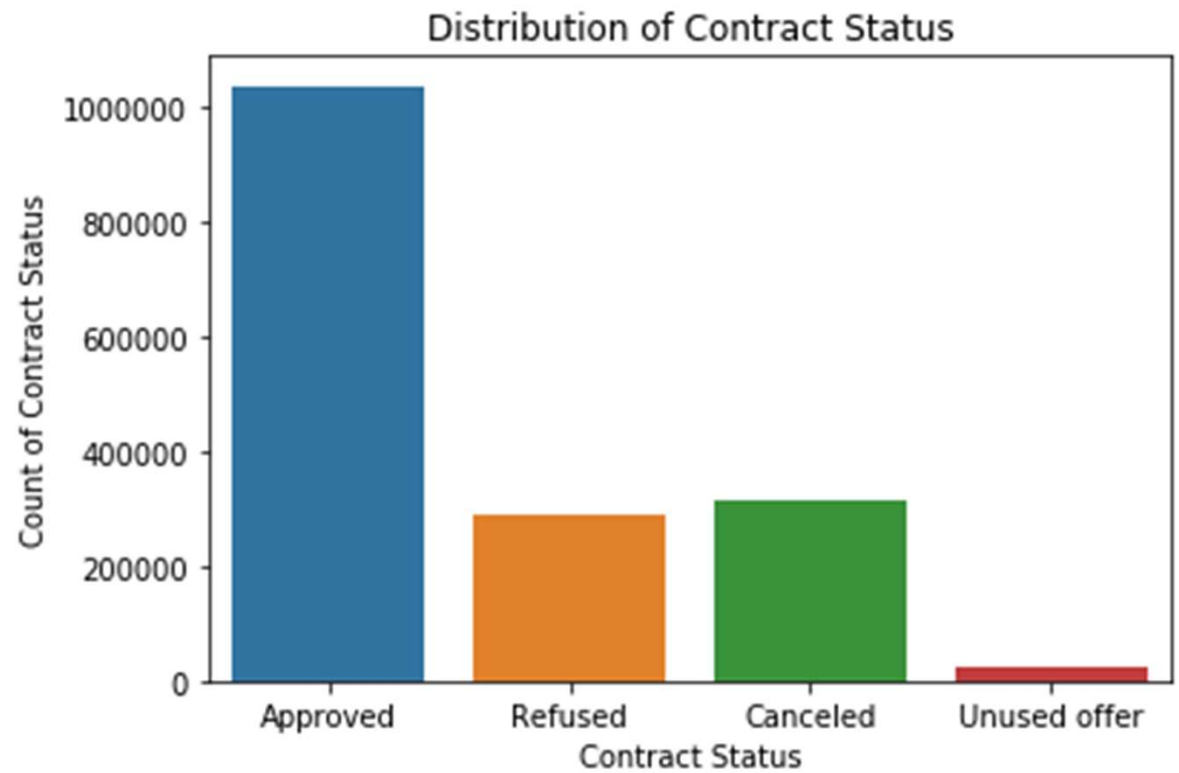


Maximum Failure to Repay is in Age Group 20-25

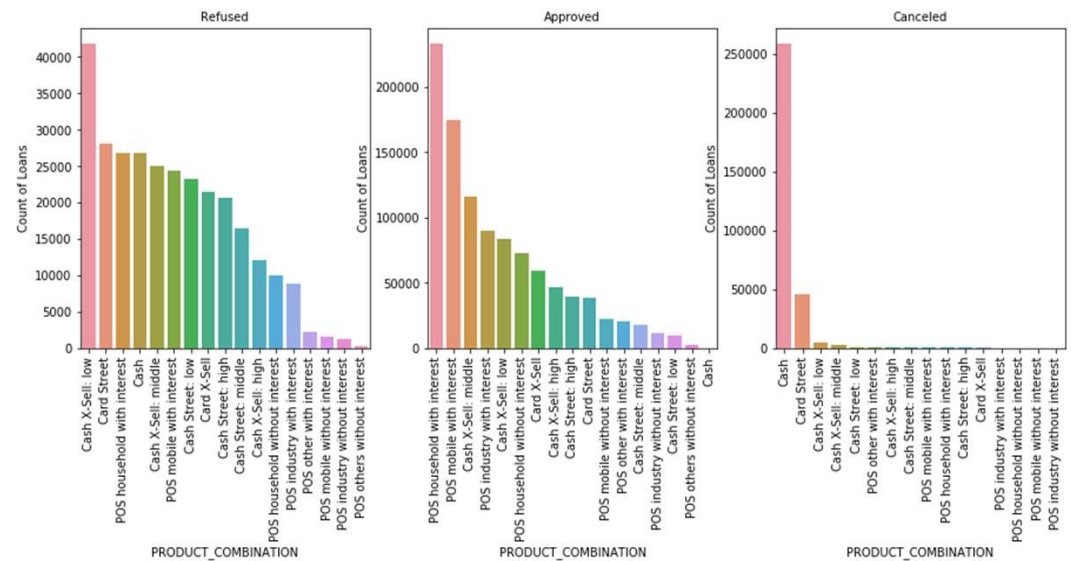




- 
- Identifying missing values and filtering out columns with high missing values

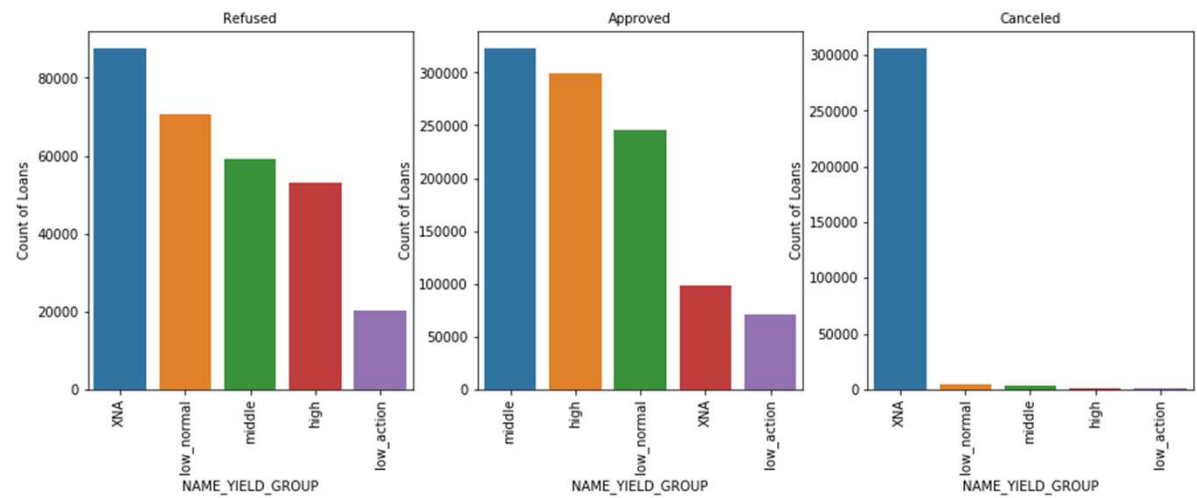


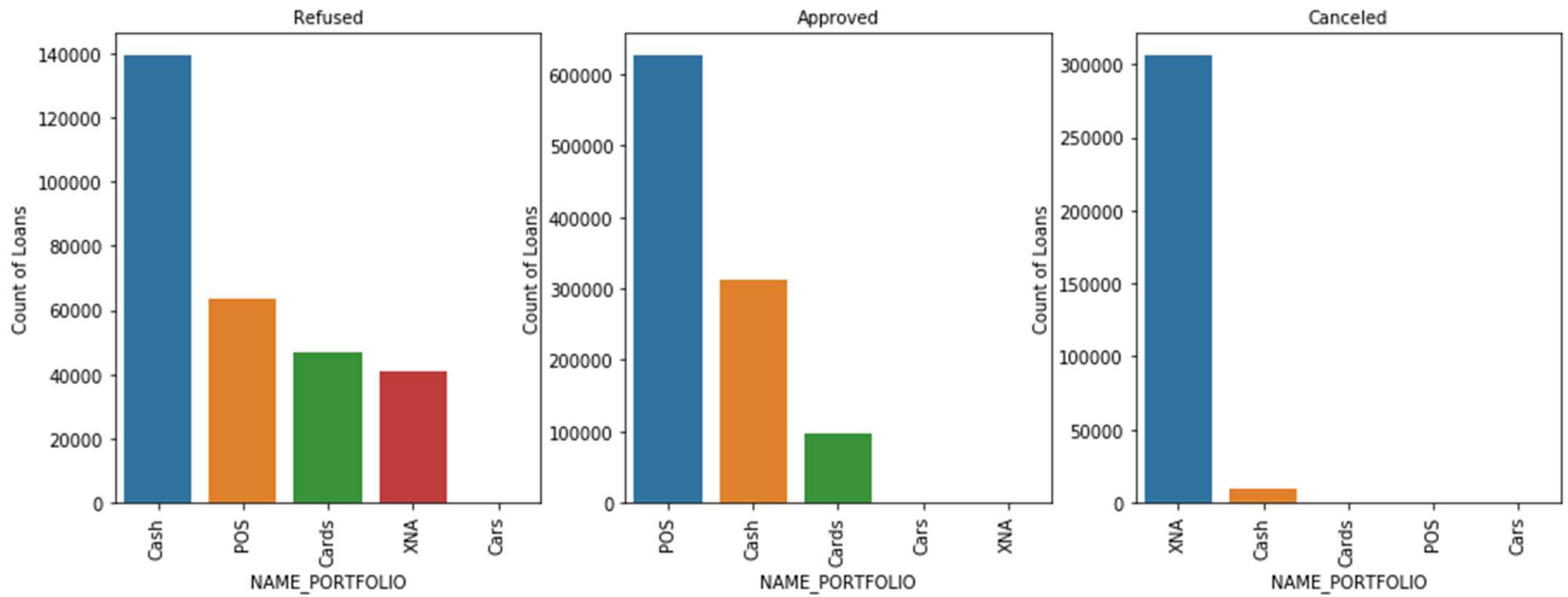
- We observe most number of loans were approved for POS household with interest.
- Most number of refused loans were of Cash X-Sell: Low Product combination
- Most Canceled loans were Cash loans





- 
- Most approved loans were from **Middle** Yield Goup
  - Most refused loans were from Yield Goups Not specified





- Most approved loans were **POS**
- Most refused loans were **Cash**

- Most approved loans were from **Country-wide** Channel
- Most refused loans were from **Credit and Cash Offices** Channel

