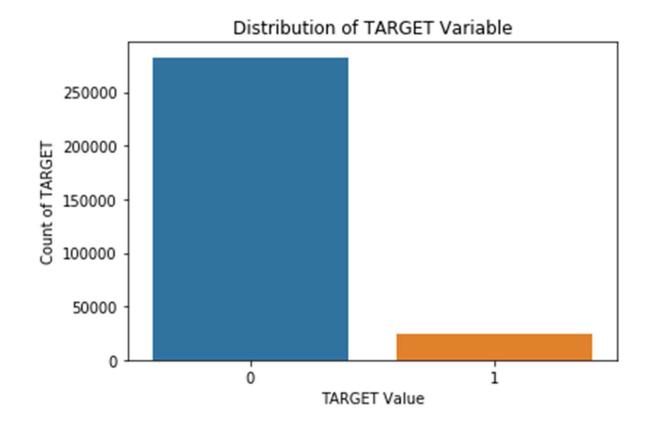
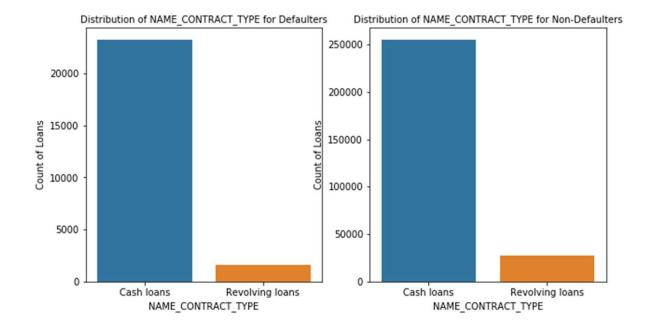
Credit EDA - Case Study

- 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.

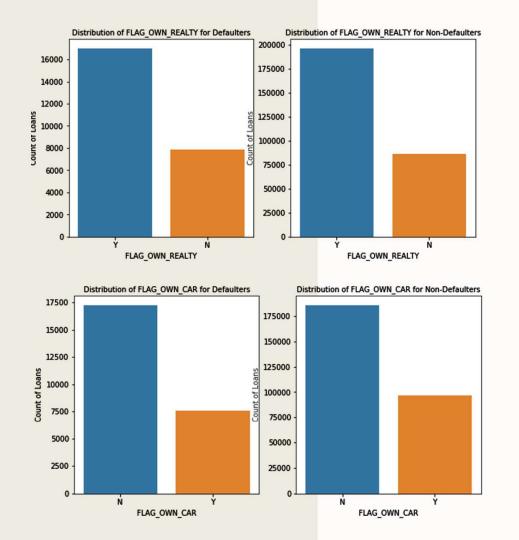


 We observe that the number of Cash loans is much higher than the number of Revolving loans for both Target = 0 and Target = 1





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



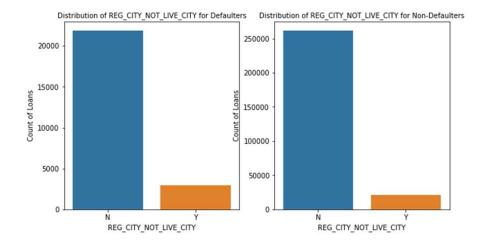


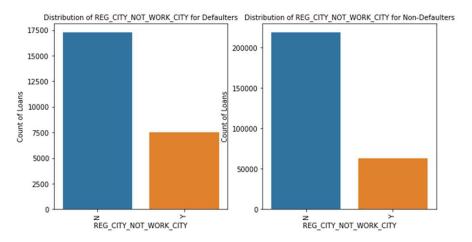
We observe that the number of most people applying for loan **do not own a car.**



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

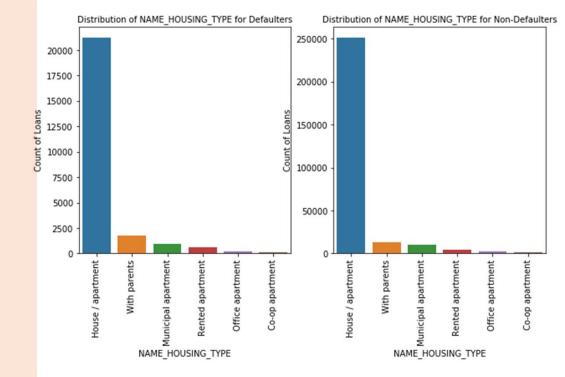
- We observe that the Ratio of people whose Registration City is not the same as live city or work city is higher in case of defaulters are compared to defaulters.
- It tells us that people who live or work in a city different than the registration city are more likely to have payment difficulties.



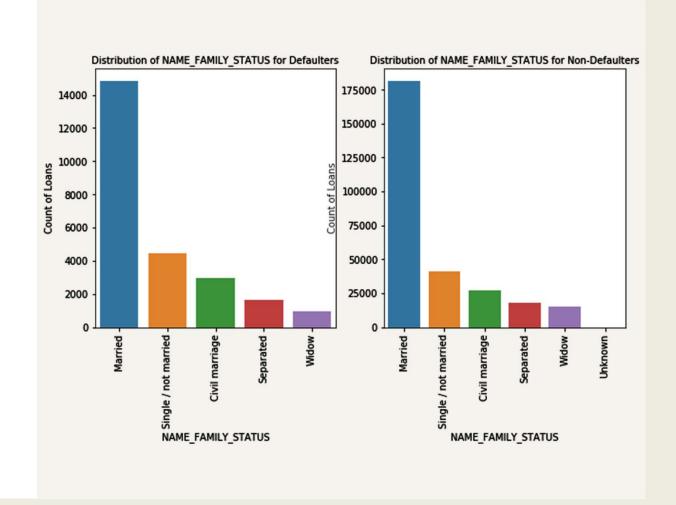


Observation:

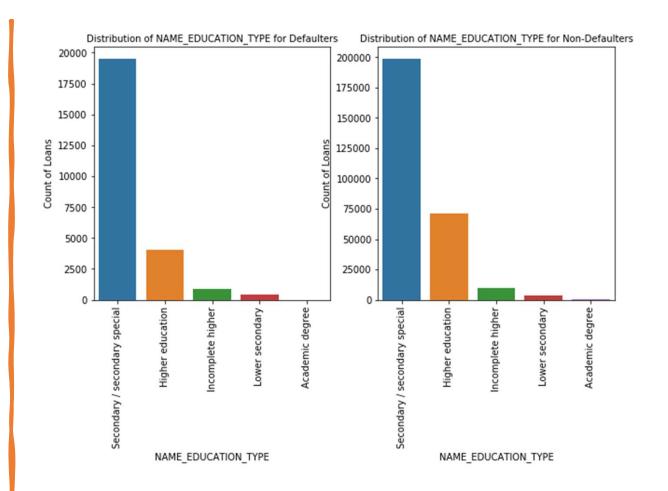
- Most people live in a House/Apartment
- Ratio of People who live **With Parents** is more for defaulter than nondefaulters. It tells us that applicant who
 live with parents have a higher chance
 of having payment difficulties.



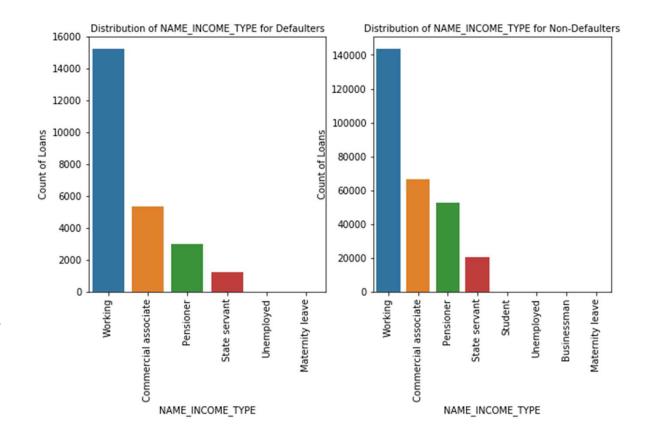
- Ratio
 of Single/Unmarried peop
 le is more in the left
 graph.
- Single/Unmarried people are more likely to have payment difficulties



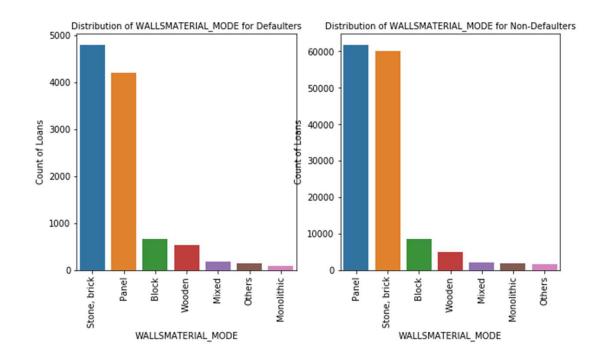
- While the category with highest count remains same.
- This chart tells us that people with Academic Degree rarely take loans and are rarely defaulters. So they are potentially good customers.
- People with higher education are less likely to have payment difficulties. The Ratio is higher for non-defaulters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

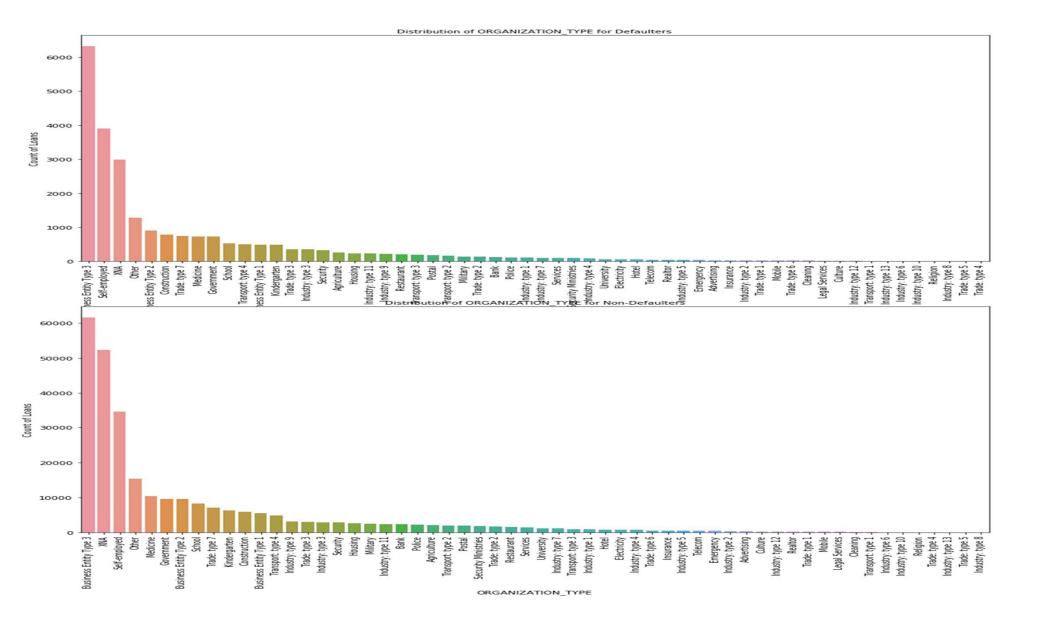


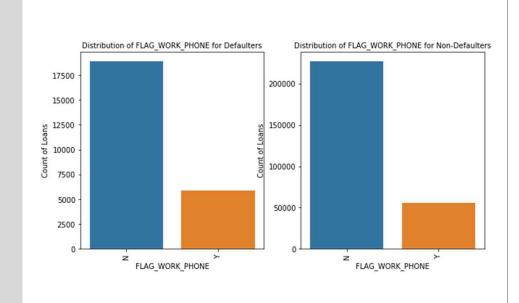
Commercial associates, Pensioner, State Servants have a higher ratio to total in non-defaulter

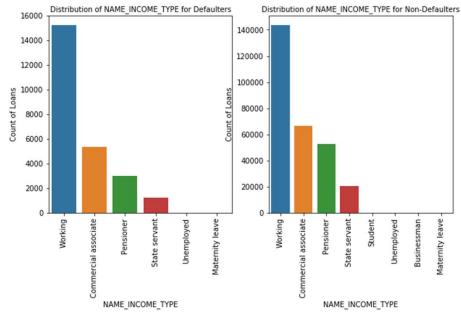


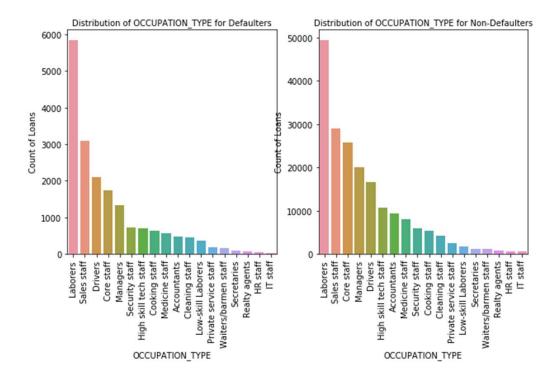
• This interesting chart tells us that most defaulters have houses made of stone and brick while most non-defaulters have houses made of Panel



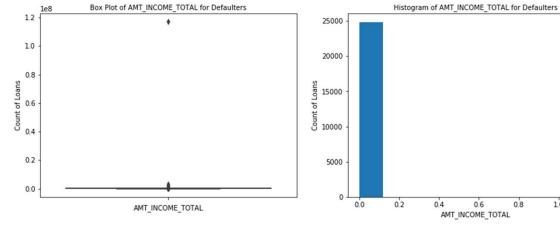








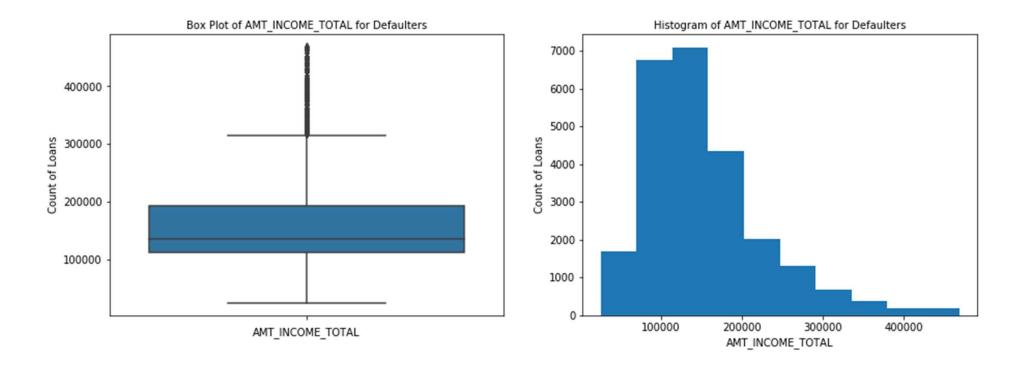
• We can see that there are some outliers and the graph looks like this to accomodate those outli

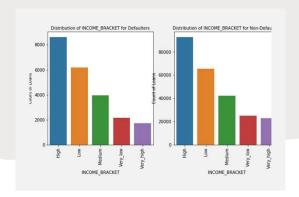


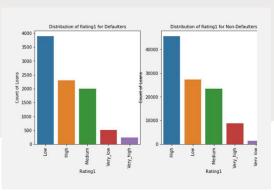
0.8

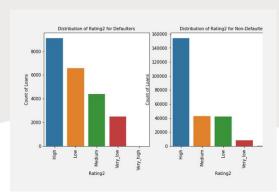
1.0

• 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



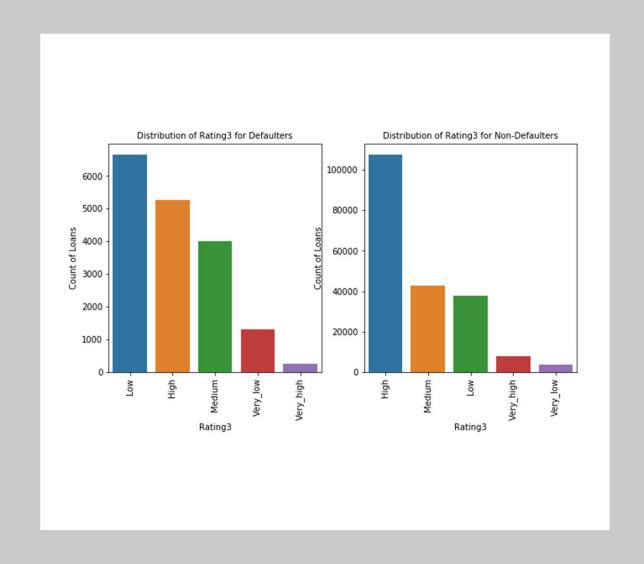




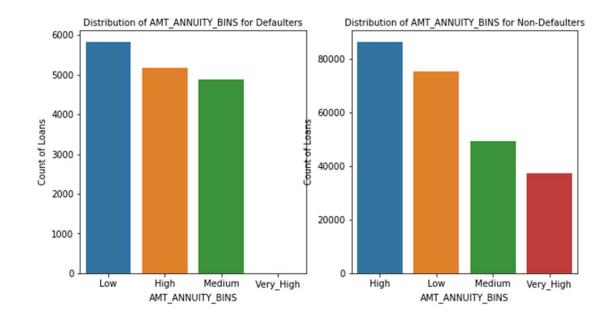


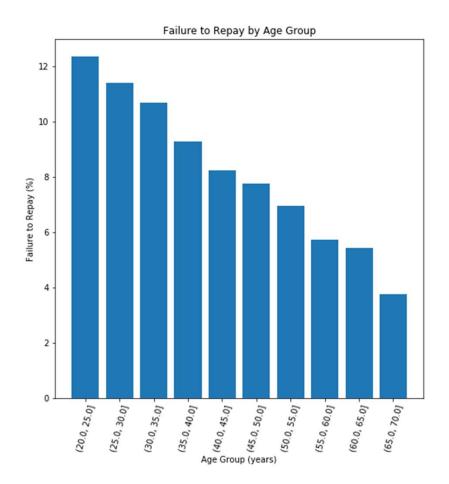
• A large number of defaulters have Low rating, while a large number of non-defaulters have a high rating

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



 maxinum number of defaulters have Low_annuity Values, while maximum number of non-defaulters have high annuity





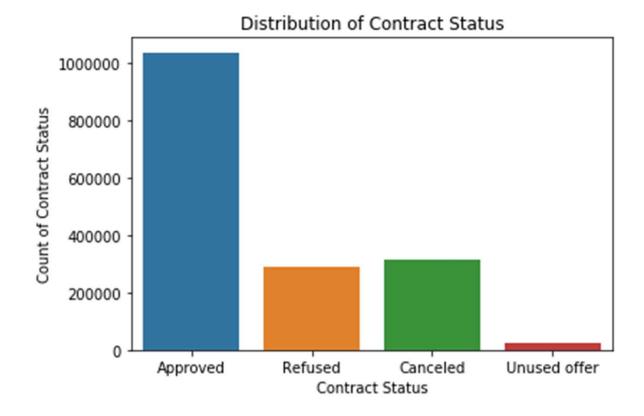
Maximum Failure to Repay is in Age Group 20-25

| EXT_SOURCE_1 - | 1 | 0.14 | 0.17 | 0.17 | 0.078 | -0.056 | -0.13 | 0.16 | 0.054 | -0.16 | 0.074 | -0.092 | 0.013 | -0.059 | -0.063 | -0.12 | -0.23 | -0.58 |
|-------------------------------|----------------|----------------|----------------|-------------------|---------------|-------------------|--------------------------|--------------|--------------------|---------------------|------------------------------|----------------|---------------------------|-------------------------------|------------------------|-------------------|-----------------|------------|
| EXT_SOURCE_3 - | 0.14 | 1 | 0.077 | 0.08 | 0.03 | -0.021 | -0.066 | 0.079 | -0.071 | -0.086 | -0.0075 | -0.02 | -0.046 | 0.0011 | 0.0029 | -0.13 | -0.09 | -0.17 |
| EXT_SOURCE_2 - | 0.17 | 0.077 | 1 | 0.11 | 0.081 | 8e-05 | -0.21 | 0.096 | 0.12 | -0.068 | 0.16 | -0.013 | 0.13 | -0.24 | -0.24 | -0.061 | -0.095 | -0.11 |
| AMT_GOODS_PRICE - | 0.17 | 0.08 | 0.11 | 1 | 0.74 | 0.048 | -0.11 | 0.98 | 0.28 | -0.038 | 0.06 | -0.0076 | 0.039 | -0.053 | -0.047 | -0.056 | -0.11 | -0.14 |
| AMT_ANNUITY - | 0.078 | 0.03 | 0.081 | | | 0.075 | -0.071 | | 0.34 | 0.031 | 0.032 | 0.031 | 0.021 | -0.044 | -0.034 | -0.014 | -0.047 | -0.016 |
| CNT_FAM_MEMBERS - | -0.056 | -0.021 | 8e-05 | 0.048 | 0.075 | 1 | -0.012 | 0.052 | -0.0017 | 0.15 | -0.033 | 0.88 | -0.026 | 0.052 | 0.045 | -0.025 | -0.0031 | 0.21 |
| DAYS_LAST_PHONE_CHANGE - | -0.13 | -0.066 | -0.21 | -0.11 | -0.071 | -0.012 | 1 | -0.1 | -0.082 | 0.07 | -0.055 | 0.01 | -0.021 | 0.018 | 0.017 | 0.13 | 0.17 | 011 |
| AMT_CREDIT - | 0.16 | 0.079 | 0.096 | 0.98 | | 0.052 | -0.1 | | 0.28 | -0.038 | 0.053 | 1.9e-05 | 0.025 | -0.047 | -0.039 | -0.052 | -0.11 | -0.14 |
| AMT_INCOME_TOTAL - | 0.054 | -0.071 | 0.12 | | | -0.0017 | -0.082 | 0.28 | 1 | 0.033 | 0.092 | -0.0073 | 0.077 | -0.16 | -0.15 | -0.002 | -0.027 | 0.004 |
| DAYS_REGISTRATION - | -0.16 | -0.086 | -0.068 | -0.038 | 0.031 | 0.15 | 0.07 | -0.038 | 0.033 | 1 | -0.055 | 0.15 | -0.031 | 0.11 | 0.11 | 0.1 | 0.14 | 0.29 |
| REGION_POPULATION_RELATIVE - | 0.074 | -0.0075 | 0.16 | 0.06 | 0.032 | -0.033 | -0.055 | 0.053 | 0.092 | -0.055 | 1 | -0.034 | 0.13 | -0.42 | -0.42 | -0.011 | -0.017 | -0.051 |
| CNT_CHILDREN - | -0.092 | -0.02 | -0.013 | -0.0076 | 0.031 | 0.88 | 0.01 | 1.9e-05 | -0.0073 | 0.15 | -0.034 | i | -0.02 | 0.045 | 0.041 | -0.028 | 0.028 | 0.26 |
| HOUR_APPR_PROCESS_START - | 0.013 | -0.046 | 0.13 | 0.039 | 0.021 | -0.026 | -0.021 | 0.025 | 0.077 | -0.031 | 0.13 | -0.02 | 1 | -0.28 | -0.29 | 0.024 | 0.0022 | 0.065 |
| REGION_RATING_CLIENT_W_CITY - | -0.059 | 0.0011 | -0.24 | -0.053 | -0.044 | 0.052 | 0.018 | -0.047 | -0.16 | 0.11 | -0.42 | 0.045 | -0.28 | | 0.96 | 0.00078 | 0.0078 | 0.034 |
| REGION_RATING_CLIENT - | -0.063 | 0.0029 | -0.24 | -0.047 | -0.034 | 0.045 | 0.017 | -0.039 | -0.15 | 0.11 | -0.42 | 0.041 | -0.29 | 0.96 | 1 | 0.0039 | 0.011 | 0.036 |
| DAYS_ID_PUBLISH - | -0.12 | -0.13 | -0.061 | -0.056 | -0.014 | -0.025 | 0.13 | -0.052 | -0.002 | 0.1 | -0.011 | -0.028 | 0.024 | 0.00078 | 0.0039 | 1 | 0.098 | 0.26 |
| DAYS_EMPLOYED - | -0.23 | -0.09 | -0.095 | -0.11 | -0.047 | -0.0031 | 0.17 | -0.11 | -0.027 | 0.14 | -0.017 | 0.028 | 0.0022 | 0.0078 | 0.011 | 0.098 | 1 | 0.31 |
| DAYS_BIRTH - | -0.58 | -0.17 | -0.11 | -0.14 | -0.016 | 0.21 | 0.11 | -0.14 | 0.004 | 0.29 | -0.051 | 0.26 | 0.065 | 0.034 | 0.036 | 0.26 | 0.31 | 1 |
| | EXT_SOURCE_1 - | EXT_SOURCE_3 - | EXT_SOURCE_2 - | AMT_GOODS_PRICE - | AMT_ANNUITY . | CNT_FAM_MEMBERS - | DAYS_LAST_PHONE_CHANGE - | AMT_CREDIT - | AMT_INCOME_TOTAL - | DAYS_REGISTRATION - | REGION_POPULATION_RELATIVE - | ONT_CHILDREN - | HOUR_APPR_PROCESS_START - | REGION_RATING_CLIENT_W_CITY - | REGION_RATING_CLIENT - | DAYS_ID_PUBLISH - | DAYS_EMPLOYED - | DAYS_BIRTH |

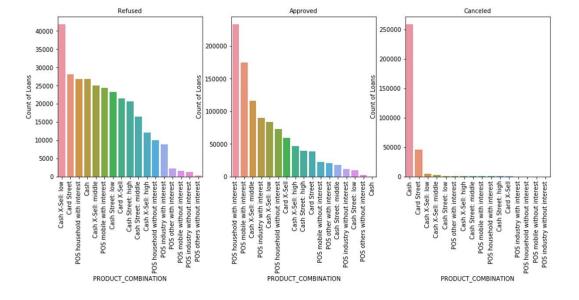
| EXT_SOURCE_1 | 1 | 0.17 | 0.19 | 0.14 | 0.075 | -0.1 | -0.13 | 0.13 | 0.052 | -0.19 | 0.077 | -0.14 | 0.02 | -0.092 | -0.093 | -0.14 | -0.25 | -0.61 |
|-----------------------------|--------------|--------------|--------------|-----------------|-------------|-----------------|------------------------|------------|------------------|-------------------|----------------------------|--------------|-------------------------|-----------------------------|----------------------|-----------------|---------------|------------|
| EXT_SOURCE_3 - | 0.17 | 1 | 0.083 | 0.021 | 0.0024 | -0.028 | -0.07 | 0.019 | -0.1 | -0.1 | -0.0091 | -0.042 | -0.045 | -0.0053 | -0.007 | -0.13 | -0.12 | -0.2 |
| EXT_SOURCE_2 | | 0.083 | 1 | 0.1 | 0.088 | -0.0026 | -0.19 | 0.094 | 0.15 | -0.053 | | -0.015 | 0.15 | -0.27 | -0.28 | -0.042 | -0.081 | -0.078 |
| AMT_GOODS_PRICE | 0.14 | 0.021 | 0.1 | 1 | | 0.057 | -0.077 | 0.98 | | 0.0049 | 0.051 | -0.0027 | 0.05 | -0.054 | -0.05 | -0.0053 | -0.087 | -0.049 |
| AMT_ANNUITY - | 0.075 | 0.0024 | 0.088 | 0.78 | 1 | 0.074 | -0.067 | 0.79 | 0.37 | 0.032 | 0.059 | 0.021 | 0.036 | -0.082 | -0.073 | 0.017 | -0.047 | 0.015 |
| CNT_FAM_MEMBERS - | -0.1 | -0.028 | -0.0026 | 0.057 | 0.074 | 1 | -0.03 | 0.058 | 0.033 | | -0.025 | 0.88 | -0.0098 | 0.032 | 0.03 | -0.017 | 0.031 | 0.29 |
| DAYS_LAST_PHONE_CHANGE - | -0.13 | -0.07 | -0.19 | -0.077 | -0.067 | -0.03 | 1 | -0.074 | -0.058 | 0.052 | -0.039 | -0.0095 | -0.014 | 0.02 | 0.021 | 0.087 | 0.14 | 0.075 |
| AMT_CREDIT - | 0.13 | 0.019 | 0.094 | 0.98 | 0.79 | 0.058 | -0.074 | 1 | 0.31 | 0.0022 | 0.047 | 0.0012 | 0.04 | -0.054 | -0.049 | -0.0032 | -0.086 | -0.051 |
| AMT_INCOME_TOTAL - | 0.052 | -0.1 | 0.15 | 0.31 | 0.37 | 0.033 | -0.058 | | 1 | 0.076 | 0.15 | 0.031 | 0.084 | -0.21 | -0.19 | 0.033 | -0.031 | 0.092 |
| DAYS_REGISTRATION - | -0.19 | -0.1 | -0.053 | 0.0049 | 0.032 | | 0.052 | 0.0022 | 0.076 | 1 | -0.05 | | -0.0077 | 0.073 | 0.078 | 0.1 | | 0.34 |
| REGION_POPULATION_RELATIVE | 0.077 | -0.0091 | | 0.051 | 0.059 | -0.025 | -0.039 | 0.047 | 0.15 | -0.05 | 1 | 0.025 | 0.15 | -0.5 | -0.5 | -0.0026 | 0.0088 | -0.027 |
| CNT_CHILDREN - | -0.14 | -0.042 | -0.015 | -0.0027 | 0.021 | 0.88 | -0.0095 | 0.0012 | 0.031 | | -0.025 | 1 | -0.0036 | 0.023 | 0.024 | -0.026 | 0.062 | 0.34 |
| HOUR_APPR_PROCESS_START | 0.02 | -0.045 | 0.15 | 0.05 | 0.036 | -0.0098 | 0.014 | 0.04 | 0.084 | -0.0077 | 0.15 | -0.0036 | 1 | -0.26 | -0.28 | 0.038 | 0.021 | 0.1 |
| REGION_RATING_CLIENT_W_CITY | -0.092 | -0.0053 | -0.27 | -0.054 | -0.082 | 0.032 | 0.02 | -0.054 | -0.21 | 0.073 | -0.5 | 0.023 | -0.26 | | 0.95 | -0.01 | -0.016 | 0.0022 |
| REGION_RATING_CLIENT - | -0.093 | -0.007 | -0.28 | -0.05 | -0.073 | 0.03 | 0.021 | -0.049 | -0.19 | 0.078 | -0.5 | 0.024 | -0.28 | 0.95 | 1 | -0.008 | -0.016 | 0.0036 |
| DAYS_ID_PUBLISH - | -0.14 | -0.13 | -0.042 | -0.0053 | 0.017 | -0.017 | 0.087 | -0.0032 | 0.033 | 0.1 | -0.0026 | -0.026 | 0.038 | -0.01 | -0.008 | 1 | 0.09 | 0.29 |
| DAYS_EMPLOYED - | -0.25 | -0.12 | -0.081 | -0.087 | -0.047 | 0.031 | 0.14 | -0.086 | -0.031 | | 0.0088 | 0.062 | 0.021 | -0.016 | -0.016 | 0.09 | 1 | 0.35 |
| DAYS_BIRTH | -0.61 | -0.2 | -0.078 | -0.049 | 0.015 | 0.29 | 0.075 | -0.051 | 0.092 | 0.34 | -0.027 | 0.34 | 0.1 | 0.0022 | 0.0036 | 0.29 | 0.35 | 1 |
| | EXT_SOURCE_1 | EXT_SOURCE_3 | EXT_SOURCE_2 | AMT_GOODS_PRICE | AMT_ANNUITY | CNT_FAM_MEMBERS | DAYS_LAST_PHONE_CHANGE | AMT_CREDIT | AMT_INCOME_TOTAL | DAY5_REGISTRATION | REGION_POPULATION_RELATIVE | CNT_CHILDREN | HOUR_APPR_PROCESS_START | REGION_RATING_CLIENT_W_CITY | REGION_RATING_CLIENT | DAYS_ID_PUBLISH | DAYS_EMPLOYED | DAYS_BIRTH |

--0.3

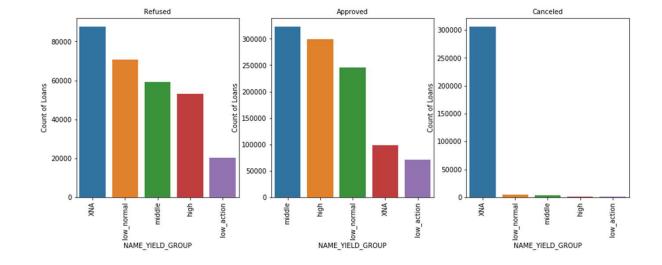
• Identifying missing values and filtering out columns with high missing values

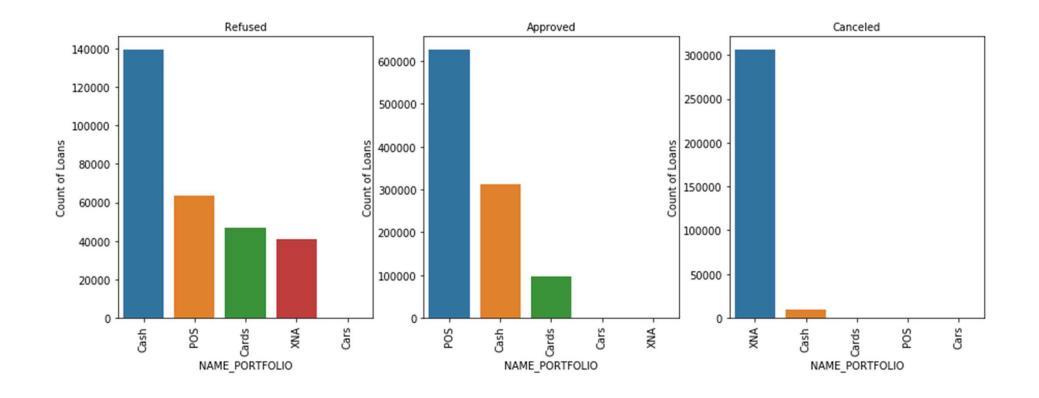


- We observe most number of loans were approved for POS household with interest.
- Most number of refused loans were of Cash X-Sell: Low Product combination
- Most Canceled loans were Cash loans



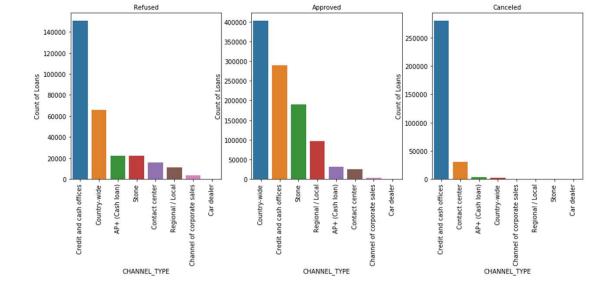
- Most approved loans were from **Middle** Yield Goup
- Most refused loans were from Yield Goups Not specified

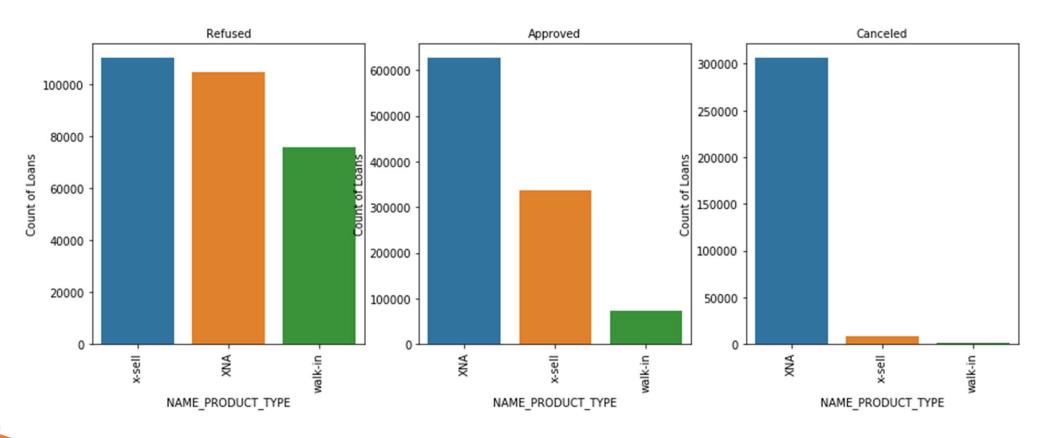


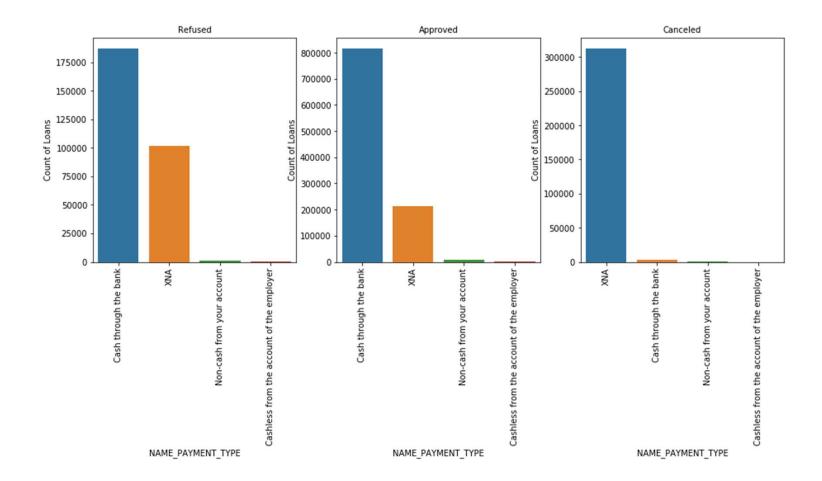


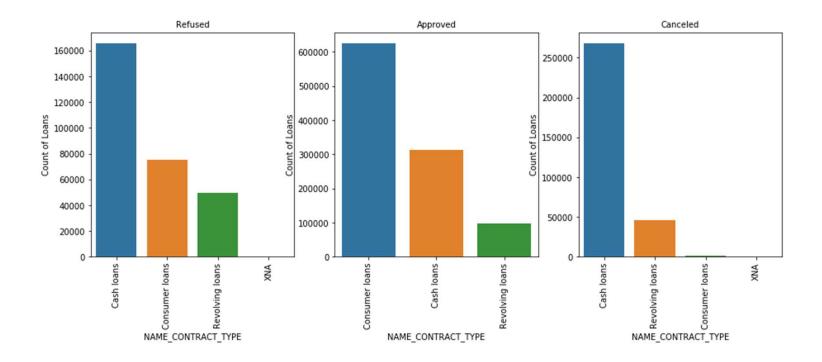
- •Most approved loans were **POS**
- •Most refused loans were Cash

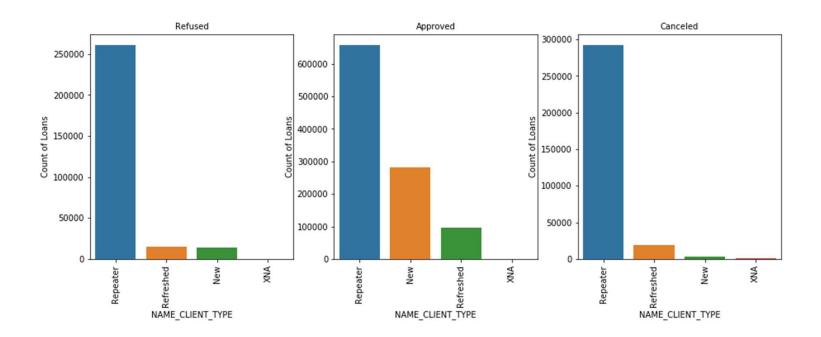
- Most approved loans were from Countrywide Channel
- Most refused loans were from Credit and Cash Offices Channel

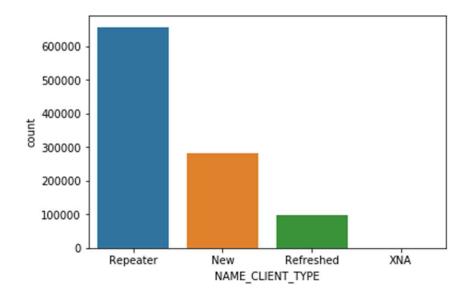


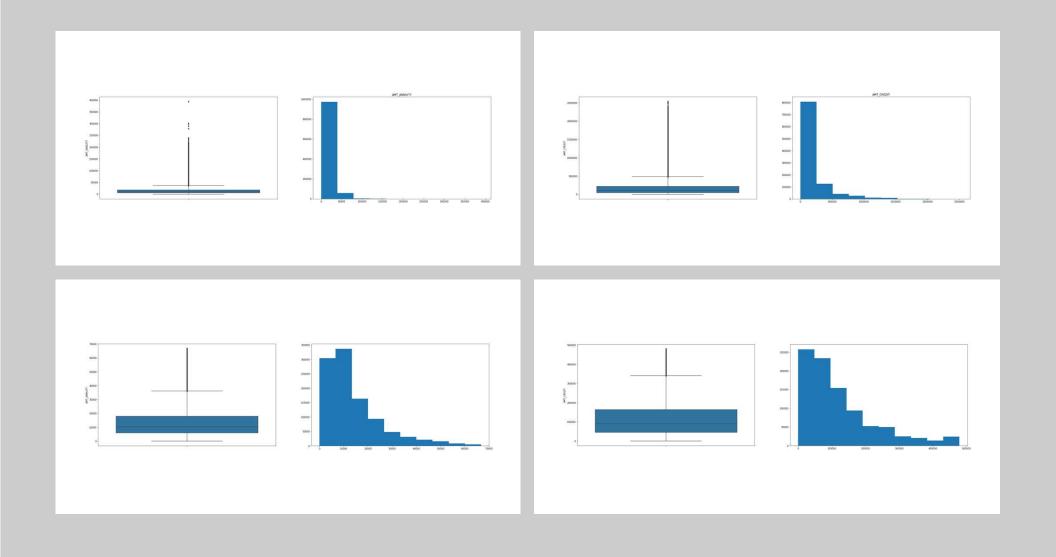


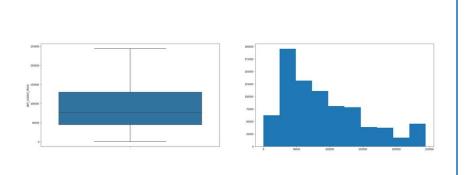


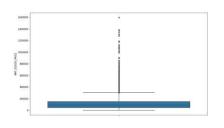


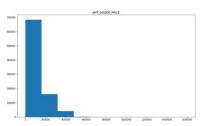












| AMT_DOWN_PAYMENT - | 1 | 0.78 | -0.079 | -0.08 | -0.0012 | -0.024 | -0.0068 | 0.25 | 0.2 | -0.15 | 0.052 | 4.07 | 0.25 |
|-----------------------------|--------------------|---------------------|--------------------|-----------------|-----------------------------|------------------|----------------------|-------------------|---------------|---------------|--------------|-----------------|-------------------|
| RATE_DOWN_PAYMENT - | | 1 | -0.15 | -0.15 | 0.011 | 0.041 | -0.0082 | -0.14 | -0.12 | -0.3 | -0.31 | -0.21 | -0.14 |
| DAYS_TERMINATION - | -0.079 | -0.15 | | 0.96 | | | -0.35 | 0.14 | -0.059 | 0.0077 | | | 0.14 |
| DAYS_LAST_DUE - | -0.08 | -0.15 | 0.96 | | | | -0.31 | | -0.062 | 0.0087 | | 0.48 | 0.14 |
| DAYS_LAST_DUE_1ST_VERSION - | -0.0012 | -0.011 | | | | | -0.74 | 0.0018 | -0.18 | -0.37 | 0.0057 | | 0.0018 |
| DAYS_FIRST_DUE - | -0.024 | -0.041 | | | | | -0.04 | -0.034 | 0.12 | -0.22 | 40.03 | | -0.034 |
| DAYS_FIRST_DRAWING - | -0.0068 | -0.0082 | -0.35 | -0.31 | -0.74 | -0.04 | 1 | -0.024 | | | -0.025 | -0.15 | -0.024 |
| AMT_GOODS_PRICE - | | -0.14 | | | 0.0018 | -0.034 | -0.024 | 1 | | | 0.97 | 0.24 | 1 |
| AMT_ANNUITY - | | -0.12 | -0.059 | -0.062 | -0.18 | -0.12 | | 0.75 | | 0.029 | | | 0.75 |
| CNT_PAYMENT - | -0.15 | 0.3 | 0.0077 | 0.0087 | -0.37 | -0.22 | 0.28 | 0.43 | 0.029 | 1 | | 0.045 | 0.43 |
| AMT_CREDIT - | 0.052 | -0.31 | | | 0.0057 | -0.03 | -0.025 | 0.97 | | 0.5 | 1 | 0.26 | 0.97 |
| DAYS_DECISION - | -0.07 | -0.21 | | | | | -0.15 | 0.24 | 0.17 | 0.045 | 0.26 | 1 | 0.24 |
| AMT_APPLICATION - | | -0.14 | | | 0.0018 | -0.034 | -0.024 | 1 | 0.75 | 0.43 | 0.97 | 0.24 | 1 |
| | AMT_DOWN_PAYMENT - | RATE_DOWN_PAYMENT - | DAYS_TERMINATION - | DAYS_LAST_DUE - | DAYS_LAST_DUE_1ST_VERSION - | DAYS_FIRST_DUE - | DAYS_FIRST_DRAWING - | AMT_GOODS_PRICE - | AMT_ANNUITY - | CUT_PAYMENT - | AMT_CREDIT - | DAYS_DECISION - | AMT_APPLICATION - |

- 0.3

-0.3

-0.6

| AMT_DOWN_PAYMENT | 1 | 0.42 | 0.54 | 0.32 | 0.095 | 0.39 | 0.0079 | 054 |
|--------------------------|------------------|-------------------|-----------------|-------------|-------------|------------|---------------|-----------------|
| RATE_DOWN_PAYMENT AMT_DO | | | -0.024 | -0.084 | -0.24 | -0.11 | -0.22 | -0.024 |
| AMT_GOODS_PRICE RATE_DO | | -0.024 | | 0.82 | 0.66 | 0.99 | 0.33 | 1 |
| AMT_ANNUITY AMT_C | 032 | -0.084 | 0.82 | | | 0.82 | 0.29 | 082 |
| CNT_PAYMENT | 0.095 | -0.24 | | 0.43 | 1 | 0.68 | 0.31 | 0.67 |
| AMT_CREDIT | 039 | -0.11 | 0.99 | 0.82 | 0.68 | 1 | 0.27 | 098 |
| DAYS_DECISION | 0.0079 | -0.22 | 0.33 | 0.29 | 0.31 | 0.27 | 1 | 0.26 |
| AMT_APPLICATION DA | 0.54 | -0.024 | 1 | 0.82 | 0.67 | 0.98 | 0.26 | 1 |
| AMT | AMT_DOWN_PAYMENT | RATE_DOWN_PAYMENT | AMT_GOODS_PRICE | AMT_ANNUITY | CNT_PAYMENT | AMT_CREDIT | DAYS_DECISION | AMT_APPLICATION |

- 0.75 - 0.50 - 0.25 - 0.00

