EDA –Bank Analysis DSC-27

By –Pramendra Pandey & Karan Babbar

DATA Set

- *'application_data.csv'* contains all the information of the client at the time of application.
  The data is about whether a **client has payment difficulties.**
- *'previous_application.csv'* contains information about the client's previous loan data. It contains the data whether the previous application had been **Approved, Cancelled, Refused or Unused offer.**
- *'columns_description.csv'* is data dictionary which describes the meaning of the variables.

Business Objectives

This case study aims to identify patterns which indicate if a client has difficulty paying their installments which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc. This will ensure that the consumers capable of repaying the loan are not rejected. Identification of such applicants using EDA is the aim of this case study.

In other words, the company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default. The company can utilize this knowledge for its portfolio and risk assessment.

To develop your understanding of the domain, you are advised to independently research a little about risk analytics - understanding the types of variables and their significance should be enough).

Analyzing Application Data

- Read the Application Data CSV file .
- It has rows count-307511 and 122 columns .
- Check for **quality control** of the Application Data file
  - It has missing values and Outliers .
  - Getting the Column name having more than 35 Percent missing Data
  - Removing Column from Data frame missing value more than 35 Percent
  - Now checking columns with less number of missing values impute them if necessary else treat them as missing .

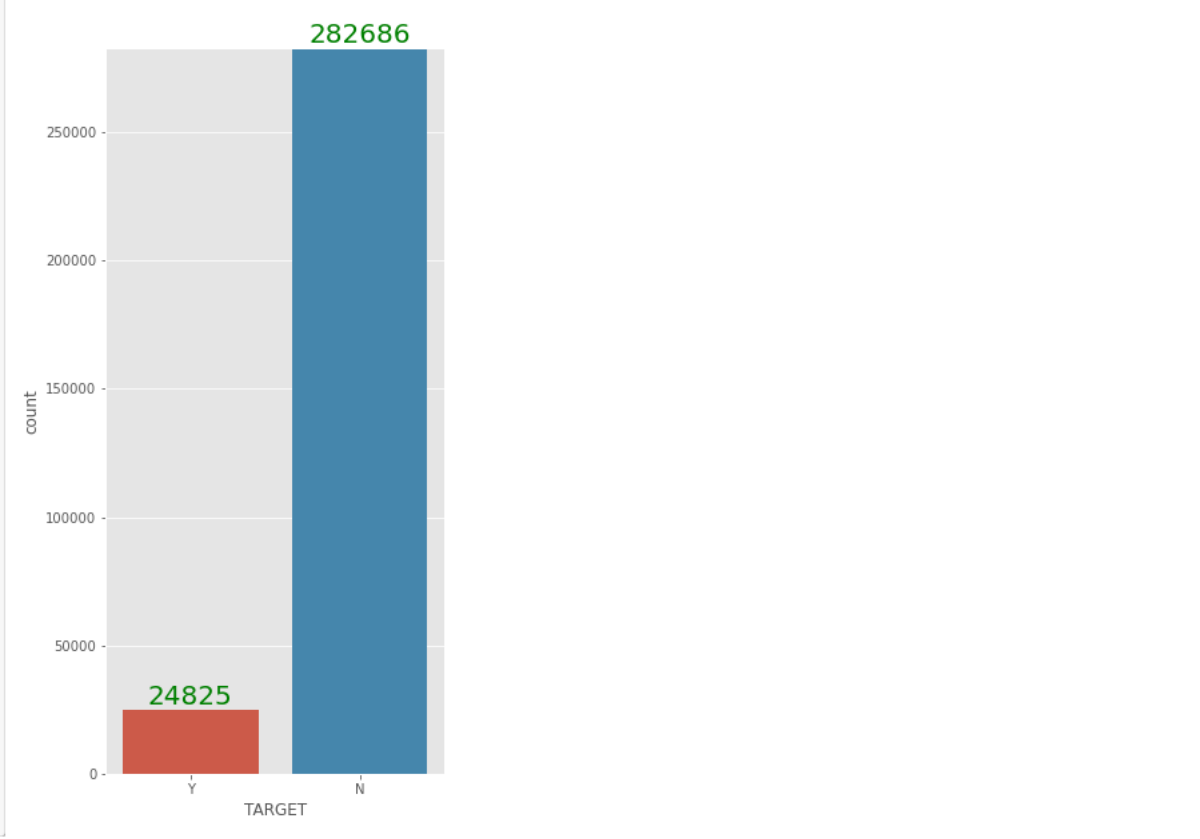
Analyzing Application Data

- **Quality control** of the Application Data file
  - There are many integer Column which just has two unique values .
  - Get the list of column having 2 unique values 0,1 .
  - Replace them as ‘N’ , ‘Y’ and change Data type to object .

Check for imbalance of data

- Number of People not paid on time 24825, Percentage of People who not paid on-time 8.07%
- Number of People paid on time 282686, Percentage of People who paid on-time 91.93% .
- Percentage Imbalance for Ratio 11.39

Imbalance Ratio –Non-Defaulter/Defaulter is 11.39%

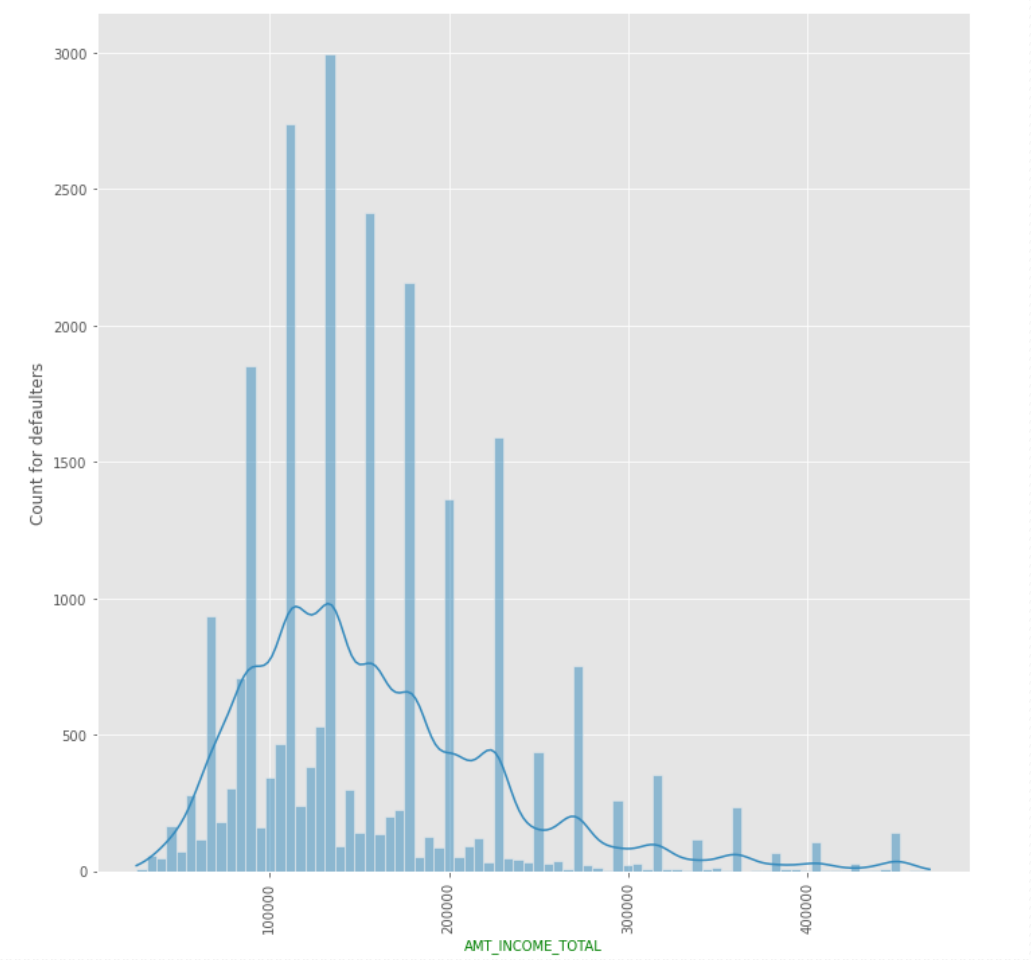


Inferences -It is Imbalaced dataset . There are more loan that were paid on time than not Paid on time .

Univariate Analysis & Bivariate Analysis Numerical columns/Categorical Columns

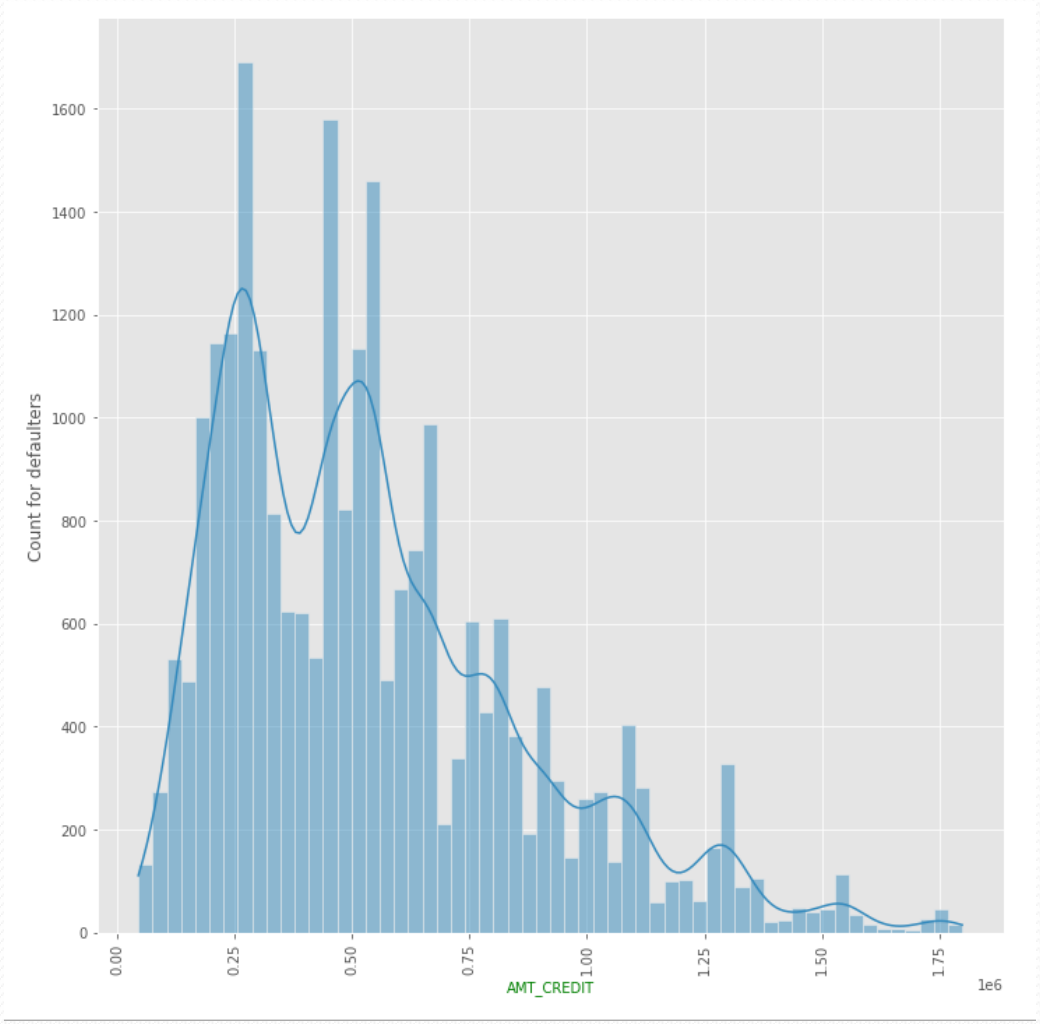
AMT_INCOME_TOTAL

For defaulters- Range for AMT_TOTAL_Income is between 100000 to 20000 and some point are also present at higher range.



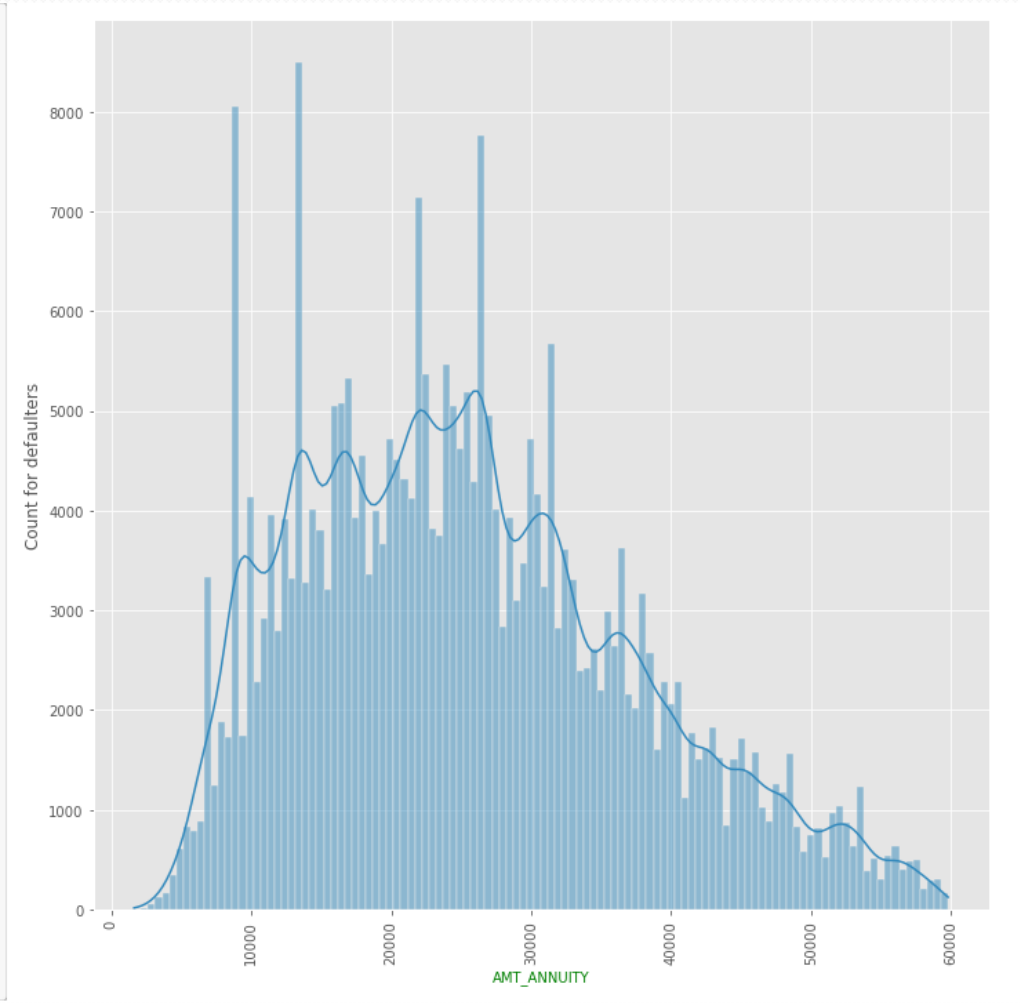
AMT_CREDIT

- From below, we can see for defaulters AMT_CREDIT lies between 200000 to 750000



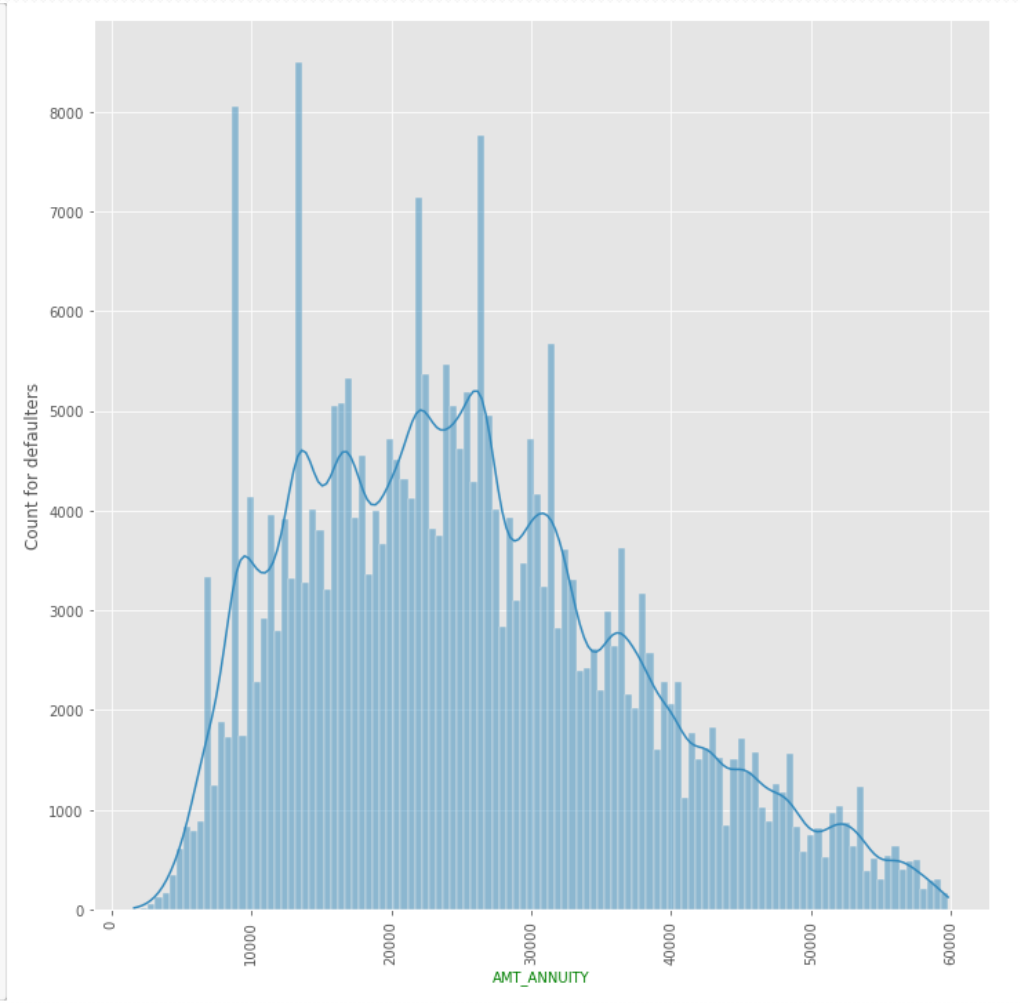
AMT_ANNUITY

- Amount of ANNUITY is on left side mostly for Defaulters lies between 10000 to 40000



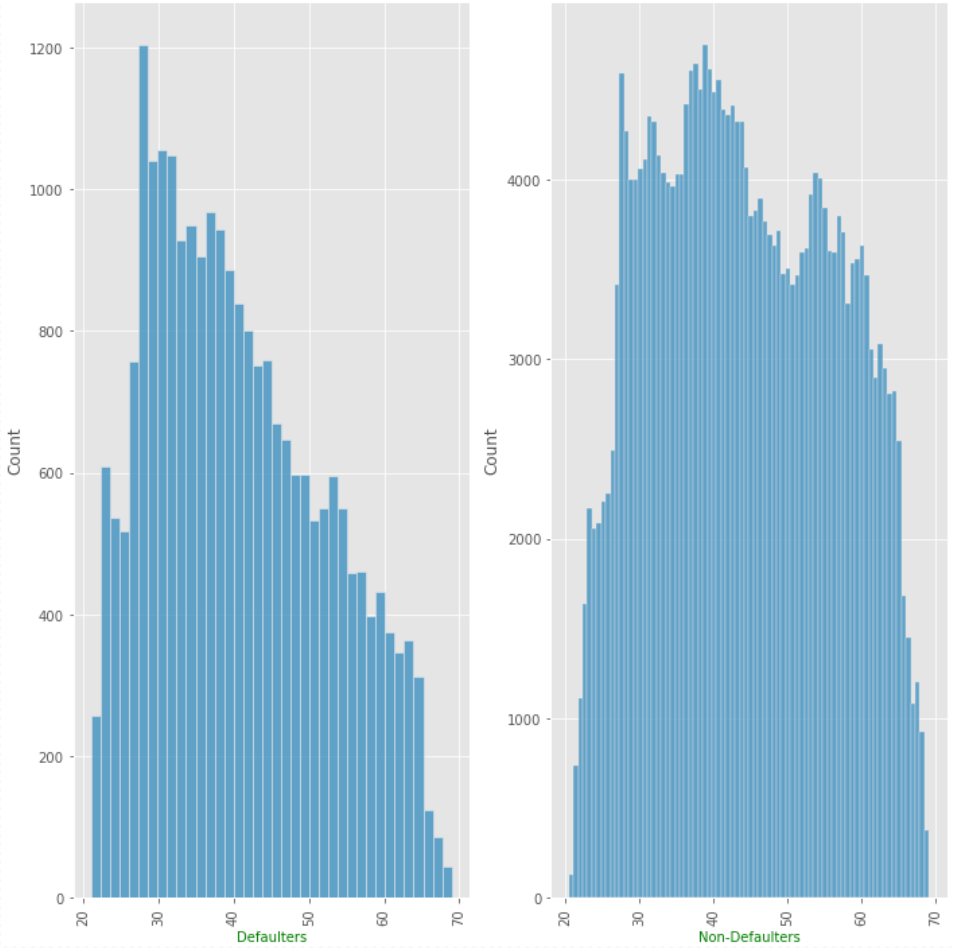
AMT_GOODS_PRICE

- Amount of ANNUITY is on left side mostly for Defaulters lies between 10000 to 40000



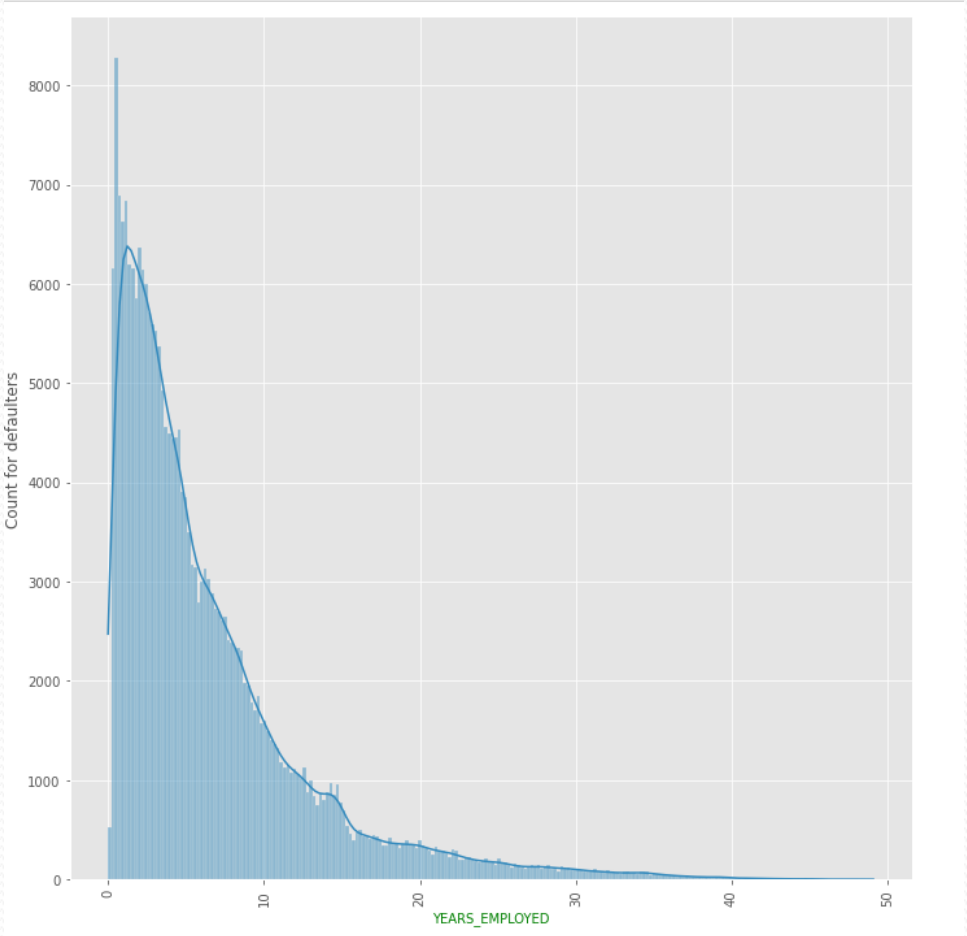
DAYS_BIRTH

- Converting Column 'DAYS_BIRTH' to New Column 'AGE‘
- In case of Both Defaulters and Non-Defaulters age seems to be fairly distributed
- We noticed by Age, defaulter count is decreasing , looks like with Age people default less .



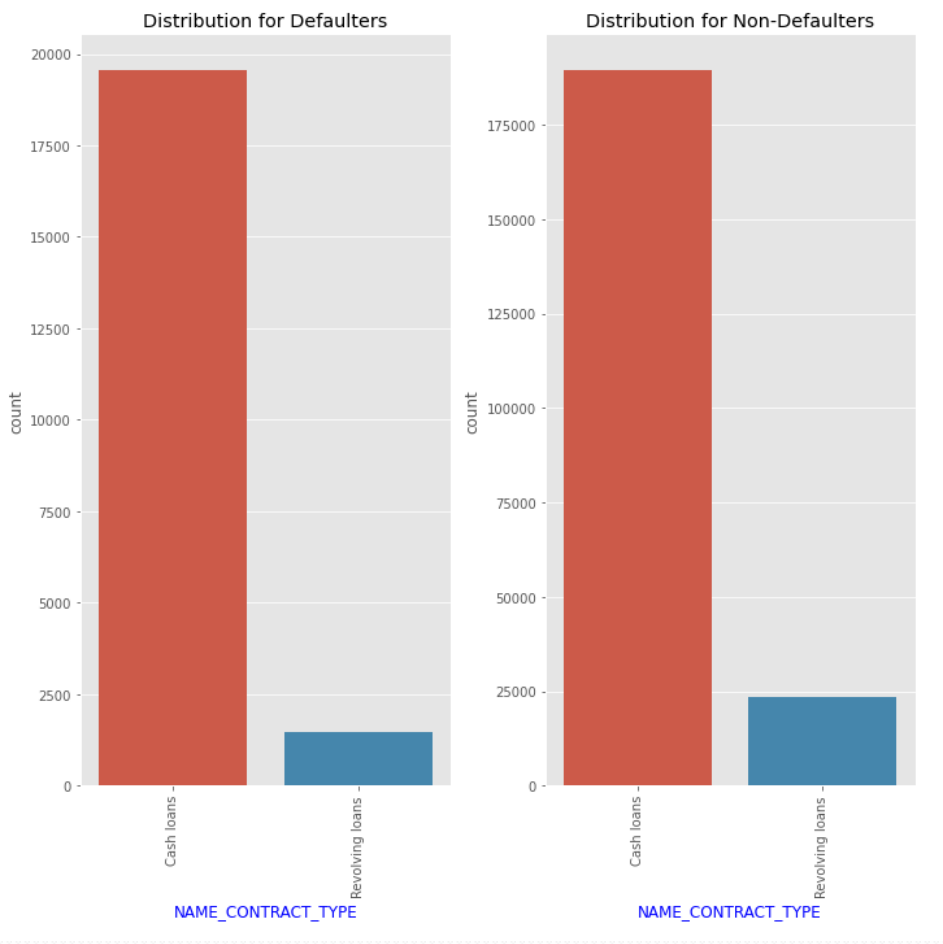
DAYS_EMPLOYED

- Converting Column DAYS_EMPLOYED ' to New Column YEARS_EMPLOYED‘
- Graph count/density is aligned left -Lots of people are having low work experience who are defaulters almost o year of service



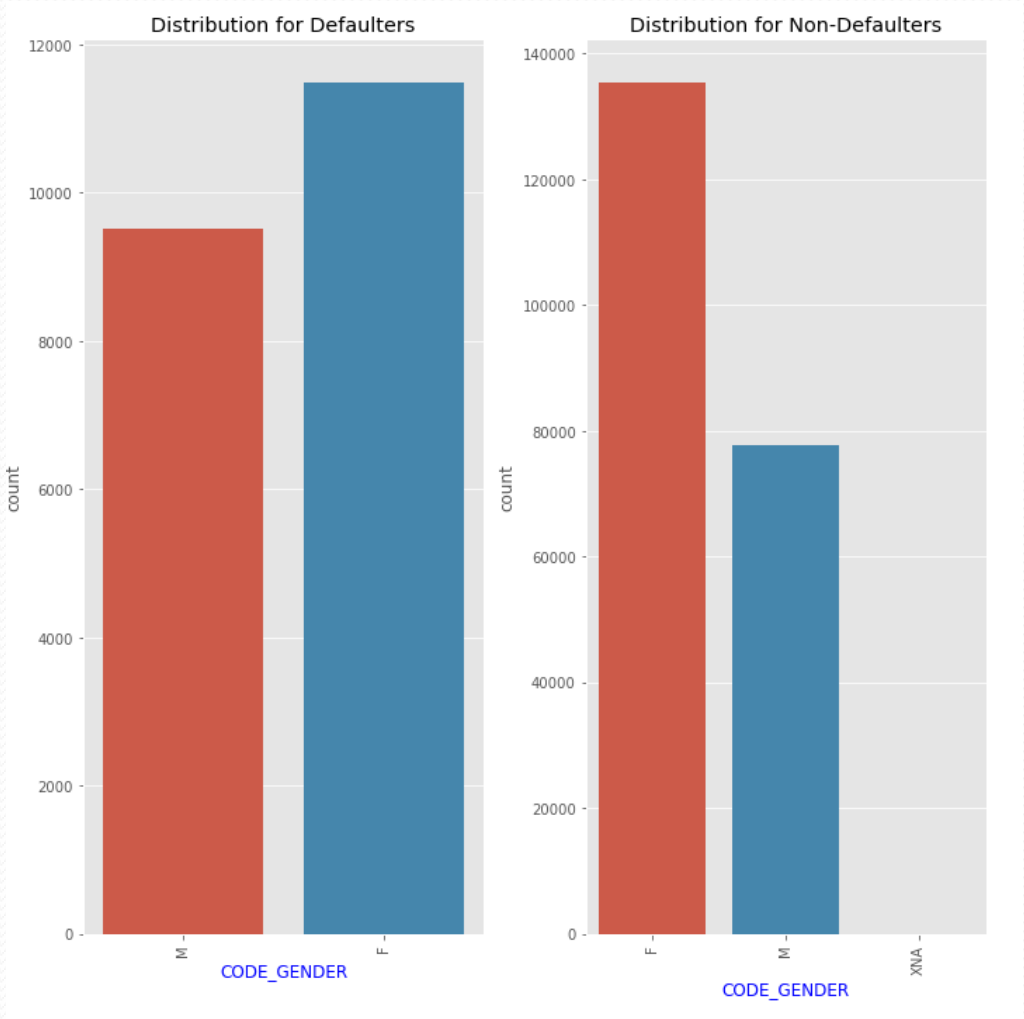
NAME_CONTRACT_TYPE

- Cash loan is way higher than revolving loans for both defaulters and non –defaulters.



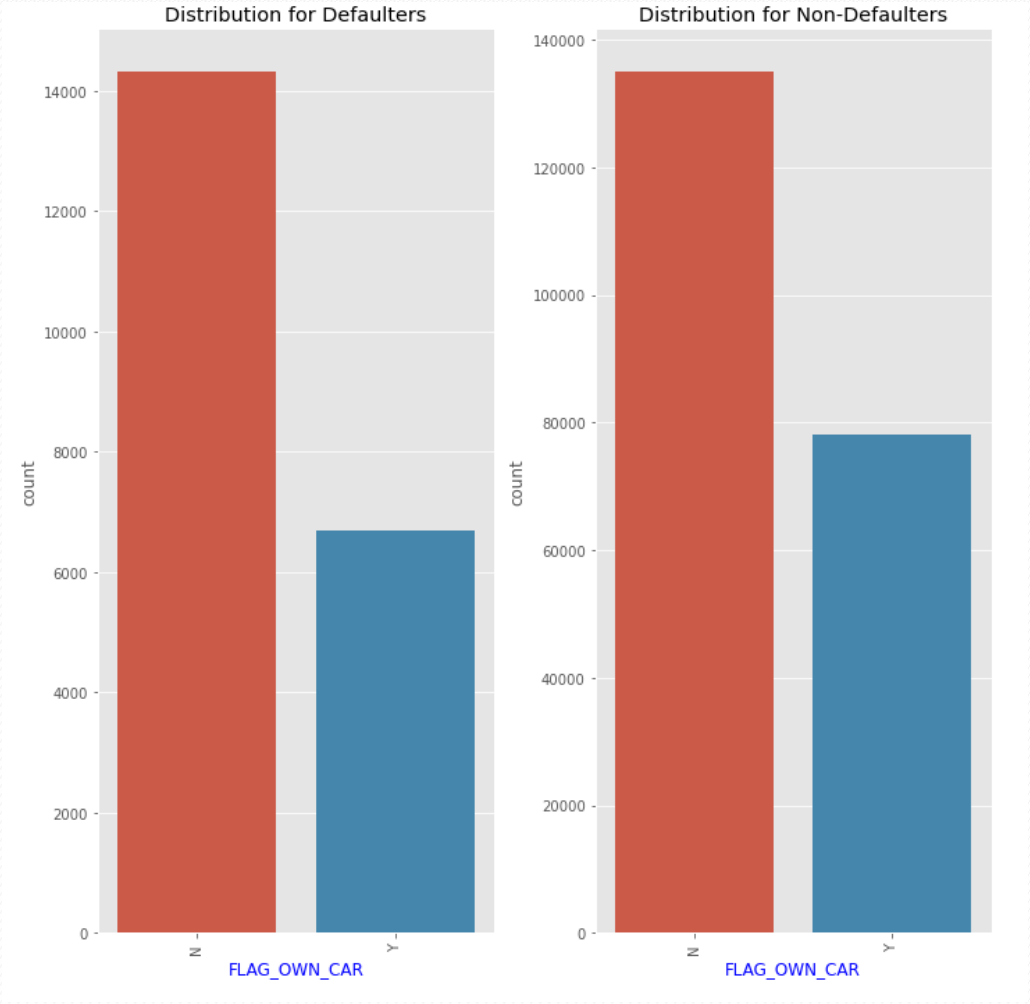
CODE_GENDER

- For both defaulter and non-defaulter , Females are taking more loan comparison to Man and Defaulting more too.



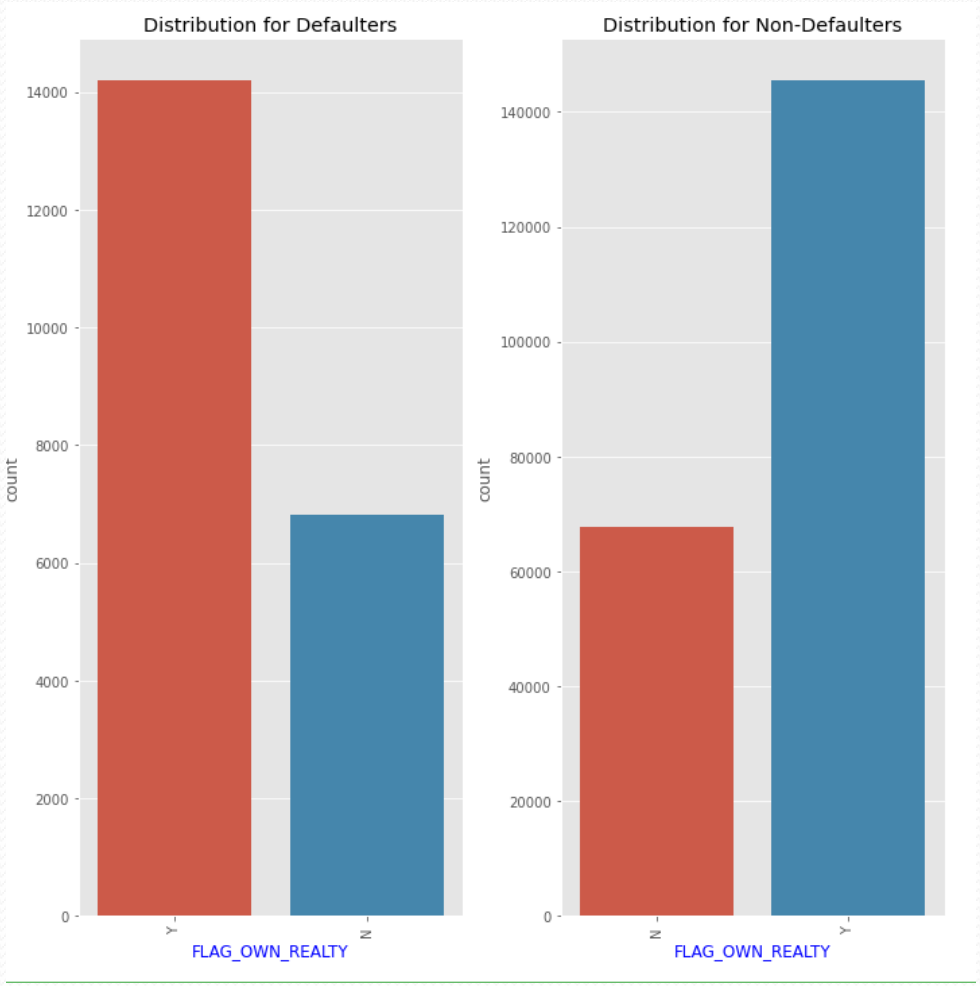
FLAG_OWN_CAR

- For Defaulters ,people not having Car tend to default more , ratio is high than people own Car- Owning Car can be used as status symbol .



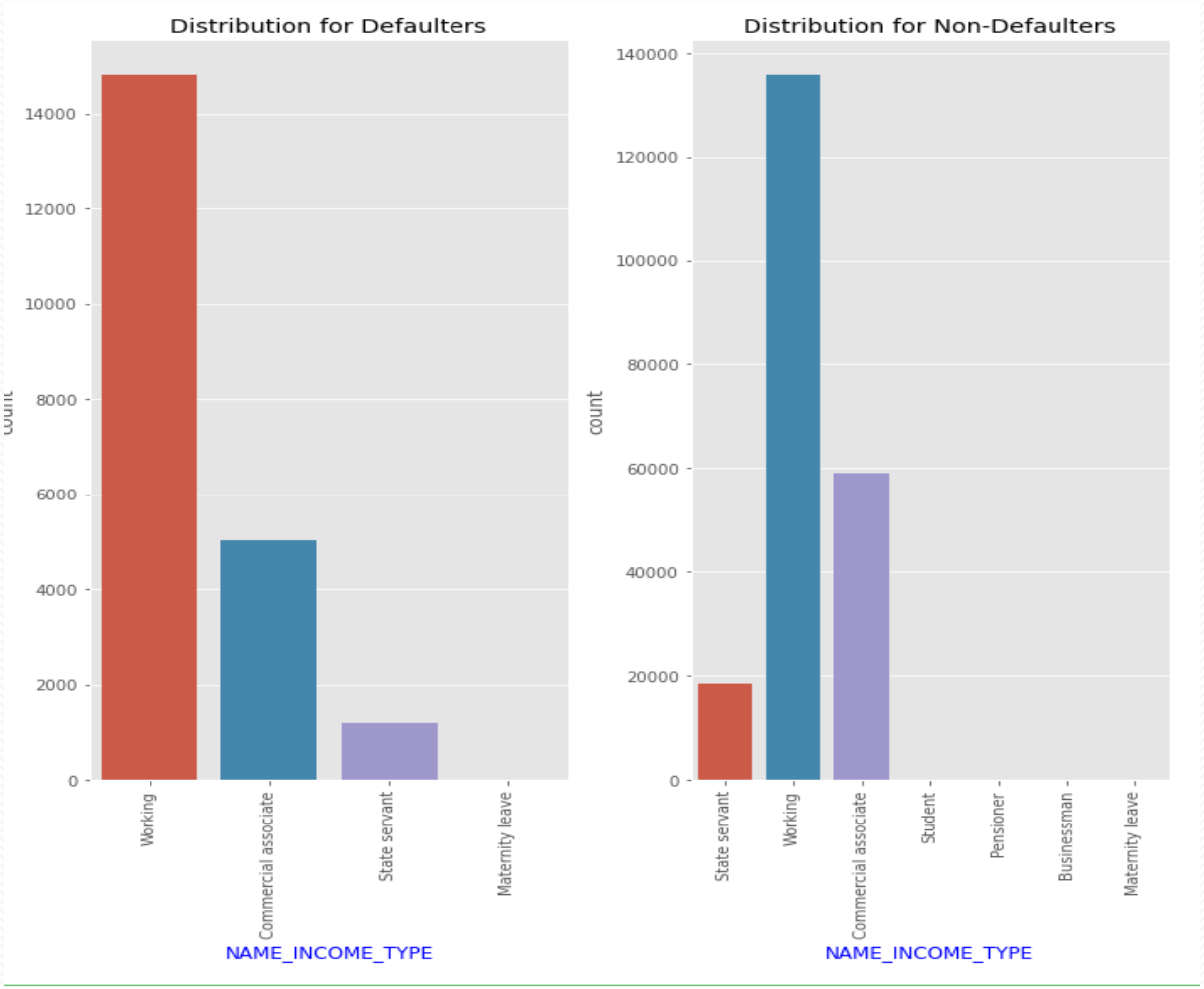
FLAG_OWN_REALTY

- In both category, people are own houses i.e. are in good numbers .



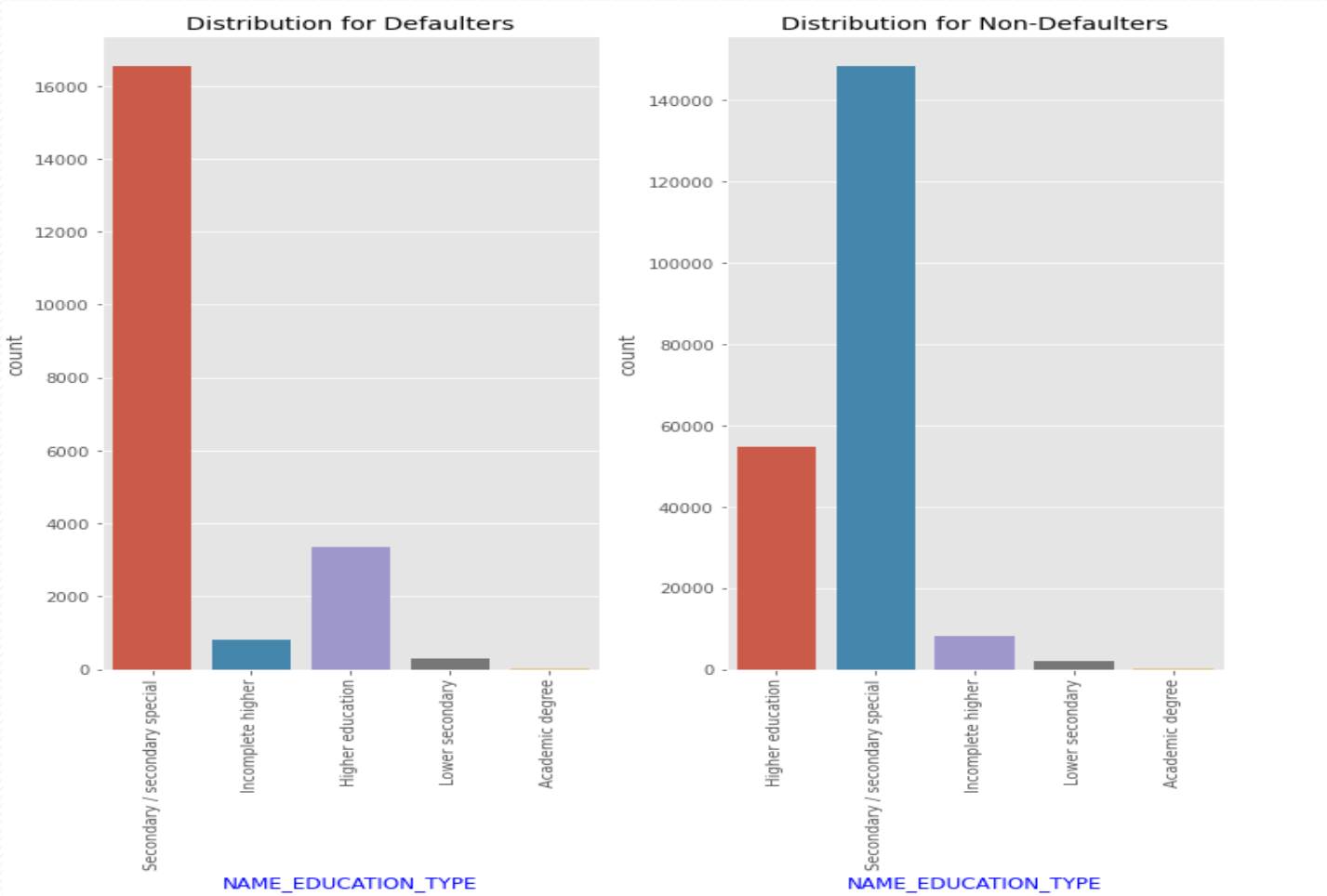
NAME_INCOME_TYPE

- Commercials Associate and State servant are in good numbers in case of non-defaulters .



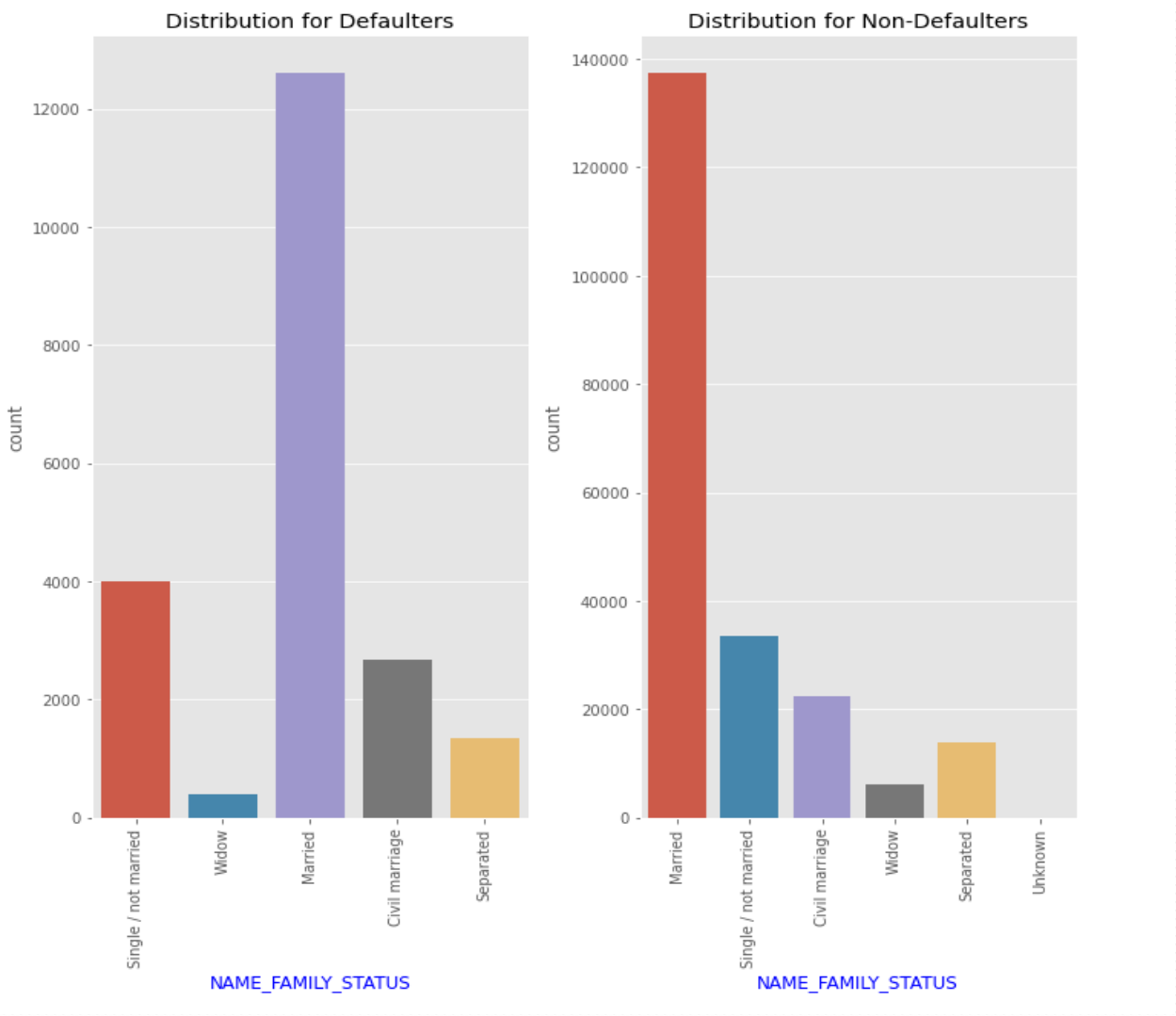
NAME_EDUCATION_TYPE

- People with Academic degree takes loan rarely and default rarely too -Good to target them
- People with Secondary/secondory special are getting more loan and defaulting more.
- People with higher education having less difficulty paying them .



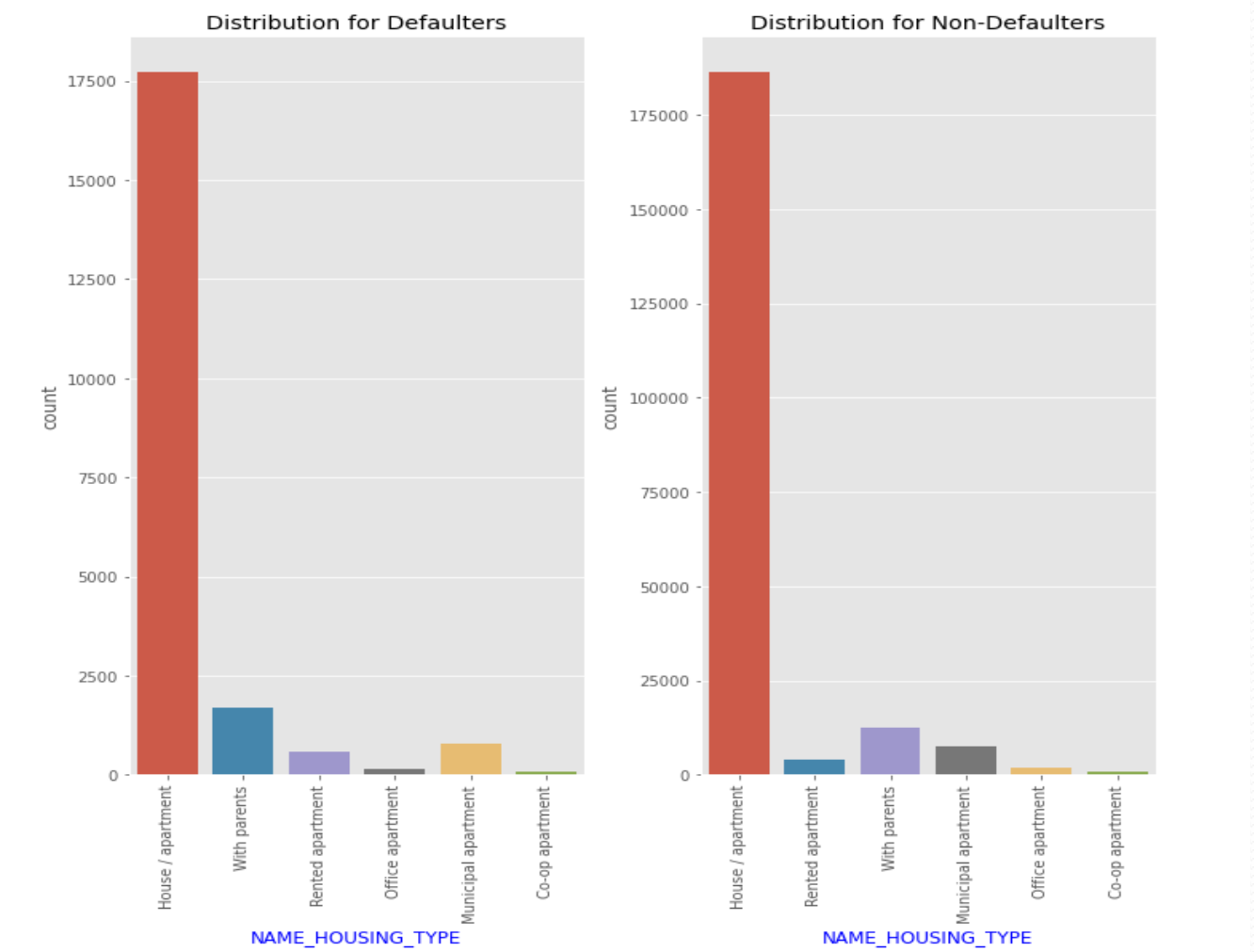
NAME_FAMILY_STATUS

- Single/Not Married People having difficulties in paying loans



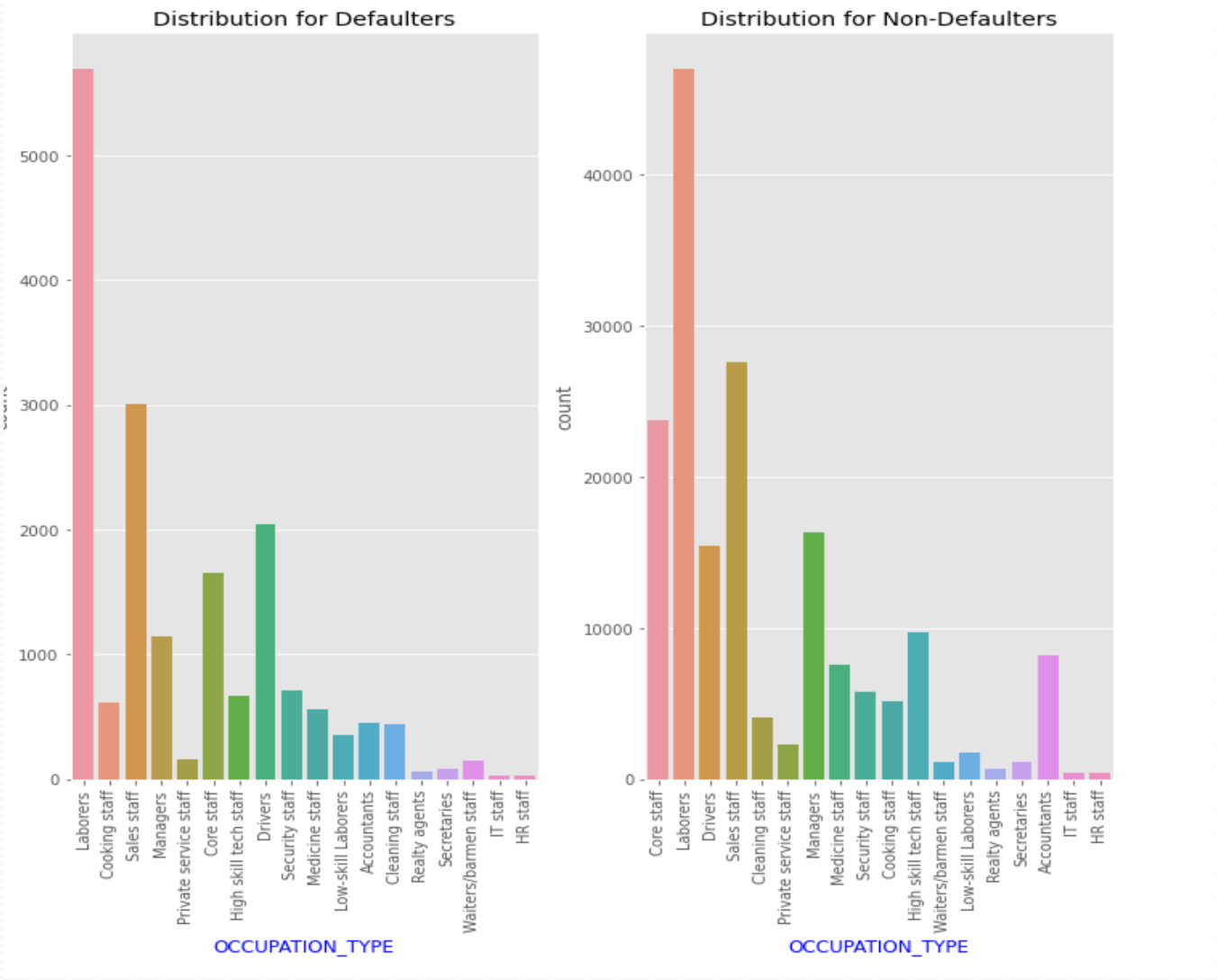
NAME_HOUSING_TYPE

- Most people lives in House/Apartments for both Defaulters and non-Defaulters .
- People living with Parents having difficulties in paying loans .



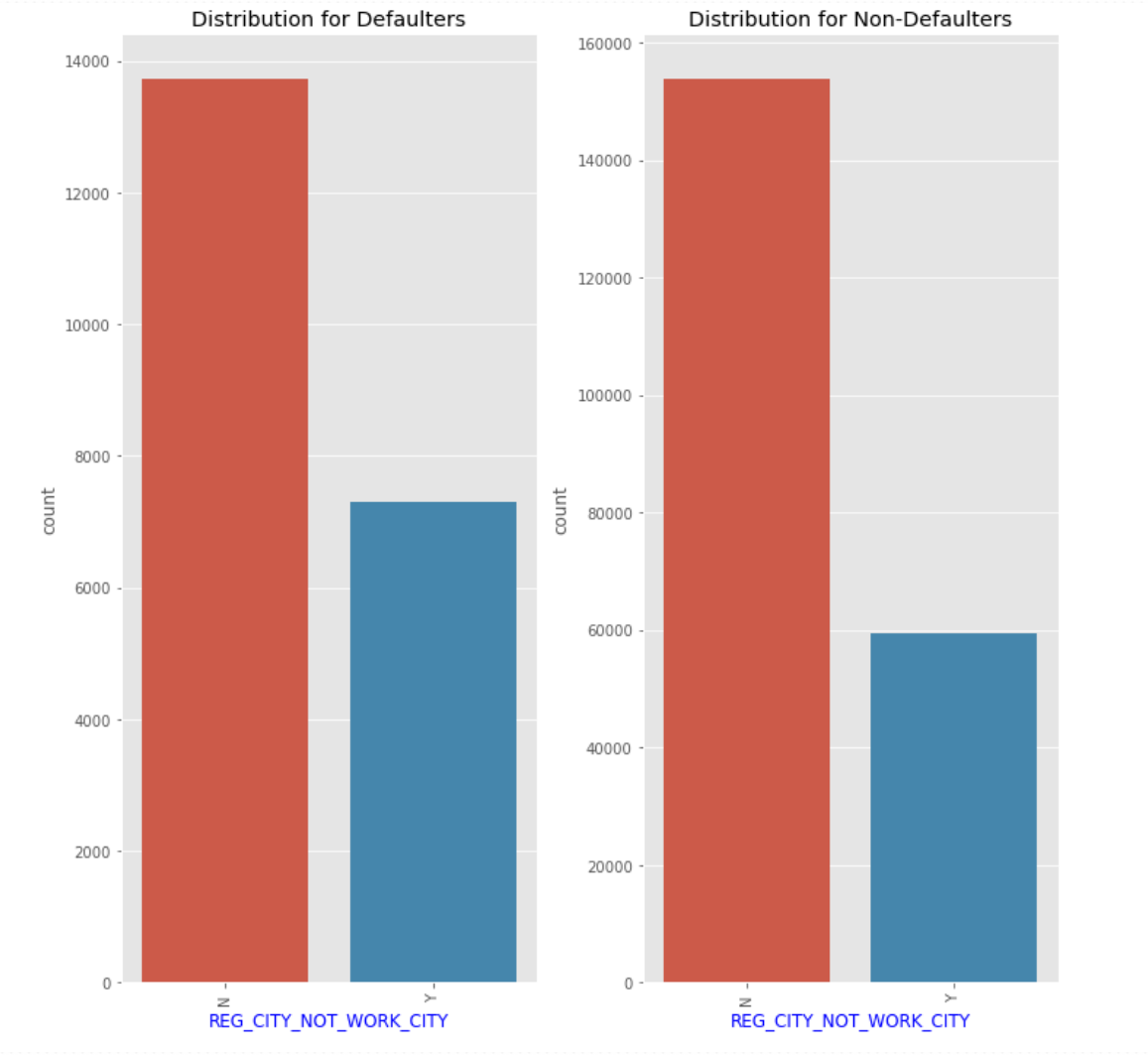
OCCUPATION_TYPE

- Laborers/Low skilled workers and Drivers are having difficulties while paying loan .



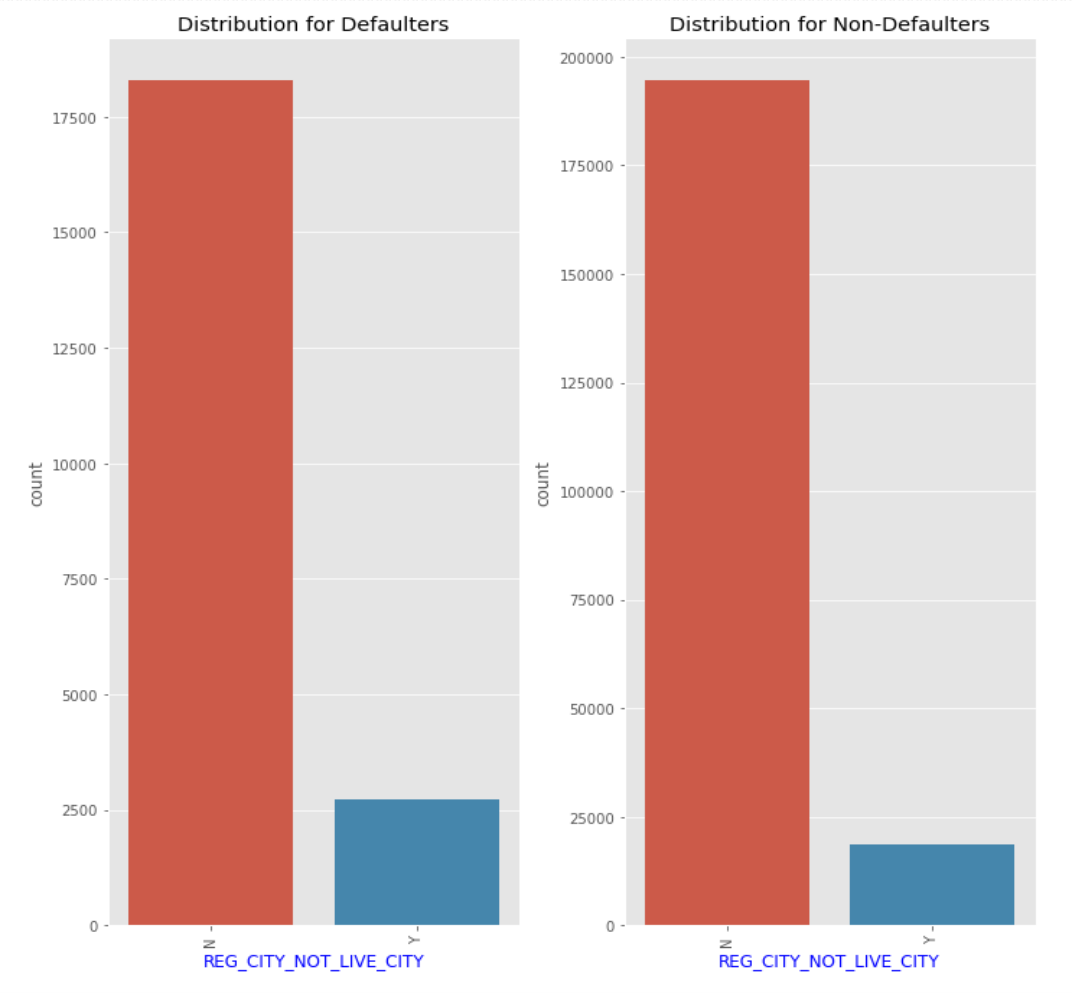
REG_CITY_NOT_WORK_CITY

- People whose reg city is not same as work city facing difficulties while paying loans

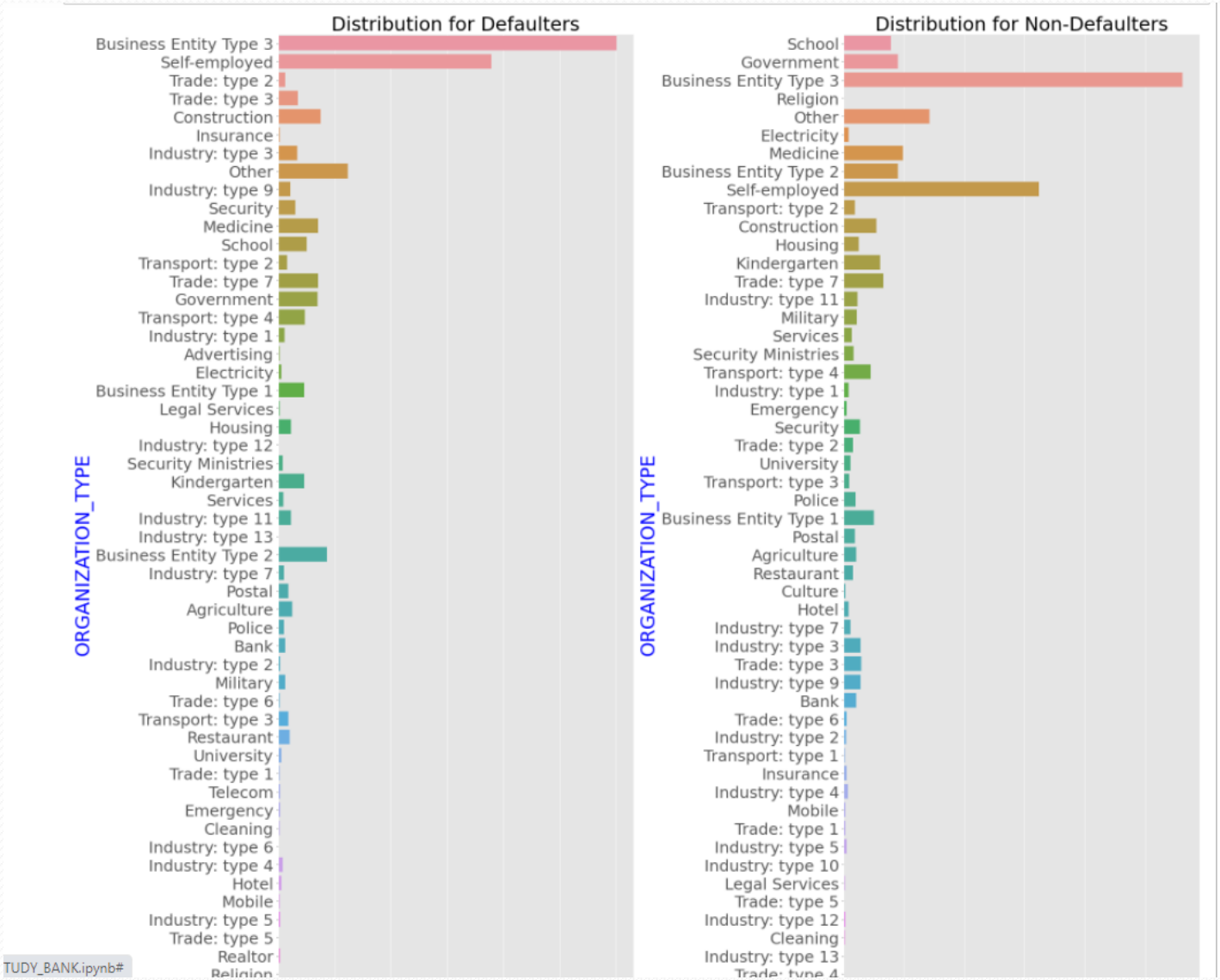


REG_CITY_NOT_LIVE_CITY

- People whose Reg city is not same as live city or work city -are facing problems while paying loans



ORGANIZATION_TYPE

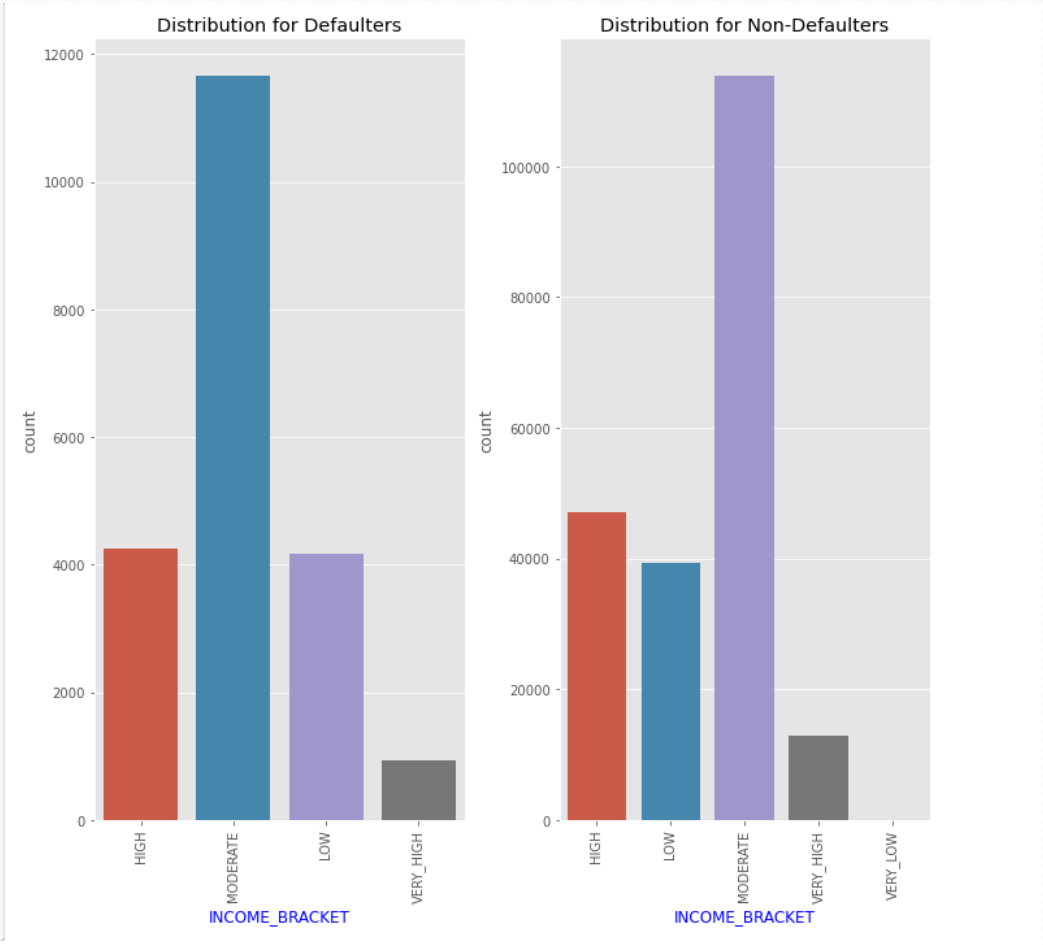


Binning of Numeric Variables-

AMT_INCOME_TOTAL, AMT_CREDIT, EXT_SOURCE_2, EXT_SOURCE_3, AGE, YEARS_EMPLOYED,

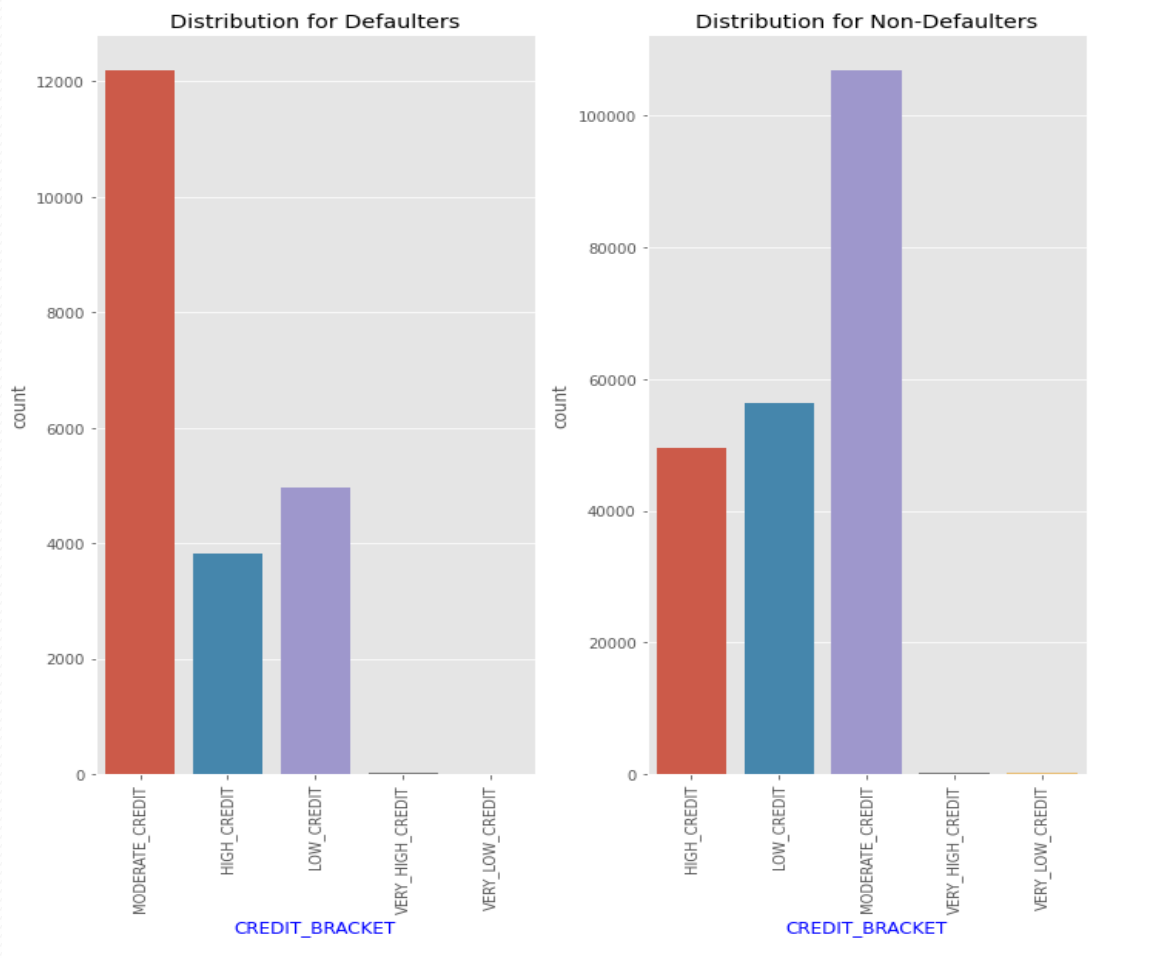
AMT_INCOME_TOTAL-Bins to INCOME_BRACKET

- People/client with Very High income less likely to default Loan .
- People with high/moderate income exist both places defaulter and non-defaulter , this is range/bracket who is applying for loan .



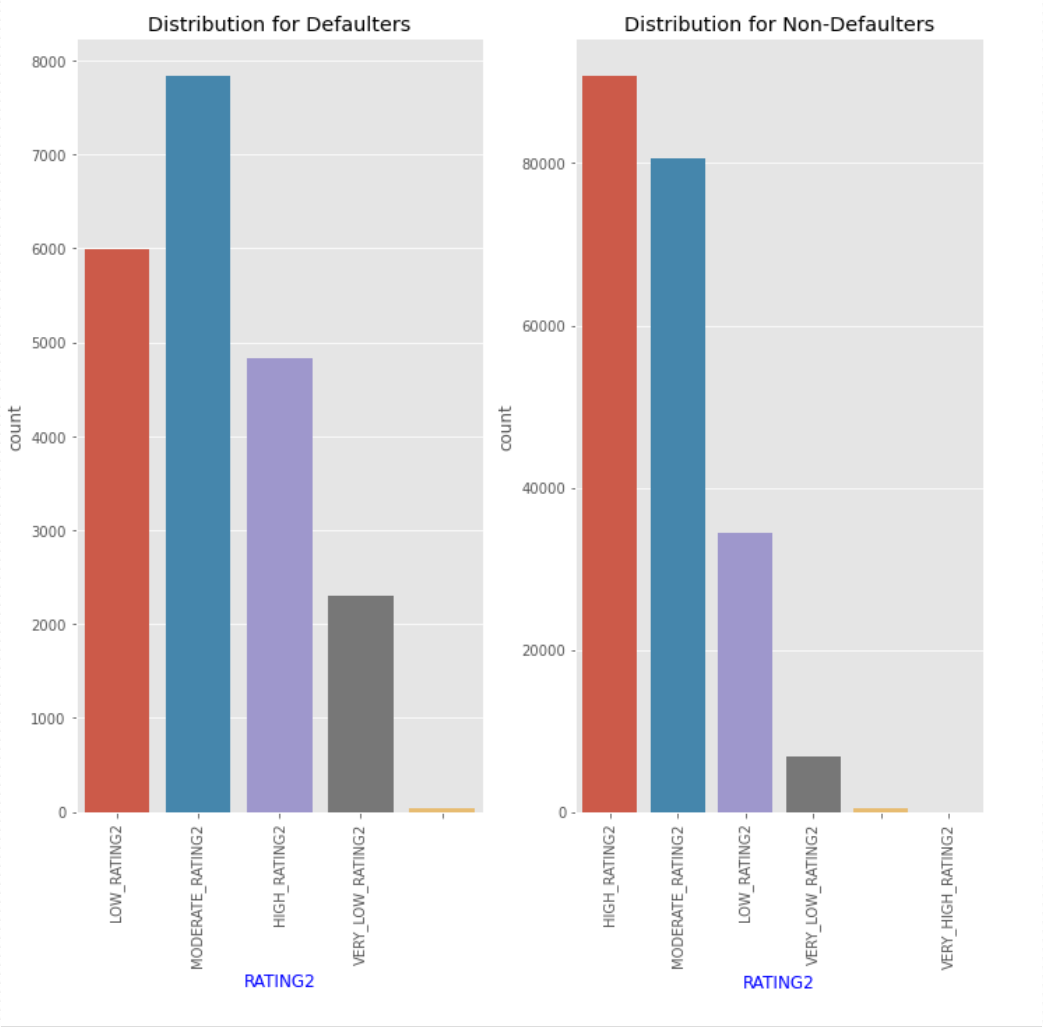
AMT_CREDIT -Bins to CREDIT_BRACKET

- At both extreme very high and very low count is less . Looks like at extreme we are ok . Moderate/High is bracket where we need to pay attention because in this segment we are crediting loan more and defaulters count is high too .



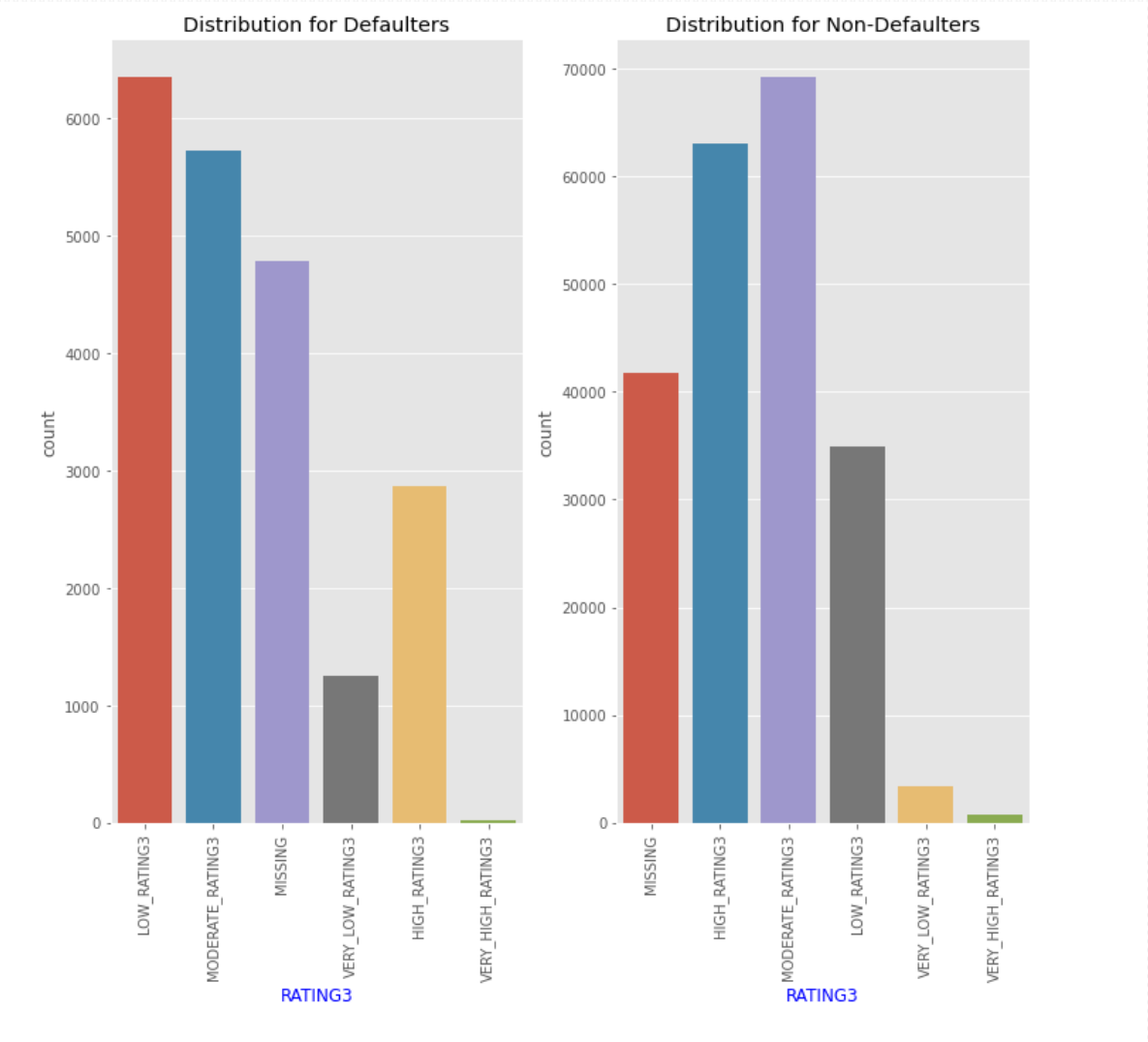
EXT_SOURCE_2 -Bins to RATING2

- Good Rating means chances for defaulting is less . People with high credit default less .



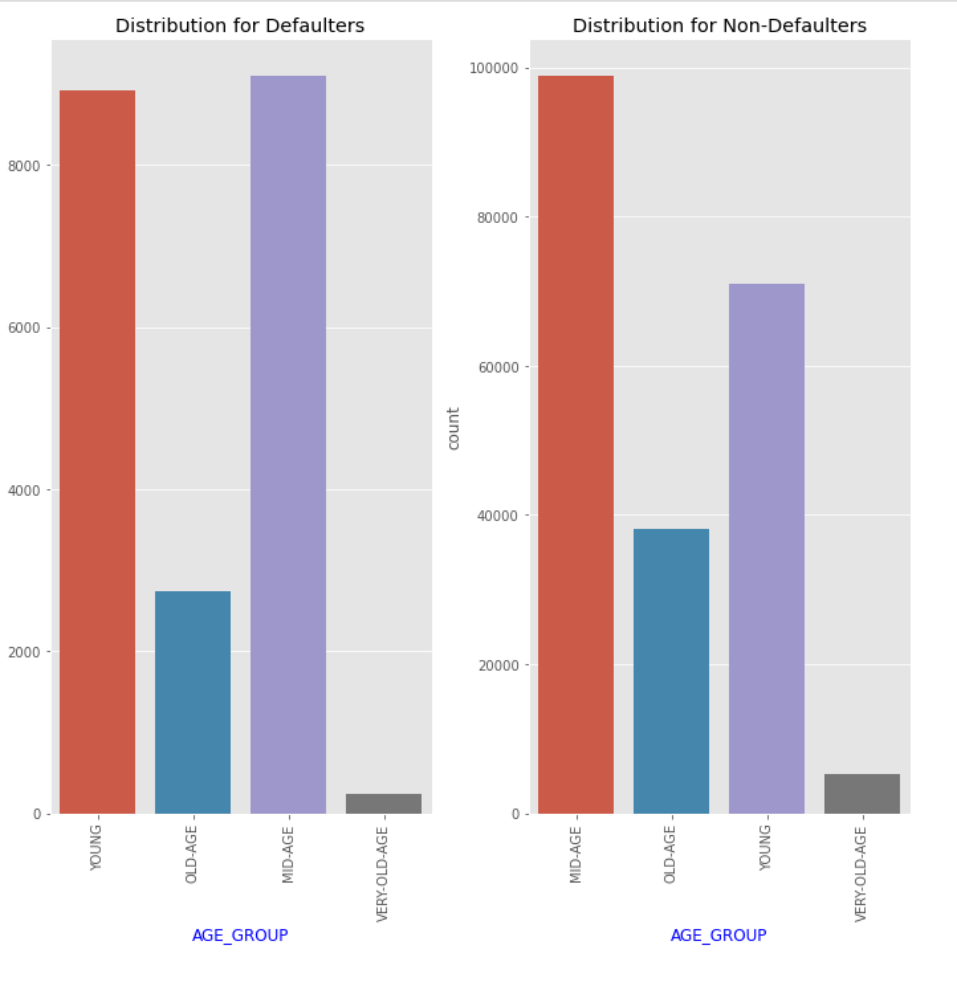
EXT_SOURCE_3 -Bins to RATING3

- Good Rating means chances for defaulting is less . People with high credit default less .



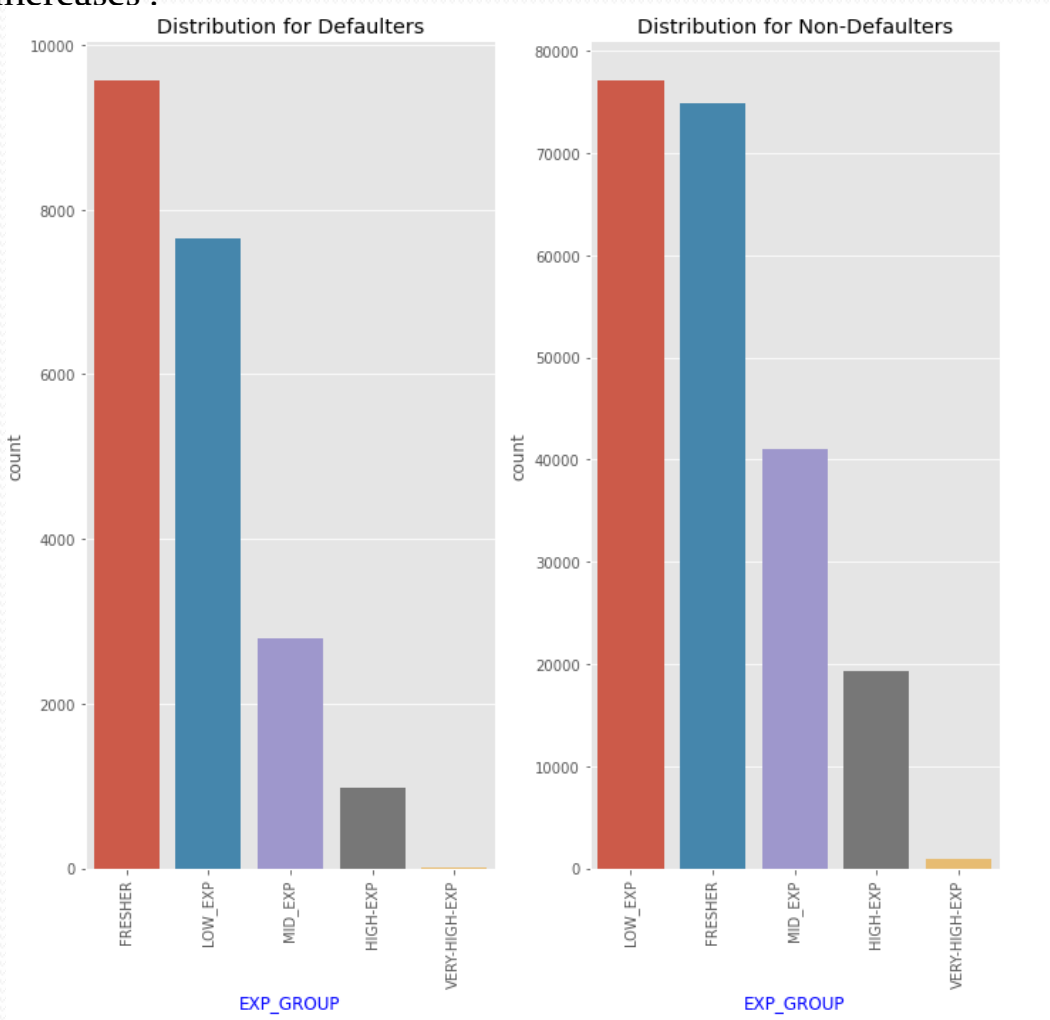
AGE-Bins to AGE_GROUP

- Young/Mid Age people count are high at both places (Defaulter/Non-Defaulter)
- Very Old Age(AGE>60) people do not apply for loan much .



YEARS_EMPLOYED -Bins to EXP_GROUP

- Very High experience (exp. >35) people is not defaulting the loan –Good Client .
- At high exp.(15 to 35) count is high in non-defaulters category means with high exp. Loan re-paying capability increases .



Heat Map –Correlations

Heat Map

- Select the Numeric and Float column for corr() functions .
- Divide Data frame into Defaulter and Non-Defaulters .
- Get Top 10 Positive and Negative correlations for Defaulters .
- Get Top 10 Positive and Negative correlations for Non-Defaulters .
- Heat Map for Defaulter and Non-Defaulters .

Heat Map- Defaulters

Top 10 most positive correlations for Defaulters

HOUR_APPR_PROCESS_START	HOUR_APPR_PROCESS_START	1.000000
AMT_GOODS_PRICE	AMT_CREDIT	0.978074
REGION_RATING_CLIENT	REGION_RATING_CLIENT_W_CITY	0.958724
CNT_FAM_MEMBERS	CNT_CHILDREN	0.893337
AMT_CREDIT	AMT_ANNUITY	0.734292
AMT_ANNUITY	AMT_GOODS_PRICE	0.733684
	AMT_INCOME_TOTAL	0.382066
AMT_GOODS_PRICE	AMT_INCOME_TOTAL	0.313932
AMT_CREDIT	AMT_INCOME_TOTAL	0.310059
YEARS_EMPLOYED	AGE	0.307084

dtype: float64

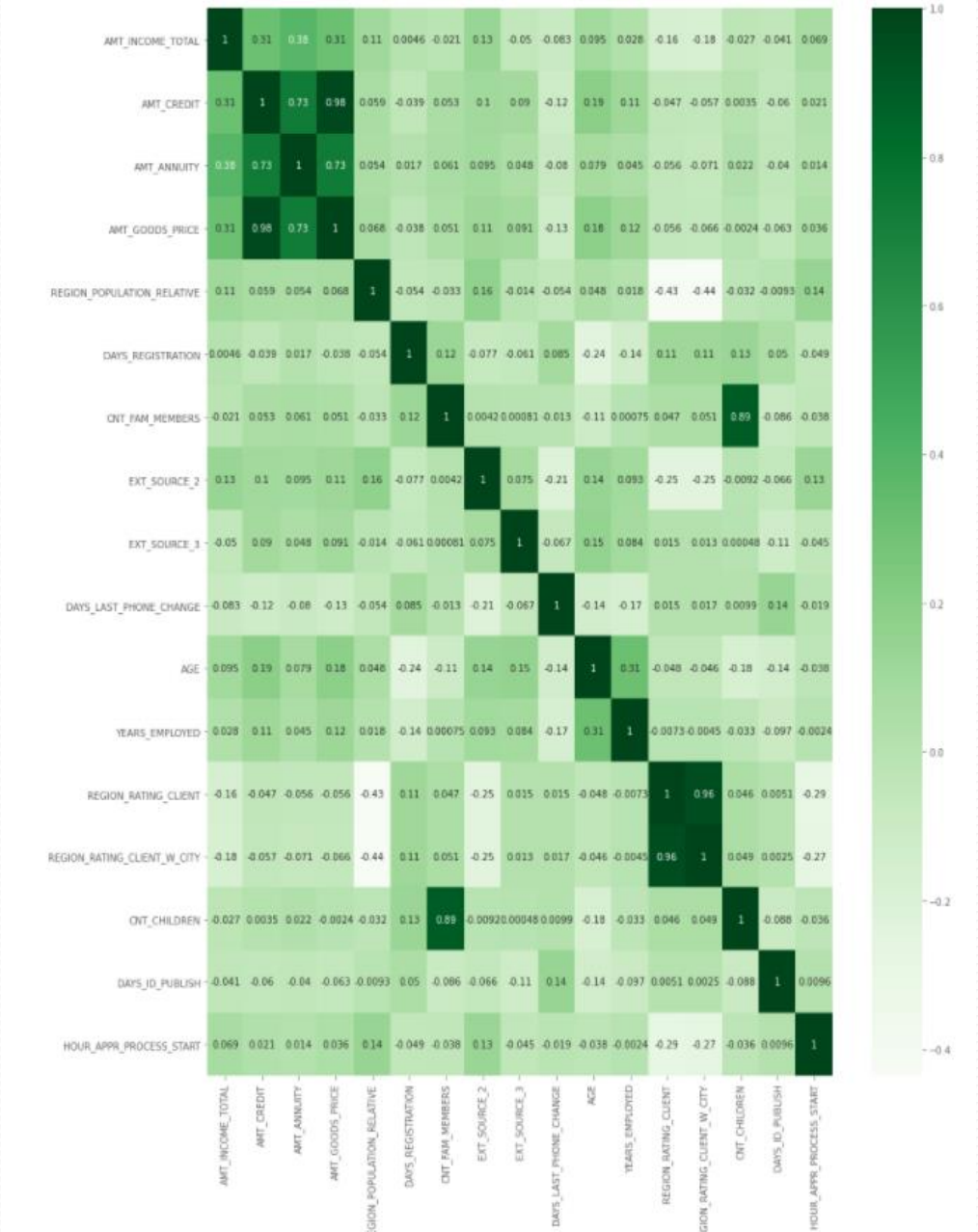
Heat Map- Defaulters

Top 10 most Negative correlations for Defaulters

AGE	CNT_CHILDREN	-0.175170
AMT_INCOME_TOTAL	REGION_RATING_CLIENT_W_CITY	-0.177599
DAYS_LAST_PHONE_CHANGE	EXT_SOURCE_2	-0.210719
AGE	DAYS_REGISTRATION	-0.241940
REGION_RATING_CLIENT_W_CITY	EXT_SOURCE_2	-0.245006
REGION_RATING_CLIENT	EXT_SOURCE_2	-0.245354
REGION_RATING_CLIENT_W_CITY	HOUR_APPR_PROCESS_START	-0.272265
REGION_RATING_CLIENT	HOUR_APPR_PROCESS_START	-0.289965
REGION_POPULATION_RELATIVE	REGION_RATING_CLIENT	-0.431929
	REGION_RATING_CLIENT_W_CITY	-0.436165

dtype: float64

Heat Map- Defaulters



Heat Map- Non-Defaulters

Top 10 most positive correlations for Non-Defaulters

HOUR_APPR_PROCESS_START	HOUR_APPR_PROCESS_START	1.000000
AMT_GOODS_PRICE	AMT_CREDIT	0.981295
REGION_RATING_CLIENT_W_CITY	REGION_RATING_CLIENT	0.950040
CNT_CHILDREN	CNT_FAM_MEMBERS	0.892382
AMT_CREDIT	AMT_ANNUITY	0.748188
AMT_GOODS_PRICE	AMT_ANNUITY	0.747675
AMT_ANNUITY	AMT_INCOME_TOTAL	0.398758
YEARS_EMPLOYED	AGE	0.353266
AMT_GOODS_PRICE	AMT_INCOME_TOTAL	0.322008
AMT_INCOME_TOTAL	AMT_CREDIT	0.317269
dtype: float64		

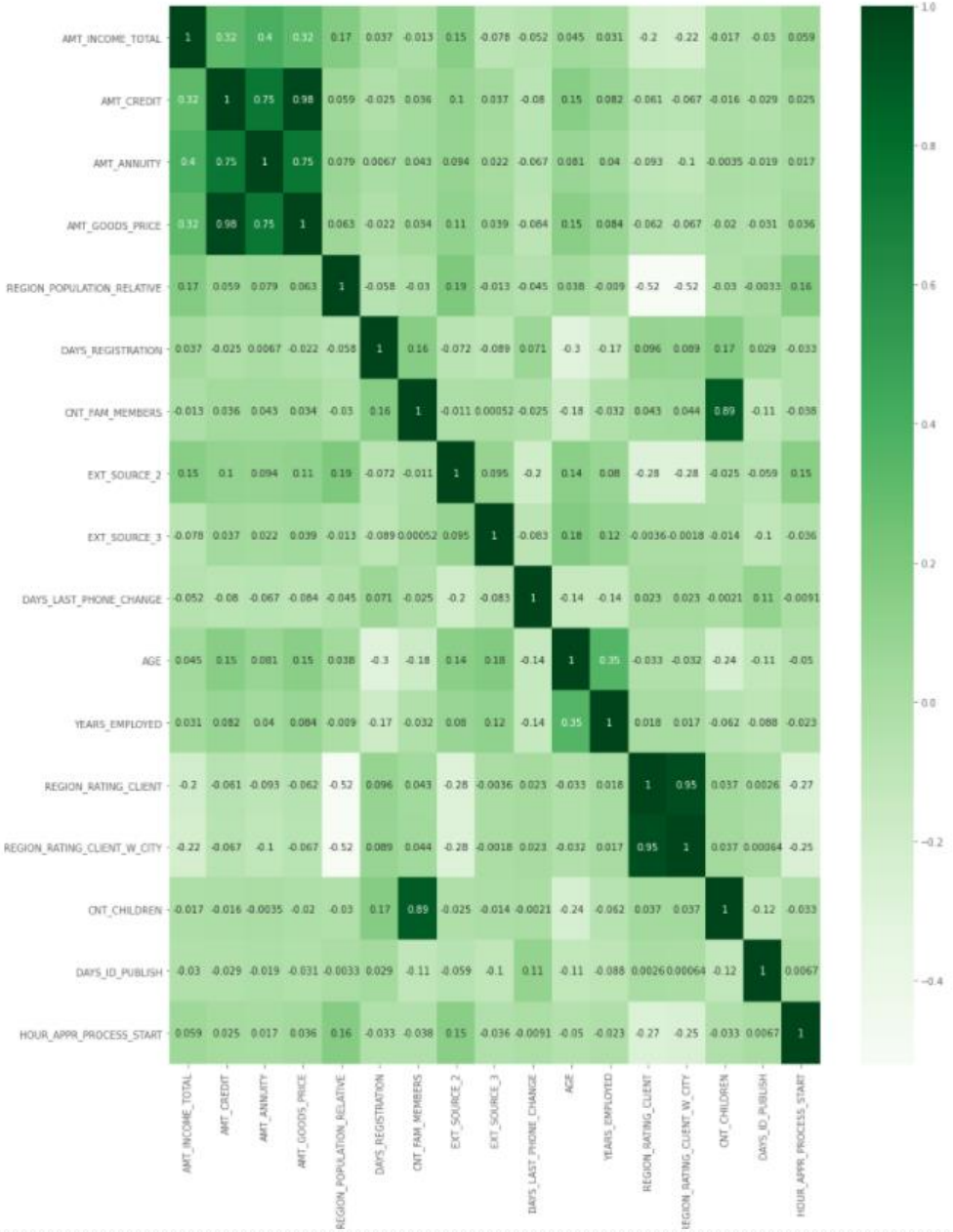
Heat Map- Non-Defaulters

Top 10 most Negative correlations for Non-Defaulters

AMT_INCOME_TOTAL	REGION_RATING_CLIENT	-0.204120
REGION_RATING_CLIENT_W_CITY	AMT_INCOME_TOTAL	-0.223490
AGE	CNT_CHILDREN	-0.239504
HOUR_APPR_PROCESS_START	REGION_RATING_CLIENT_W_CITY	-0.253230
REGION_RATING_CLIENT	HOUR_APPR_PROCESS_START	-0.274244
EXT_SOURCE_2	REGION_RATING_CLIENT_W_CITY	-0.279049
	REGION_RATING_CLIENT	-0.284411
AGE	DAYS_REGISTRATION	-0.301832
REGION_RATING_CLIENT_W_CITY	REGION_POPULATION_RELATIVE	-0.518915
REGION_RATING_CLIENT	REGION_POPULATION_RELATIVE	-0.520460

dtype: float64

Heat Map- Non-Defaulters



Analyzing Previous Applications Data Set

- Read the Previous Application Data CSV file .
- It has rows count-1670214 and 37 columns .
- Check for **quality control** of the Previous Application Data file .
  - It has missing values and Outliers .
  - Getting the Column name having more than 35 Percent missing Data
  - Removing Column from Data frame missing value more than 55 Percent .
  - Now checking columns with less number of missing values impute them if necessary else treat them as missing .

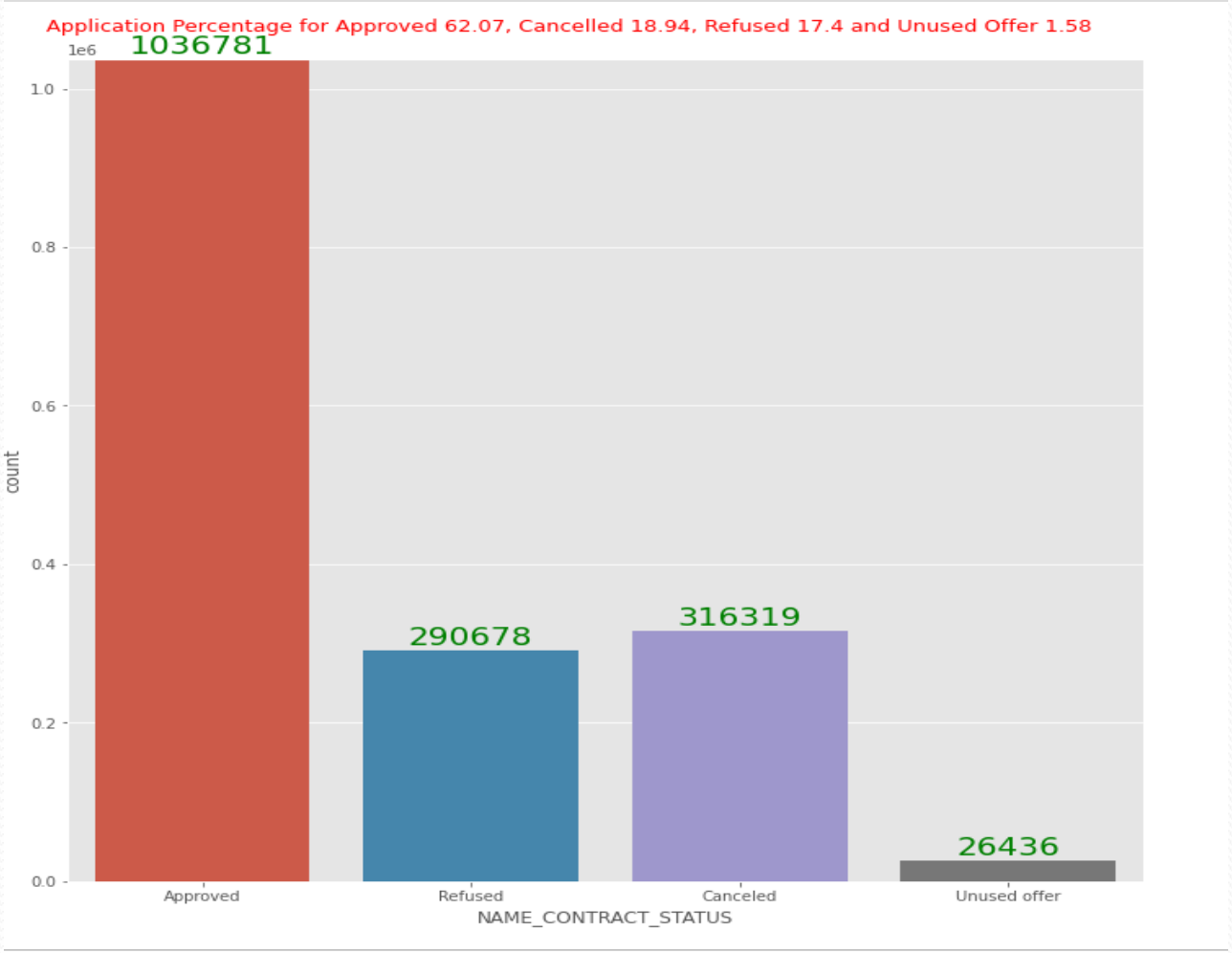
Analyzing Previous Applications Data Set

- **Quality control** of the Previous Application Data file
  - There are two integer Column which just has two unique values .
  - Get the list of column having 2 unique values 0,1 .
  - Replace them as ‘N’ , ‘Y’ and change Data type to object .

Univariate Analysis & Bivariate Analysis Numerical columns/Categorical Columns

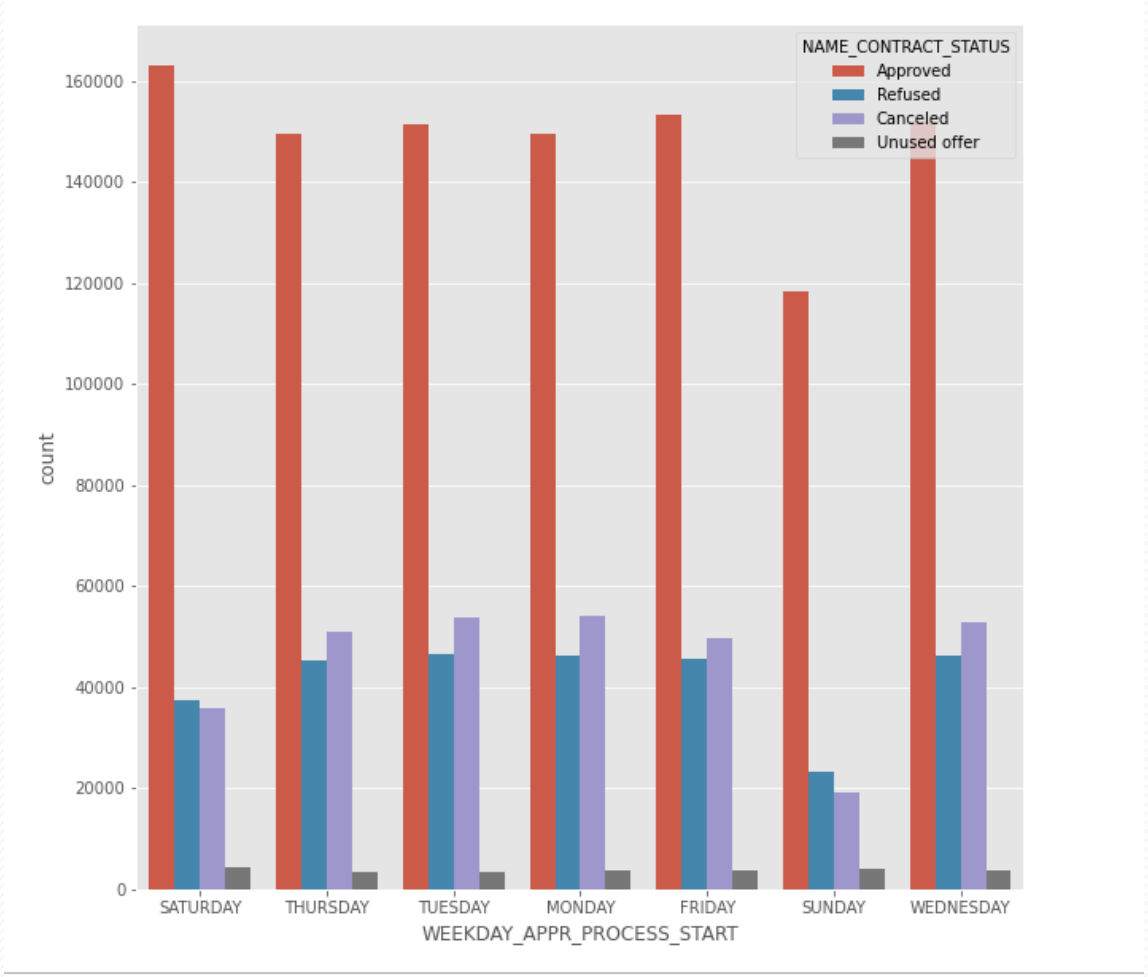
Count of Applications Status/Contract Status

- 62 Percent applications were **approved** , 18 percent cancelled and around 17 percent were **rejected** . Most of previous applications are getting approved .¶



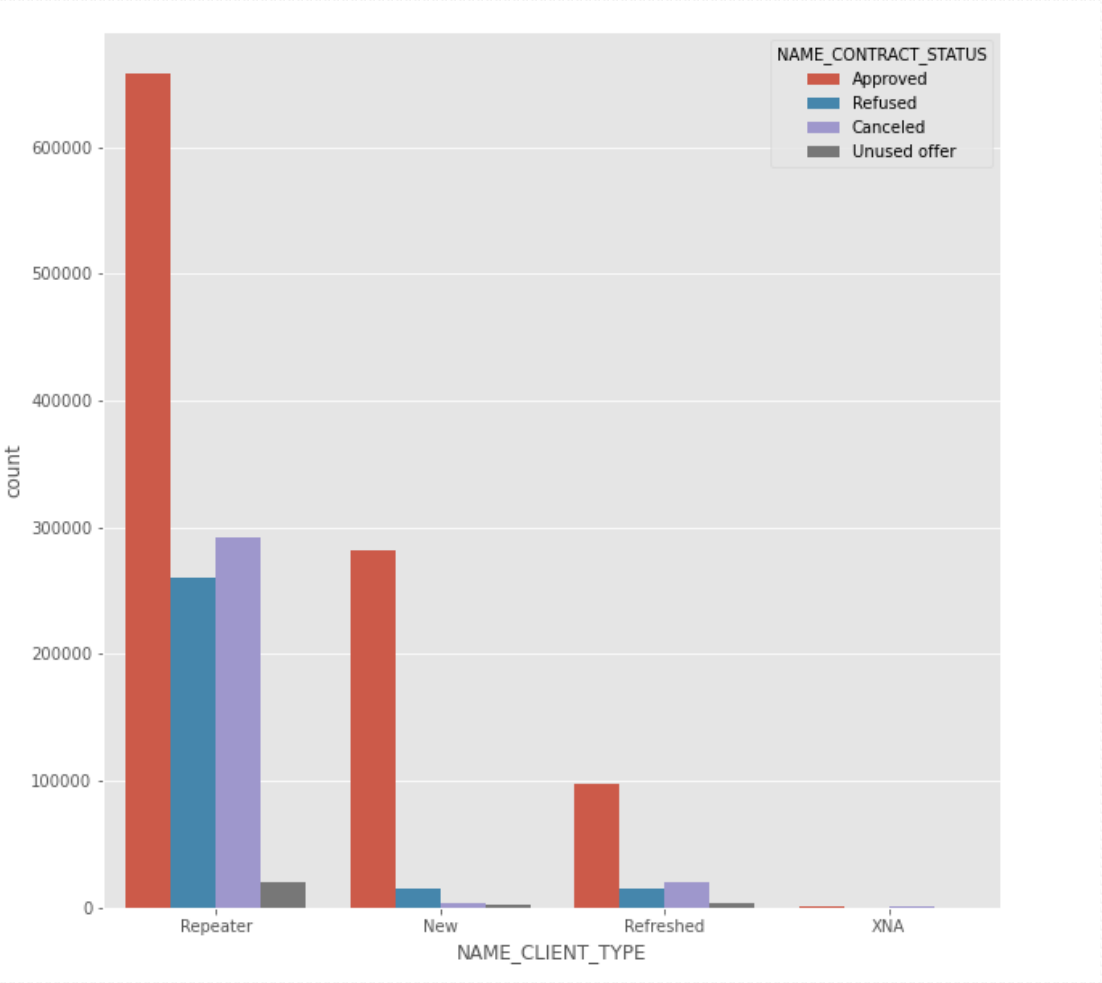
WEEKDAY_APPR_PROCESS_START

- Application approval count is almost same for all day except Sunday ,count is less for Sunday .



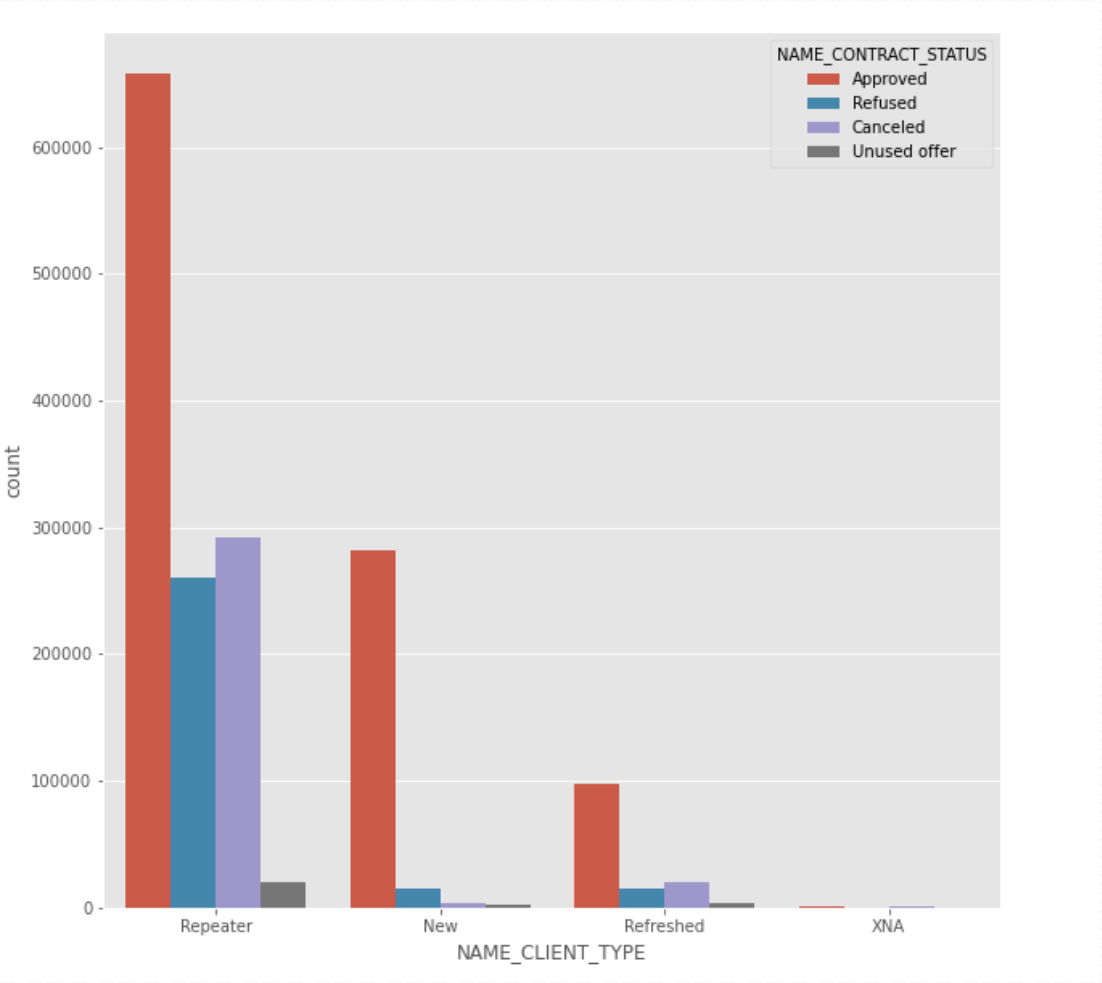
NAME_CLIENT_TYPE

- If you are repeater , chances are high that your loan gets approved .



NAME_CLIENT_TYPE

- If you are repeater , chances are high that your loan gets approved .



NAME_GOODS_CATEGORY

- Top 5 goods category for which loan were refused are XNA,Mobile ,Computers , Consumer Electronics. Audio/Video
- Top 5 goods category for which loan were approved are XNA, Mobile ,Computers , Consumer Electronics. Audio/Video
- No Major difference found on the basis of NAME_GOODS_CATEGORY , top5 product for which load were refused/accepted are same .¶

```
: by_goods[by_goods.NAME_CONTRACT_STATUS=='Refused'].head(5)
```

:

	NAME_CONTRACT_STATUS	NAME_GOODS_CATEGORY	count	mean
2	Refused	XNA	223788	447464.268105
11	Refused	Mobile	20473	53199.087490
13	Refused	Computers	13534	117870.776520
14	Refused	Consumer Electronics	9100	111998.036819
15	Refused	Audio/Video	9080	129350.105812

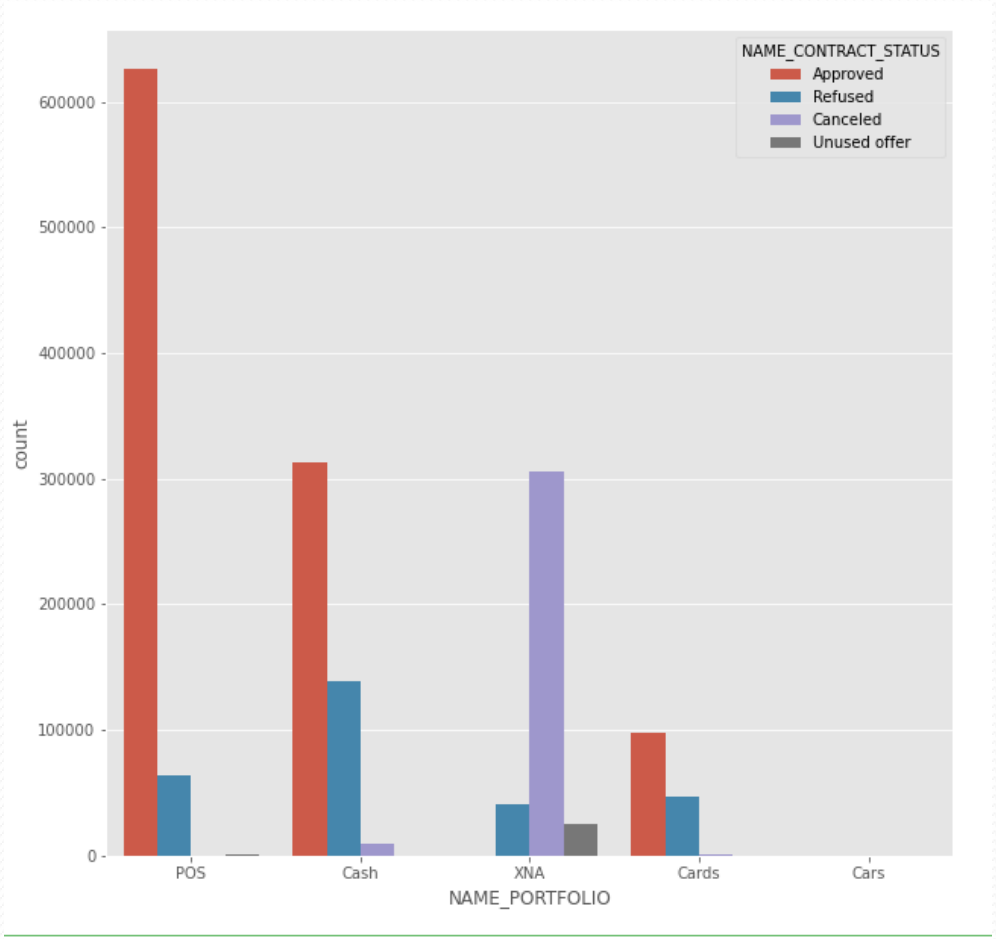
```
: by_goods[by_goods.NAME_CONTRACT_STATUS=='Approved'].head(5)
```

:

	NAME_CONTRACT_STATUS	NAME_GOODS_CATEGORY	count	mean
0	Approved	XNA	410409	375542.711455
3	Approved	Mobile	186174	44184.933640
4	Approved	Consumer Electronics	111525	88494.762669
5	Approved	Audio/Video	89394	101411.330460
6	Approved	Computers	88050	104607.511476

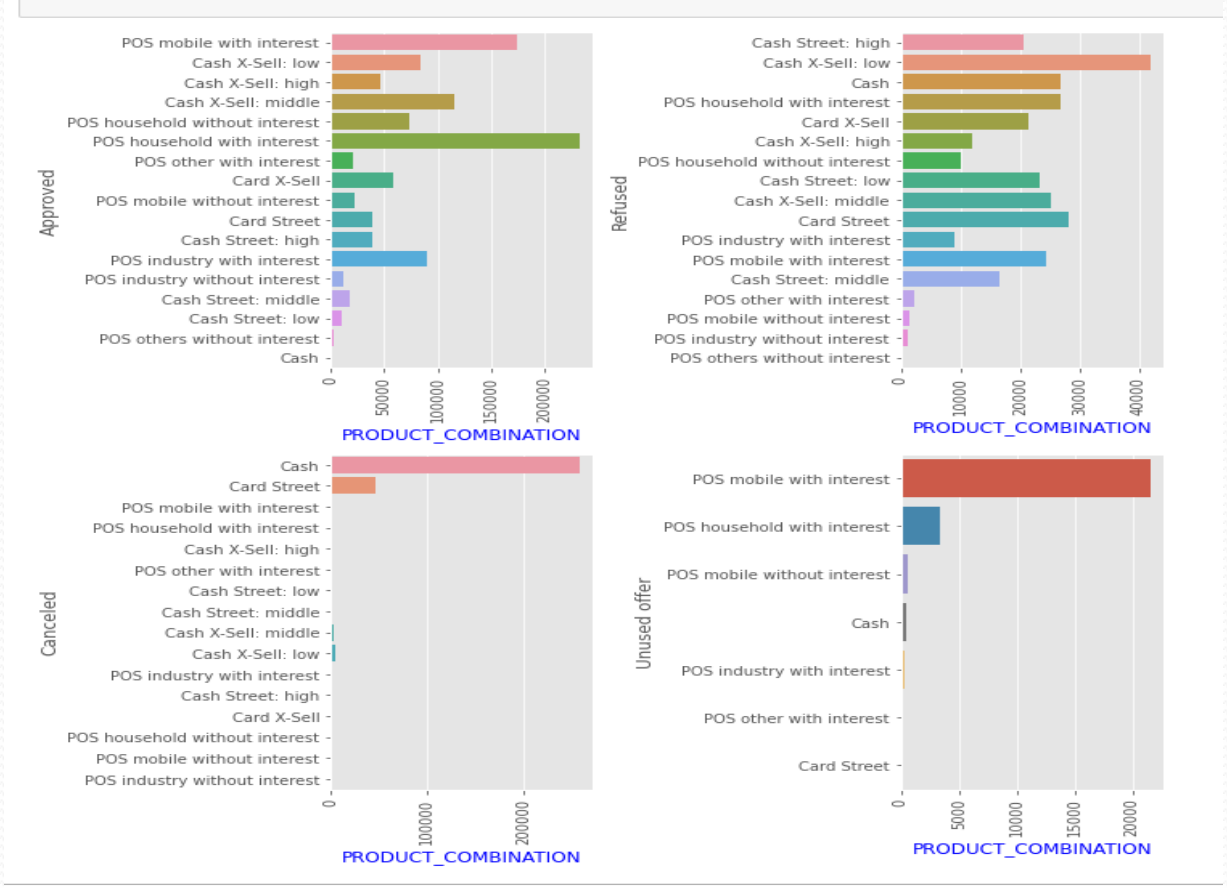
NAME_PORTFOLIO

- Most approved Portfolio is POS followed by Cash
- Most refused Portfolio is Cash .



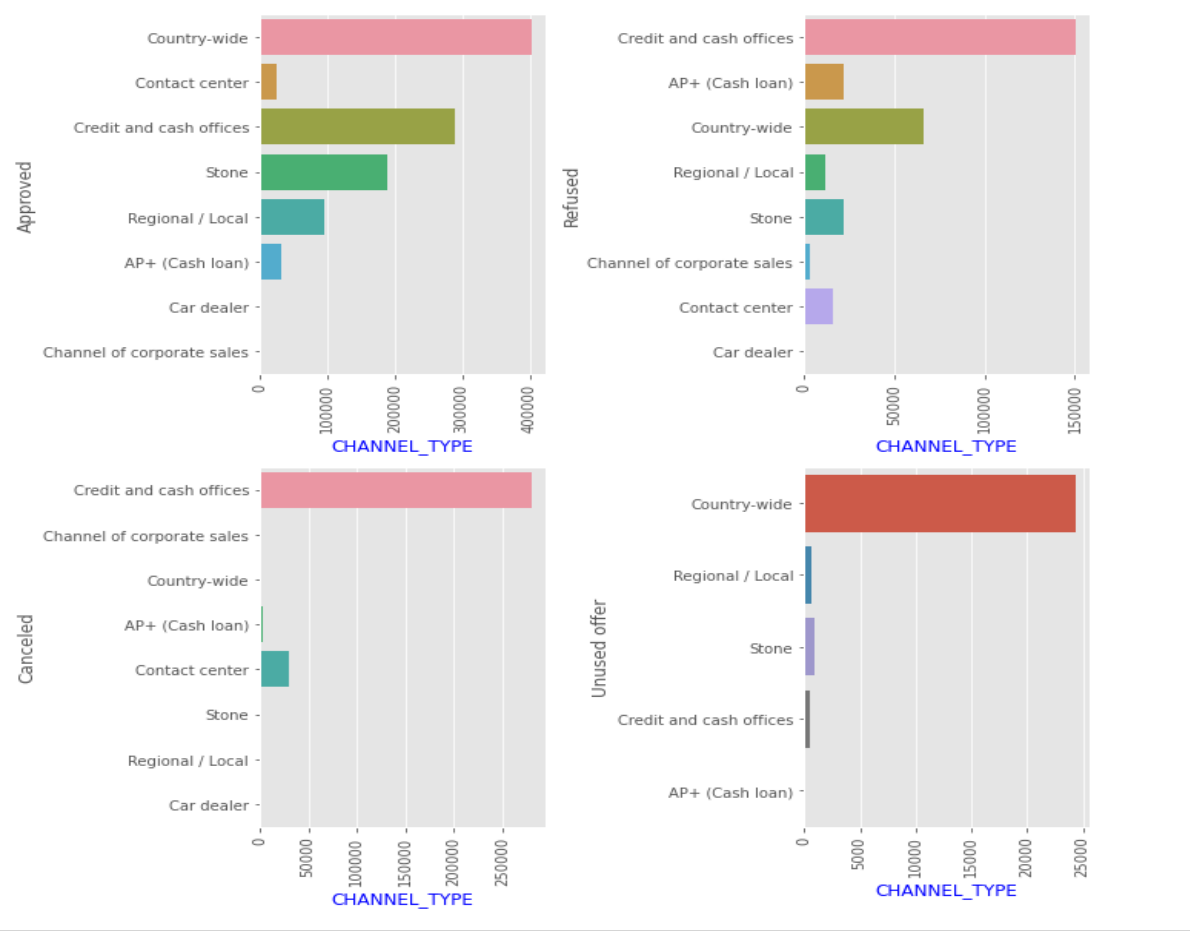
PRODUCT_COMBINATION

- for Product Combination most approved combination is POS household with Interest followed by POS mobile with interest
- Most number for refused loans were Cash x-Sell: Low followed by Card Street , Cash..
- Most cancelled Product Combination is Cash followed by Card Street .



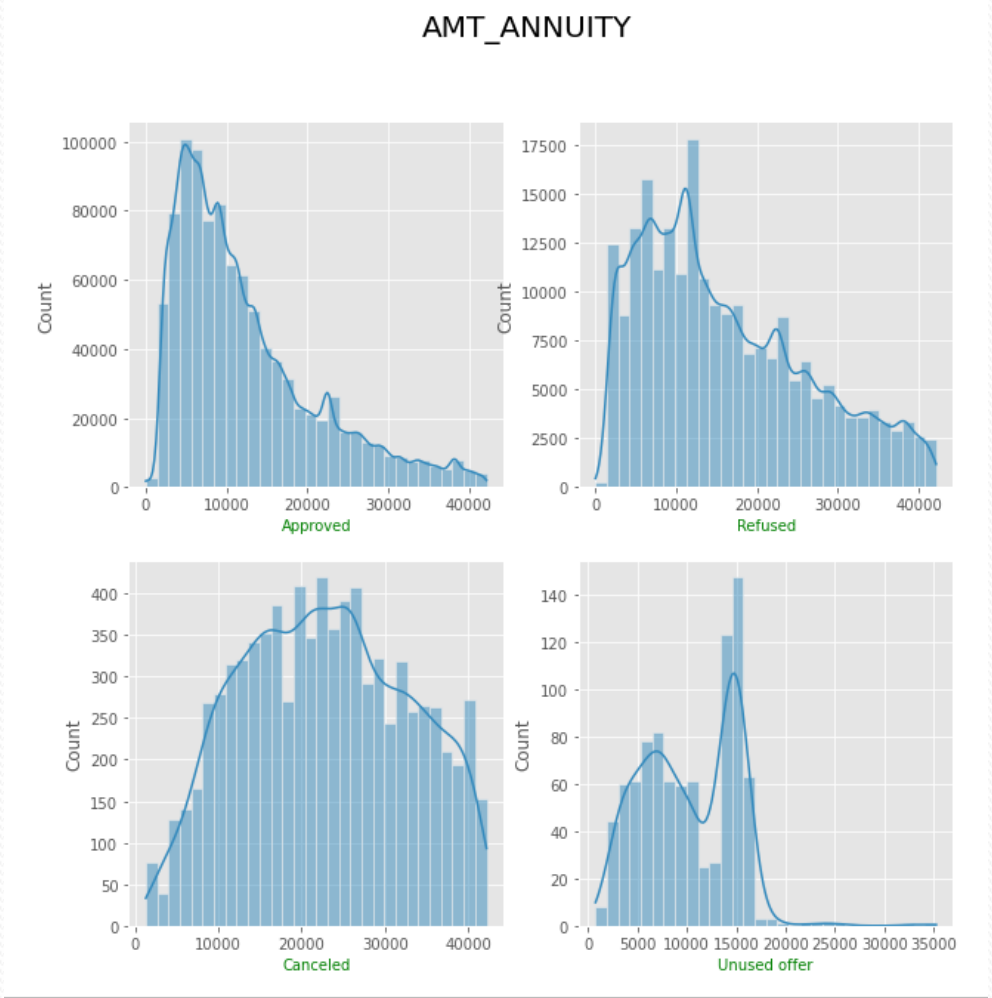
CHANNEL_TYPE

- Most Approved Loans were from Country-wide followed by Credit and Cash offices
- Most refused Loans were from Credit and cash Offices followed by Country wide .
- Source/Channel Type -cash Offices and Country wide are in top most at both approval/rejection , makes most top -2 source overall all applications are coming .



AMT_ANNUITY

- Most Approved Density for AMT_ANNUITY is 10000 and onwards it is decreasing .
- Refusal count density is high around 10000 and then it started decreasing .



AMT_APPLICATION

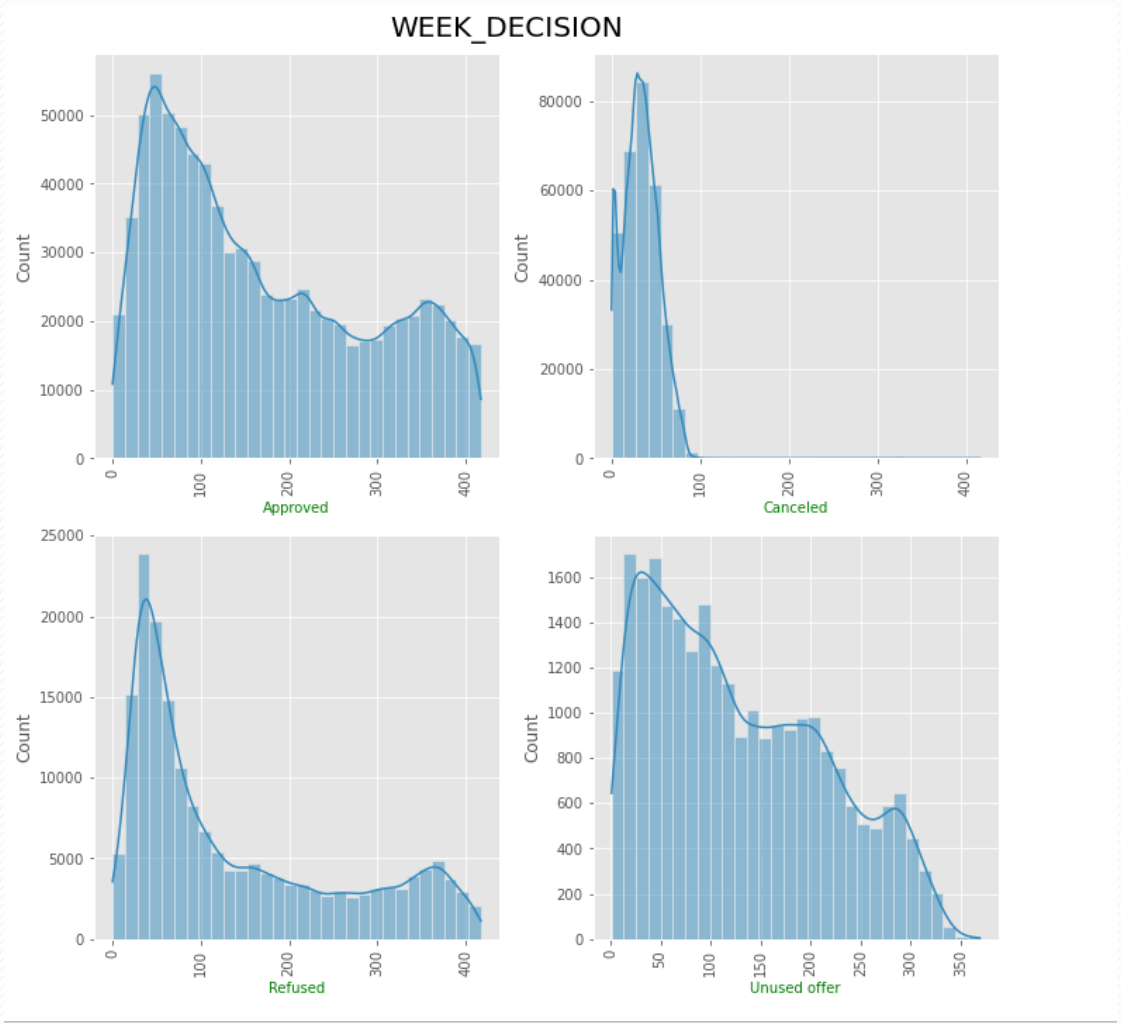
- Application Amount <**100000** has higher approval count after that approval count is decreasing .

AMT_CREDIT

- Most of AMT_CREDIT approved is between 50000 to 100000
- Most of AMT_CREDIT refused is also 50000 .

WEEK_DECISION

- If you Refused graph - high count were at left side under -50 weeks and after that refusal count was decreasing. It means longer it will take probabality getting it approved is better . Most of application were refused within 50 weeks .



Heat Map –Correlations

Previous Application Data

Heat Map-Previous Applications

- Select the Numeric and Float column for corr() functions .
- Divide Data frame into Approved and Refused Applications
- Get Top 10 Positive and Negative correlations for Approved Applications .
- Get Top 10 Positive and Negative correlations for Non-Defaulters .
- Heat Map for Approved and Refused Applications Category .

Heat Map-for Approved Applications

Top 10 most positive factors for approved Applications

CNT_PAYMENT	CNT_PAYMENT	1.000000
AMT_CREDIT	AMT_GOODS_PRICE	0.982787
AMT_APPLICATION	AMT_CREDIT	0.889051
AMT_DOWN_PAYMENT	RATE_DOWN_PAYMENT	0.750255
AMT_GOODS_PRICE	AMT_ANNUITY	0.749529
AMT_ANNUITY	AMT_APPLICATION	0.735833
	AMT_CREDIT	0.725190
AMT_APPLICATION	CNT_PAYMENT	0.479731
CNT_PAYMENT	AMT_GOODS_PRICE	0.436147
	AMT_CREDIT	0.415907

dtype: float64

Heat Map-for Approved Applications

Top 10 most Negative factors for approved Applications

```
CNT_PAYMENT       AMT_DOWN_PAYMENT    -0.183626
AMT_ANNUITY       WEEK_DECISION       -0.194873
AMT_CREDIT        WEEK_DECISION       -0.250491
WEEK_DECISION     AMT_GOODS_PRICE     -0.263017
AMT_ANNUITY       RATE_DOWN_PAYMENT   -0.266060
AMT_APPLICATION   WEEK_DECISION       -0.289400
CNT_PAYMENT       RATE_DOWN_PAYMENT   -0.318429
AMT_APPLICATION   RATE_DOWN_PAYMENT   -0.341041
WEEK_DECISION     DAYS_TERMINATION    -0.387488
AMT_CREDIT        RATE_DOWN_PAYMENT   -0.414858
dtype: float64
```

Heat Map-for Approved Applications

Heat Map-for Refused Applications

Top 10 most positive factors for Refused Applications

```
CNT_PAYMENT          CNT_PAYMENT          1.000000
AMT_APPLICATION      AMT_GOODS_PRICE      0.999270
AMT_CREDIT           AMT_GOODS_PRICE      0.938092
AMT_APPLICATION      AMT_CREDIT           0.900644
RATE_DOWN_PAYMENT    AMT_DOWN_PAYMENT     0.768046
AMT_ANNUITY          AMT_CREDIT           0.664031
                     AMT_GOODS_PRICE      0.652383
                     AMT_APPLICATION      0.644866
CNT_PAYMENT          AMT_APPLICATION      0.369707
                     AMT_CREDIT           0.341546
dtype: float64
```

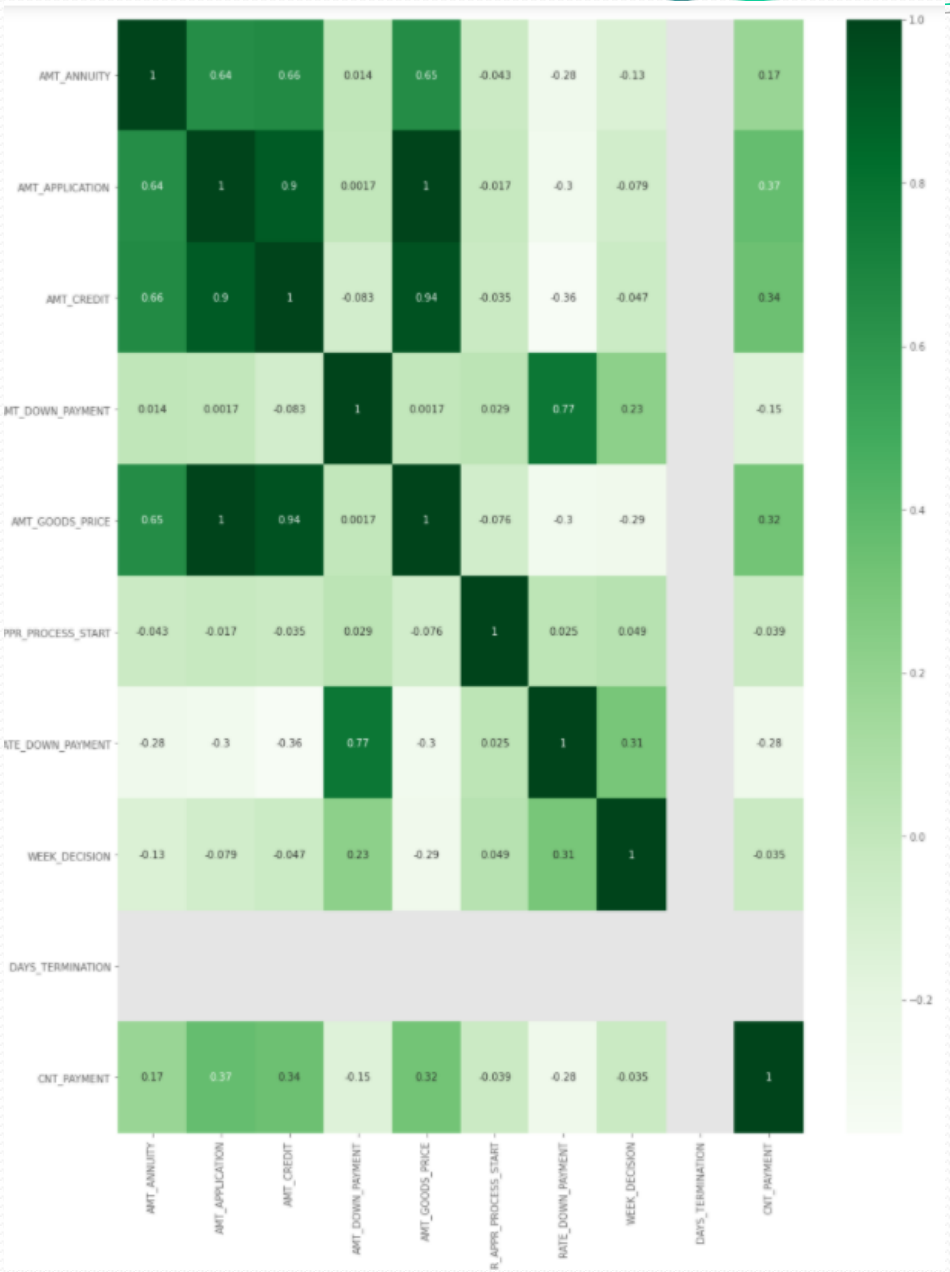
Heat Map-for Refused Applications

Top 10 most Negative factors for Refused Applications

WEEK_DECISION	AMT_APPLICATION	-0.078638
AMT_DOWN_PAYMENT	AMT_CREDIT	-0.082908
AMT_ANNUITY	WEEK_DECISION	-0.129544
CNT_PAYMENT	AMT_DOWN_PAYMENT	-0.152723
RATE_DOWN_PAYMENT	CNT_PAYMENT	-0.281675
	AMT_ANNUITY	-0.283395
AMT_GOODS_PRICE	WEEK_DECISION	-0.291002
AMT_APPLICATION	RATE_DOWN_PAYMENT	-0.299715
AMT_CREDIT	RATE_DOWN_PAYMENT	-0.362955
AMT_ANNUITY	DAYS_TERMINATION	nan

dtype: float64

Heat Map-for Rejected Applications



Merging Application Data & Previous Application Data

- Do the left Join on table Application Data on filed - SK_ID_CURR
- It has rows count-83074 and 116 columns .
- Get all the Numeric and Float columns in list and then use that list for Heat Map analysis for combined data for defaulter and Non-Defaulter .

Heat Map –Correlations

Merged Data –Application +Previous Application Data

Heat Map –for Defaulter on combined Data

Top 10 most positive correlation in case of Defaulter

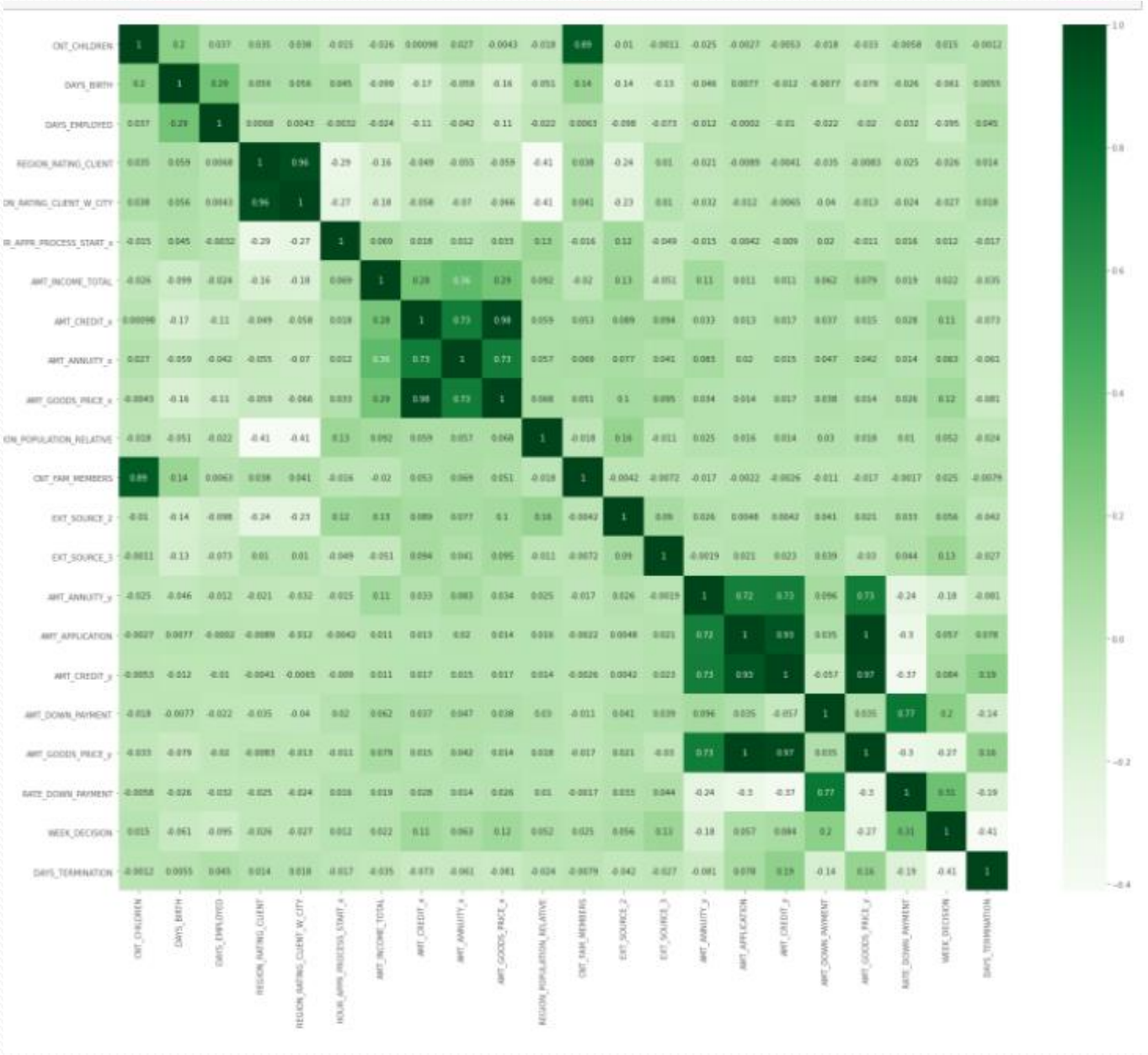
DAYS_TERMINATION	1.000000
AMT_GOODS_PRICE_y	0.999958
AMT_CREDIT_x	0.978016
AMT_CREDIT_y	0.972518
REGION_RATING_CLIENT_W_CITY	0.958780
AMT_CREDIT_y	0.926571
CNT_CHILDREN	0.894874
RATE_DOWN_PAYMENT	0.766375
AMT_ANNUITY_y	0.733742
AMT_ANNUITY_x	0.731086
dtype: float64	

Heat Map –for Defaulter on combined Data

Top 10 most negative correlation in case of Defaulter

EXT_SOURCE_2	REGION_RATING_CLIENT	-0.236747
AMT_ANNUITY_y	RATE_DOWN_PAYMENT	-0.244502
HOUR_APPR_PROCESS_START_x	REGION_RATING_CLIENT_W_CITY	-0.266427
WEEK_DECISION	AMT_GOODS_PRICE_y	-0.270622
REGION_RATING_CLIENT	HOUR_APPR_PROCESS_START_x	-0.286143
RATE_DOWN_PAYMENT	AMT_GOODS_PRICE_y	-0.301402
	AMT_CREDIT_y	-0.371066
REGION_POPULATION_RELATIVE	REGION_RATING_CLIENT	-0.405853
REGION_RATING_CLIENT_W_CITY	REGION_POPULATION_RELATIVE	-0.409806
DAYS_TERMINATION	WEEK_DECISION	-0.411353
dtype: float64		

Heat Map-for Defaulter Combined Data



Heat Map –For Non-Defaulter on combined Data

Top 10 most positive correlation in case of Non-Defaulter

```
final_corr_Non_defaulter.unstack().sort_values(ascending=False).drop_duplicates().head(10)
```

```
DAYS_TERMINATION       DAYS_TERMINATION             1.000000
AMT_GOODS_PRICE_y      AMT_APPLICATION              0.999843
AMT_GOODS_PRICE_x      AMT_CREDIT_x                 0.981538
AMT_CREDIT_y           AMT_GOODS_PRICE_y            0.972773
REGION_RATING_CLIENT   REGION_RATING_CLIENT_W_CITY  0.946807
AMT_APPLICATION        AMT_CREDIT_y                 0.917660
CNT_CHILDREN           CNT_FAM_MEMBERS              0.894690
AMT_DOWN_PAYMENT       RATE_DOWN_PAYMENT            0.756335
AMT_ANNUITY_x          AMT_CREDIT_x                 0.745633
AMT_GOODS_PRICE_x      AMT_ANNUITY_x                0.744524
dtype: float64
```

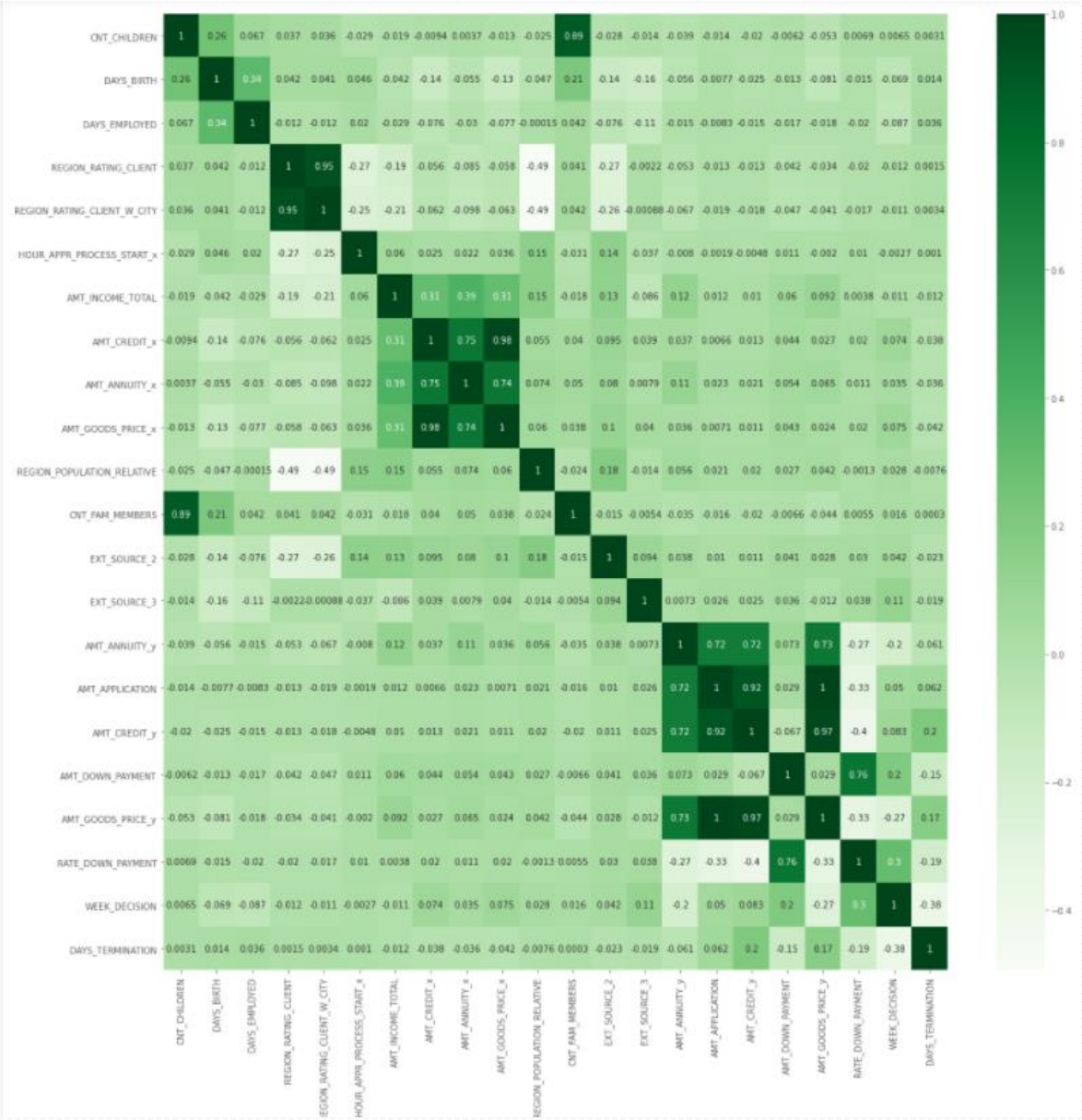
Heat Map –for Non-Defaulter on combined Data

Top 10 most negative correlation in case of Non-Defaulter

```
final_corr_Non_defaulter.unstack().sort_values(ascending=False).drop_duplicates().tail(10)
```

```
REGION_RATING_CLIENT_W_CITY  EXT_SOURCE_2                -0.264920
WEEK_DECISION                AMT_GOODS_PRICE_y           -0.265931
RATE_DOWN_PAYMENT            AMT_ANNUITY_y               -0.269501
REGION_RATING_CLIENT         HOUR_APPR_PROCESS_START_x   -0.269926
EXT_SOURCE_2                 REGION_RATING_CLIENT        -0.271129
RATE_DOWN_PAYMENT            AMT_APPLICATION             -0.327158
WEEK_DECISION                DAYS_TERMINATION            -0.381829
RATE_DOWN_PAYMENT            AMT_CREDIT_y                -0.398689
REGION_POPULATION_RELATIVE   REGION_RATING_CLIENT_W_CITY -0.493339
REGION_RATING_CLIENT         REGION_POPULATION_RELATIVE  -0.493939
dtype: float64
```

Heat Map-for Non-Defaulter Combined Data



AMT_CREDIT_x-Application Data

```
##  function to get corr values of specified column and return into dataframe
def getCorrelatedFeature(corrdata, threshold):
    feature=[]
    value=[]
    for i, index in enumerate(corrdata.index):
        if corrdata[index] > threshold :
            if corrdata[index]!=1:
                feature.append(index)
                value.append(corrdata[index])
    df=pd.DataFrame(data=value,index=feature,columns=['corr value'])
    return df

threshold=0.05
corr_df=getCorrelatedFeature(corrmat['AMT_CREDIT_x'], threshold)
corr_df.sort_values(by=['corr value'],ascending=False)
```

	corr value
AMT_GOODS_PRICE_x	0.981194
AMT_ANNUITY_x	0.744030
AMT_INCOME_TOTAL	0.306432
EXT_SOURCE_2	0.095629
WEEK_DECISION	0.078461
REGION_POPULATION_RELATIVE	0.055936

AMT_CREDIT_y-Application Data

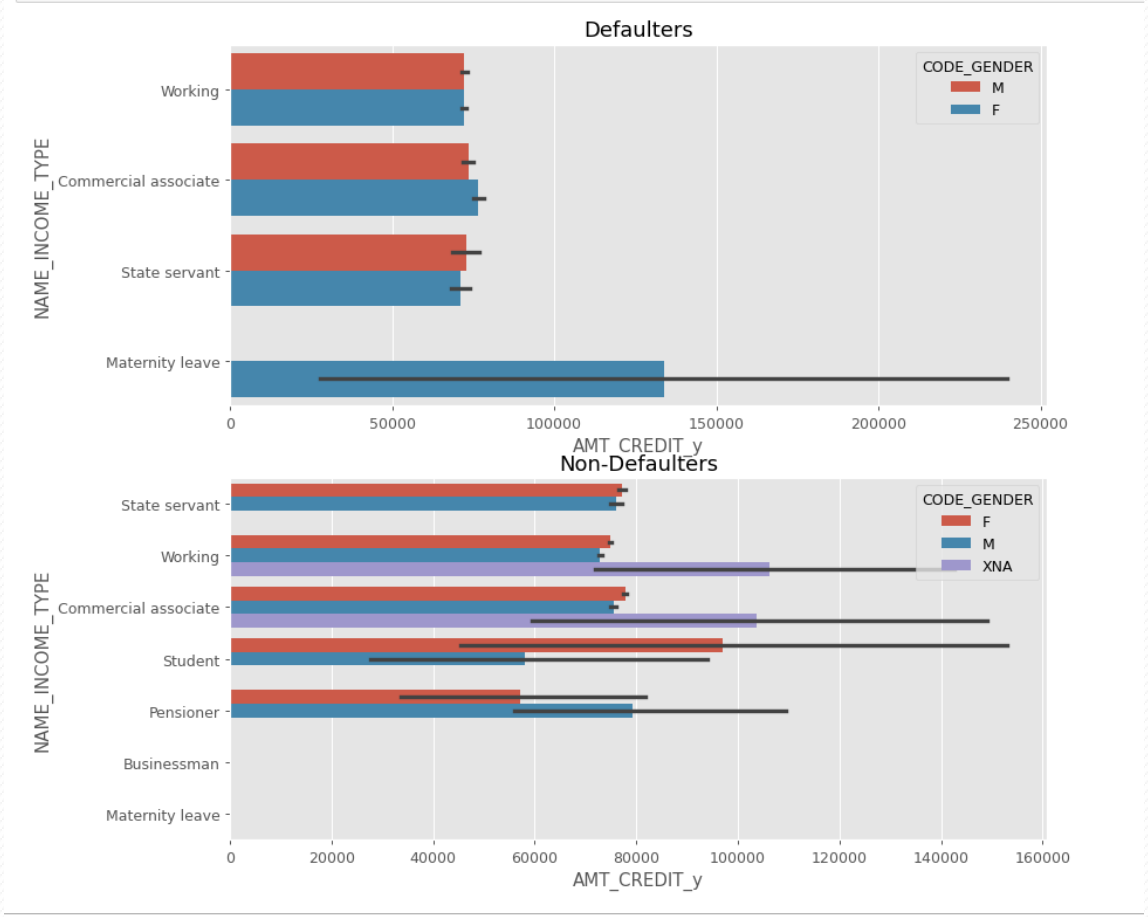
```
##  function to get corr values of specified column and return into dataframe
def getCorrelatedFeature(corrdata, threshold):
    feature=[]
    value=[]
    for i, index in enumerate(corrdata.index):
        if corrdata[index] > threshold :
            if corrdata[index]!=1:
                feature.append(index)
                value.append(corrdata[index])
    df=pd.DataFrame(data=value,index=feature,columns=['corr value'])
    return df

threshold=0.05
corr_df=getCorrelatedFeature(corrmat['AMT_CREDIT_y'], threshold)
corr_df.sort_values(by=['corr value'],ascending=False)
```

	corr value
AMT_ANNUITY_y	0.715807
AMT_APPLICATION	0.918549
AMT_GOODS_PRICE_y	0.972719
WEEK_DECISION	0.083721
DAYS_TERMINATION	0.199109

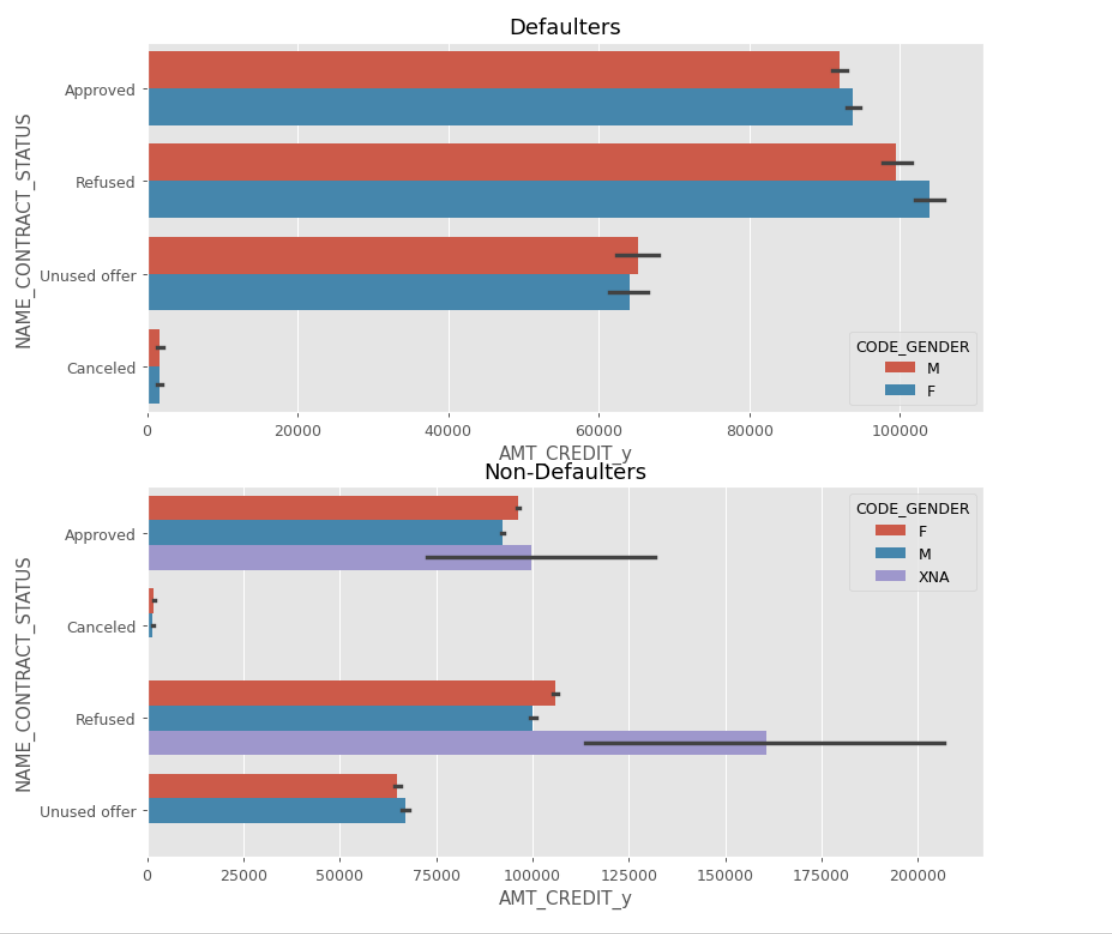
NAME_INCOME_TYPE

- For Defaulter , Women who goes on Maternity leave defaulter max .
- for Defaulter category- Student are missing whereas Non_defaulter they have good amount of Amount credit . In student , female students have larger amount distributed -Good customers .



NAME_CONTRACT_STATUS

- **For Defaulter , most of Application were Previously refused .**



More Observations –For Defaulters

- For Defaulter , People with housing type -Co-op apartment should not be targeted , look like they are facing problems while repaying the loans .
- Buying a Holiday home/land is purpose were loan were defaulted in high range and people are facing difficulty while paying it .
- People working in Industry type13 default much -Bad Customers .
- People with high rating2 likely to default less

Thanks for Watching

By –Pramendra Pandey & Karan Babbar