

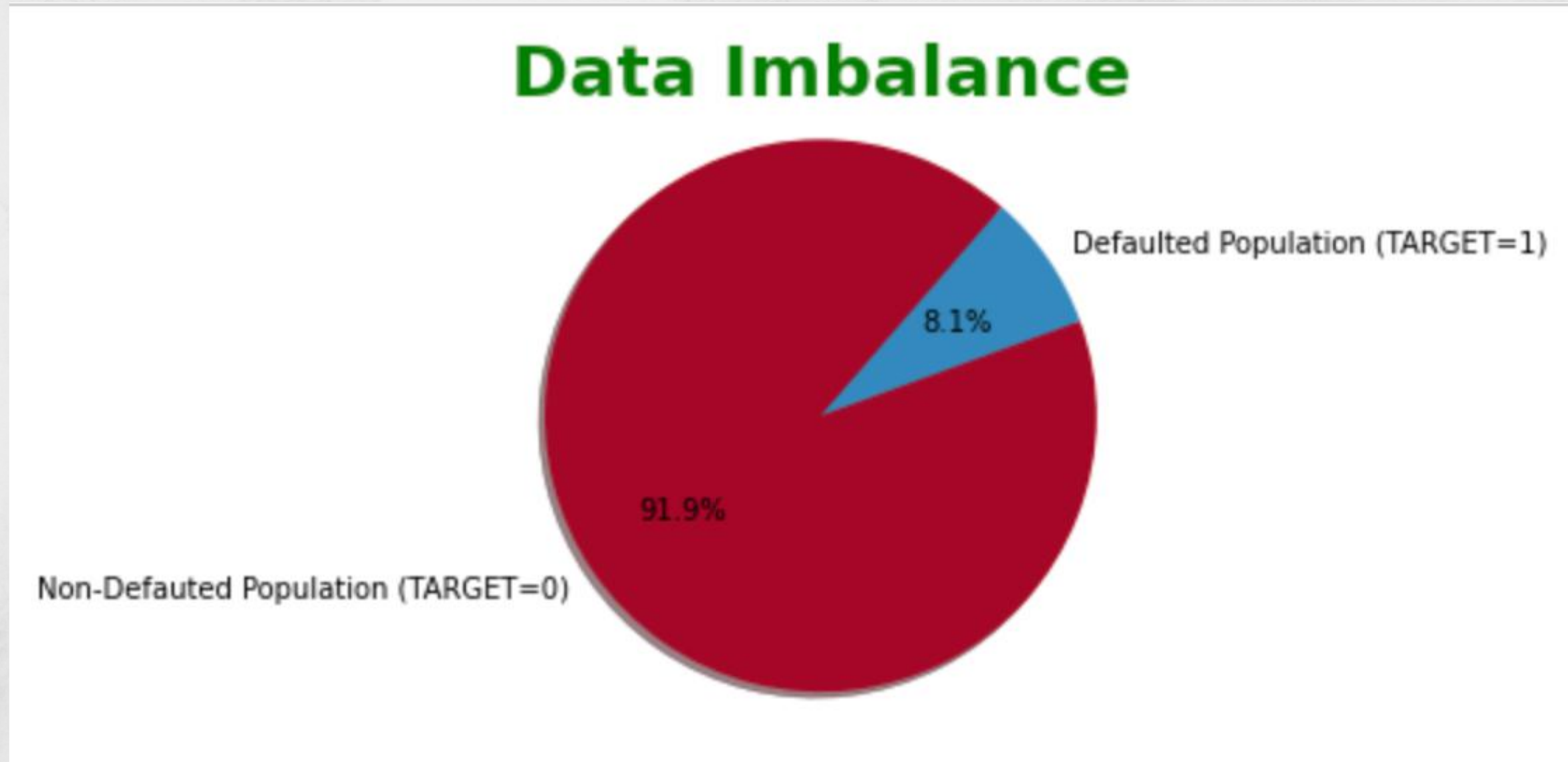
# Credit EDA Case Study

BY :-

AISHWARYA GYANJYOTI  
SHIVANG NAIK

# APPLICATION DATA

# Checking Data Imbalance

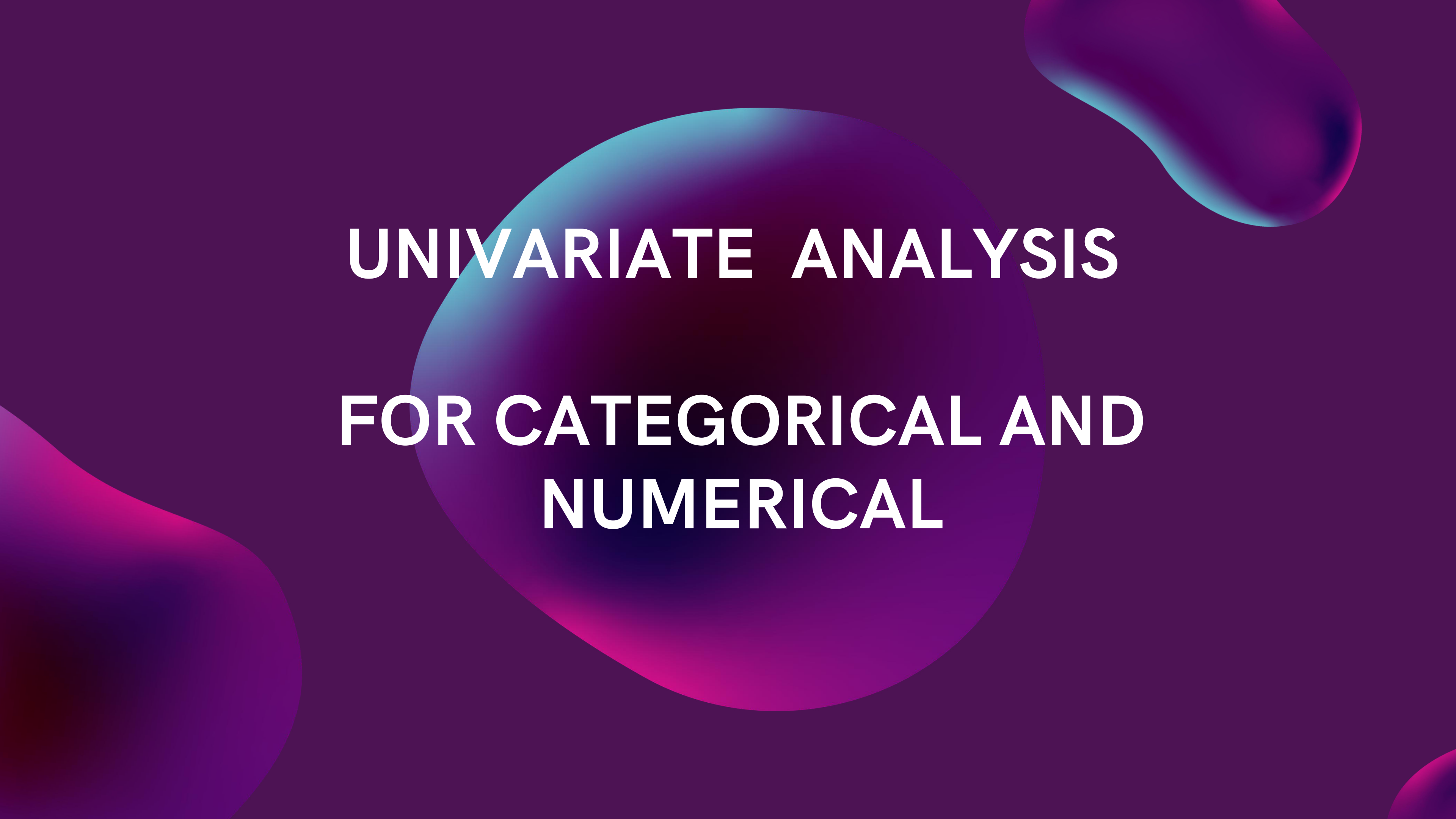


As per the pie chart, 91.9% population are non-defaulter in a loan payment.

And 8.1% are defaulters.

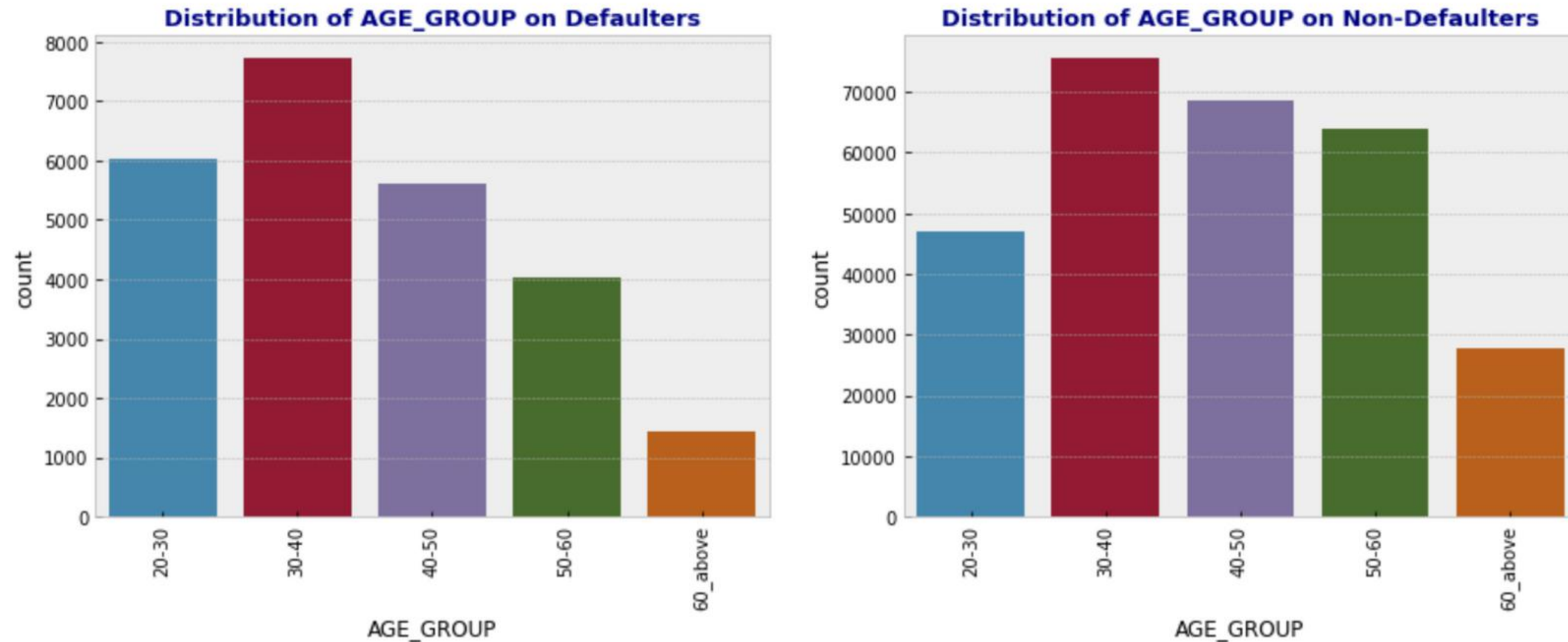
Ratio of the imbalance is 11.4.



The background features a dark purple gradient with several large, overlapping organic shapes in lighter shades of purple and blue. A large, semi-transparent circle is centered behind the text.

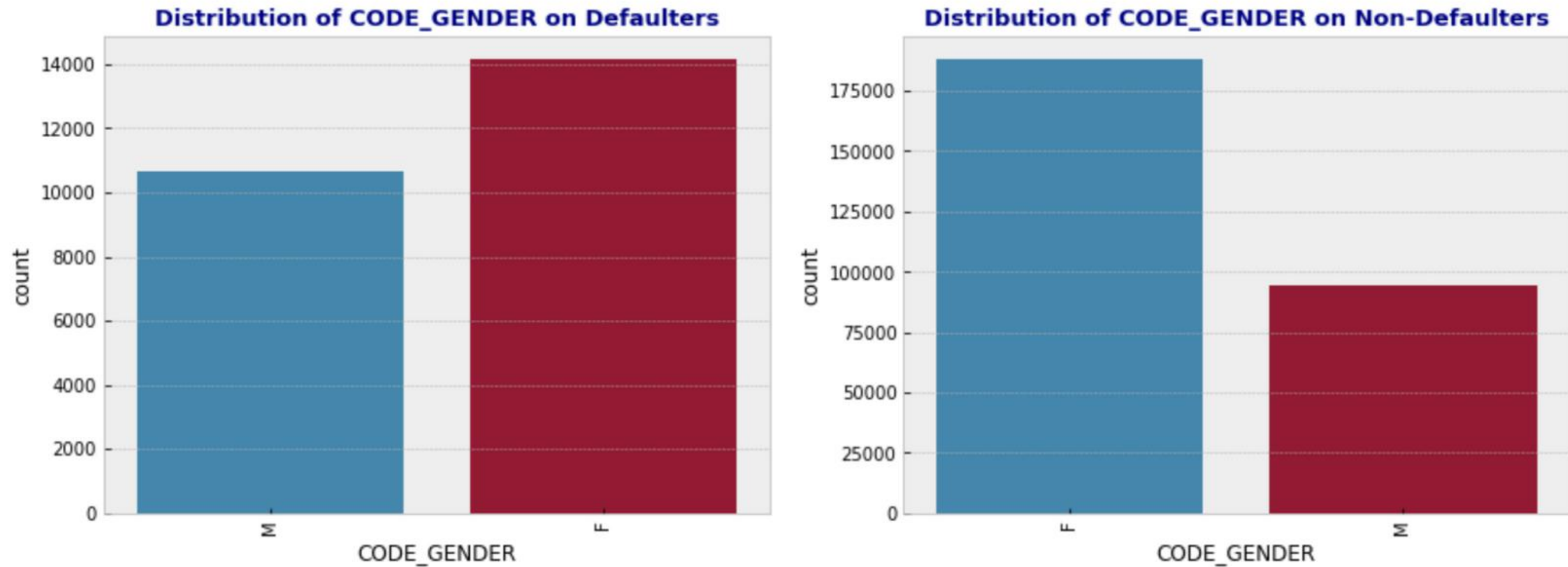
# UNIVARIATE ANALYSIS FOR CATEGORICAL AND NUMERICAL

# AGE GROUP



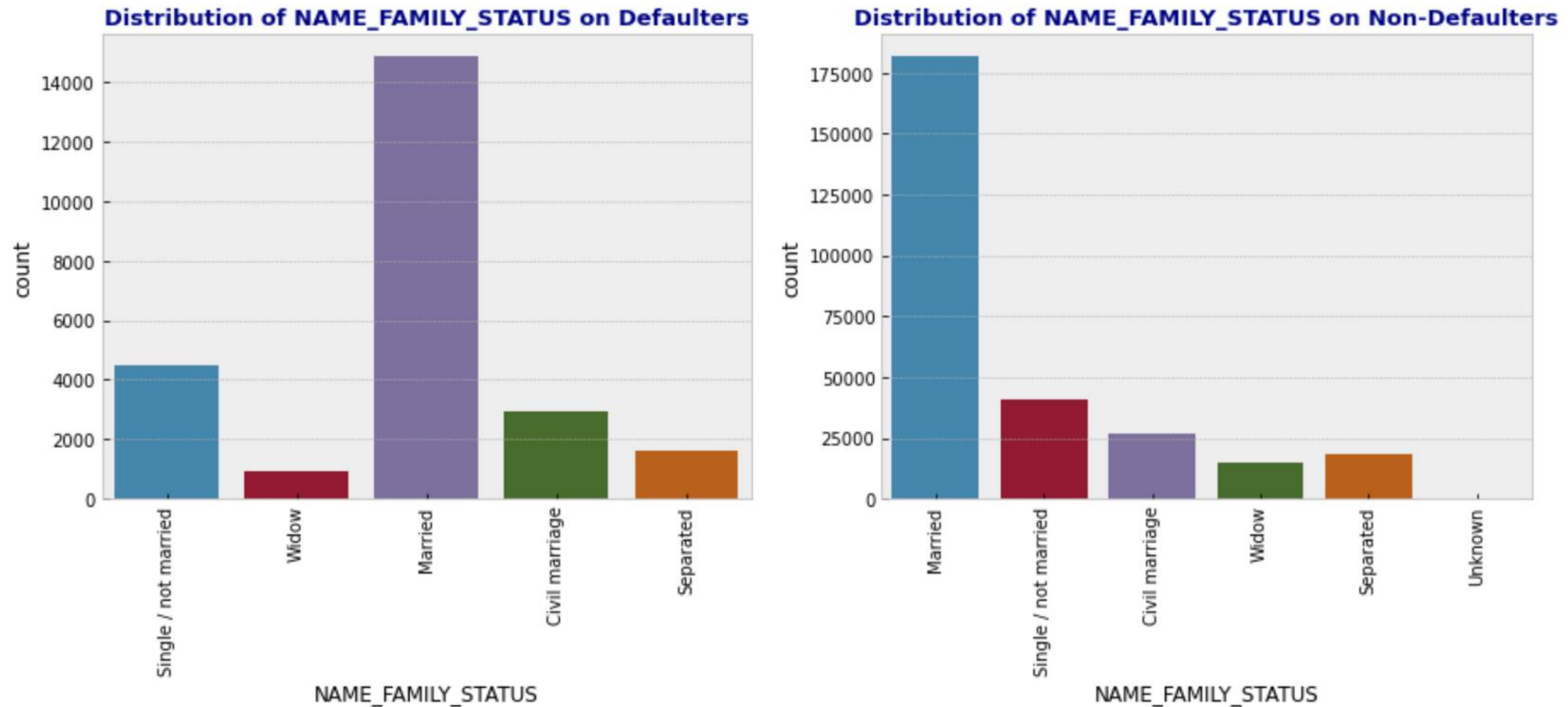
As per the above plot, there seems to be an increase in the number of people in the age group 20-30 that to defaulters (facing payment difficulties in loan) while comparing with non-defaulters.

# GENDER



Females are the majority in both Defaulters and Non-Defaulters. Undoubtedly, most females have applied for loan than males.

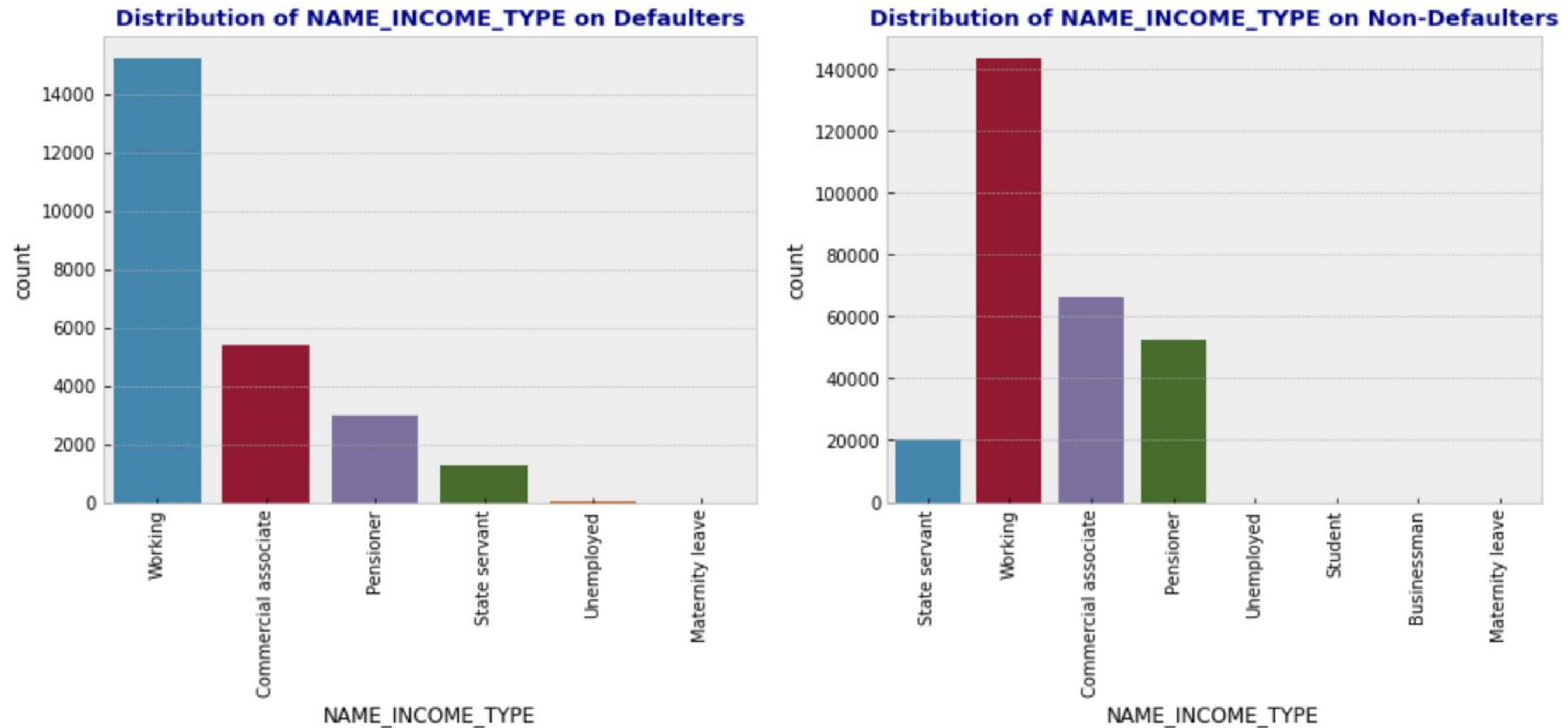
# FAMILY STATUS



Most of the people are Married. Single / not married is proportionally higher in defaulter population as compared to non defaulted population. This concludes that Single applicants have higher defaults.



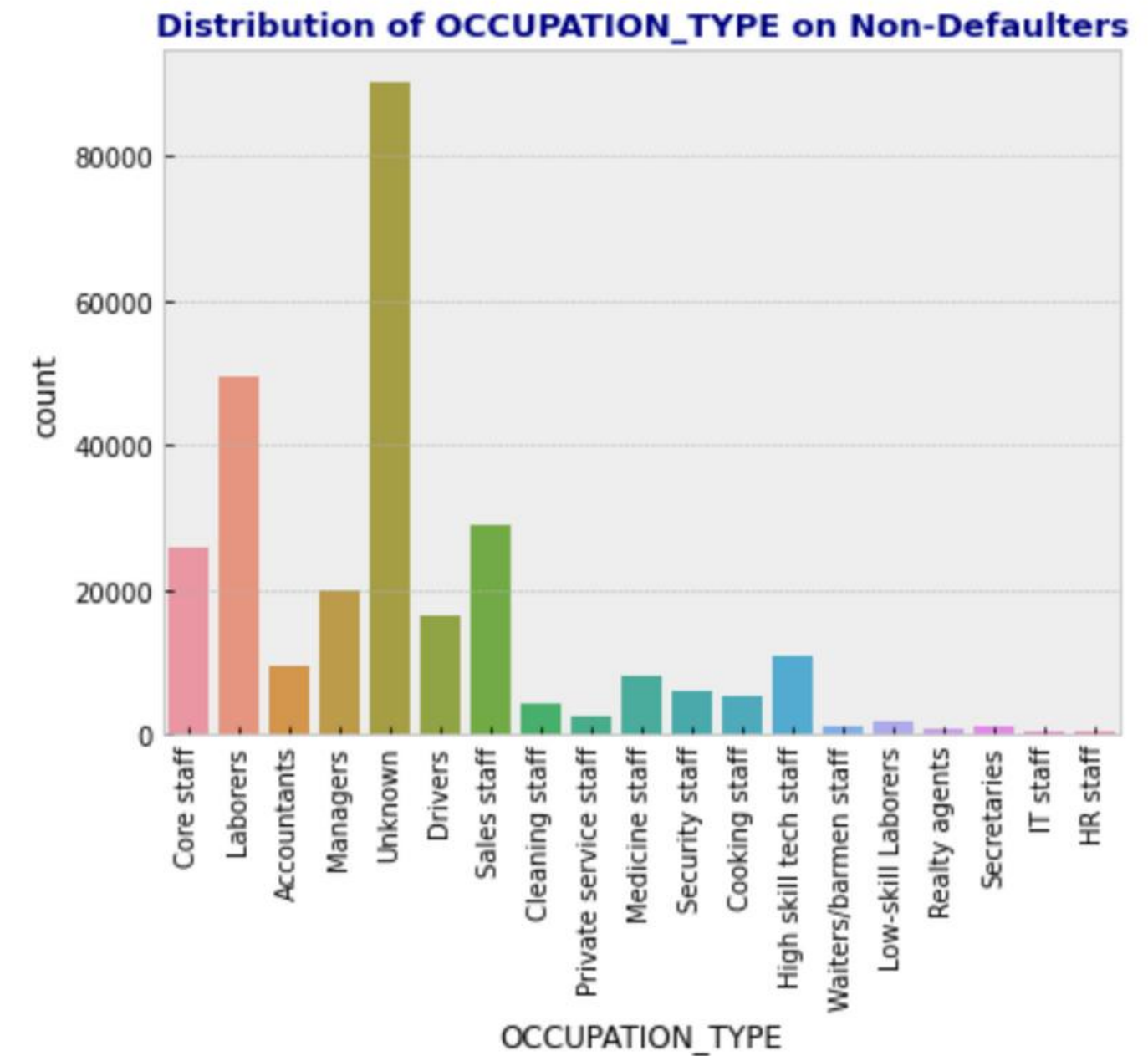
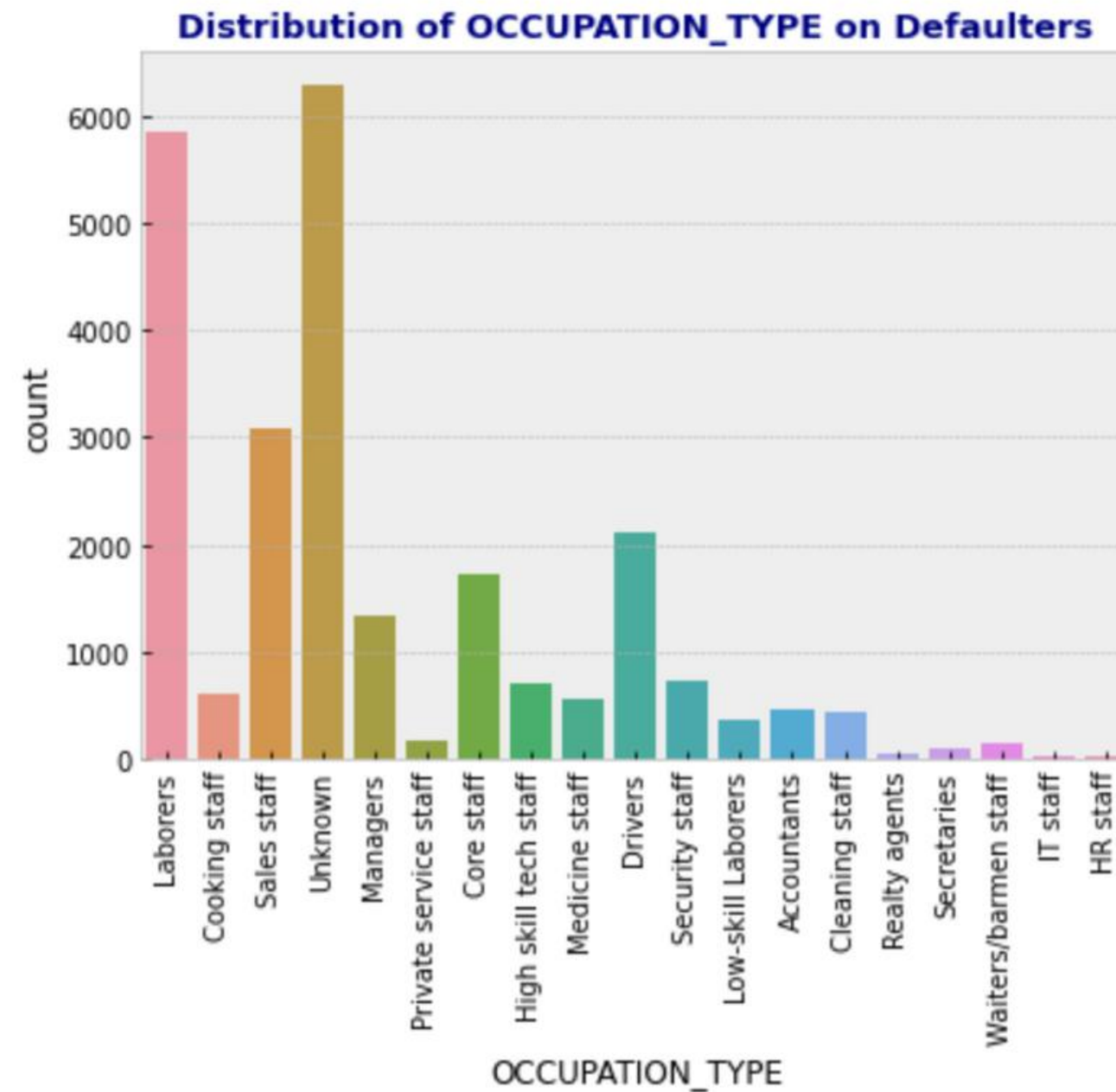
# INCOME TYPE



- It seems like most of the defaulters are the working people, so as in Non-Defaulter. And there seems to be a moderate rate in Commercial associate, Pensioner, and state servant.
- Least count to be seen in Unemployed, Maternity Leave, Businessman and Student.



# OCCUPATION TYPE

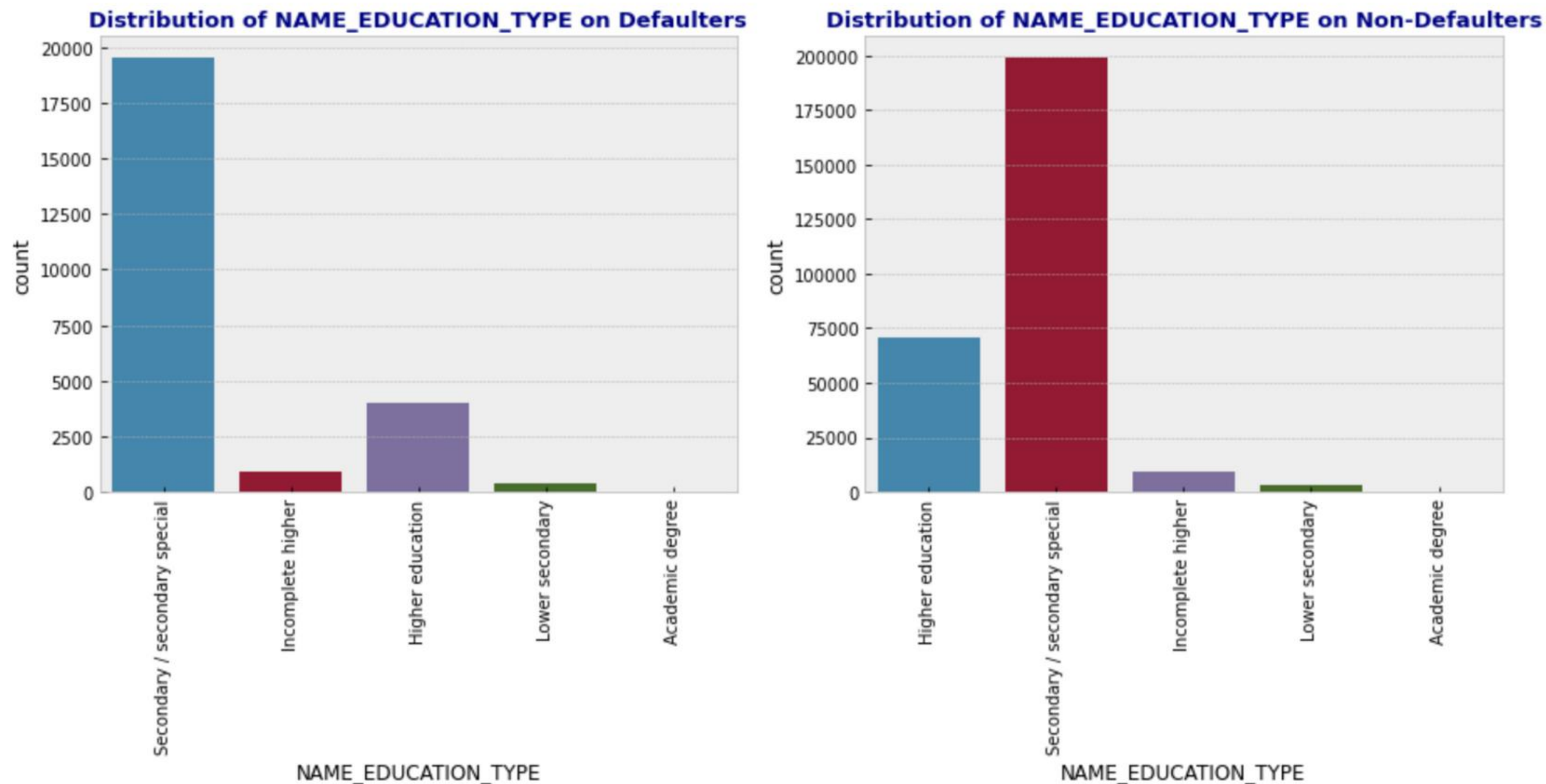




In the above plot,

- Sales, core staff has the highest number of loan applications but when it comes to Non-repayment of the loan then the highest number can be seen in Low-skill laborer's, Drivers, Sales staff and so on.
- People having other occupations and laborers are highest in both cases.
- Sales Staff, Core staff has highest number of applicants in Defaulters.

# EDUCATION TYPE



Higher education count is proportionally lesser in defaulted population as compared to non defaulted population. Hence higher the education level, the lower the default rate.



# ORGANIZATION TYPE



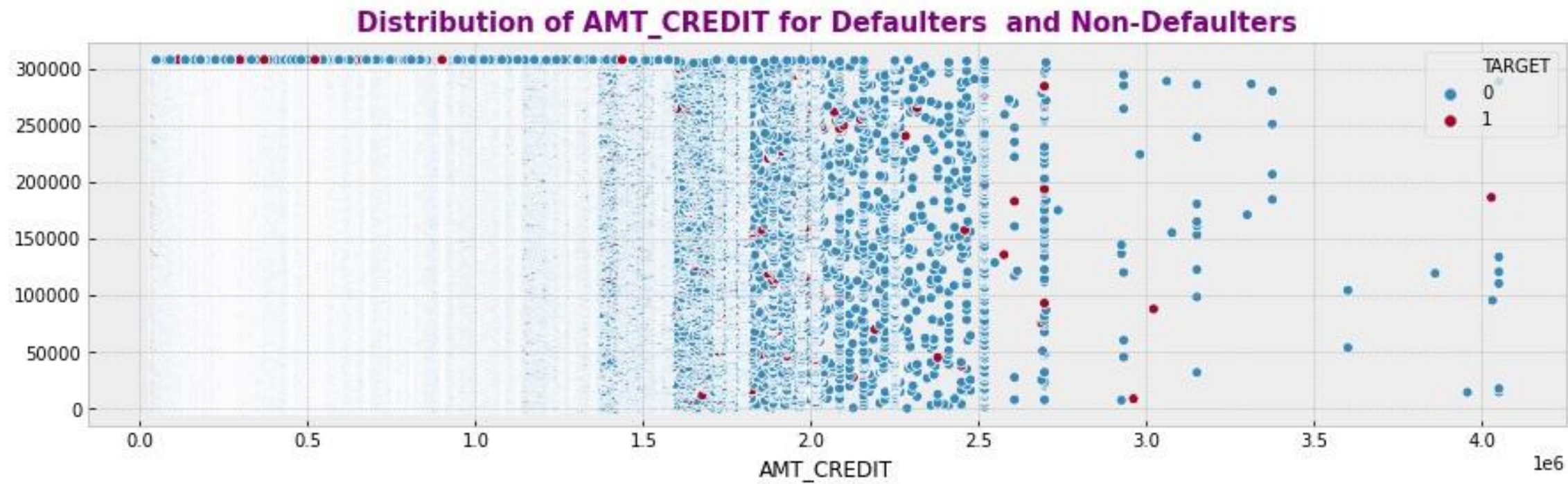
## **For Non - Defaulters:**

- The highest number of loan applications are from most of the organization type 'Other', 'Business entity Type 3', 'Self-employed', 'Medicine' and 'Government'.
- Fewer applications are from Industry type 8, type 5, type 13, type 4, and religion.

## **For Defaulters:**

- The highest number of loan applications are from most of the organization type 'Business entity Type 3', 'Other', 'Self-employed', 'Business entity Type 2', and 'Construction'.
- Less applications are from Industry type 4, type 8, type 5, type 10 and religion.

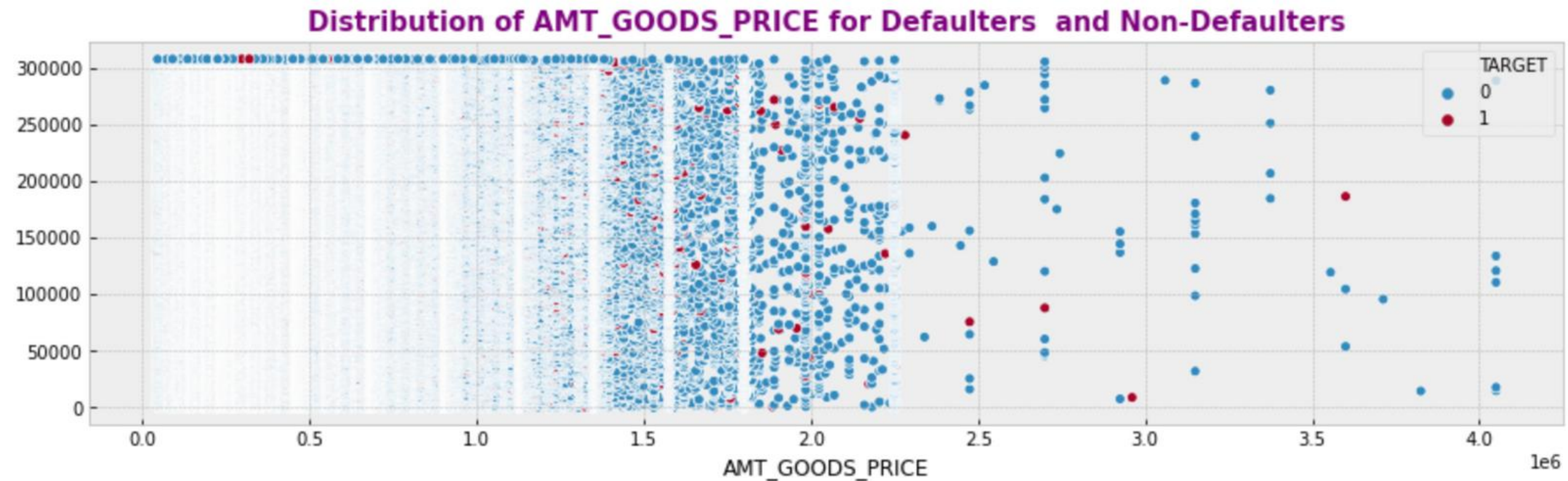
# AMT\_CREDIT



In the above scatterplot, there's a visibility of outliers and most of the values are in between (0.1 lakh to 2 lakh).



# AMT\_GOODS\_PRICE



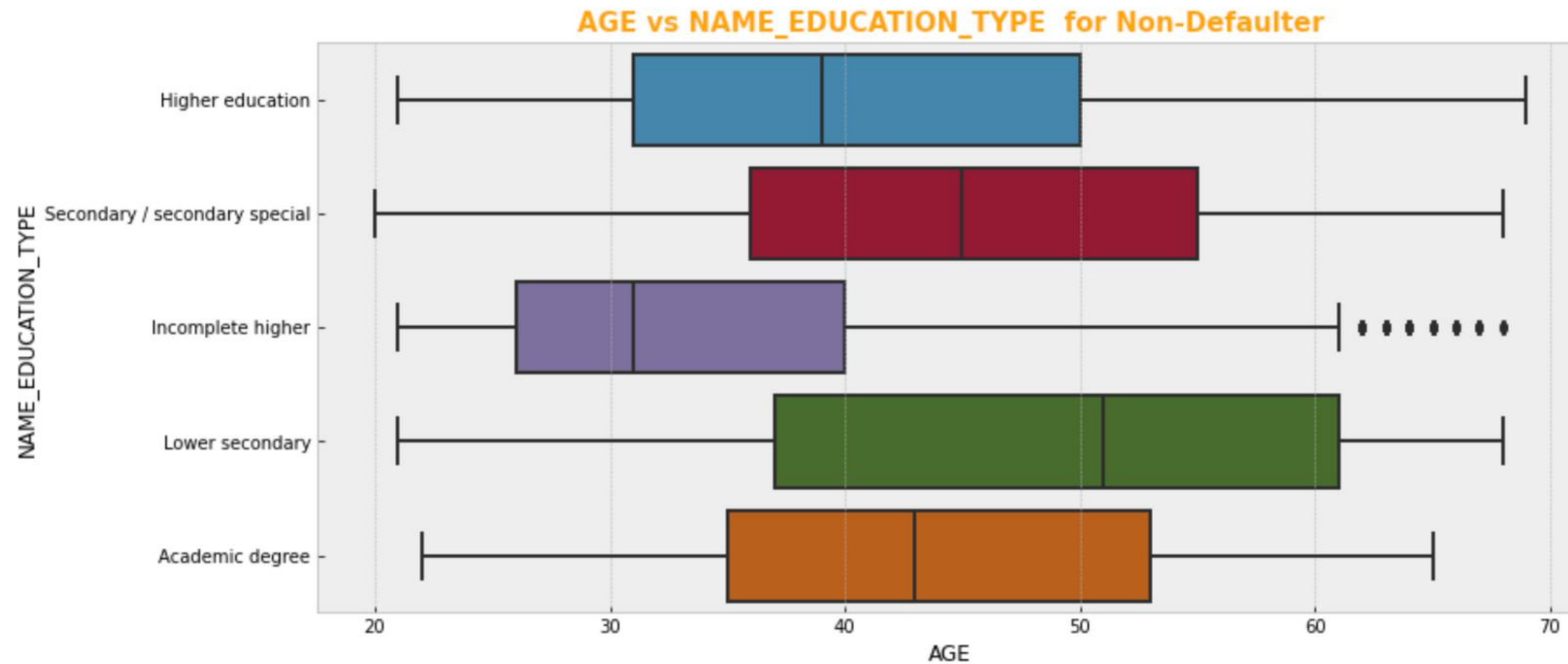
In the above scatterplot, there's a visibility of outliers and most of the values are up to 1.5 lakh.



# BIVARIATE ANALYSIS AND CORRELATION

# AGE Vs EDUCATION TYPE

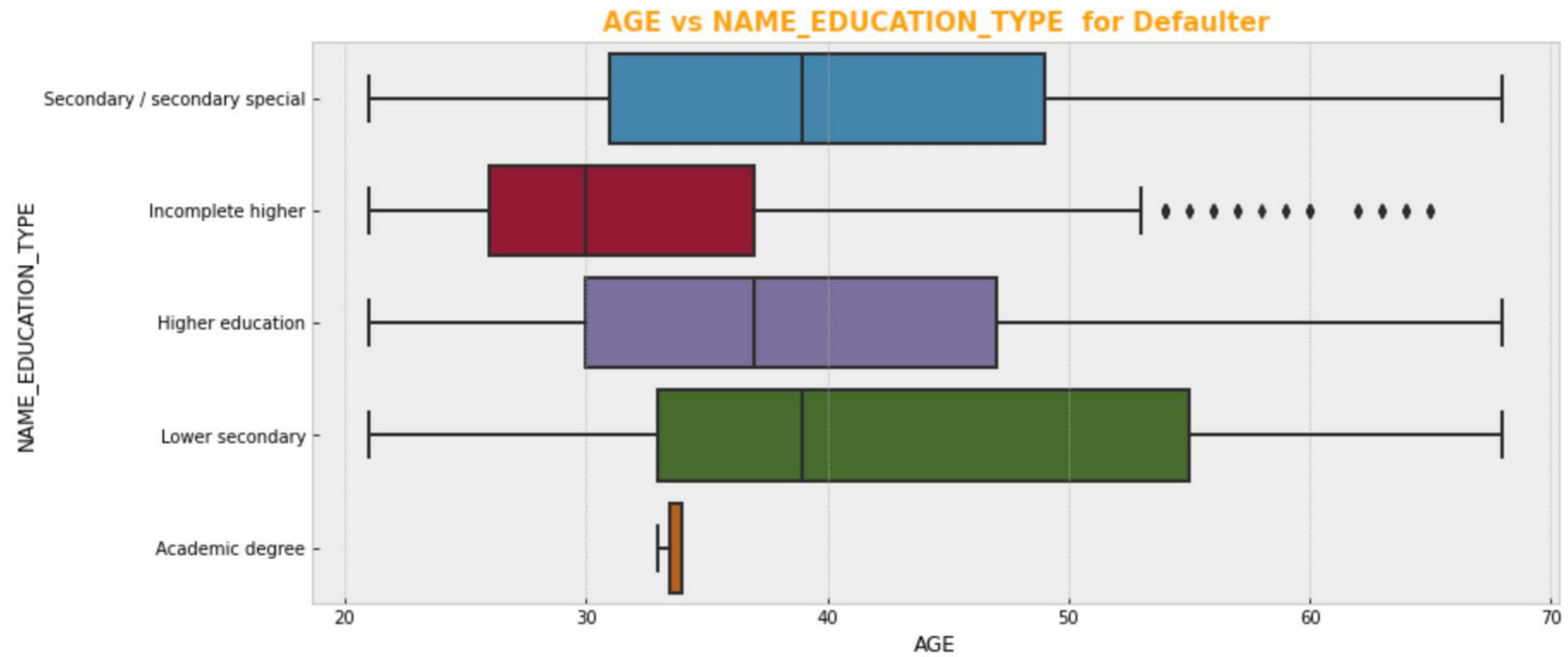
For Non- Defaulter





# AGE Vs EDUCATION TYPE

For Defaulter





For Non-Defaulter,

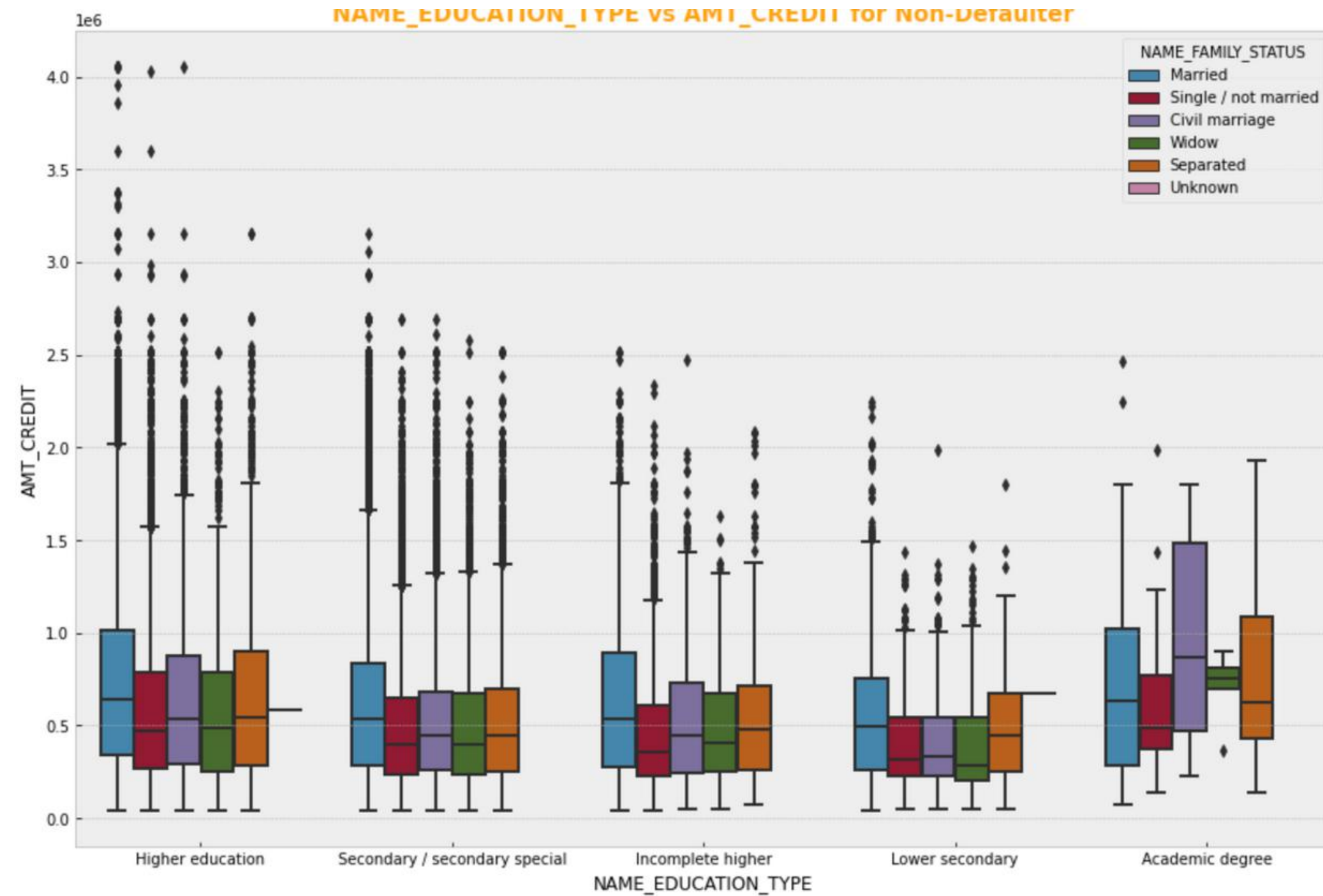
- There is an outliers in Incomplete higher in between age 60-70.
- Apart from Incomplete higher where the first quartile lies at age 40, other education types aren't facing much difficulties in loan repayment.

For Defaulter,

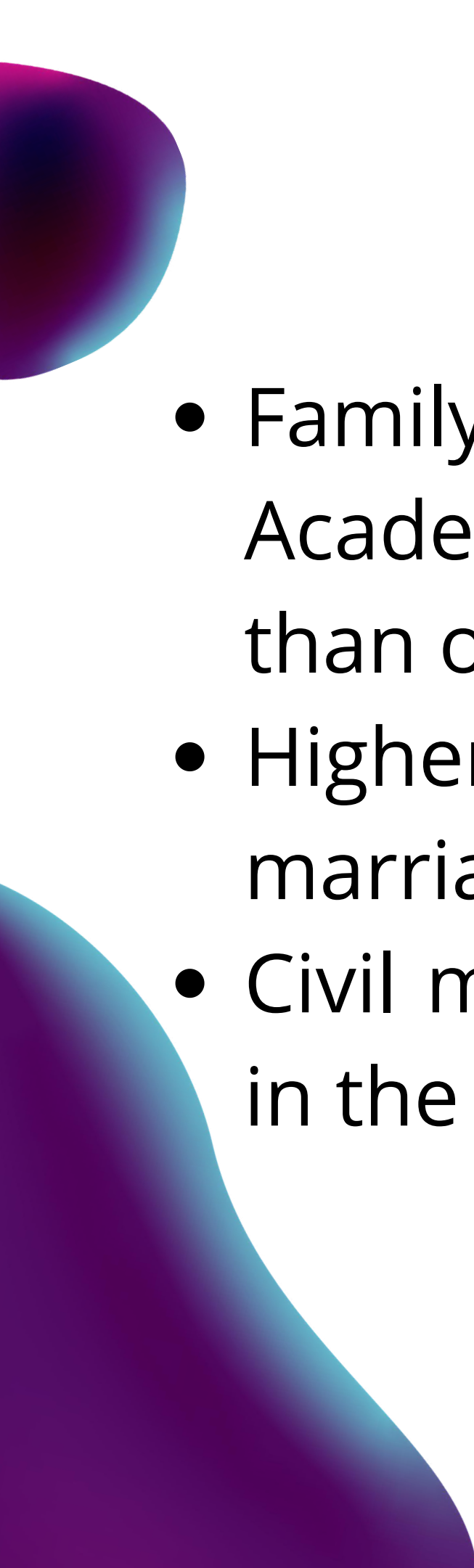
- There is the visibility of an outliers in Incomplete higher from the age group between 50-70.
- People with an age group 30-40 and education type as Academic degree and Incomplete higher seem to be facing difficulties in loan repayment.

# EDUCATION TYPE Vs AMT\_CREDIT

For Non-Defaulter

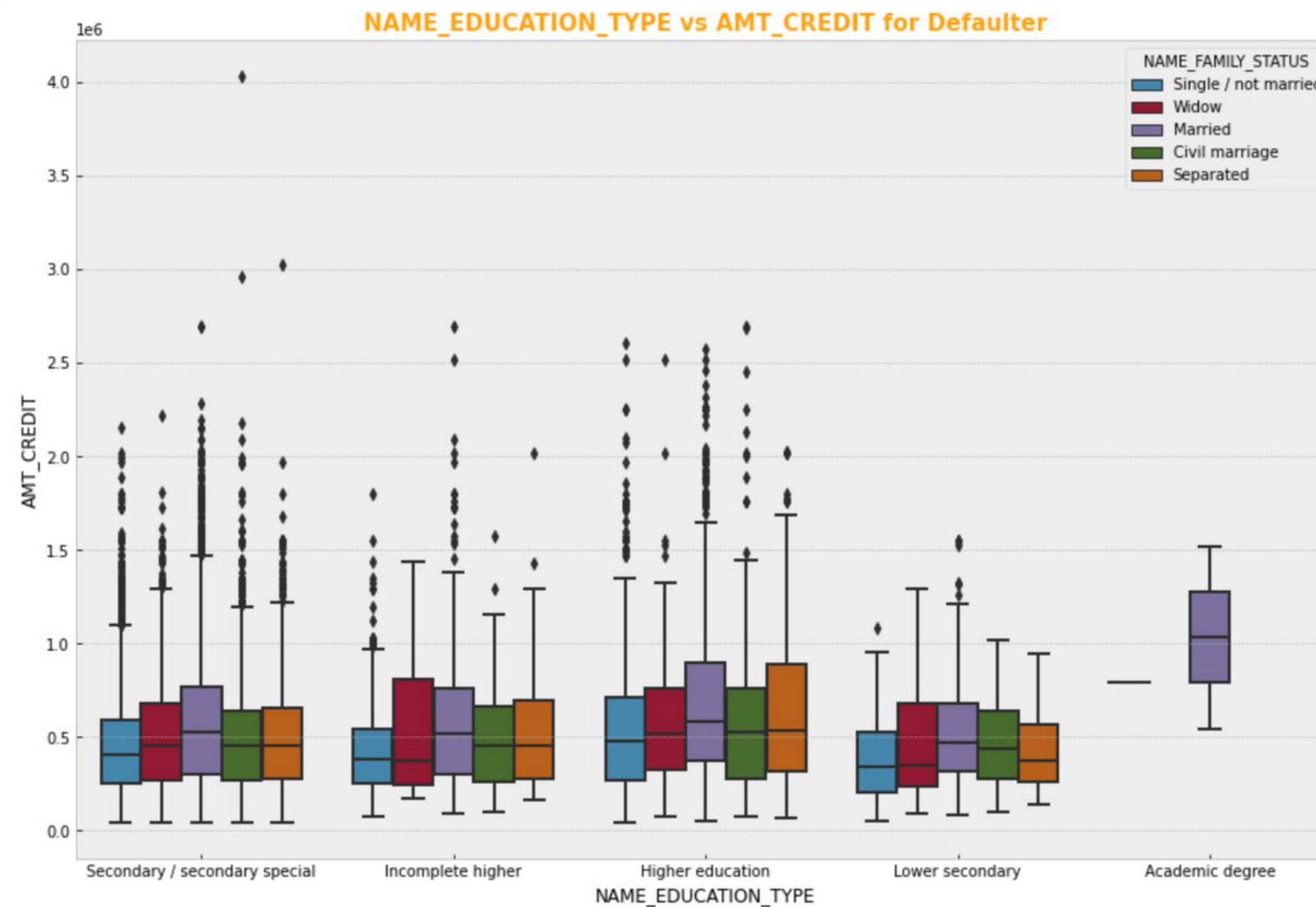


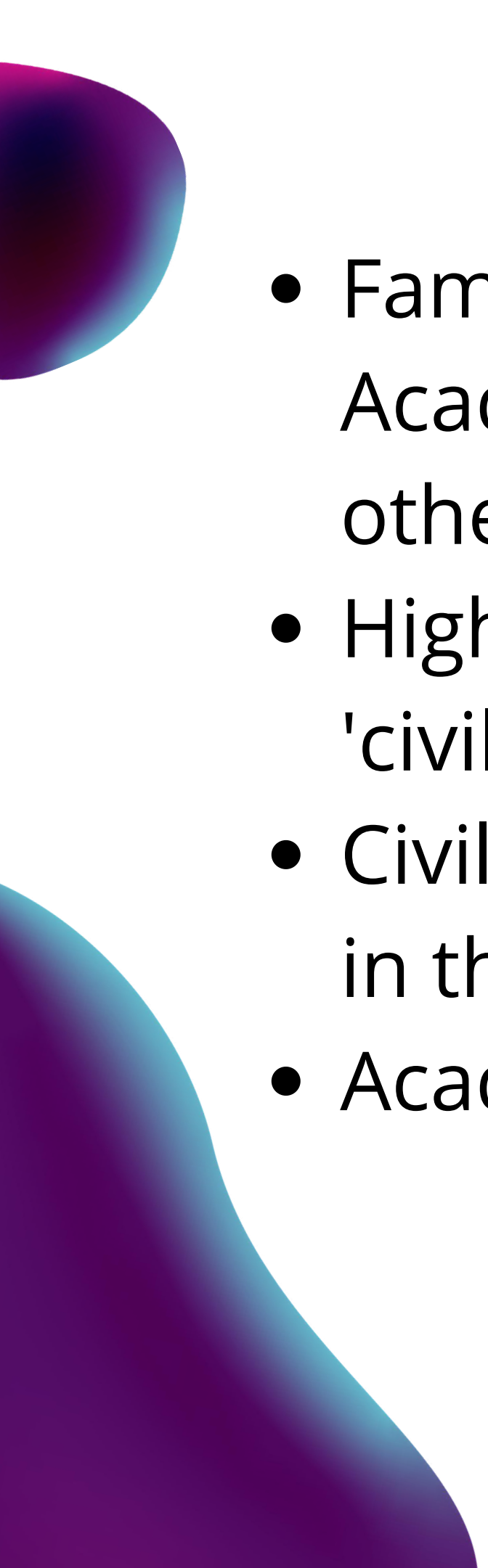


- 
- Family status of 'civil marriage', 'marriage' and 'separated' of Academic degree education are having higher credit amount than others.
  - Higher education of family status of 'marriage', 'single' and 'civil marriage' are having more outliers.
  - Civil marriage for Academic degree is having most of the credits in the third quartile.

# EDUCATION TYPE Vs AMT\_CREDIT

For Defaulter

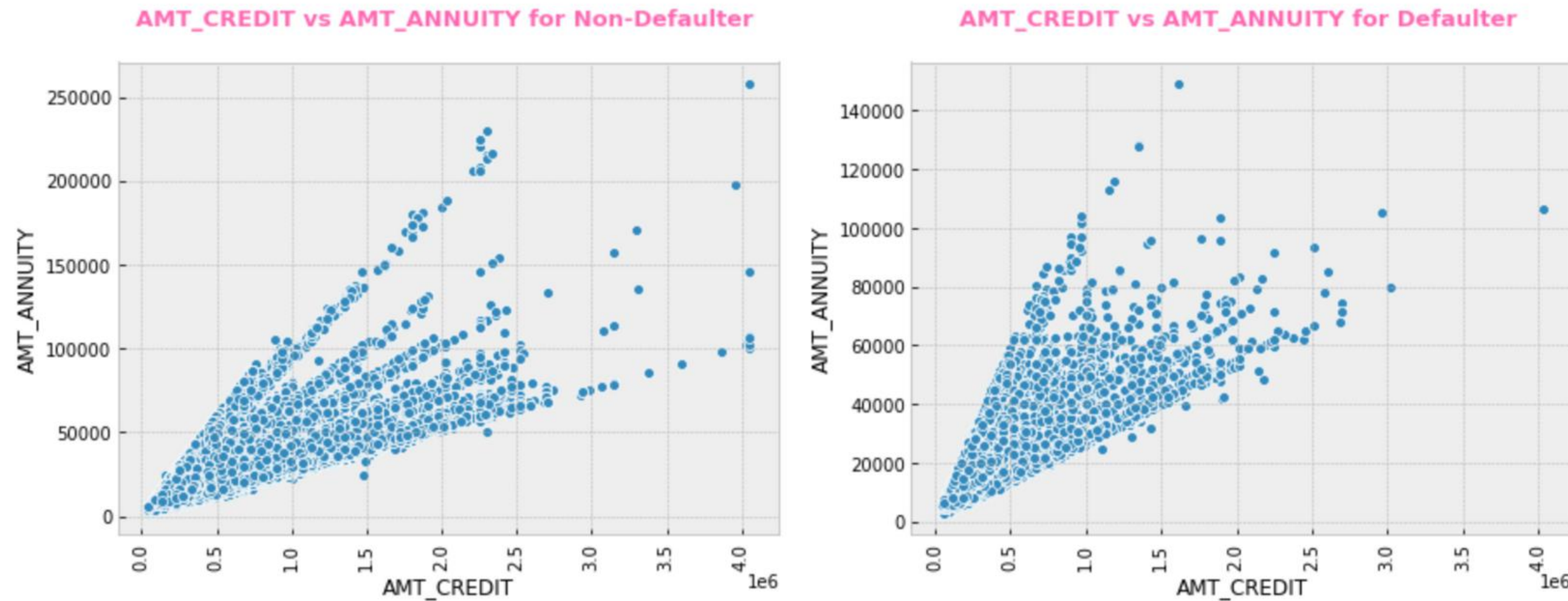


- 
- Family status of 'civil marriage', 'marriage' and 'separated' of Academic degree education have higher credit amount than others.
  - Higher education of family status of 'marriage', 'single' and 'civil marriage' seems to have more outliers.
  - Civil marriage for Academic degree has most credit amount in the third quartile.
  - Academic education category has less number of defaulters.



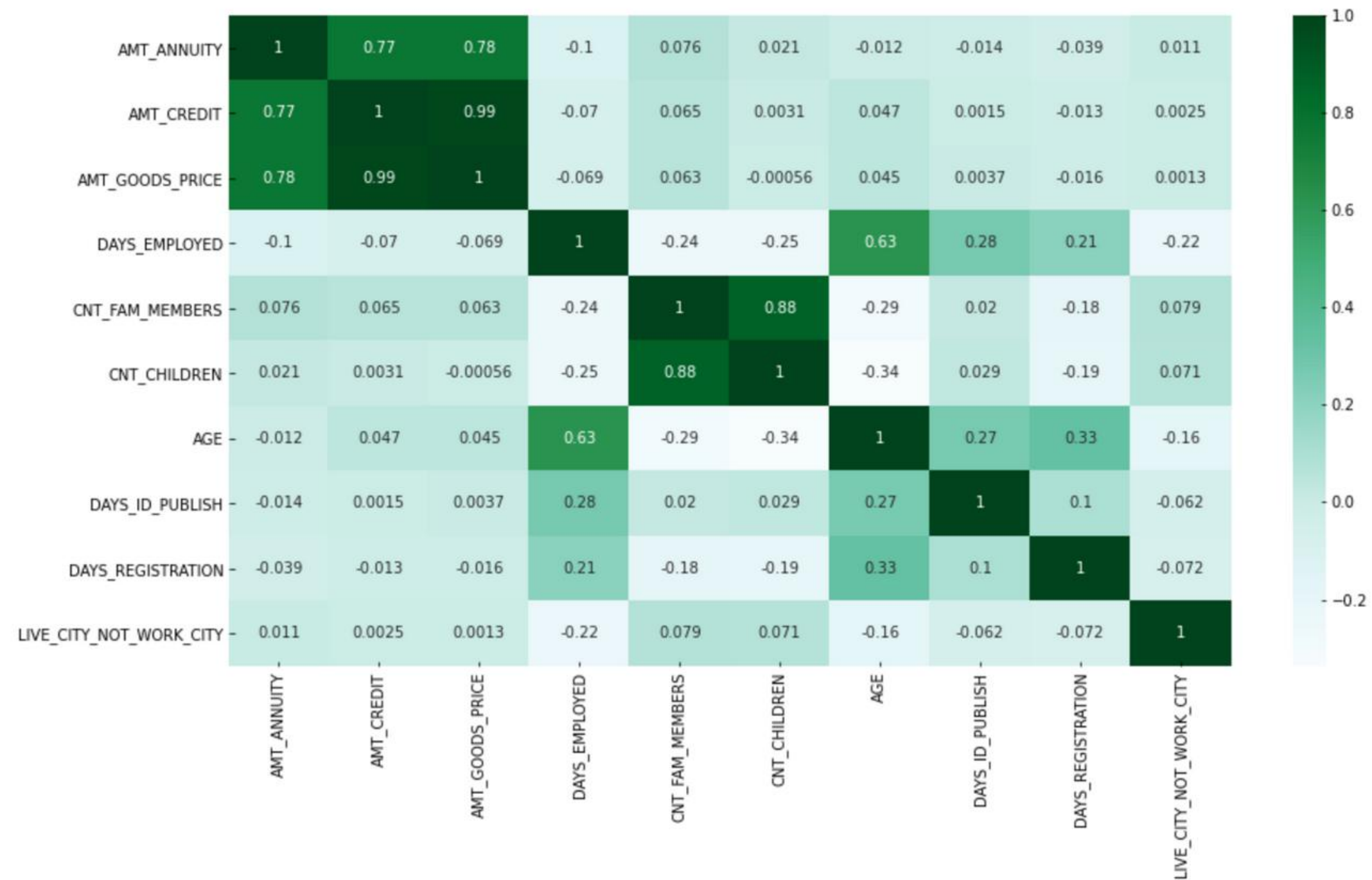
# AMT\_CREDIT Vs AMT\_ANNUITY

For Non-Defaulter and Defaulter

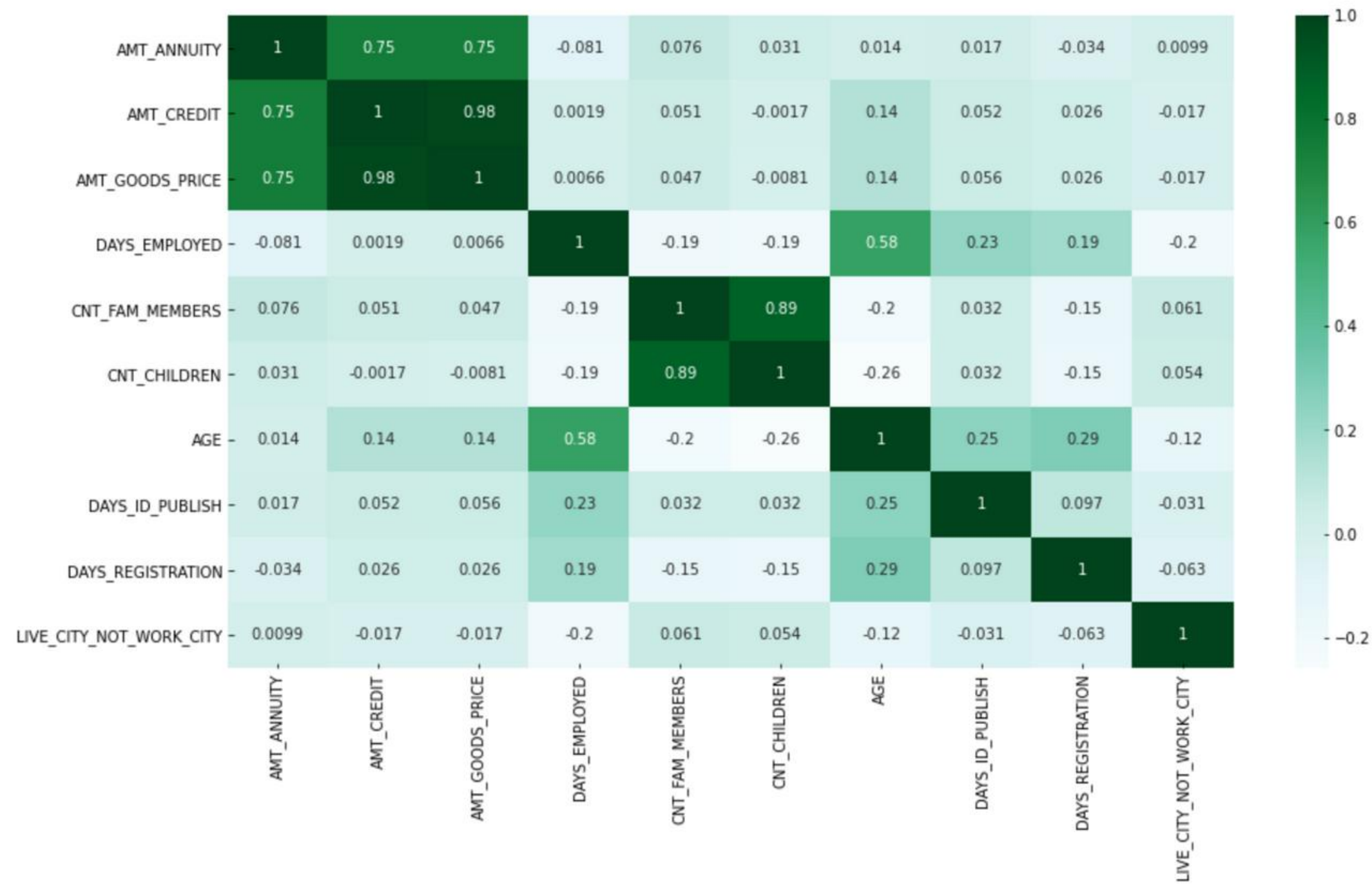


In the above scatter plot, the credit amount increases with the increase in the annuity amount. The slope seems to be a bit higher in Defaulters.

# CORRELATION FOR TARGET = 0



# CORRELATION FOR TARGET = 1

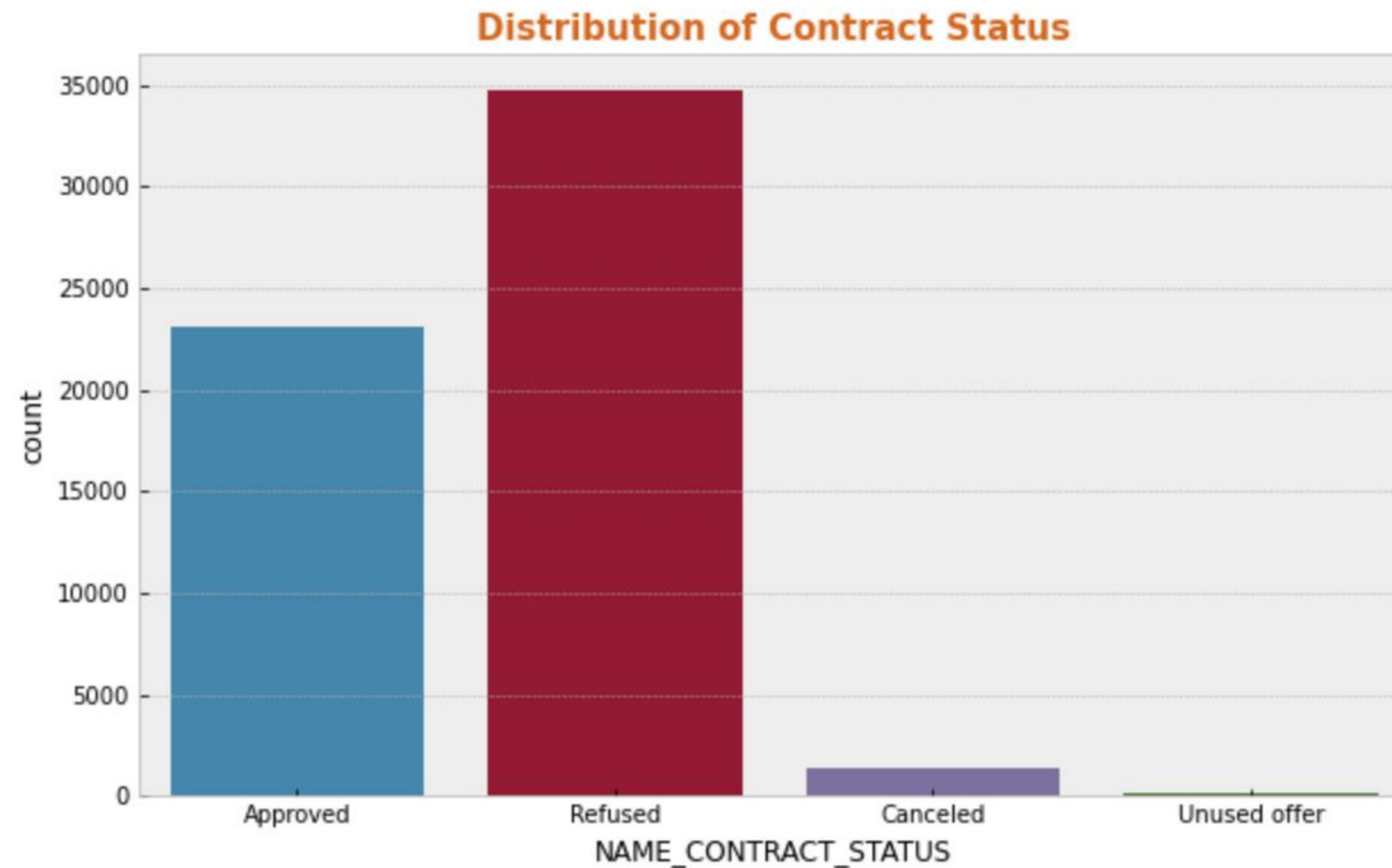




The background features a dark purple gradient with three prominent organic, fluid shapes. A large, light purple shape on the left side curves around the text. A smaller, elongated shape is positioned at the top center, and a circular shape is located at the bottom center. All shapes have a subtle gradient from light purple to dark purple.

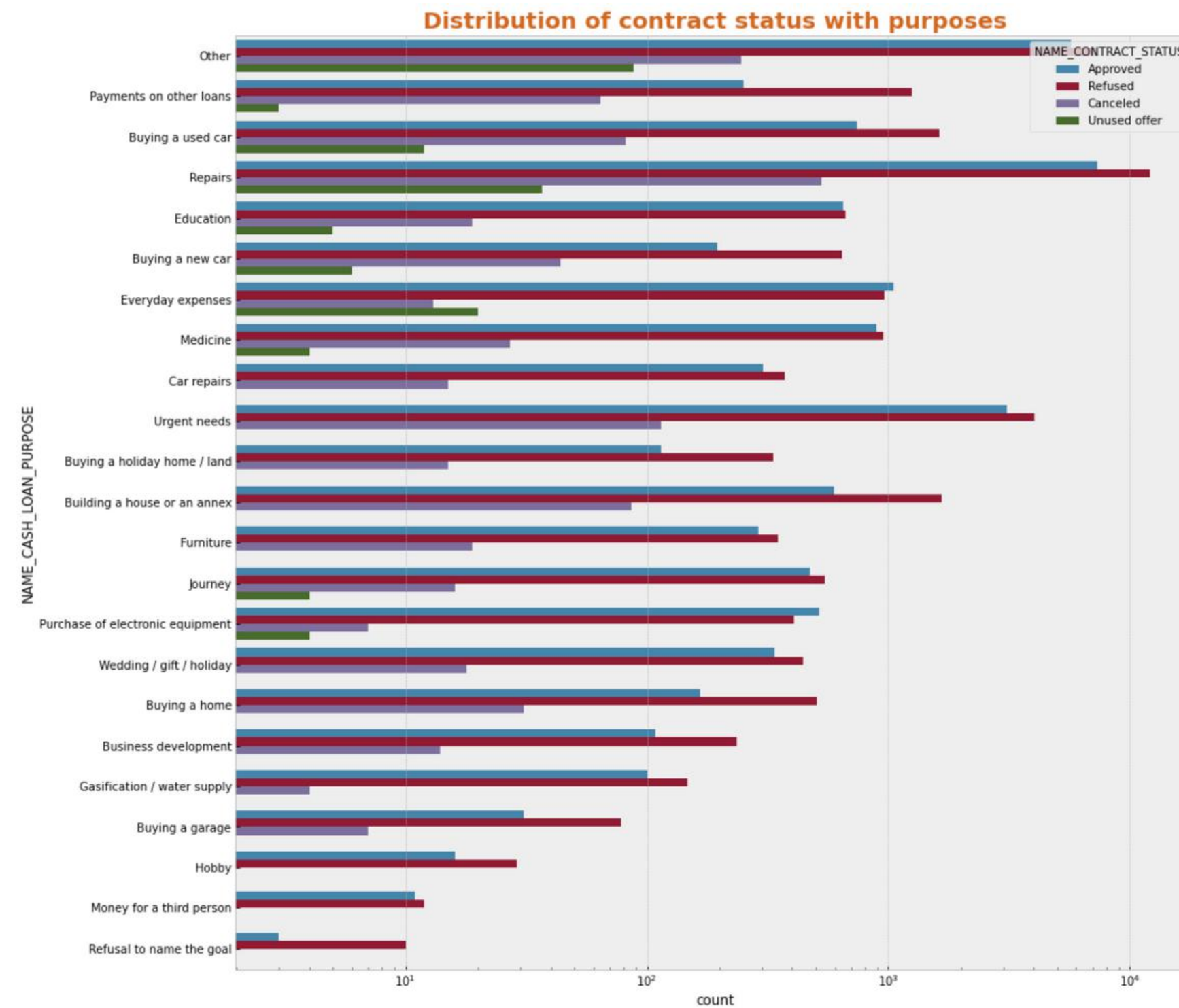
**PREVIOUS APPLICATION**

# DISTRIBUTION OF CONTRACT STATUS

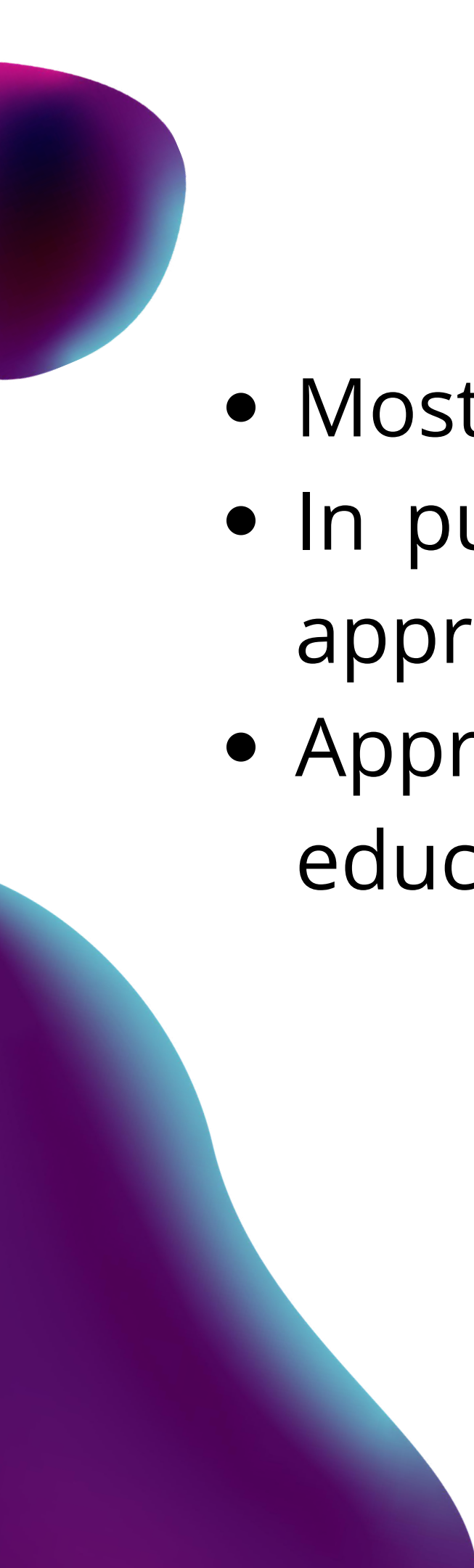


In the above plot, the refused application seems to be more than other status of applications.

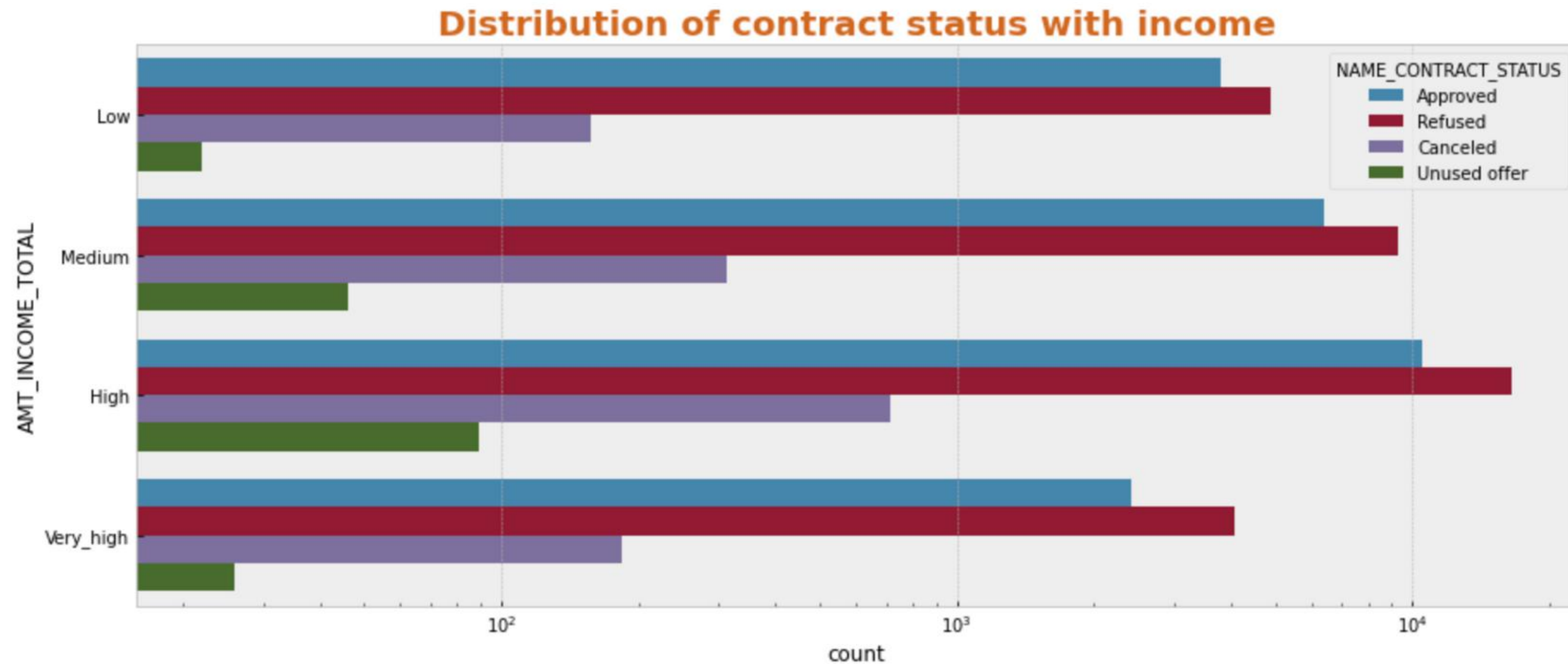
# DISTRIBUTION OF CONTRACT STATUS WITH PURPOSES





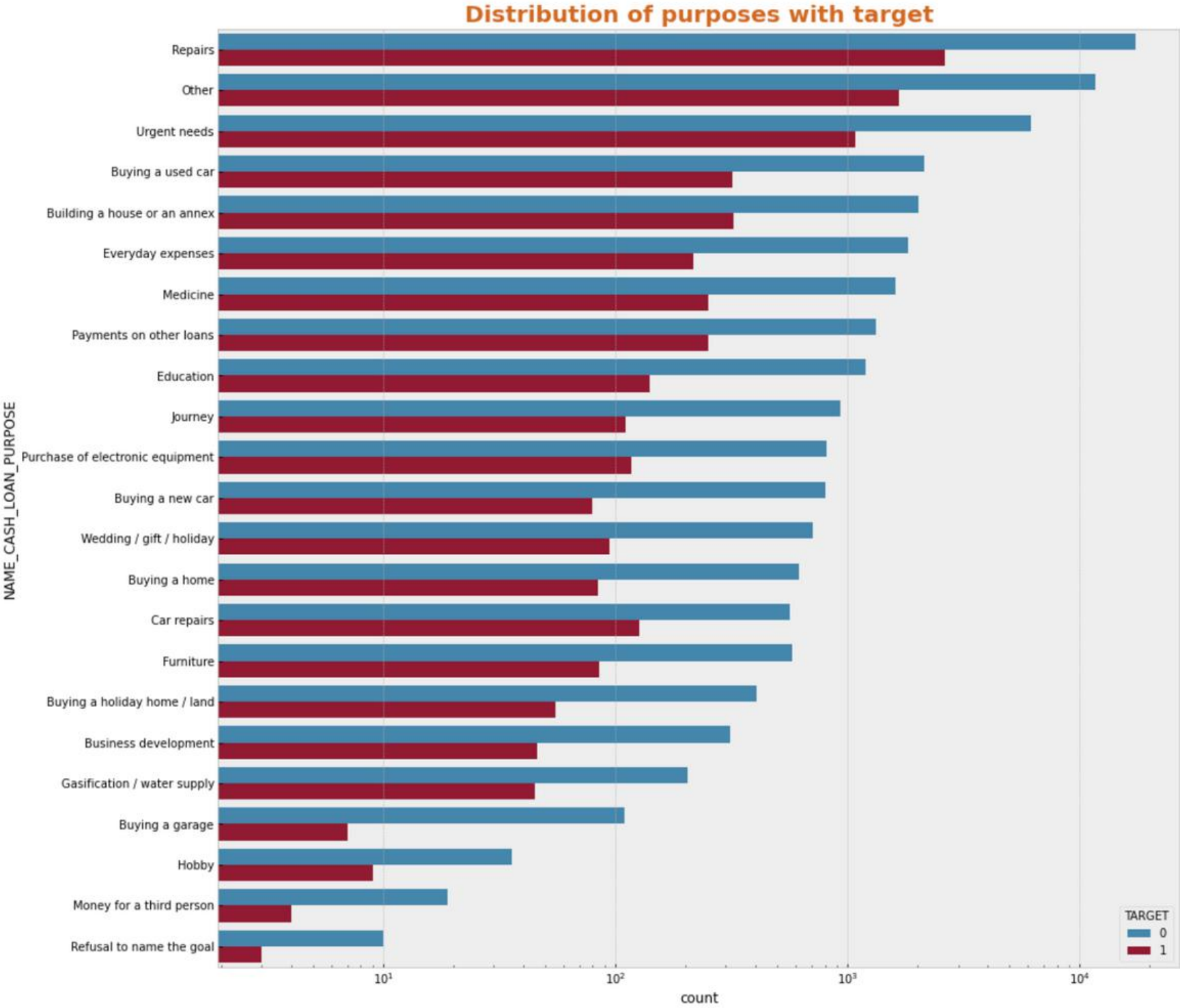
- 
- Most of the refused applications is from Repairs.
  - In purchase of electronic equipment and everyday expenses, approved application status is more than refused one.
  - Approved and Refused status seems to be almost equal in education.

# DISTRIBUTION OF CONTRACT STATUS WITH INCOME

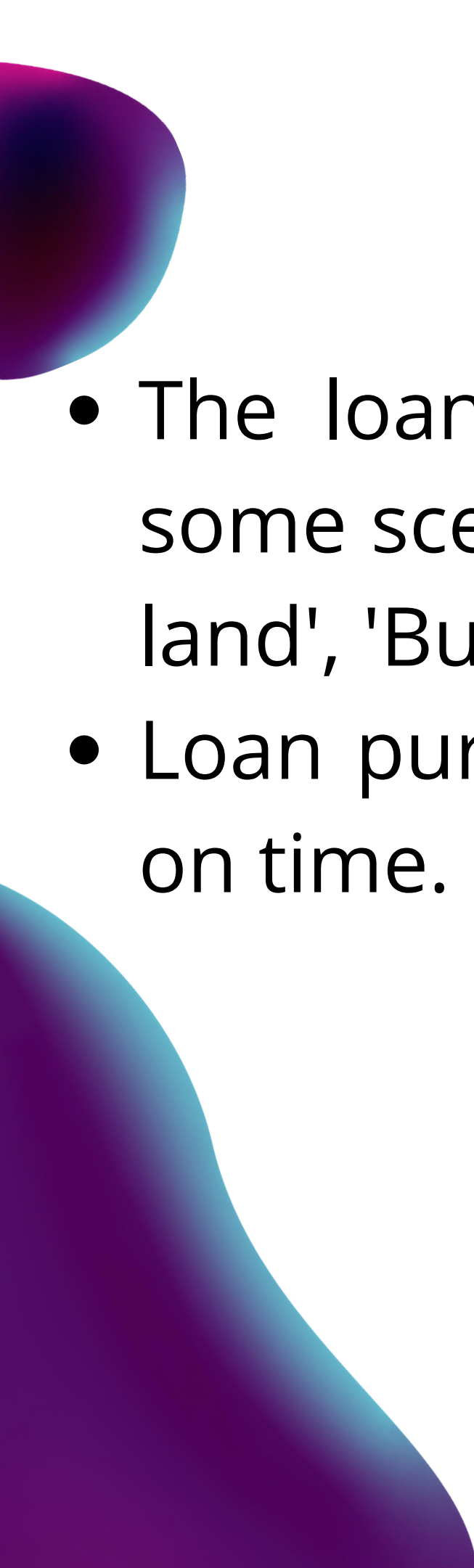


Application status got refused for the people earning high salary .

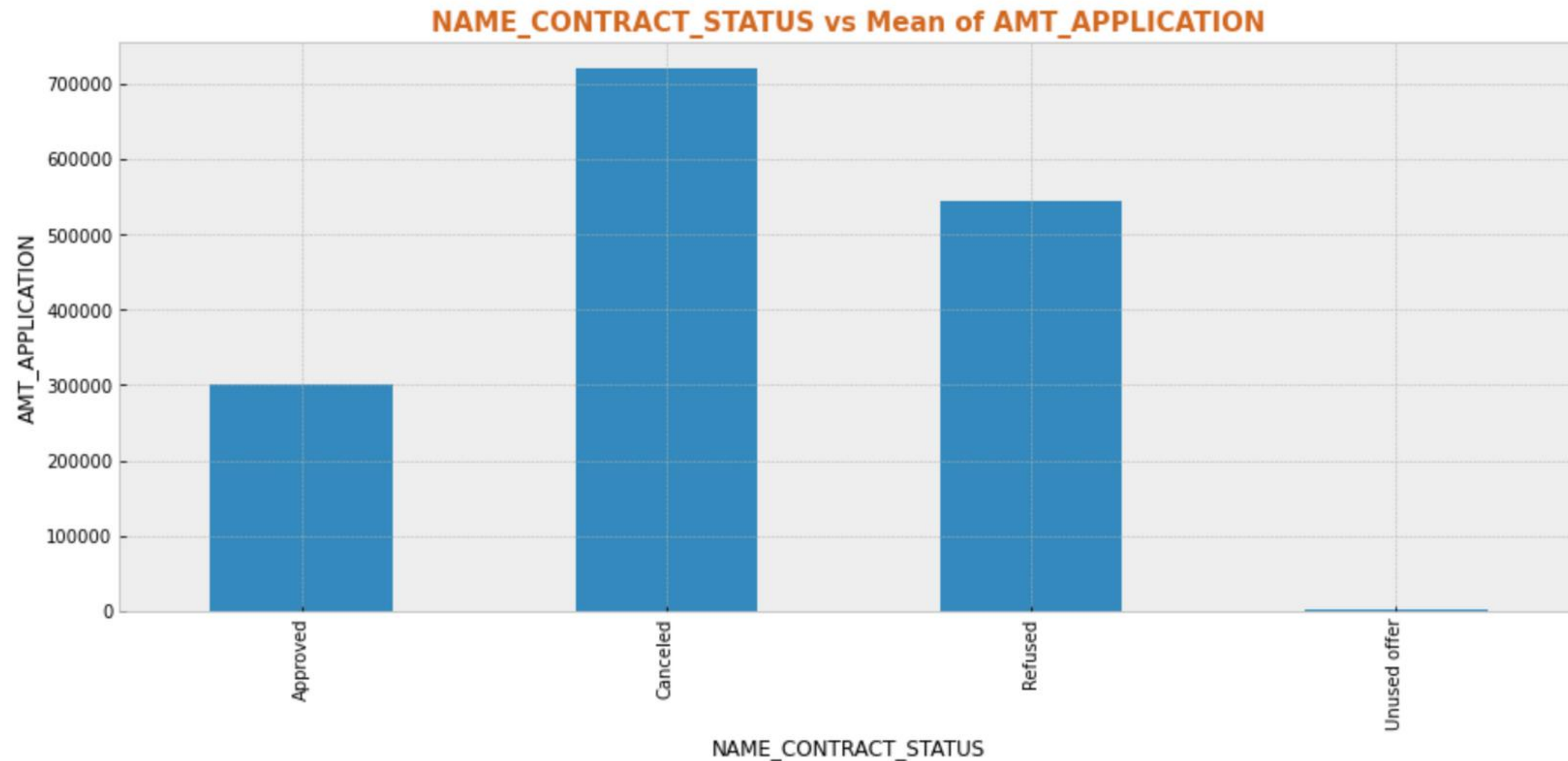
# DISTRIBUTION OF CONTRACT STATUS WITH TARGET





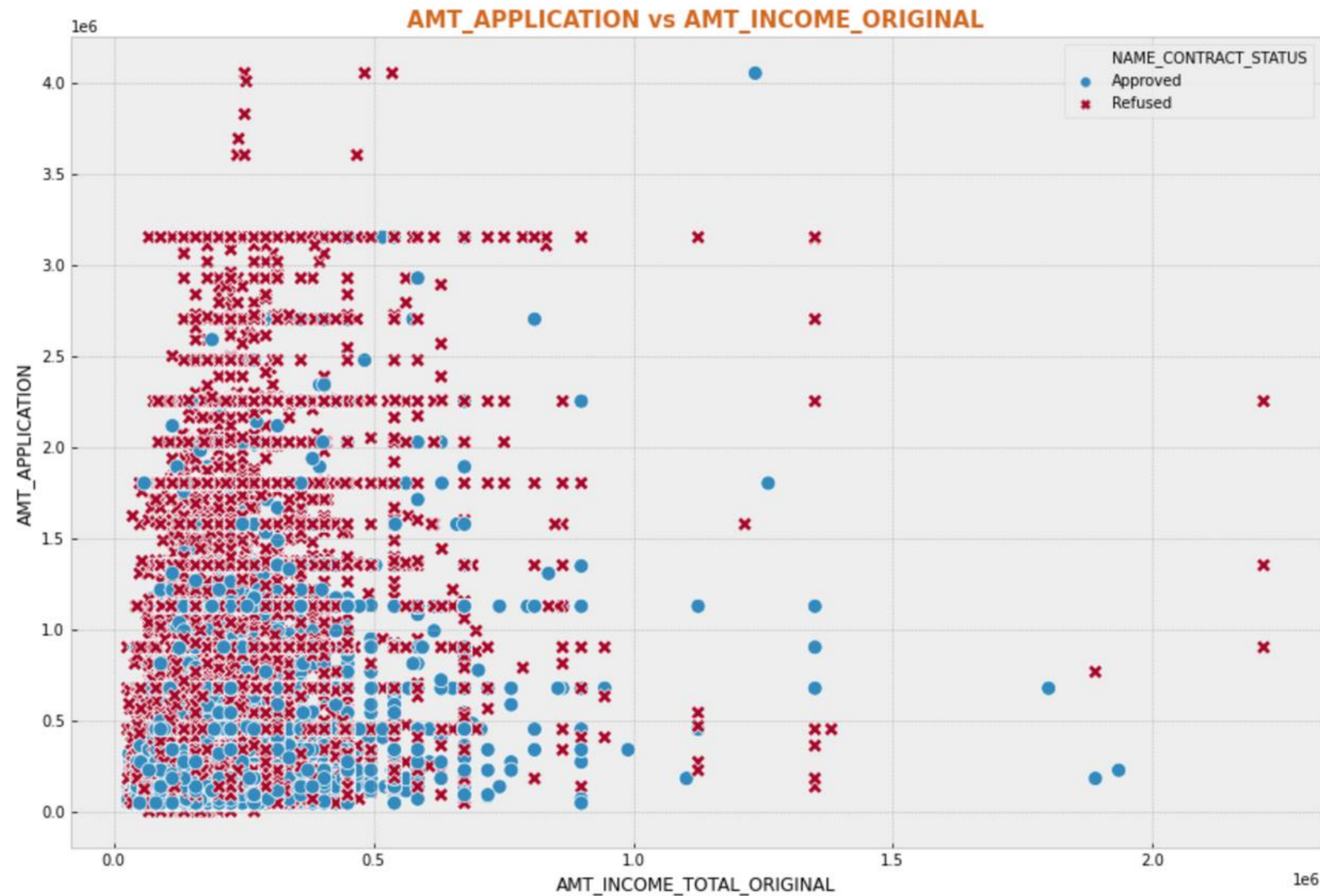
- 
- The loan payment is significantly higher than facing difficulties in some scenarios i.e, 'Buying a garage', 'Business development', 'Buying land', 'Buying a new car', and 'Education'.
  - Loan purposes with 'Repairs' are facing more difficulties in payment on time.

# CONTRACT STATUS Vs MEAN OF AMT\_APPLICATION



In the above plot, the application amount is quite high (300k) in the loan application status that is canceled.

# APPLICATION AMOUNT Vs TOTAL INCOME



- Loan requests higher than 200k had a higher rejection rate.
- The loan rejection rate was much lower if the income was higher than 500k.



# CONCLUSION

- Bank should refuse loan application if loan amount is higher than 300k and income is less than 100k.
- Banks should focus less on income type 'Working' as they are having most number of unsuccessful payments.
- Loan purpose 'Repair' is having higher number of unsuccessful payments on time.
- Bank should refuse loan application of Labors because they found as most defaulter.
- Banks should focus more on contract type 'Student' , 'pensioner' and 'Businessman' with housing 'type other than 'Co-op apartment' for successful payments.

THANK YOU