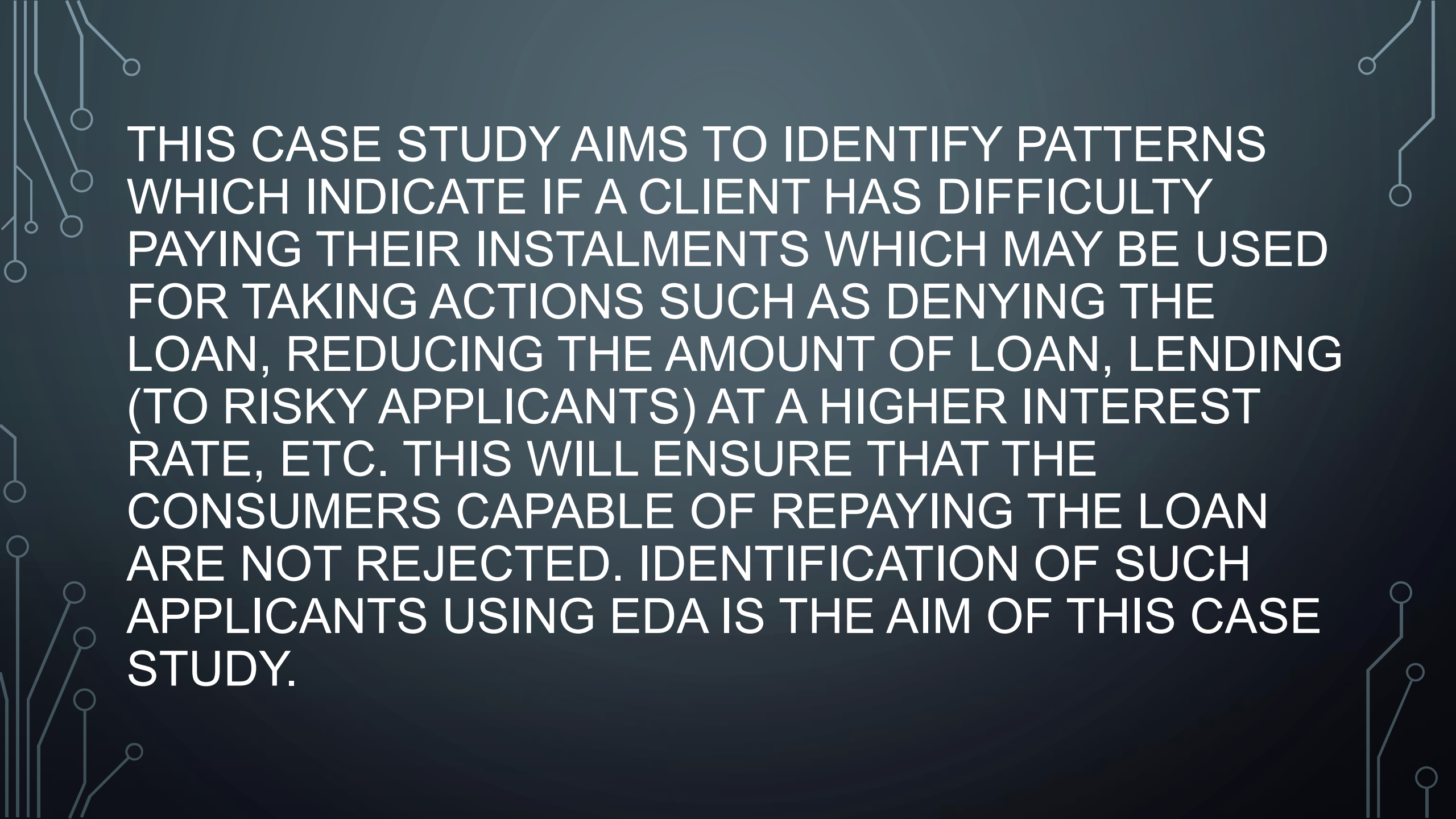




CREDIT EDA – LOAN CASE STUDY

SUBMITTED BY: ASHISH PAWAR (DS – 60)

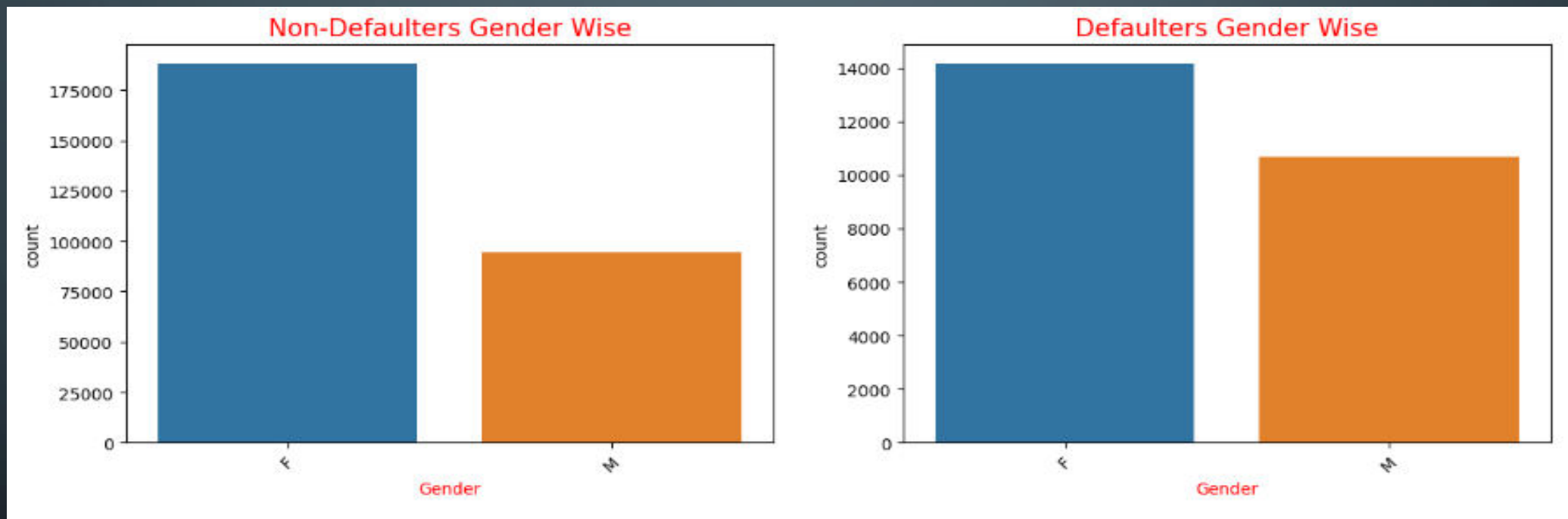
The background is a dark blue-grey color. It features decorative circuit-like lines in a lighter blue-grey shade. These lines are composed of straight segments and small circles, resembling a stylized electronic circuit board. They are positioned along the left and right edges of the image, framing the central text.

THIS CASE STUDY AIMS TO IDENTIFY PATTERNS WHICH INDICATE IF A CLIENT HAS DIFFICULTY PAYING THEIR INSTALMENTS 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.

ANALYSIS DONE

- Data Cleaning and handling missing values
- Imputing the missing values
- Check and handling Outliers
- Check data Imbalance and ratio of Imbalance
- Check for Univariate (Categorical and Continuous Univariate analysis) and Bivariate analysis and correlation.
- Merging both data set and perform analysis.
- Recommendation and Risk

UNIVARIATE ANALYSIS FOR CATEGORICAL VARIABLE

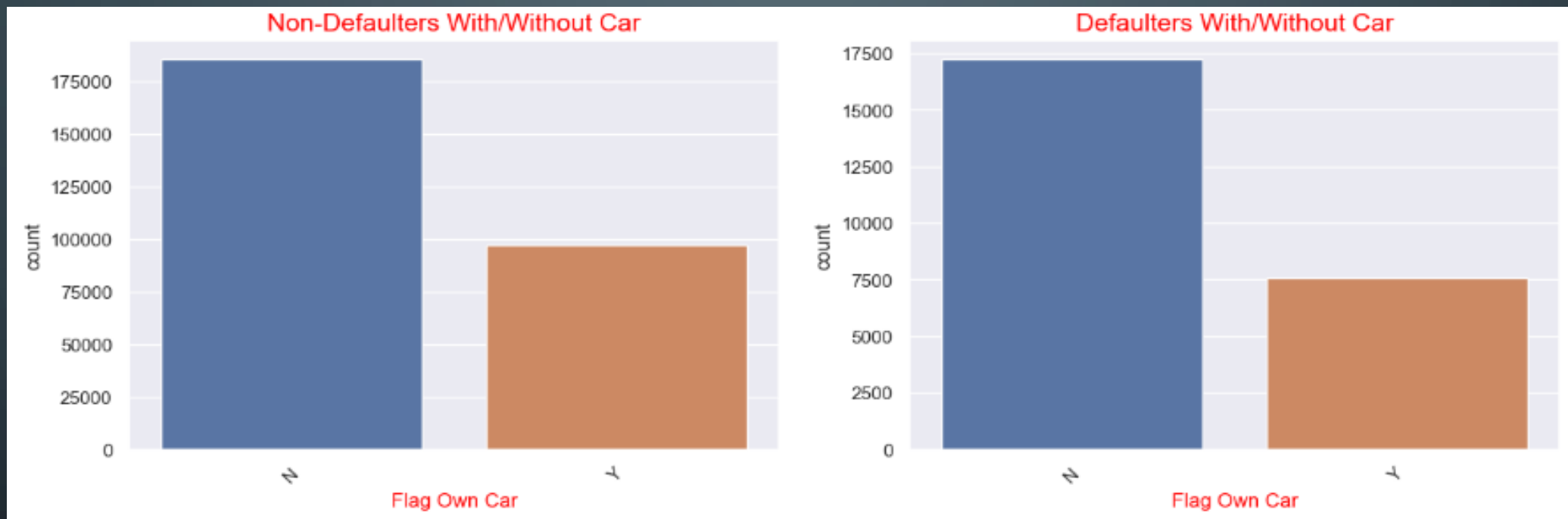


Observation:

As per the above graph, we can find that the number of females applicant are higher in non defaulter category.

Also, we can check the same trend for the defaulters where females are high number of defaulters.

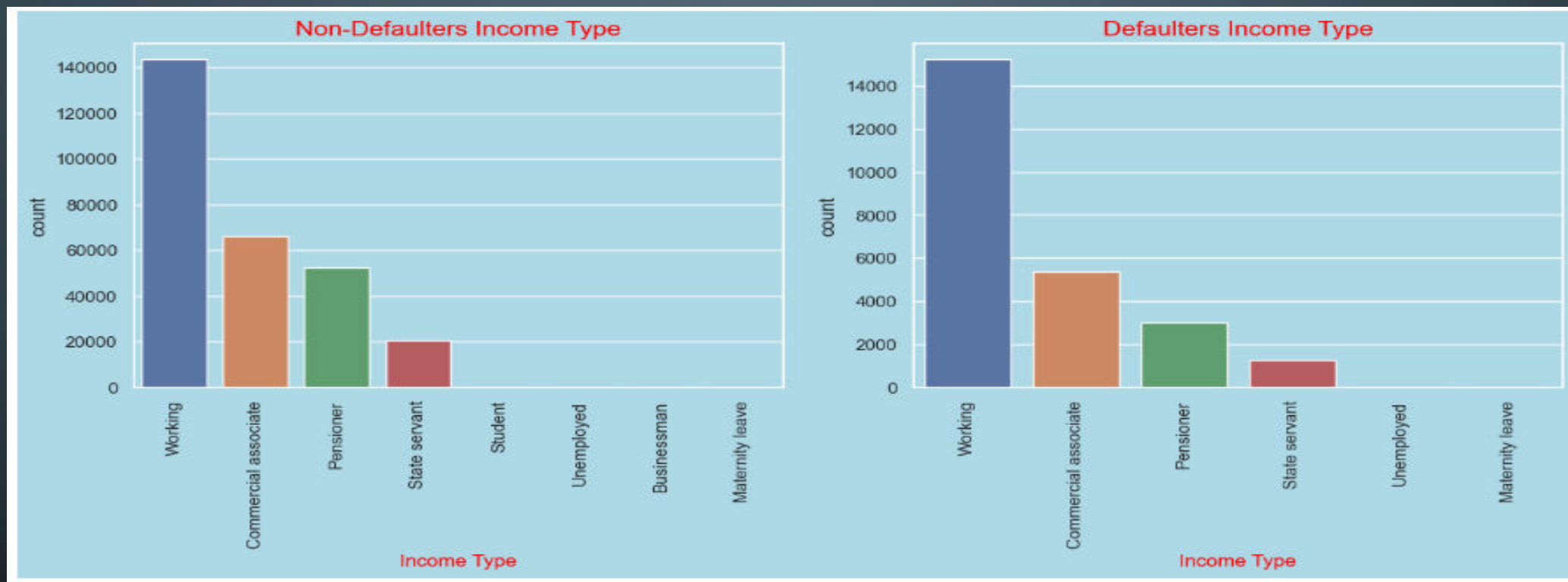
UNIVARIATE ANALYSIS FOR CATEGORICAL VARIABLE



Observation:

It is clearly visible that maximum number of Non-defaulters are not having their own a car. However, Defaulters also show the same trend where maximum defaulters does not own a car.

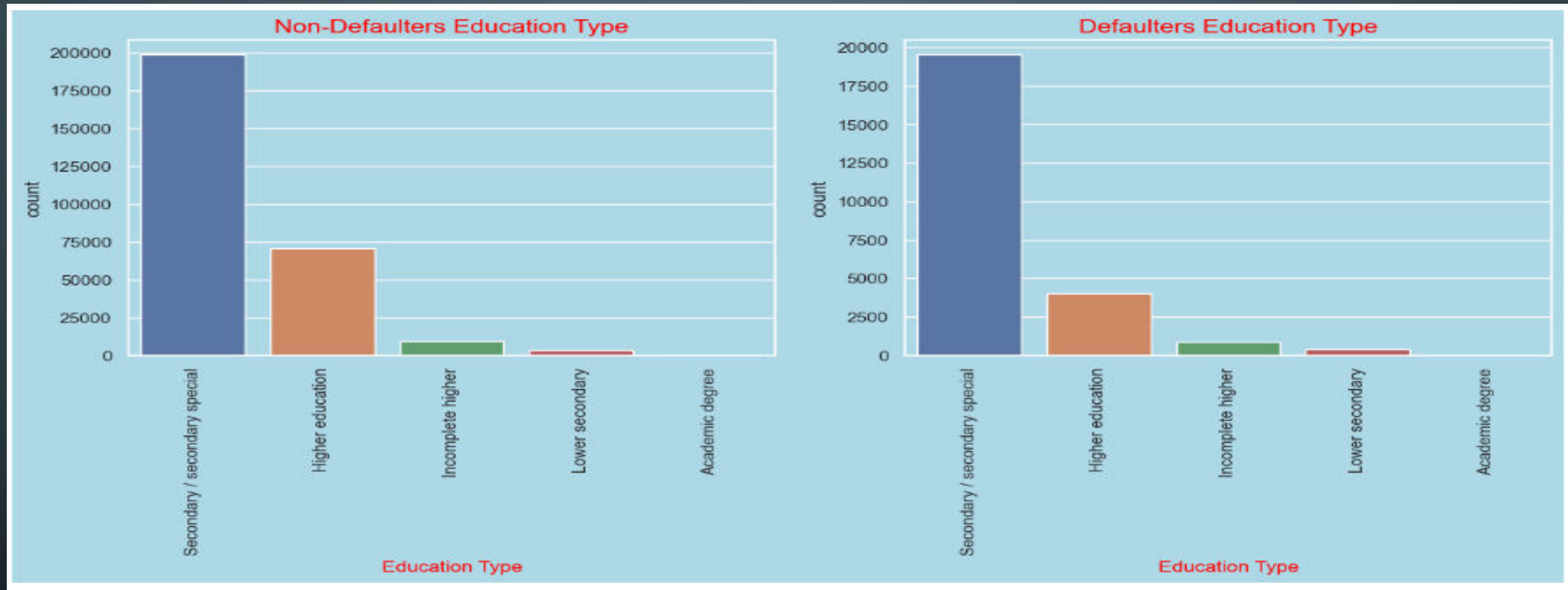
UNIVARIATE ANALYSIS FOR CATEGORICAL VARIABLE



Observation:

The Defaulters are mostly working professionals followed by commercial Associates.

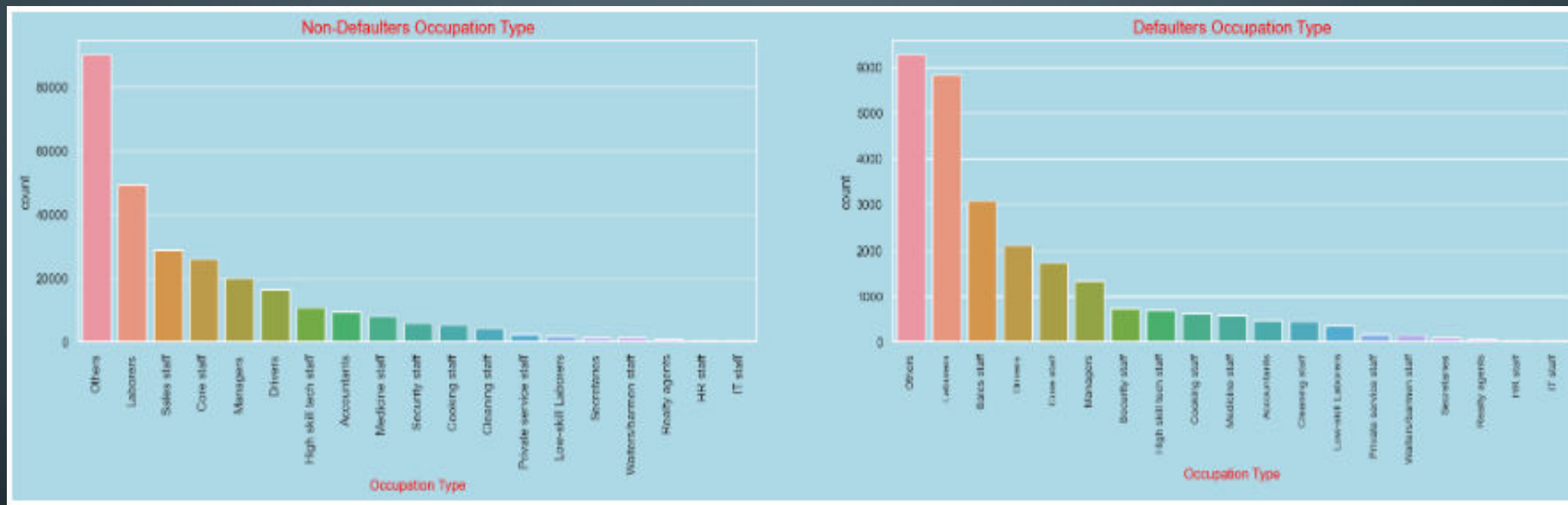
UNIVARIATE ANALYSIS FOR CATEGORICAL VARIABLE



Observation:

The defaulter are mostly having Secondary/Secondary special kind of education.

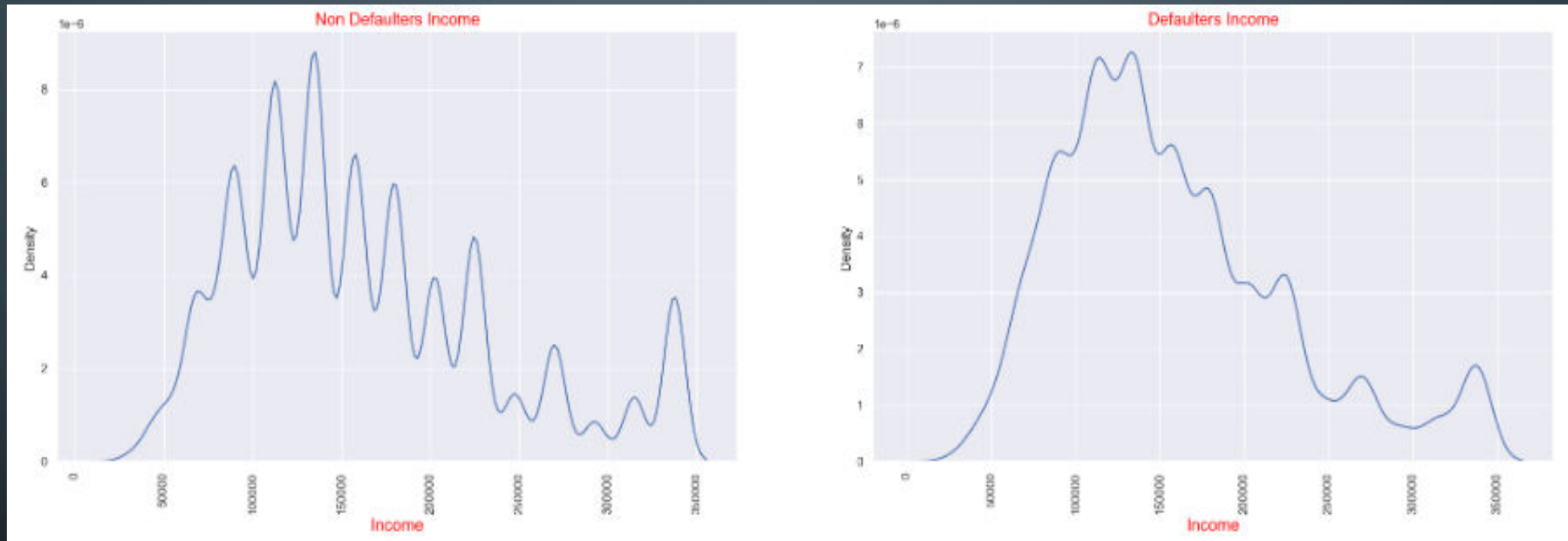
UNIVARIATE ANALYSIS FOR CATEGORICAL VARIABLE



Observation:

Most of the defaulters are labours by their occupation. The labours are the one who is lower skilled people so, they are struggling with the loan bill payment and becoming defaulters.

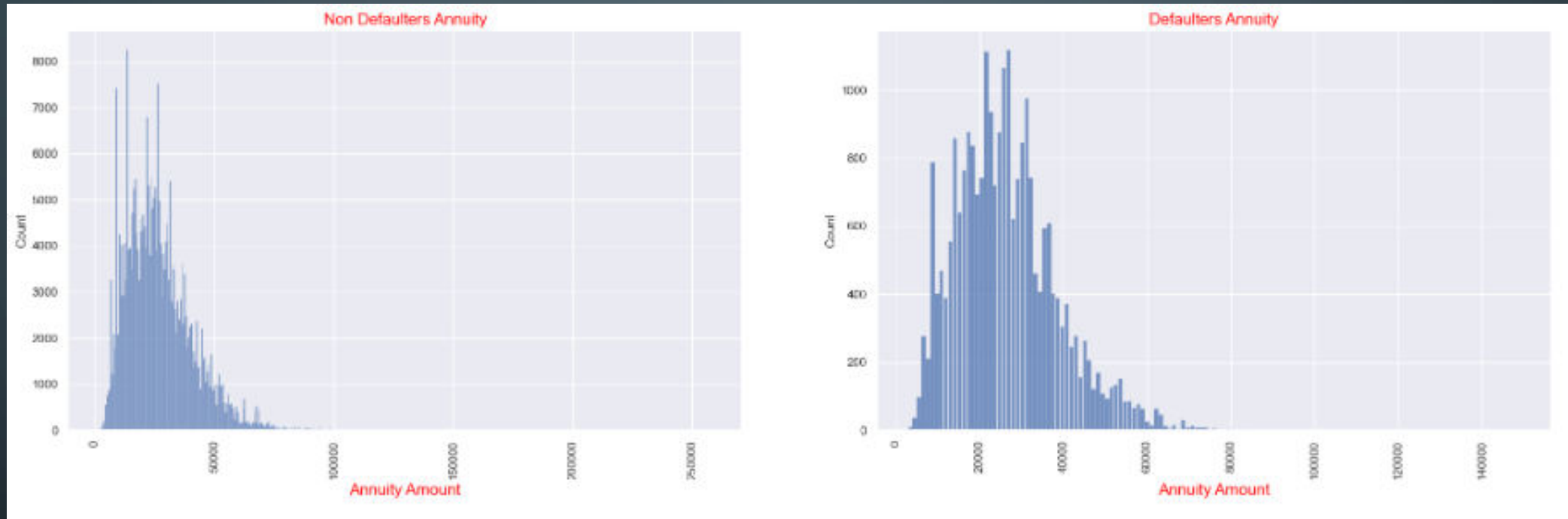
UNIVARIATE ANALYSIS FOR CONTINUOUS VARIABLE



Observation:

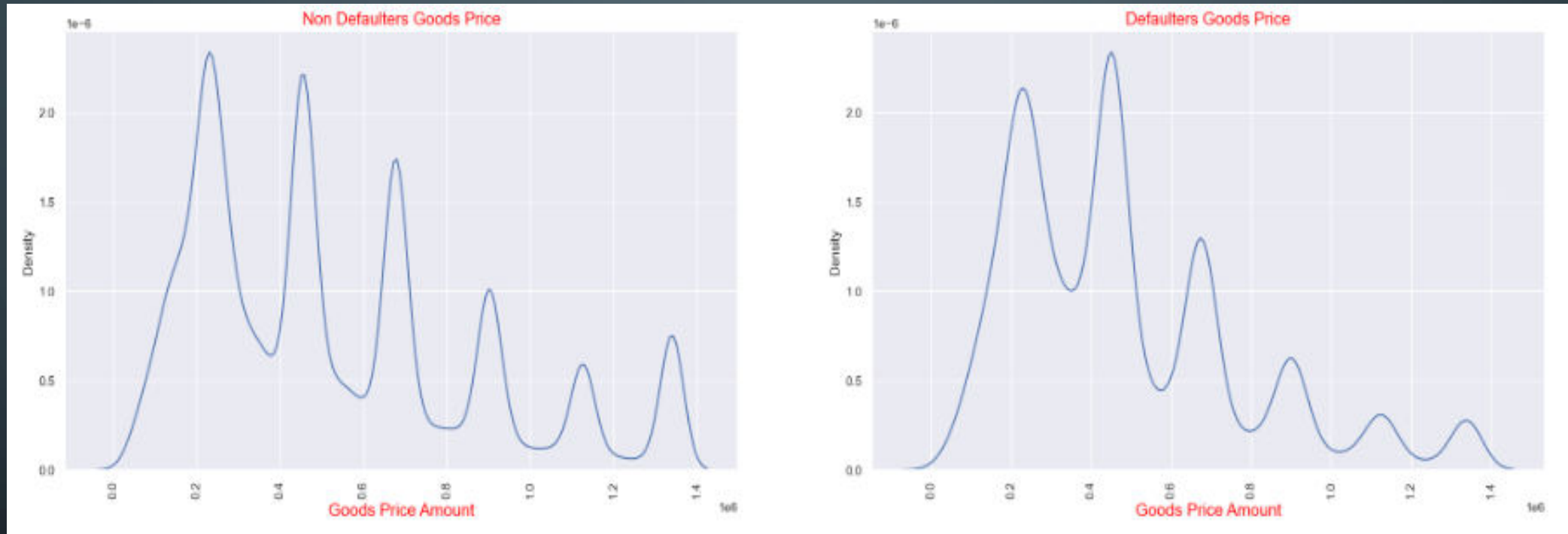
Most of the defaulters are earning between 1 lakhs to 1.7 lakhs annually.

UNIVARIATE ANALYSIS FOR CONTINUOUS VARIABLE



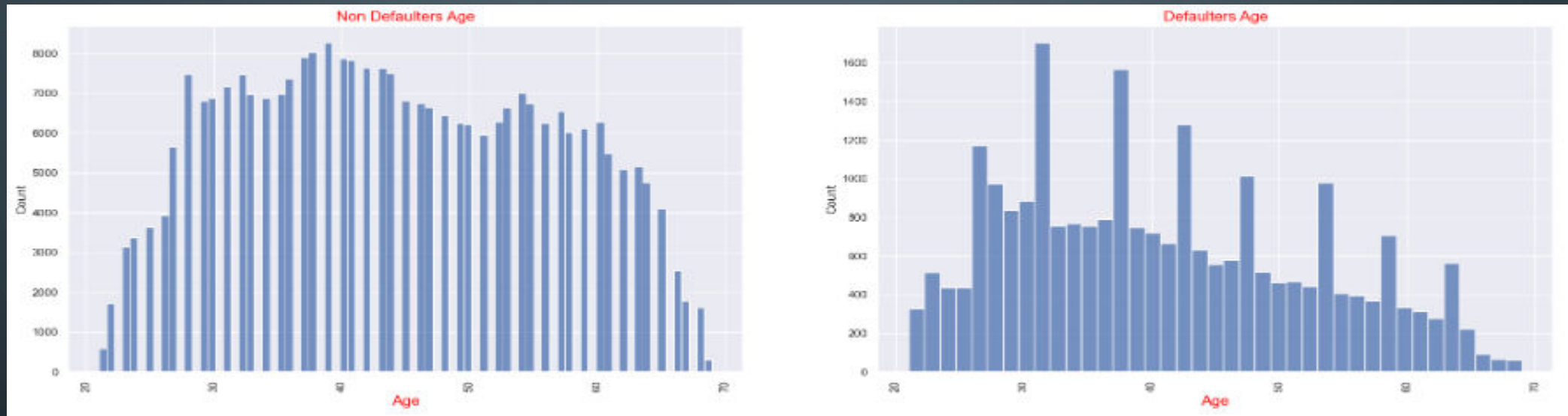
Observation:
Defaulters are having loan annuity between 15000 to 30000.

UNIVARIATE ANALYSIS FOR CONTINUOUS VARIABLE



Observation:
Defaulters Goods price amount is largely lying between 0.2(200000) to 0.5(500000).

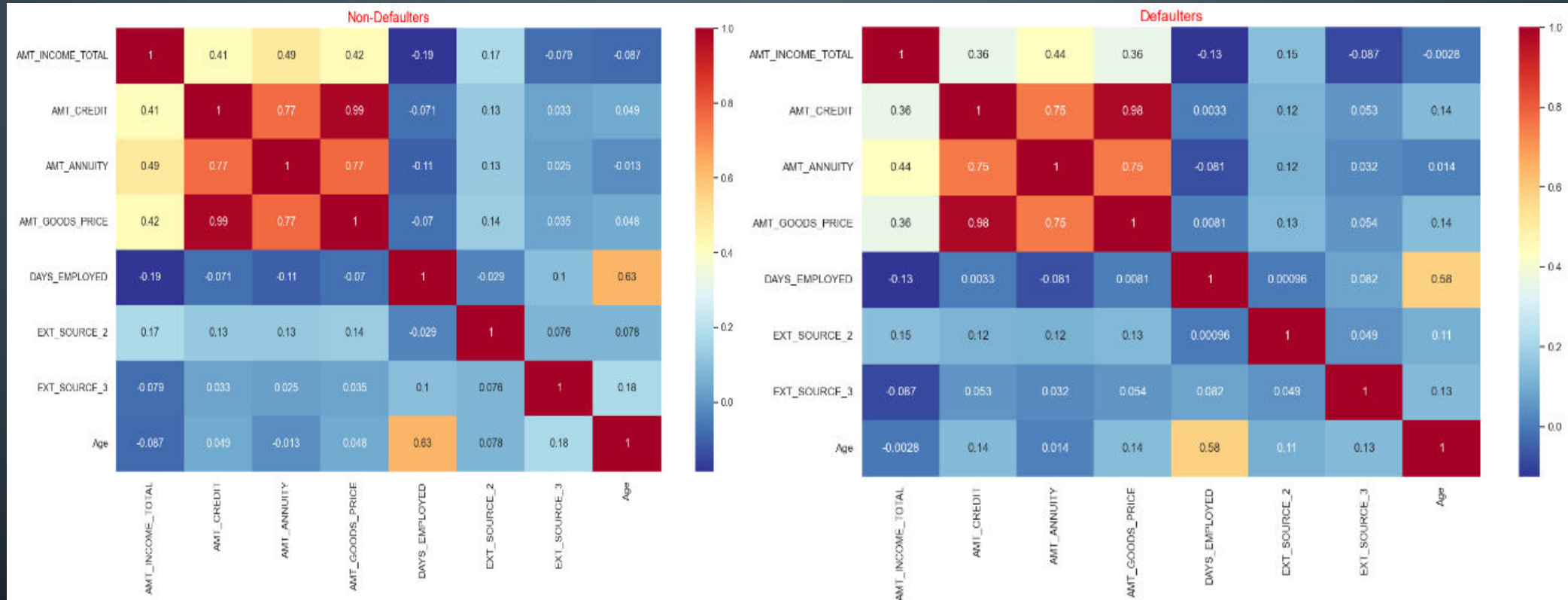
UNIVARIATE ANALYSIS FOR CONTINUOUS VARIABLE



Observation:

The defaulters age is lying mostly between 25 to 45. Middle age people are getting defaulters more compared to other age people.

BIVARIATE ANALYSIS



Observation:

We can come to conclusion that there is a similarity between columns when it comes to defaulters and non-defaulters data correlation. The column relation which are showing higher percentage of correlation in Non-defaulters are the one showing in defaulters correlation.

BIVARIATE ANALYSIS

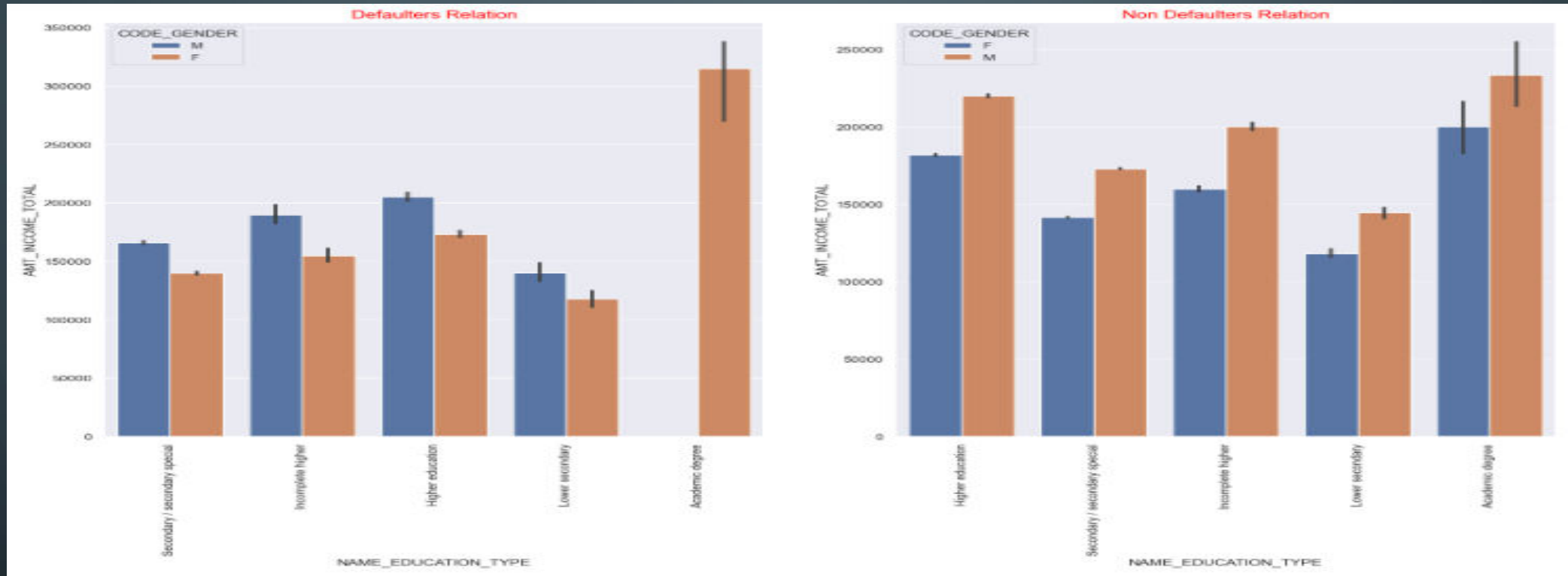


Observation:

For Non-Defaulters: The values are evenly distributed AMT_CREDIT and AMT_INCOME_TOTAL. There is no distinguish between Male and female. It looks like all are having same ratio in this plot. So, It's hard to figure it out here.

For Defaulters: There are more congested values in initial or lower areas of both AMT_CREDIT and AMT_INCOME_TOTAL. With the Income the LOAN value also increases.

BIVARIATE ANALYSIS

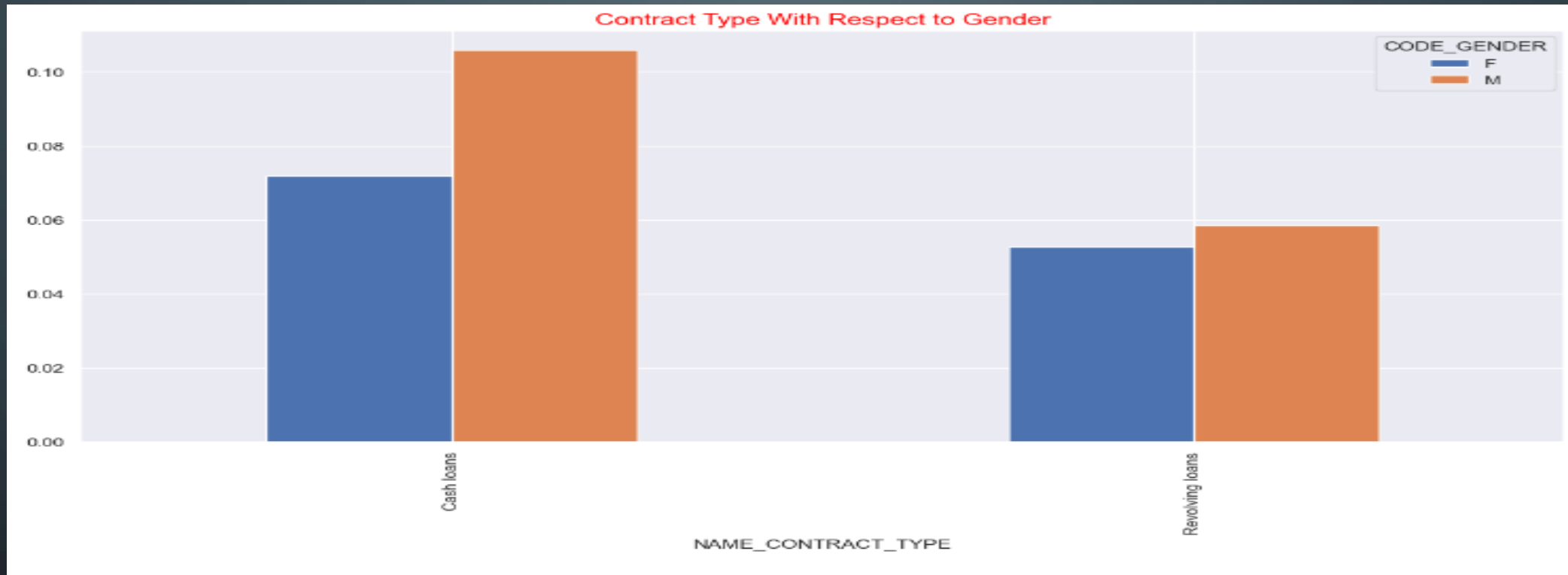


Observation:

Defaulters: Female Academic Degree holders are earning more but still there in defaulter list more. Whereas male Academic degree holders are not in defaulter list. But for rest of the education type we can see that males are earning more as well being the more defaulters.

Non Defaulters: Here also Female Academic degree holders are more and they are in high number when it comes to non defaulters. Nearly same amount non defaulter are there for guys who studied Higher education. Applied to both male and female. Female in all educational aspects are more non-defaulters than male.

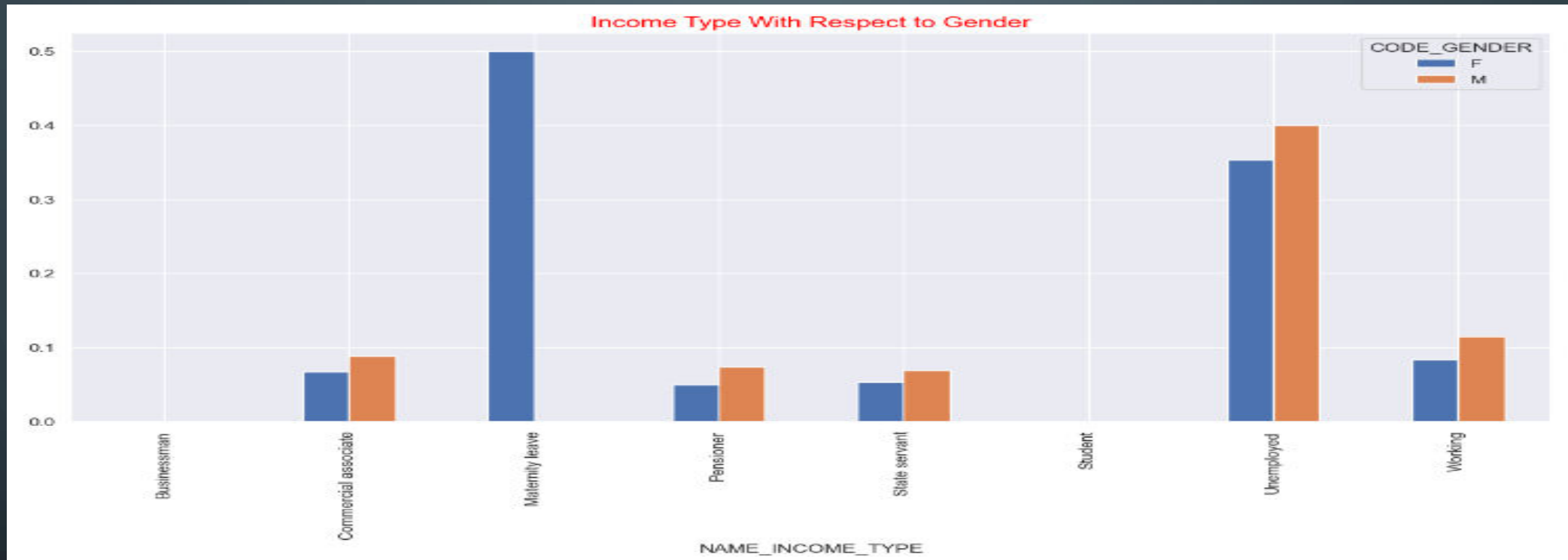
BIVARIATE ANALYSIS



Observation:

Cash loans by males are more as compared to female and they are defaulters as well. However, Revolving loans are also taken by male more and they are most defaulters.

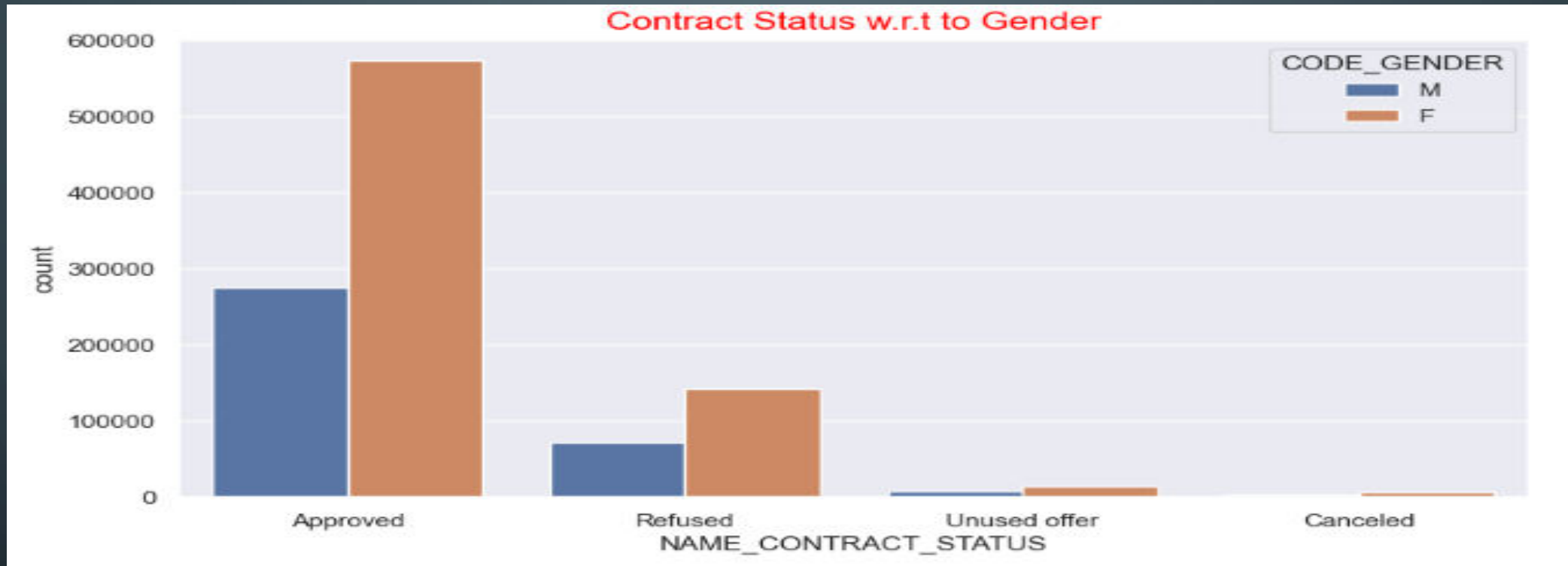
BIVARIATE ANALYSIS



Observation:

- We can see that Unemployed people are more defaulters in both male and female case.
- Males are more unemployed than female.
- Maternity leave females are also in higher no in defaulters list.
- Male numbers are more compare to female in defaulters list.

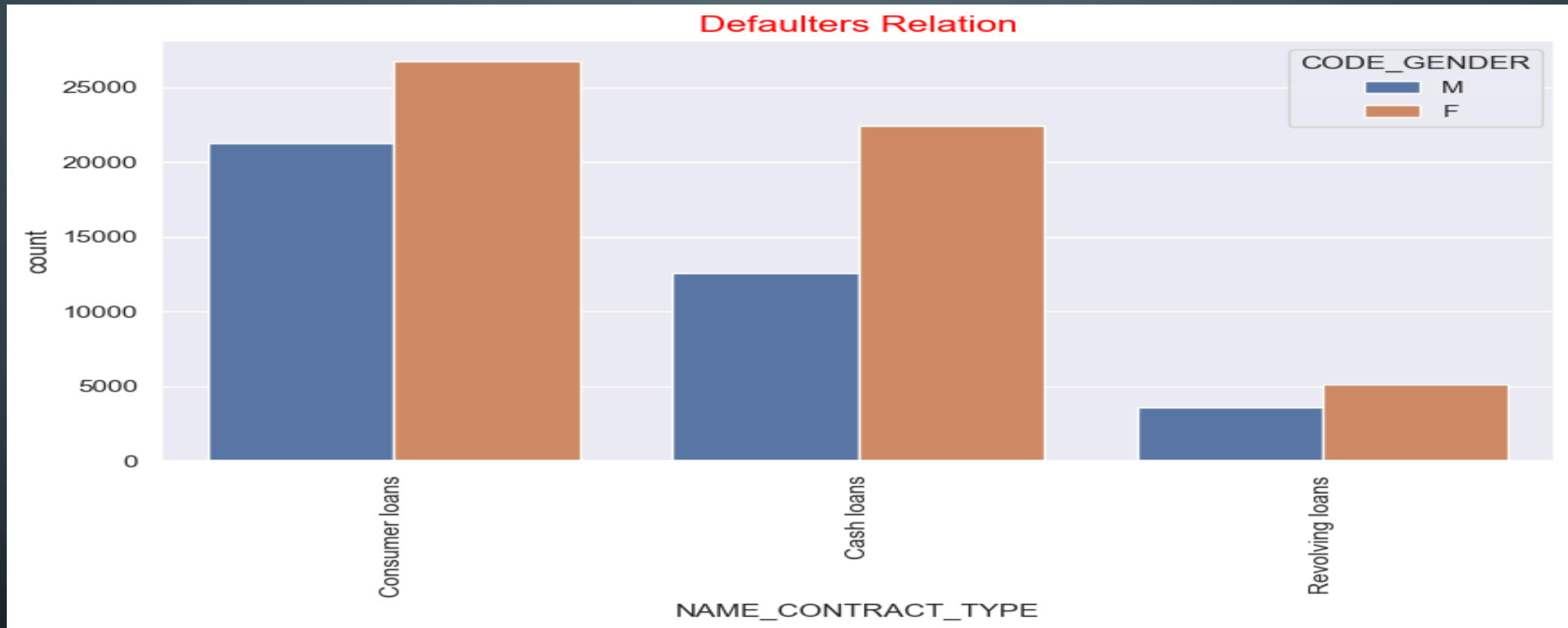
BIVARIATE ANALYSIS



Observation:

Female contracts got approved more compared to male.
Female contracts also got refused more compare to male but that is nearly negligible and Less in numbers compare to approval.

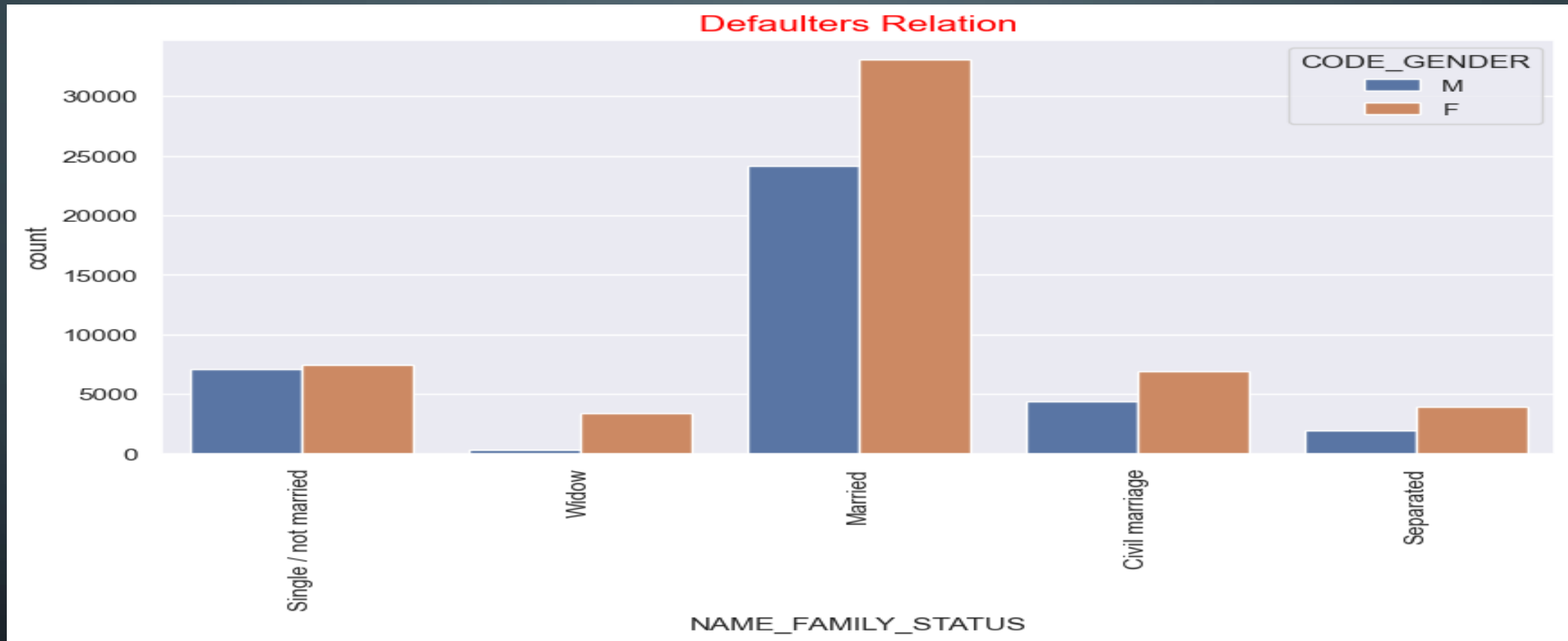
BIVARIATE ANALYSIS USING TARGET VARIABLE



Observation:

Here, female with consumer loans type are the highest defaulters. However, here, females are defaulters in all three categories

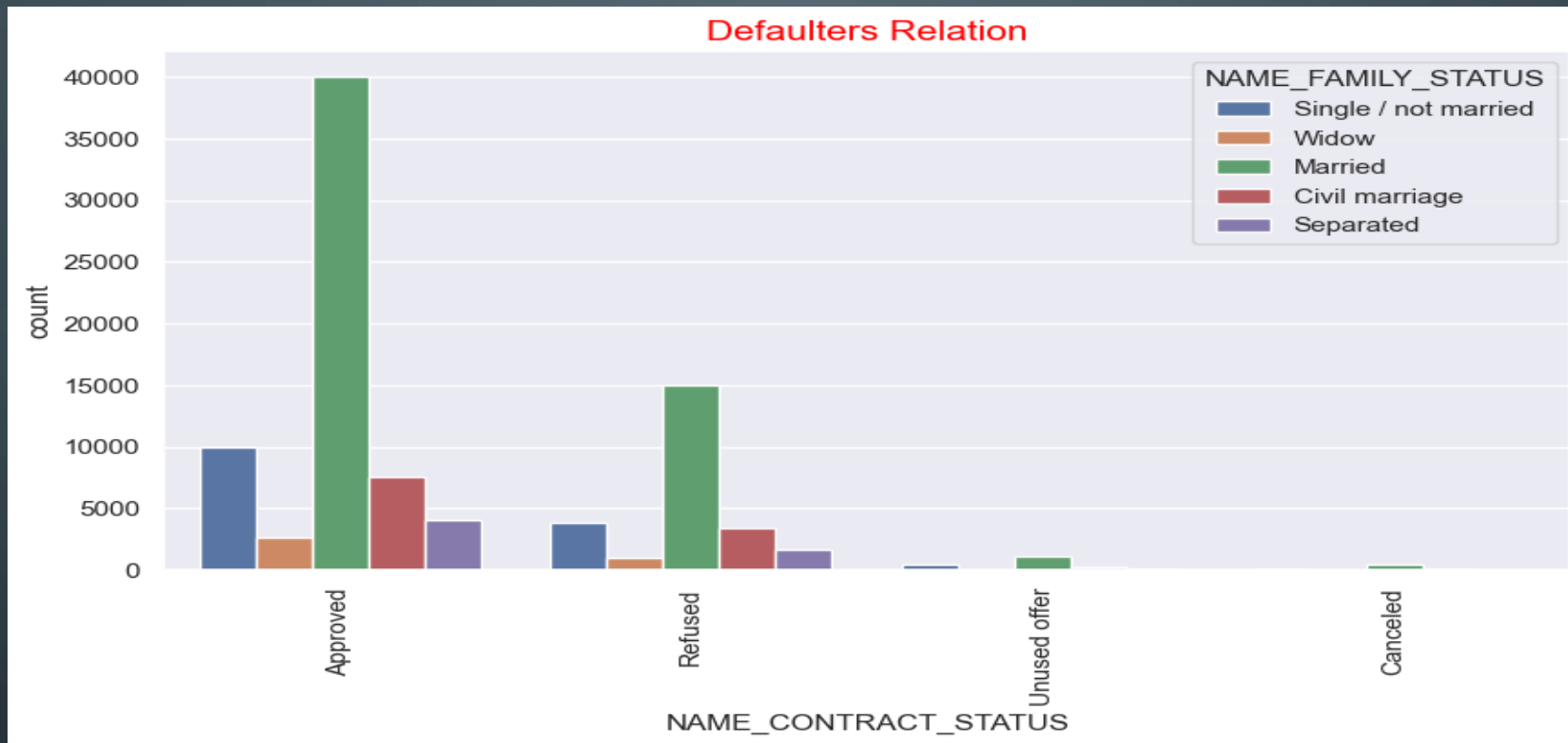
BIVARIATE ANALYSIS USING TARGET VARIABLE



Observation:

Married people are struggling to pay the loan bill compared to single and separated people.

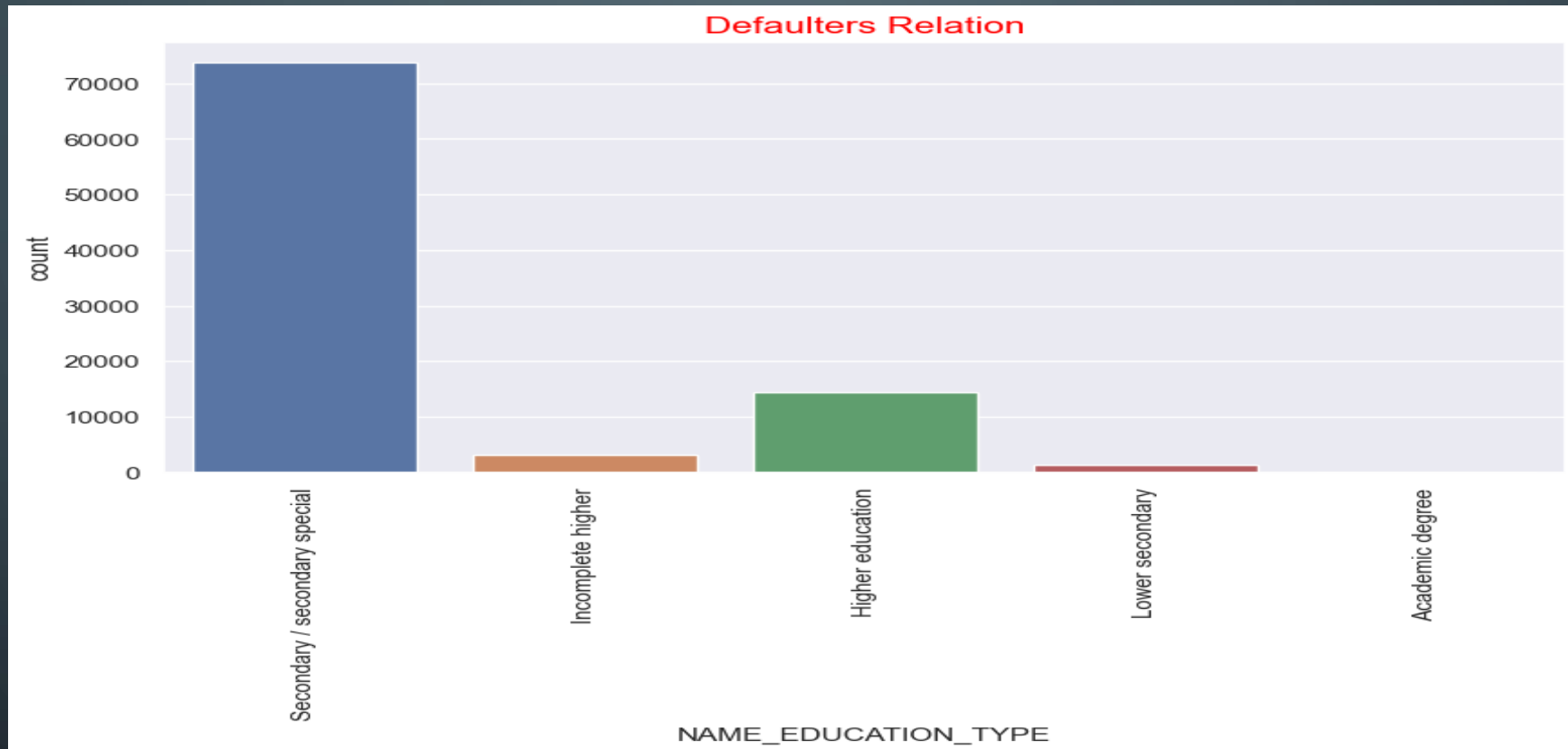
BIVARIATE ANALYSIS USING TARGET VARIABLE



Observation:

More number of approved people are married. From next time onwards we need to give loan to those who are either single or divorced.

BIVARIATE ANALYSIS USING TARGET VARIABLE



Observation:

Secondary/ secondary special people with education are struggling to pay their loan bill and becoming defaulters.

CONCLUSION

- The banks should focus more on contract type student, pensioner, and businessman with income type other than co-op apartment, office apartment for housing type for successful payments.
- The banks should focus less on income type Working category as they are having most number unsuccessful payments.
- CODE GENDER female is having higher rate in both defaulters and non-defaulters. So, bank should focus giving loans to female gender.
- In Name Family Status Married followed by Single/Not Married has high rate in defaulters & Non-Defaulters.
- With loan purpose Repair is having higher number of unsuccessful payments on time.