EDA CREDIT - ASSIGNMENT

By: Abhishek Hota
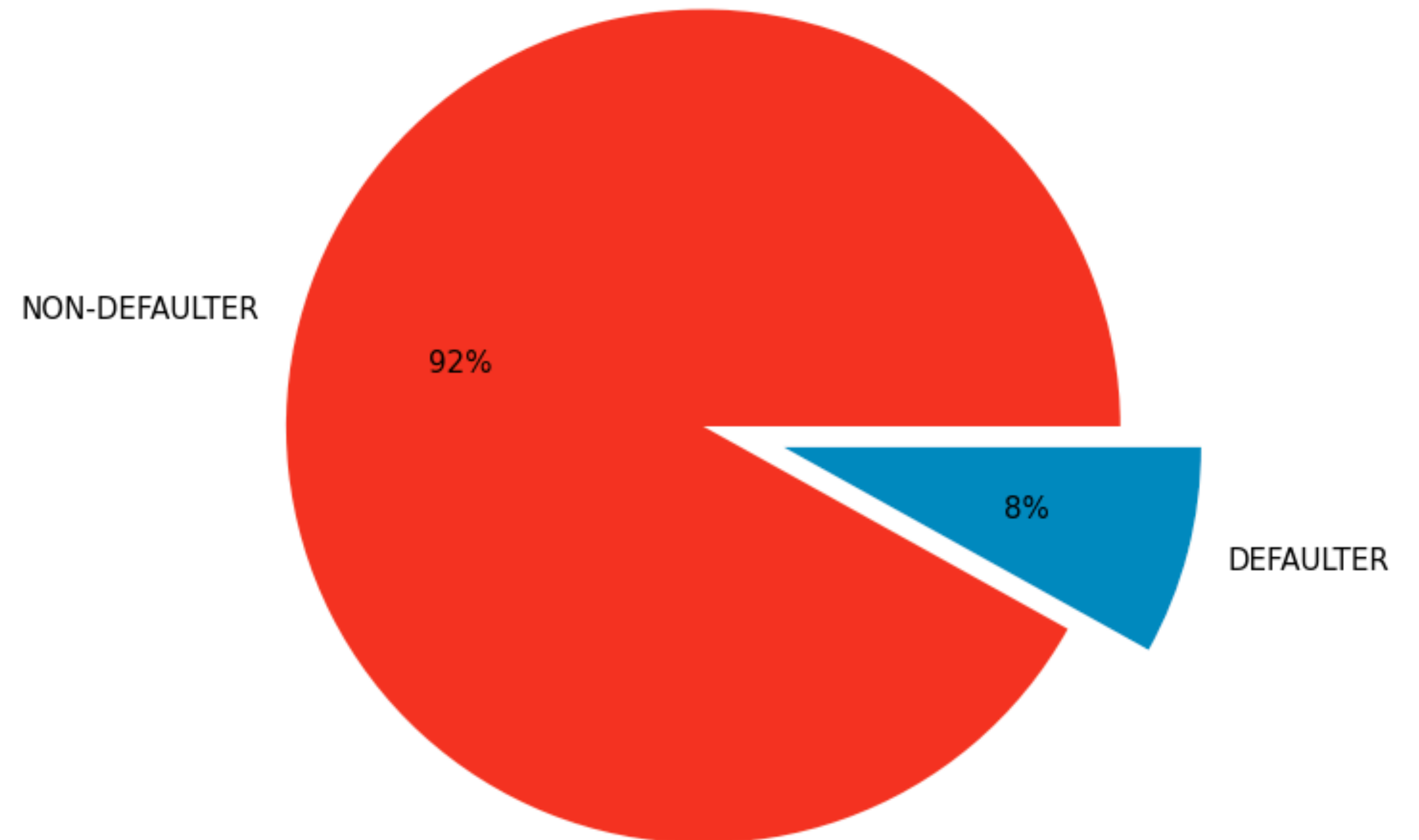
UNIVARIATE ANALYSIS

TARGET Imbalance Check 1

- 8% of the total applicant are Defaulter.

TARGET Variable - DEFAULTER Vs NONDEFAULTER

NON-DEFAULTER

92%
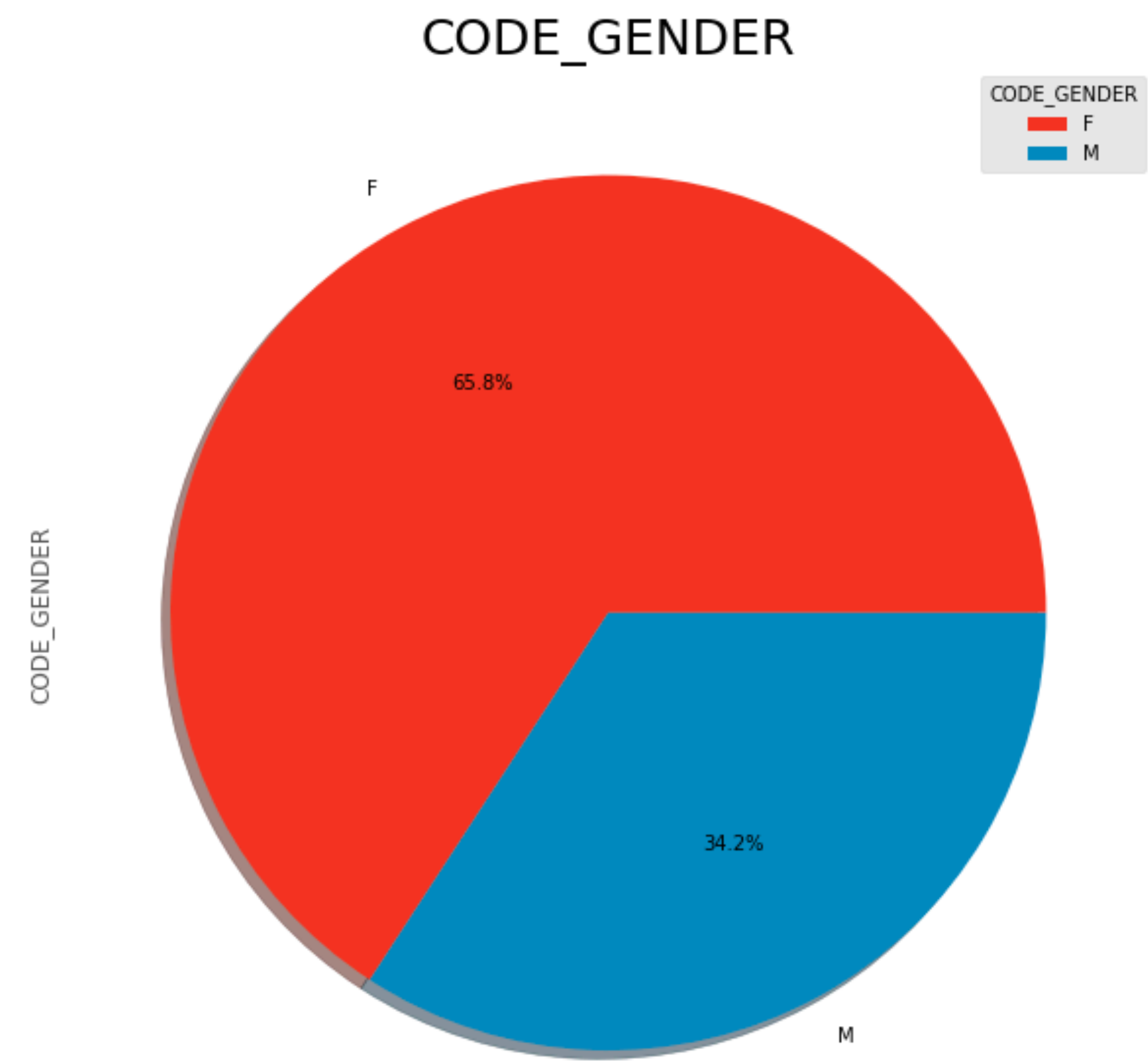
8%

DEFAULTER

TARGET Imbalance Check 2

65.8% from the total Applicants are Female. Whereas **34.2%** of the Applicants are Male.

CODE_GENDER
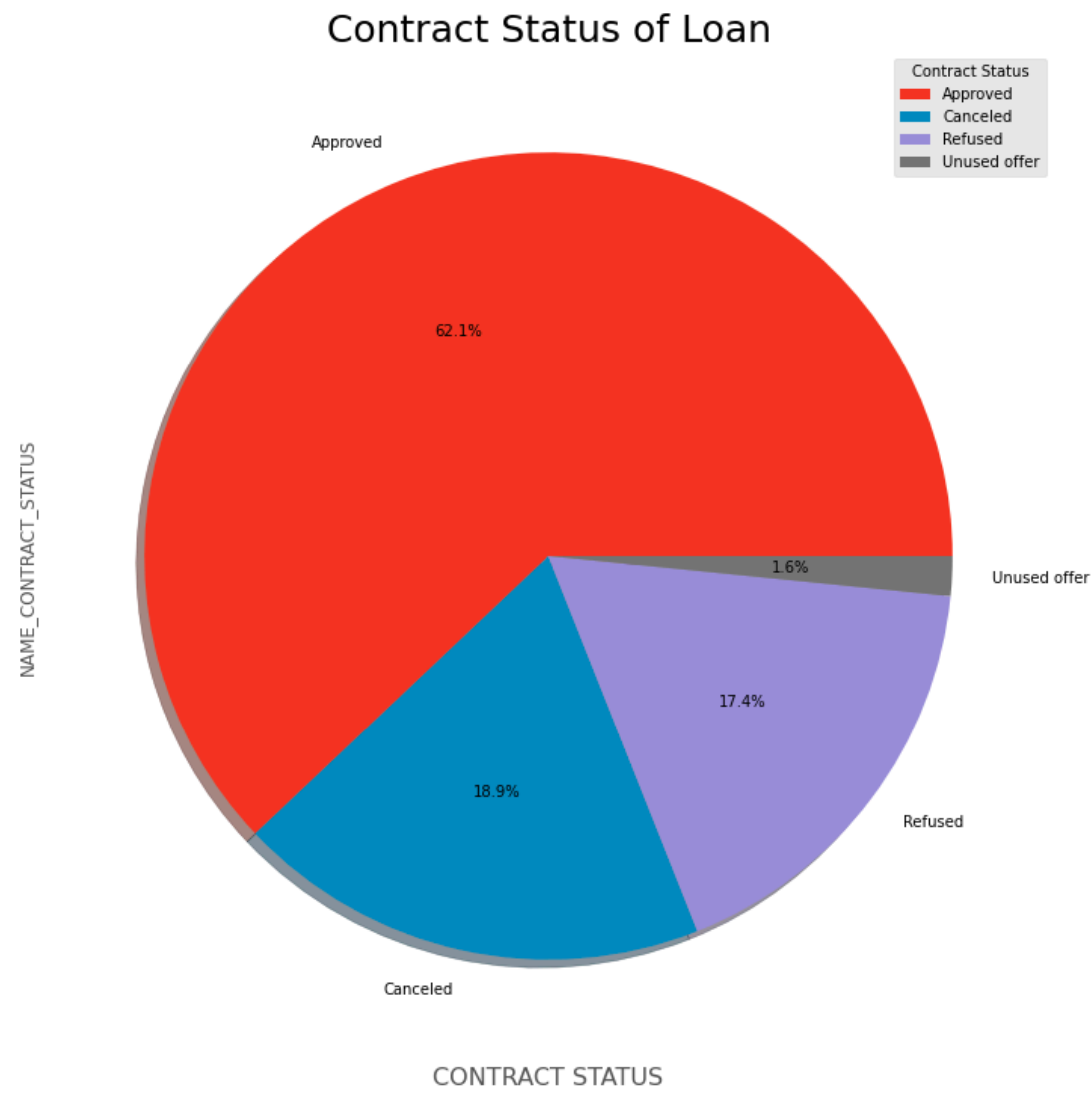
CODE_GENDER
F
M

65.8%

34.2%

F

M

CODE_GENDER

Contract Status of loans from Previous Application data
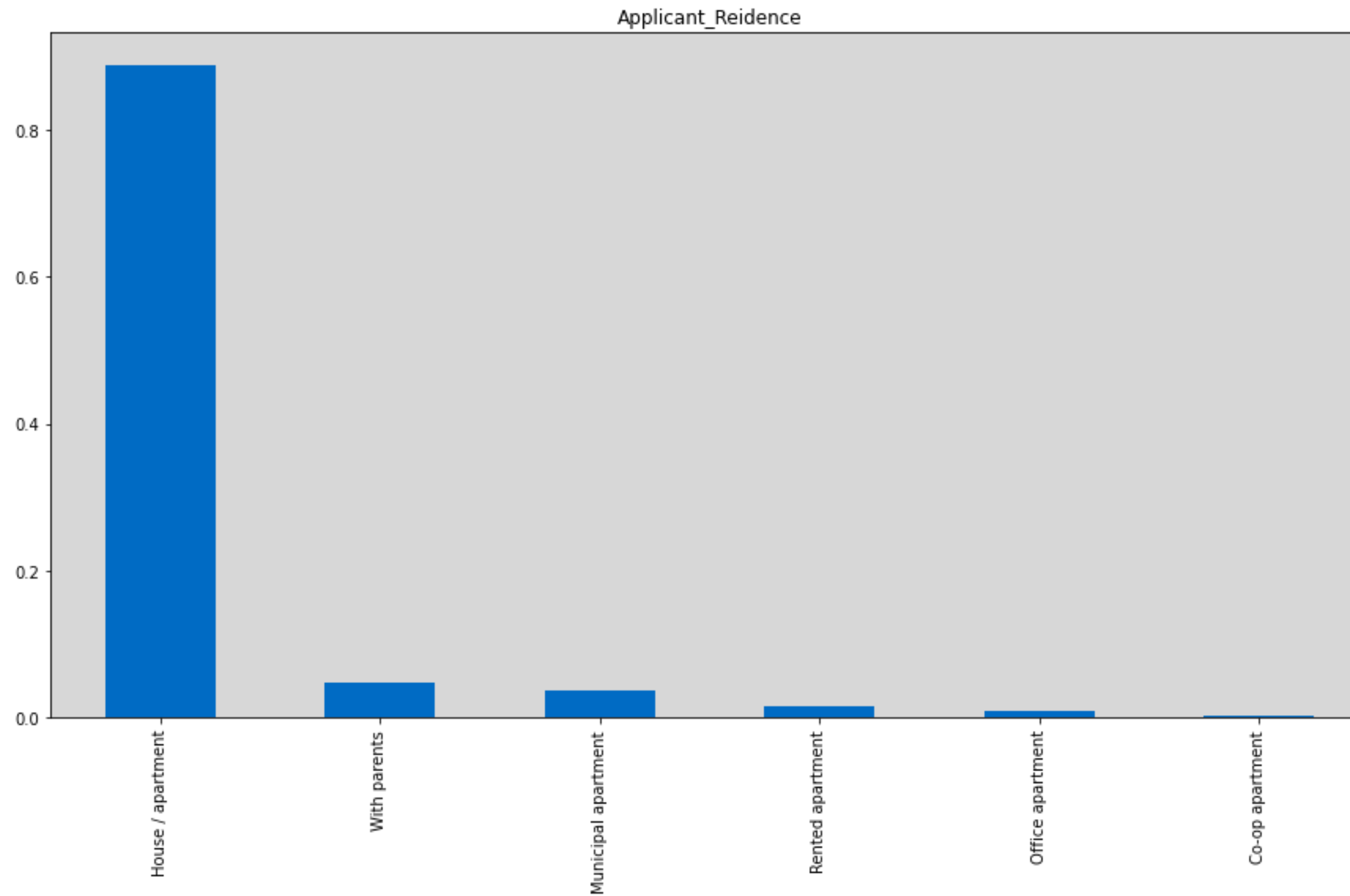
Contract Status of Loan

OBSERVATION:

As we can see the pie plot of previous data 62.1% applications got apporved and 17% application was refused. But almost 19% of the applications got cancelled.
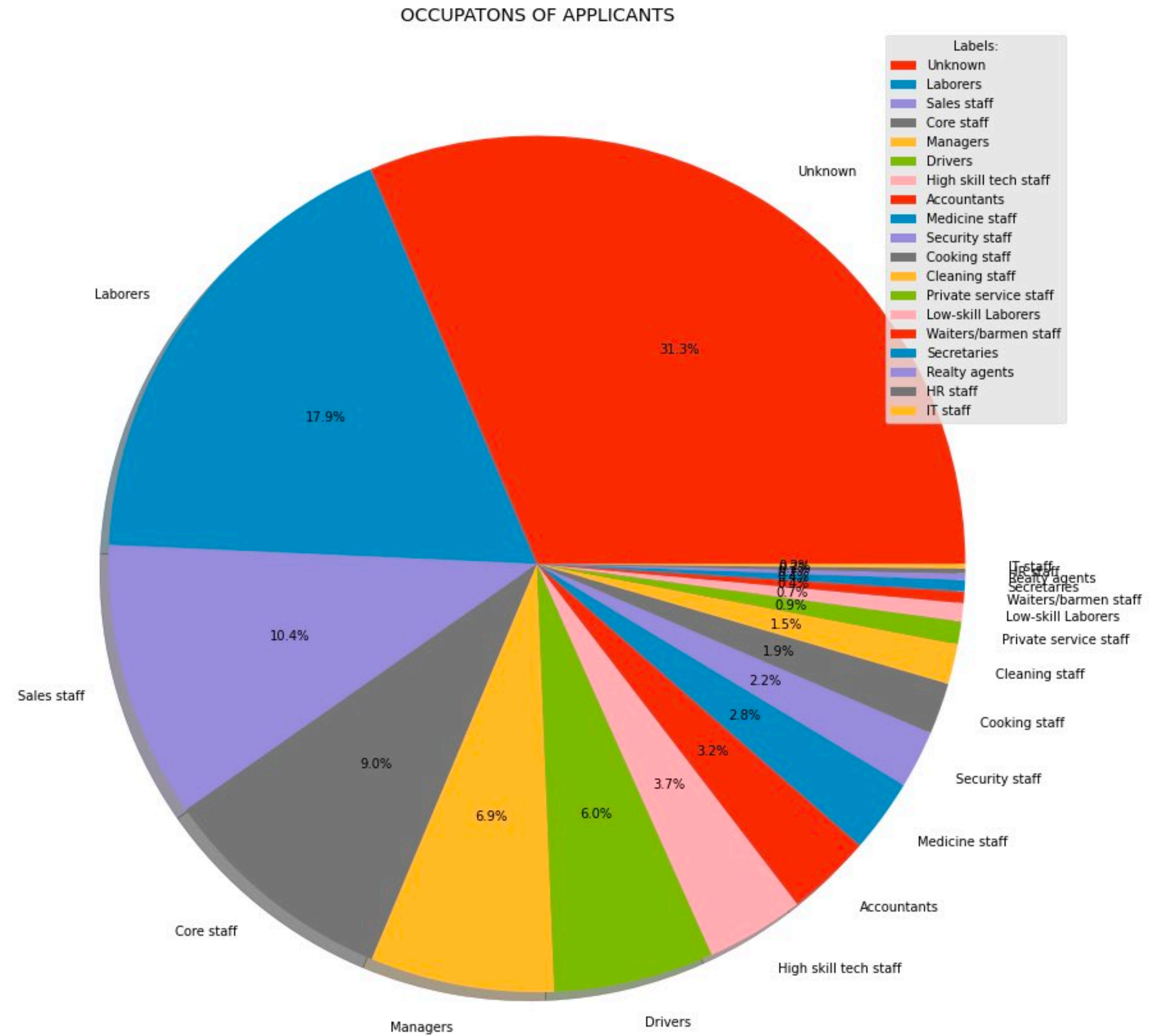
AMT_CREDIT ANALYSIS

Applicant_Reidence

OBSERVATION:

88.73% of the Applicant's are living in House/Apartment

OCCUPATION ANALYSIS

OCCUPATONS OF APPLICANTS

31.3% of the applicant whose Occupation is Unknown.

Labels:
- Unknown
- Laborers
- Sales staff
- Core staff
- Managers
- Drivers
- High skill tech staff
- Accountants
- Medicine staff
- Security staff
- Cooking staff
- Cleaning staff
- Private service staff
- Low-skill Laborers
- Waiters/barmen staff
- Secretaries
- Realty agents
- HR staff
- IT staff

EDUCATION ANALYSIS

OBSERVATION:

71% of the applicant's education type is Secondary or secondary special.
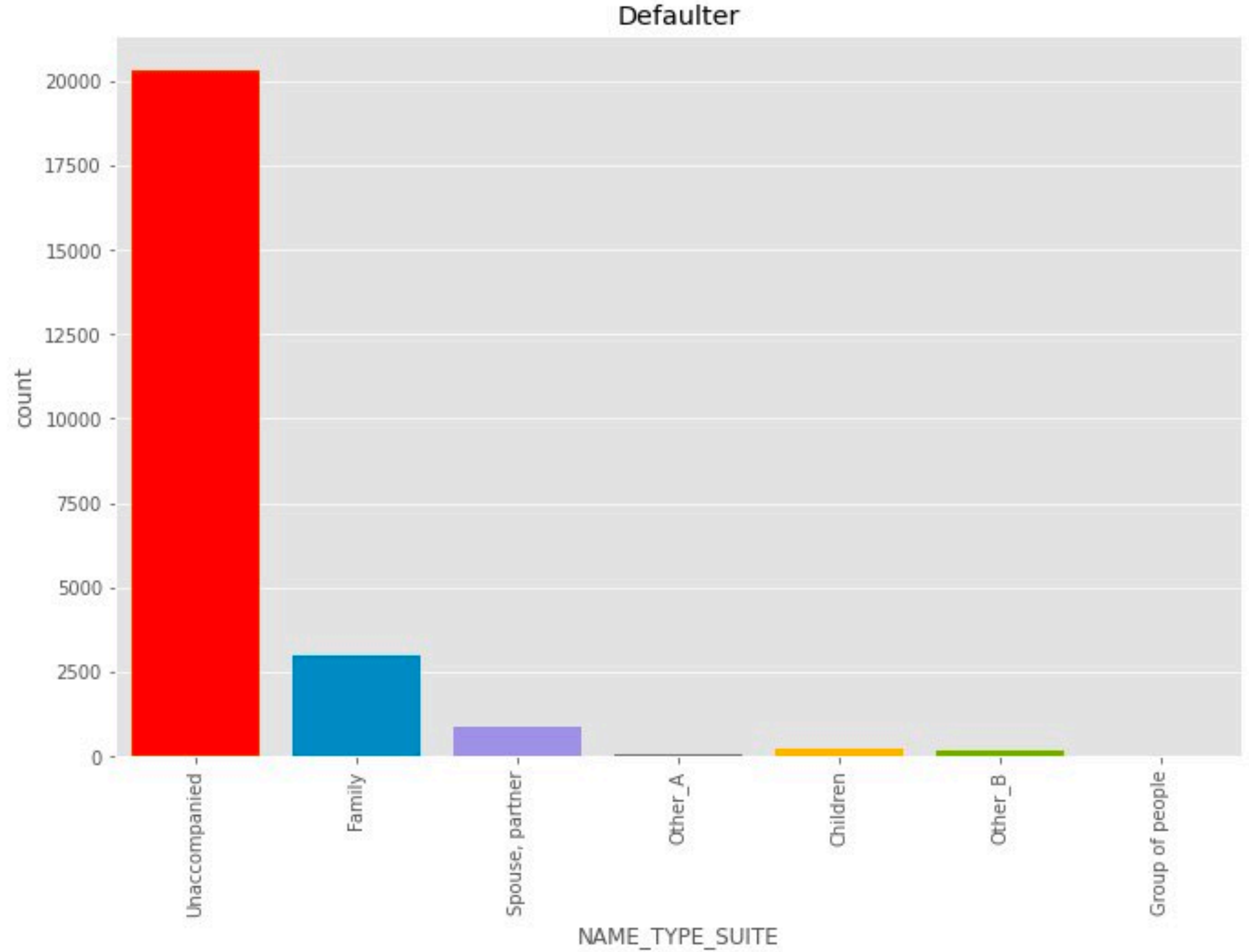

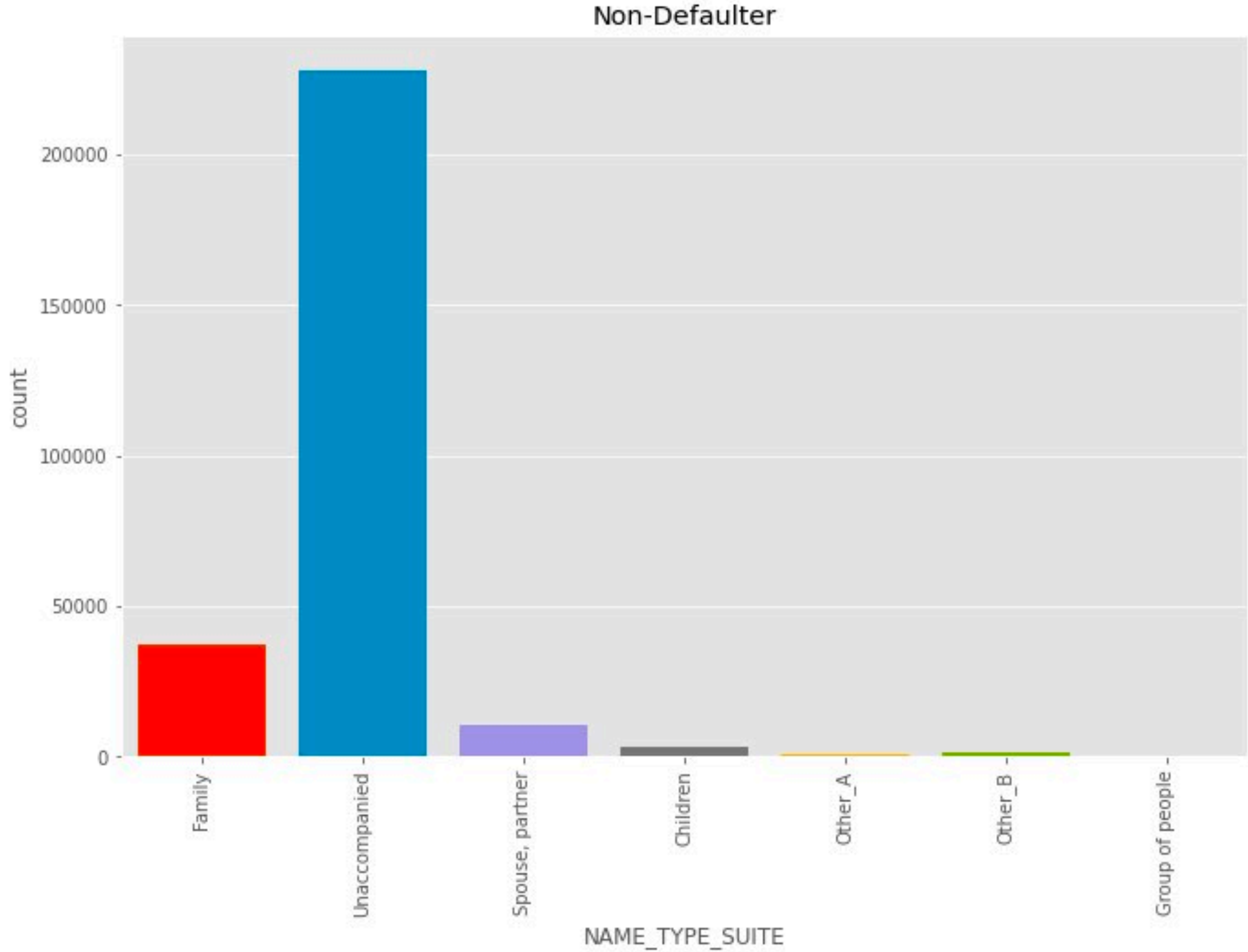TARGET Variable - DEFAULTER Vs NONDEFAULTER
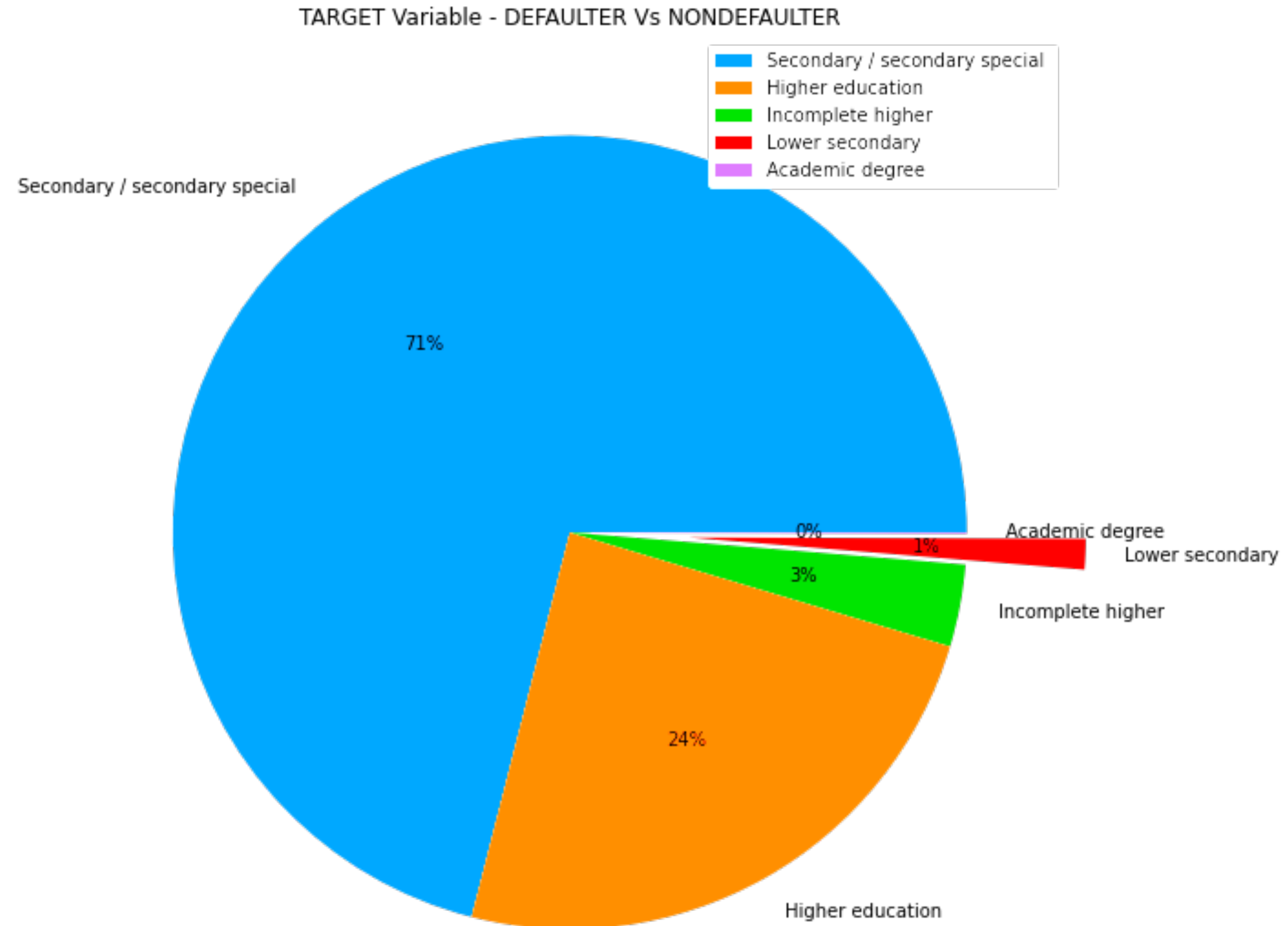
BIVARIATE ANALYSIS

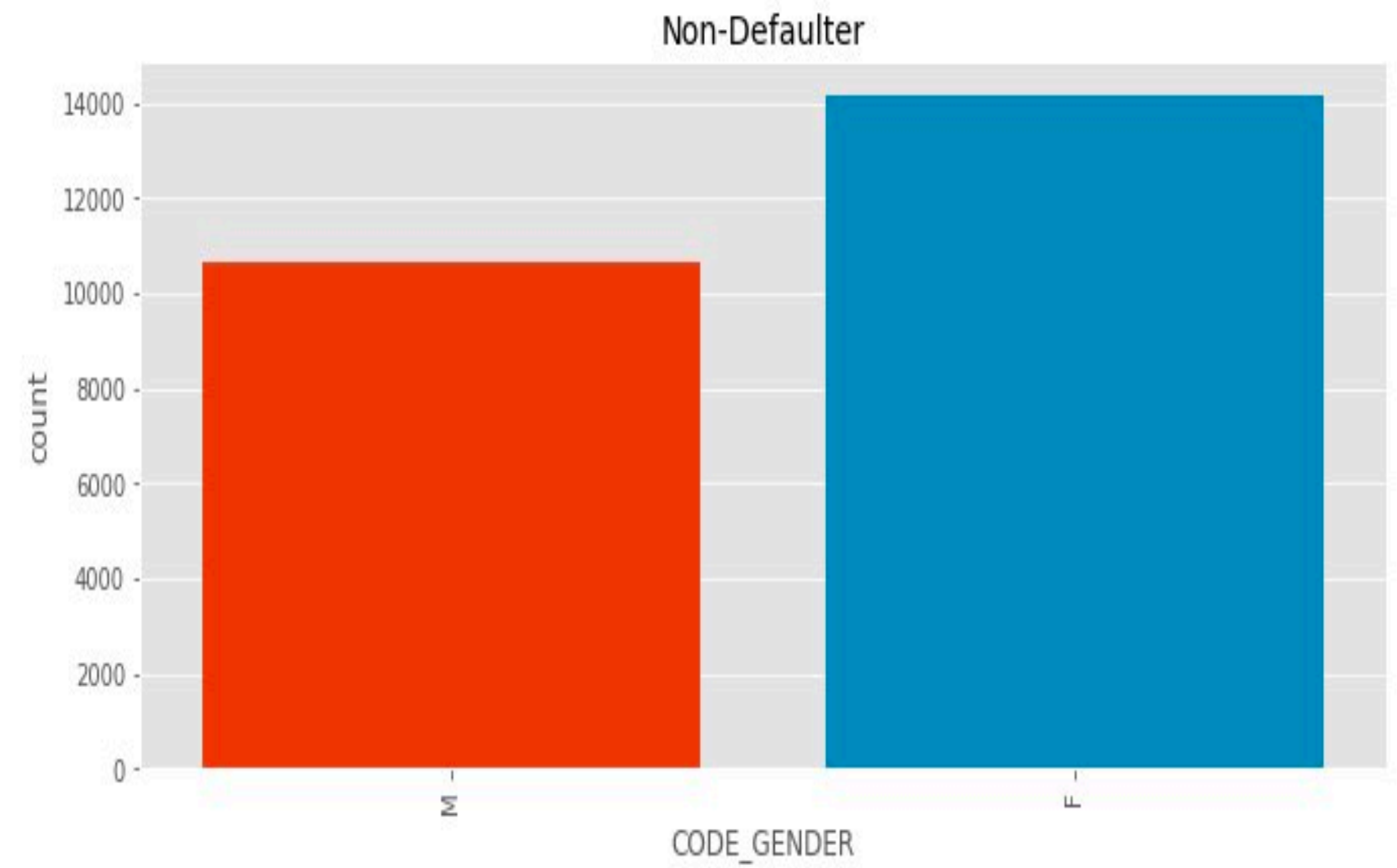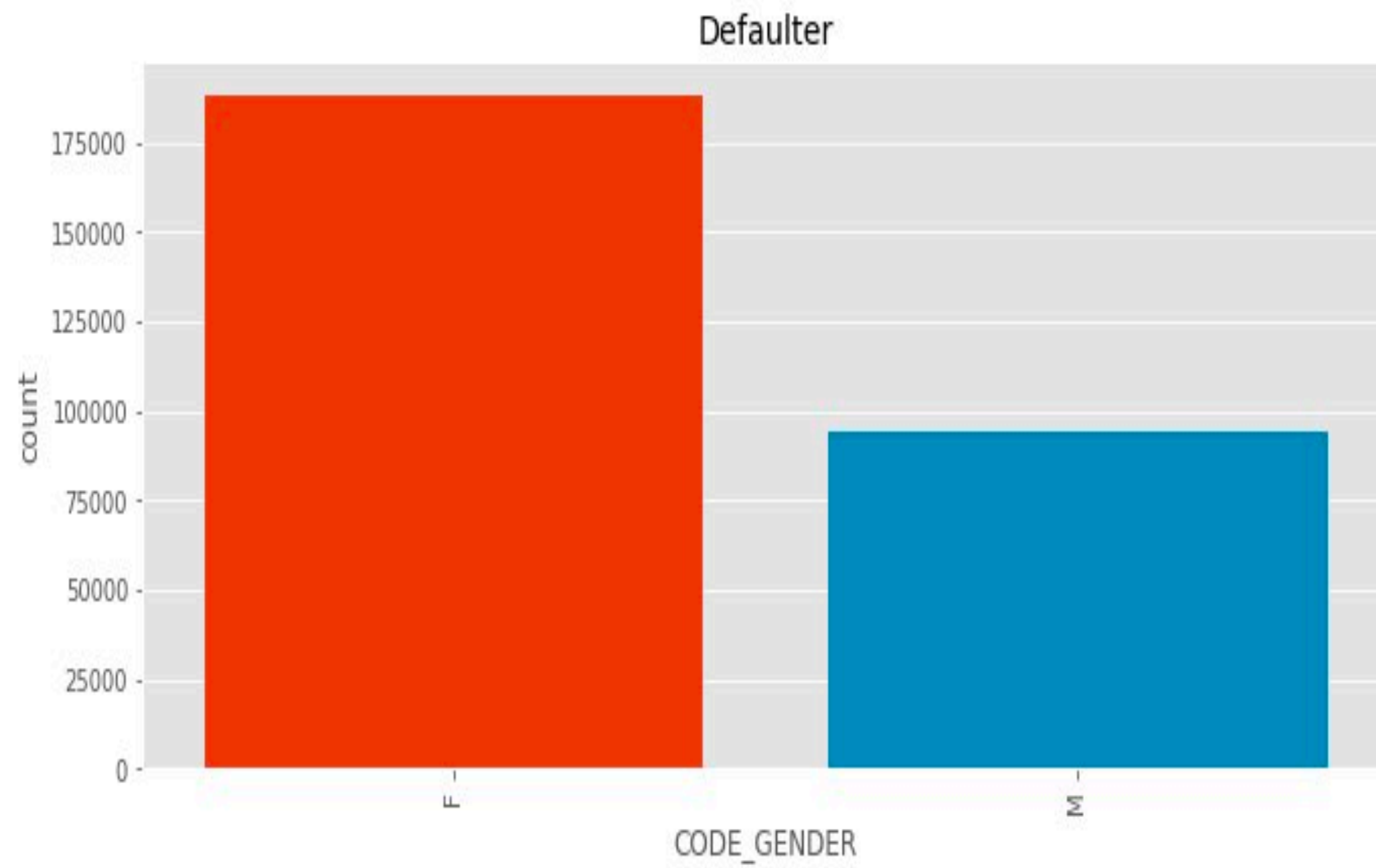# NAME_TYPE_SUITE Analysis w.r.t. Defaulter and Non-Defaulter data.

OBSERVATION:

We can see most of the applicant were unaccopanied while appling for a loan. So, we can also say that Unaccompanied can not be a suitable criteria to define defaulter's data
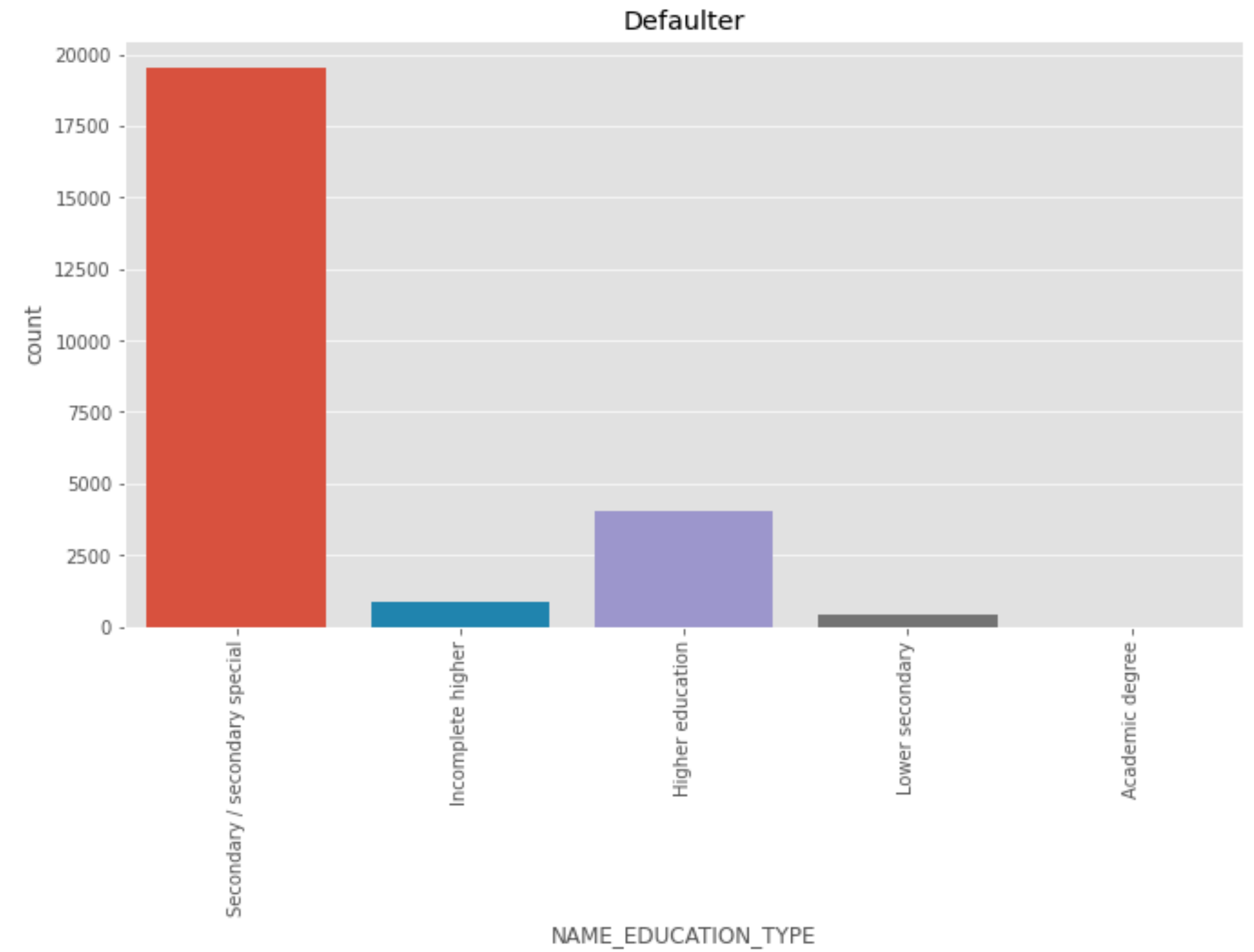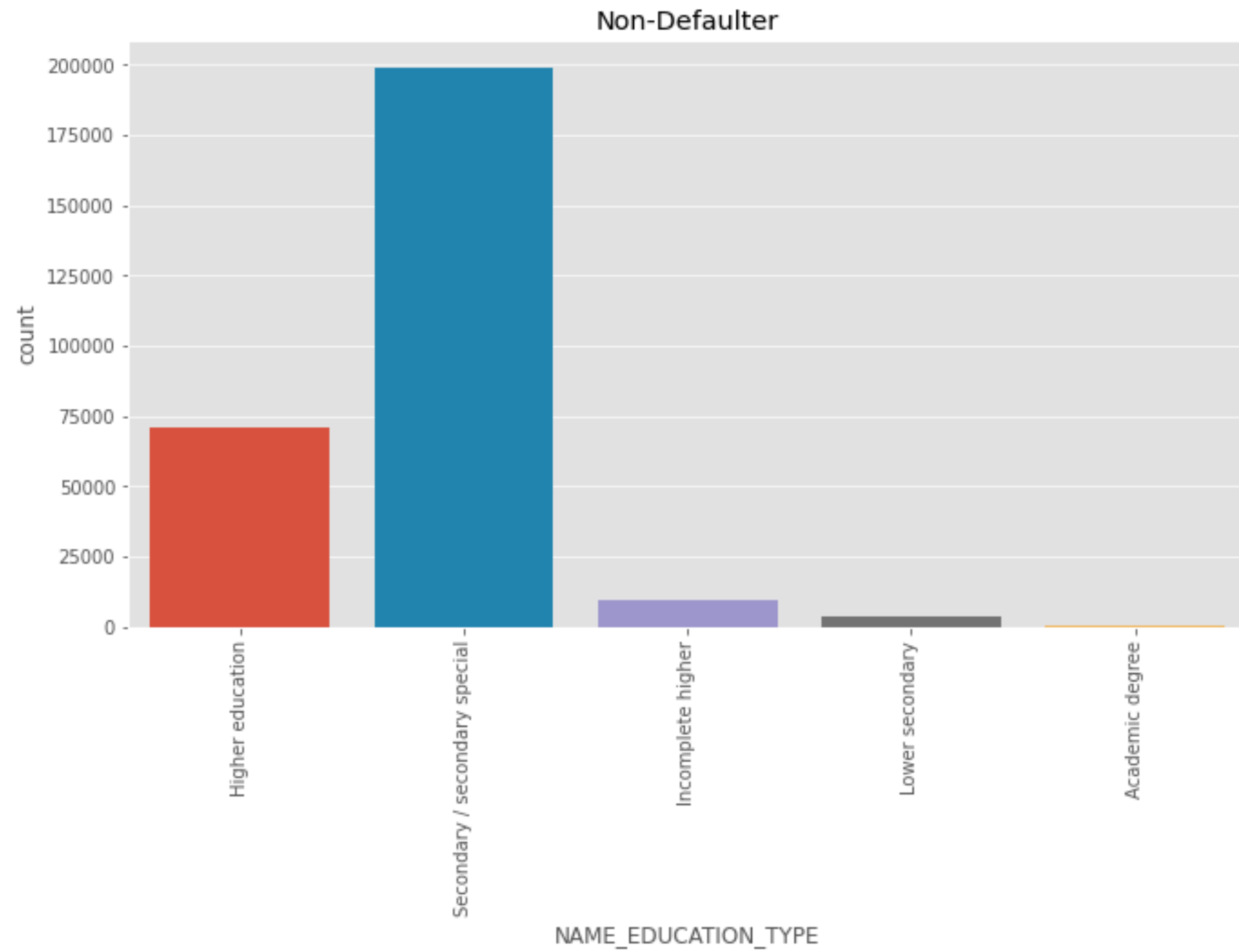
CODE_GENDER Analysis w.r.t. Defaulter and Non-Defaulter data.

- Females are applying for loans more than males and hence the more number of female defaulters.
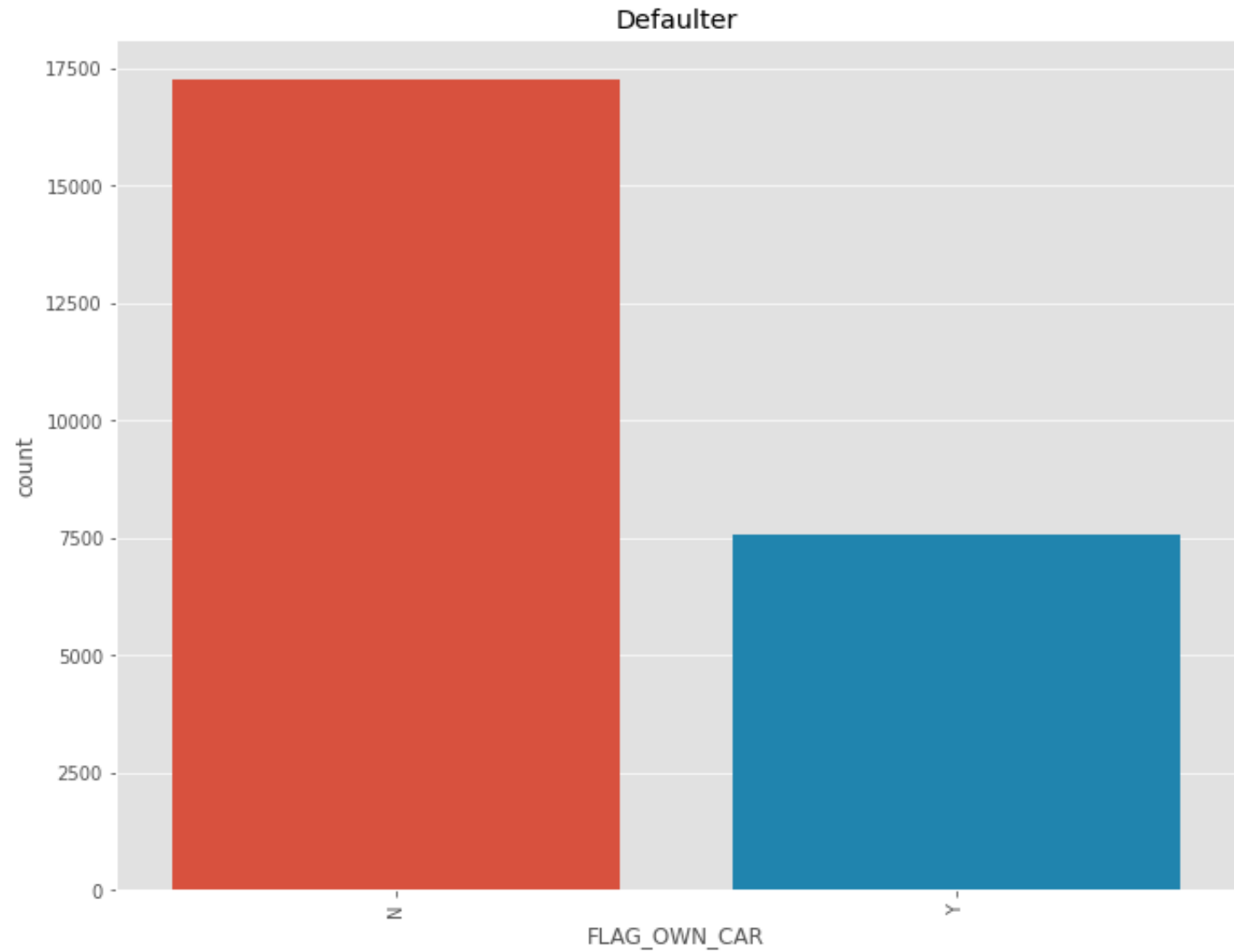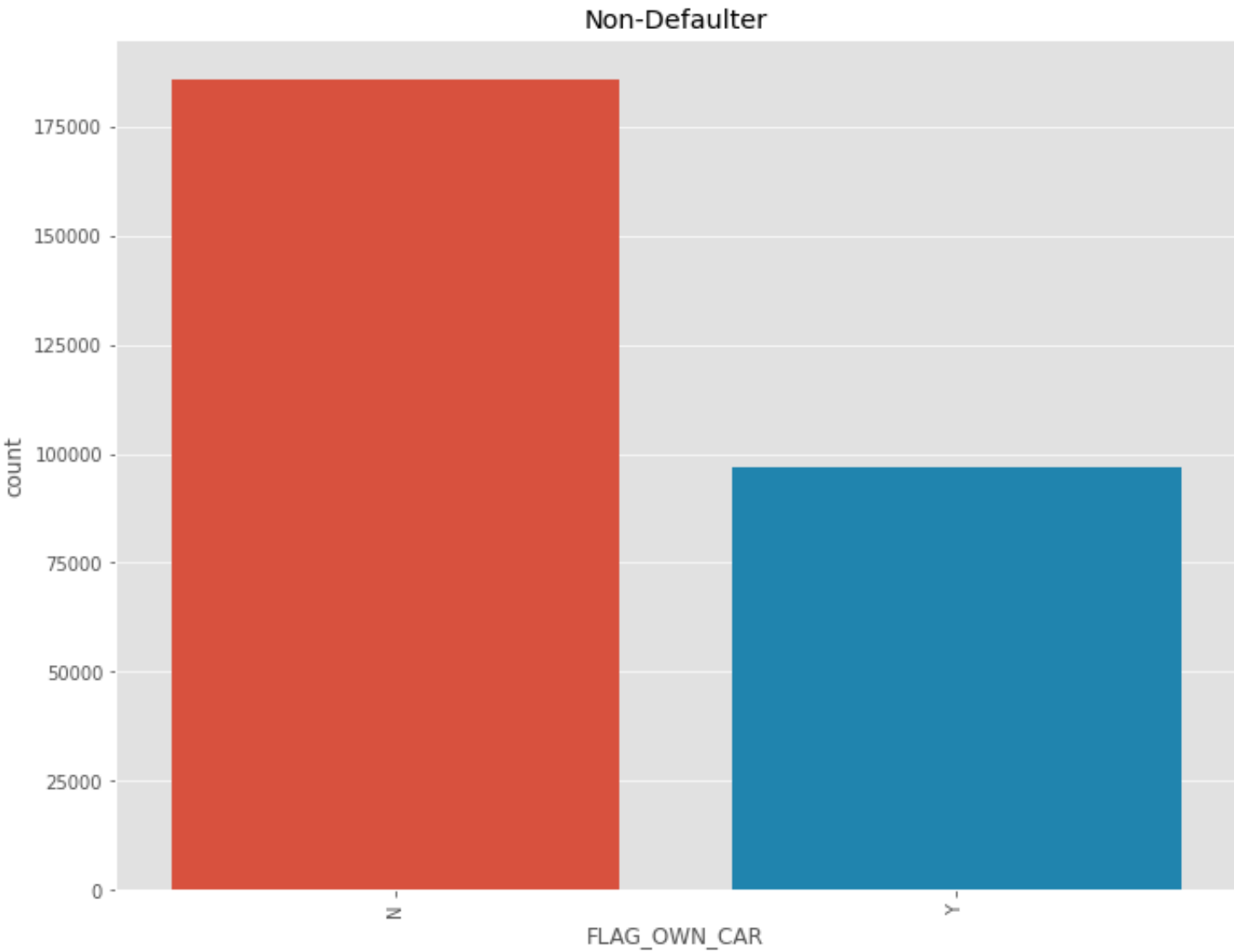
NAME_EDUCATION_TYPE Analysis w.r.t. Defaulter and Non-Defaulter data.
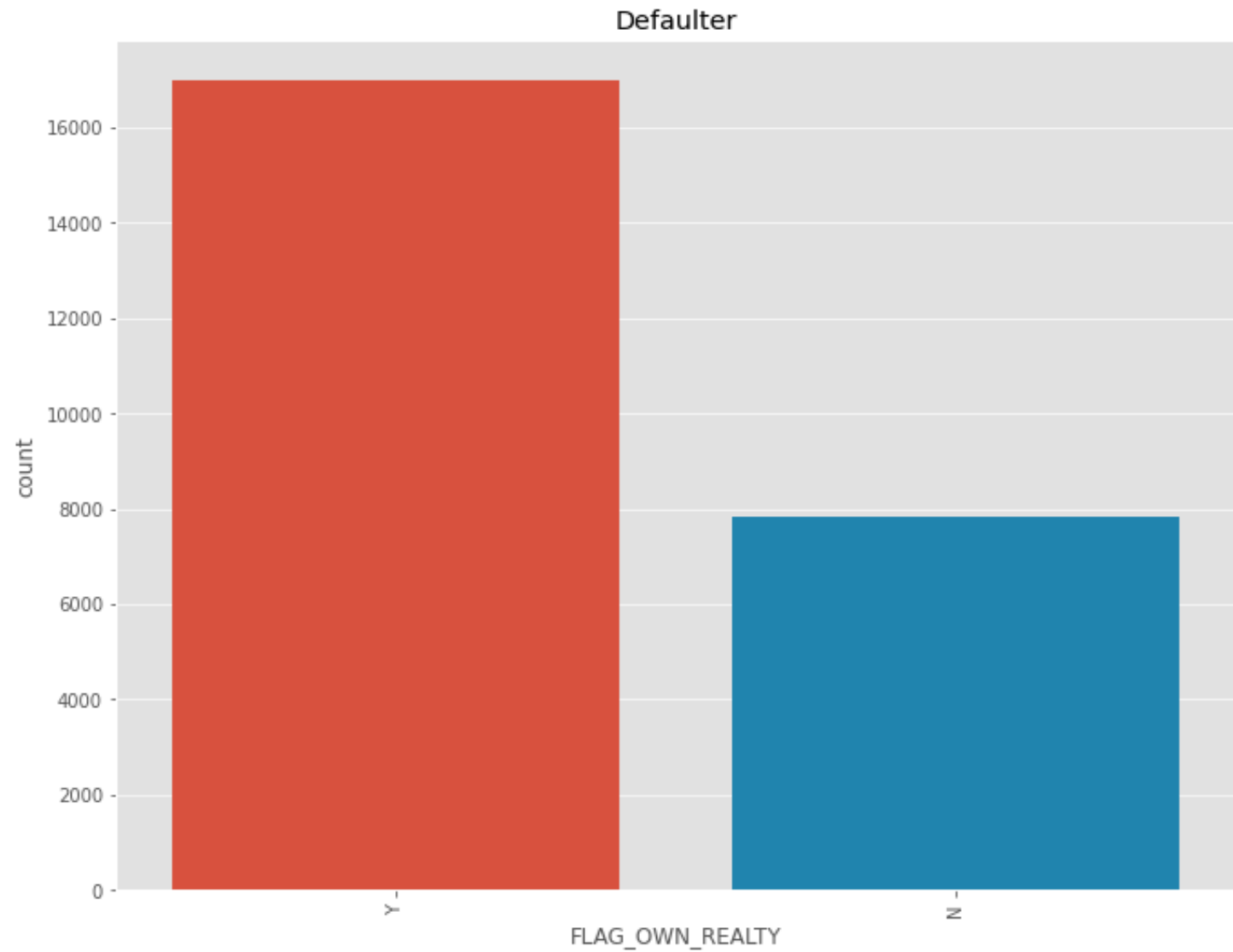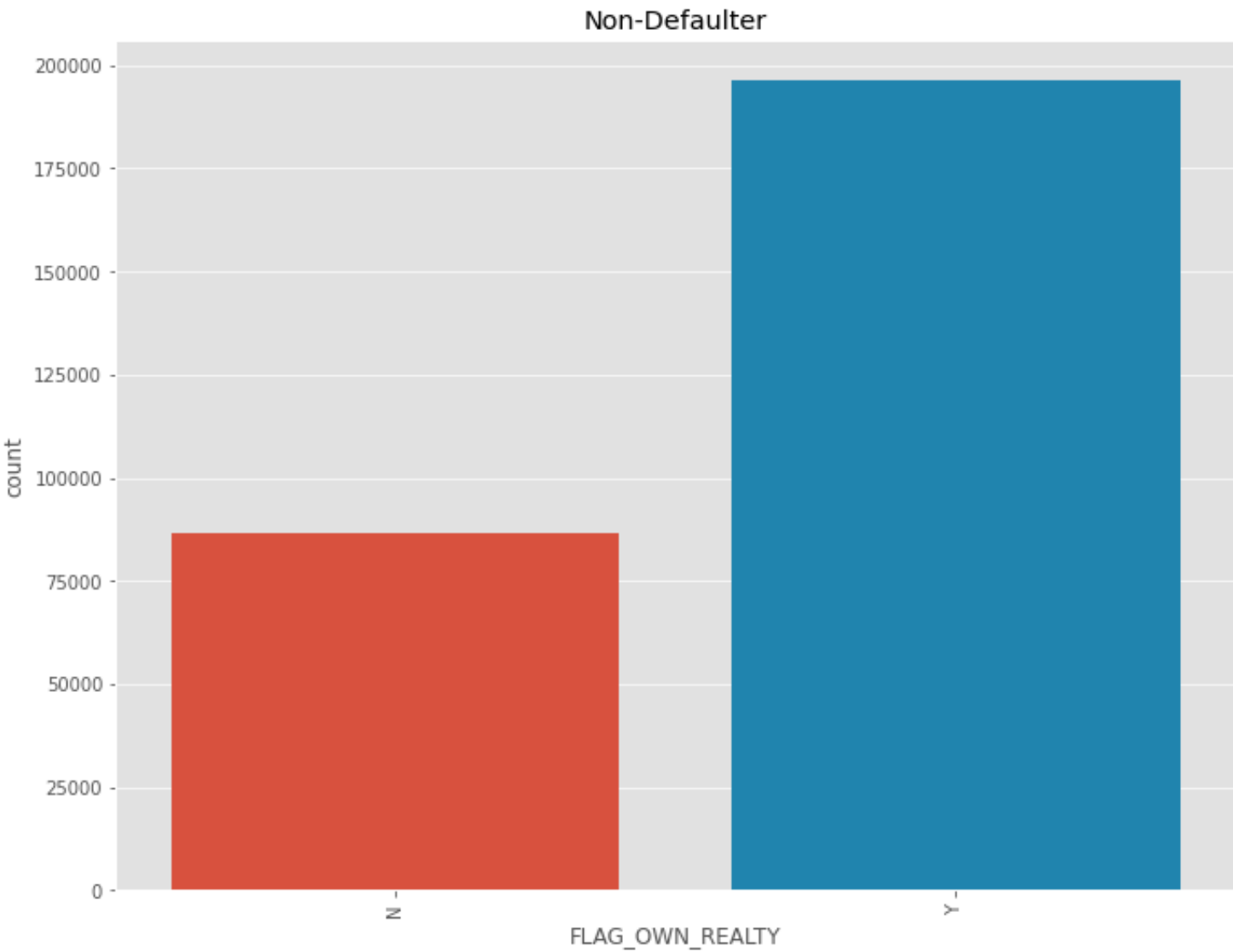

OBSERVATION:

- **71%** of the applicant's education type is Secondary or secondary special, in case of the same Education type Defaulters rate will be high

APPLICANT WITH CAR & REALESTATE w.r.t. Defaulter and Non-Defaulter data.
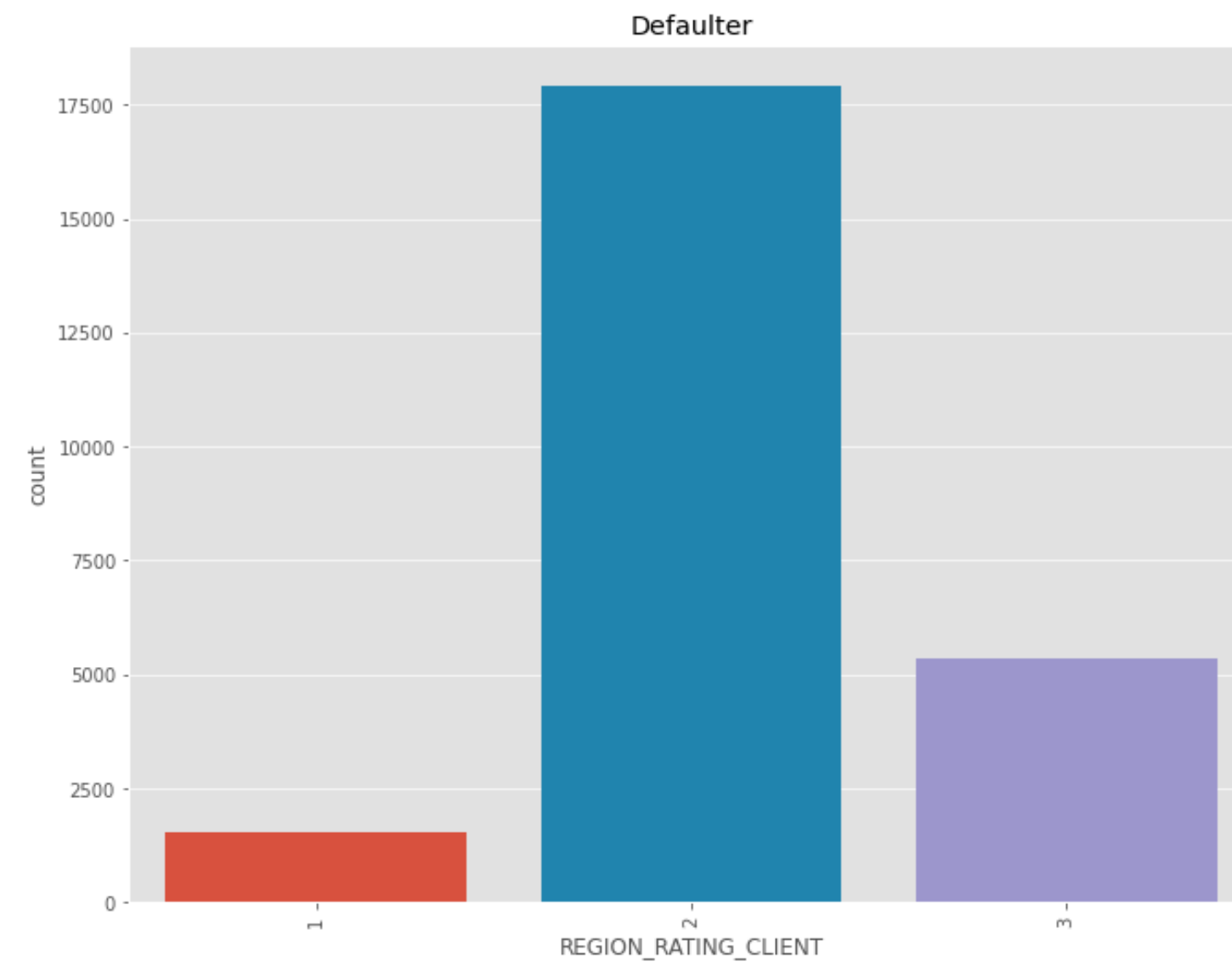

OBSERVATION:

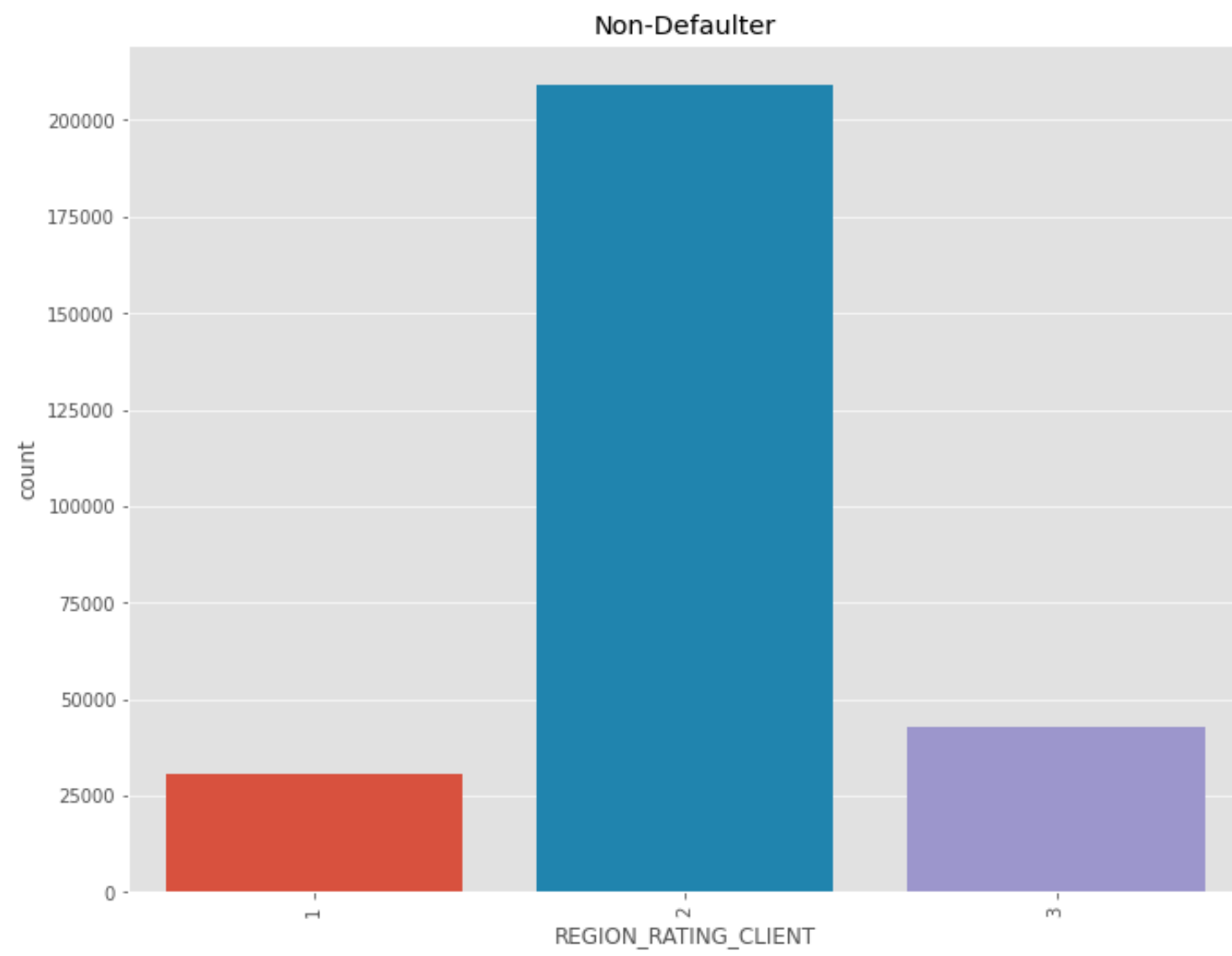- 8% of the total applicant are Defaulter.

OBSERVATION:

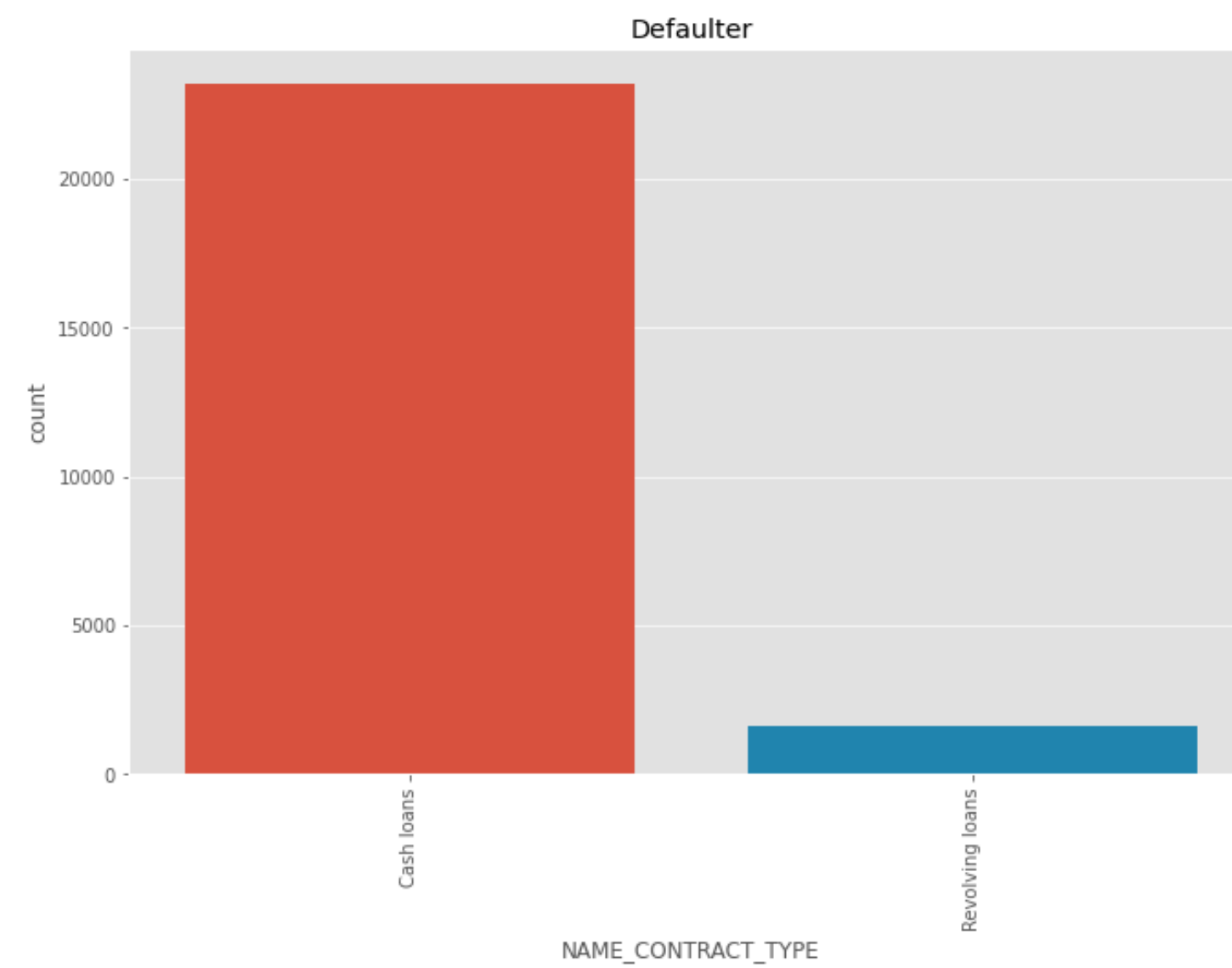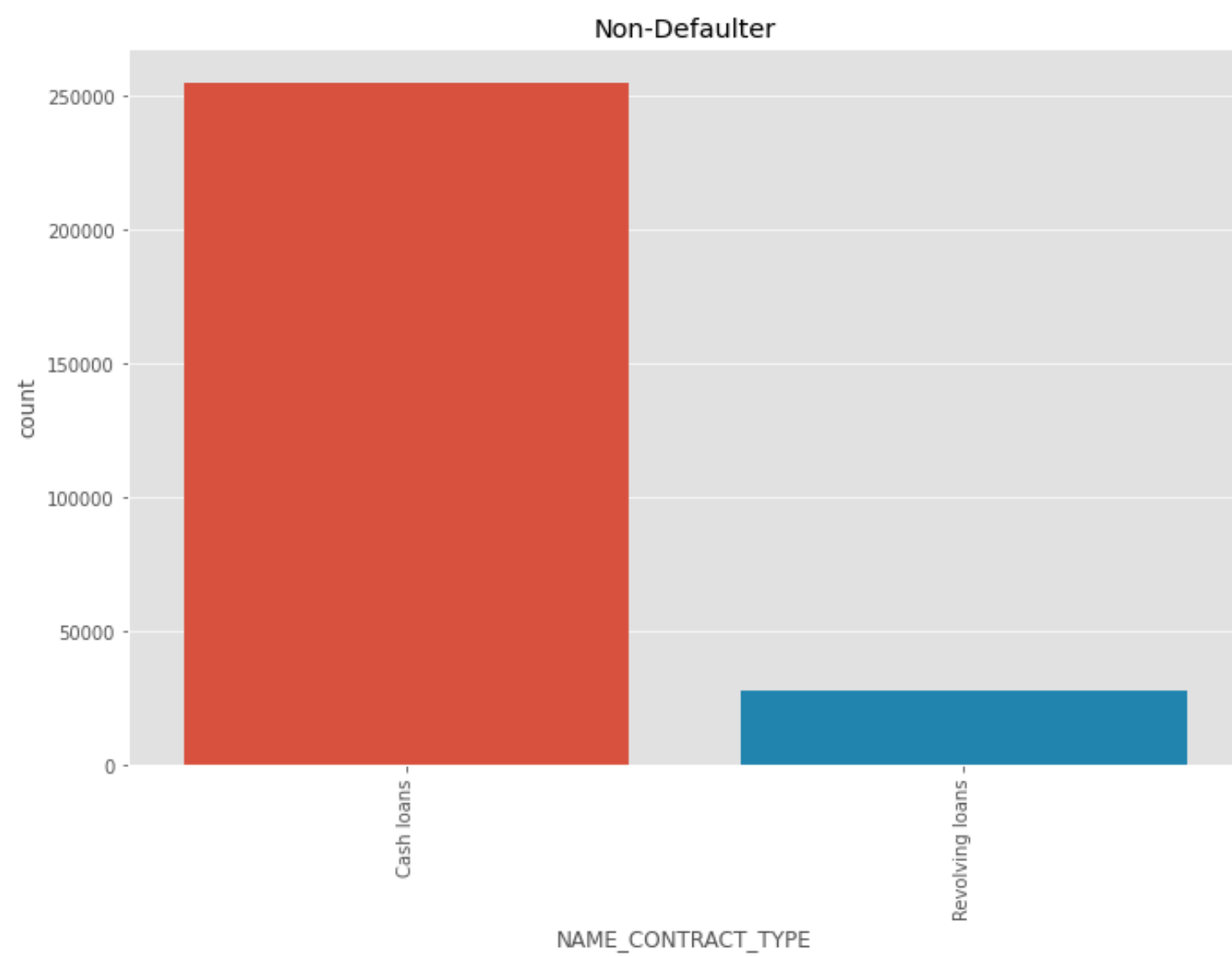- 8% of the total applicant are Defaulter.

REGION & LOAN TYPE Analysis w.r.t. Defaulter and Non-Defaulter data.

OBSERVATION:

- applicants from Type-3 city tend to default more than others as per their weightage.

OBSERVATION:

- Rate of defaults also will be higher than the Revolving loans in Cash loan as issuing rate is higher.

INCOME TYPE Analysis w.r.t. Defaulter/Non-Defaulter & GENDER data.

OBSERVATION:

- Working Class, Commercial associate are having the most payment difficulties.(column 2 in fig_1)

- Maternity Leave is very less and it also has maximum % of payment difficulties. Hence, client with income type as 'Maternity leave' are the driving factors for Loan Defaulters.
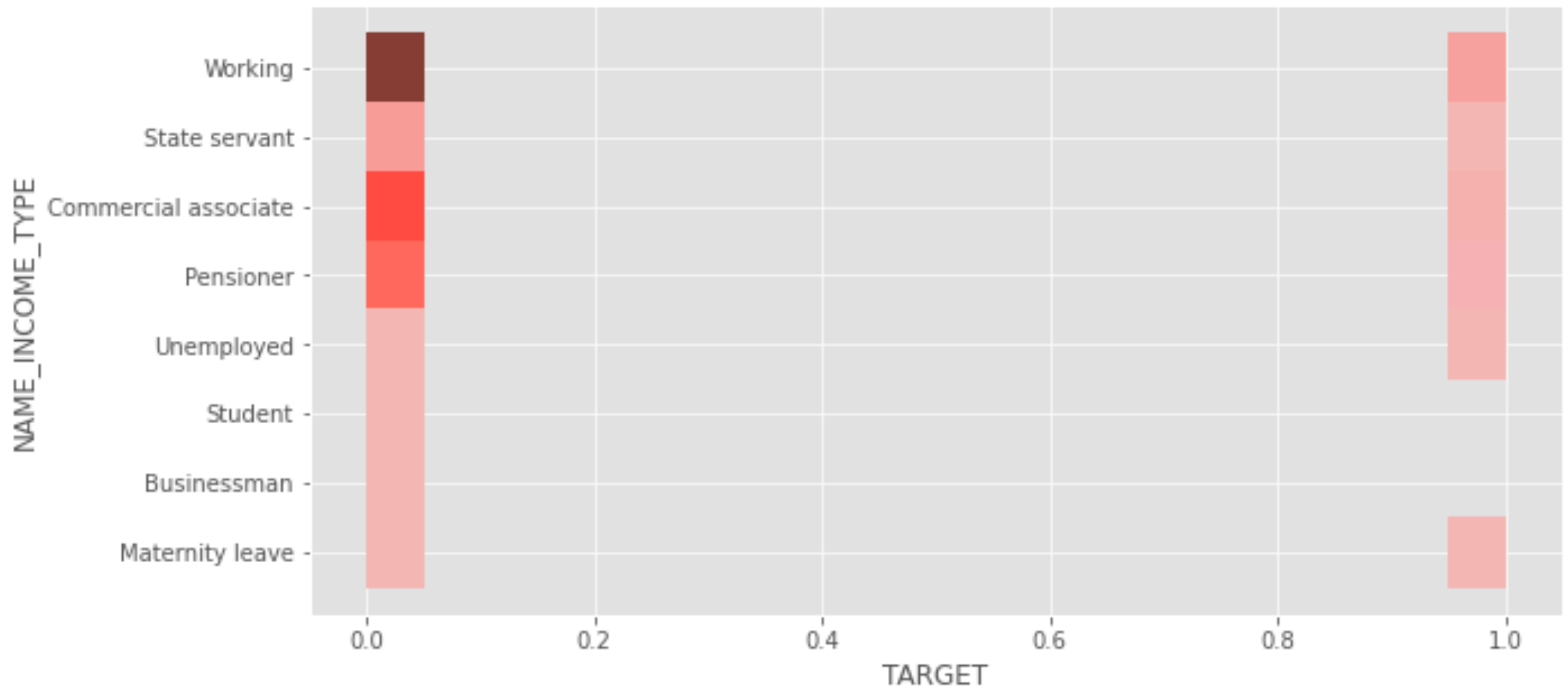
fig_1

fig_2

MULTIVARIATE ANALYSIS

Income Type Vs. Contact Status
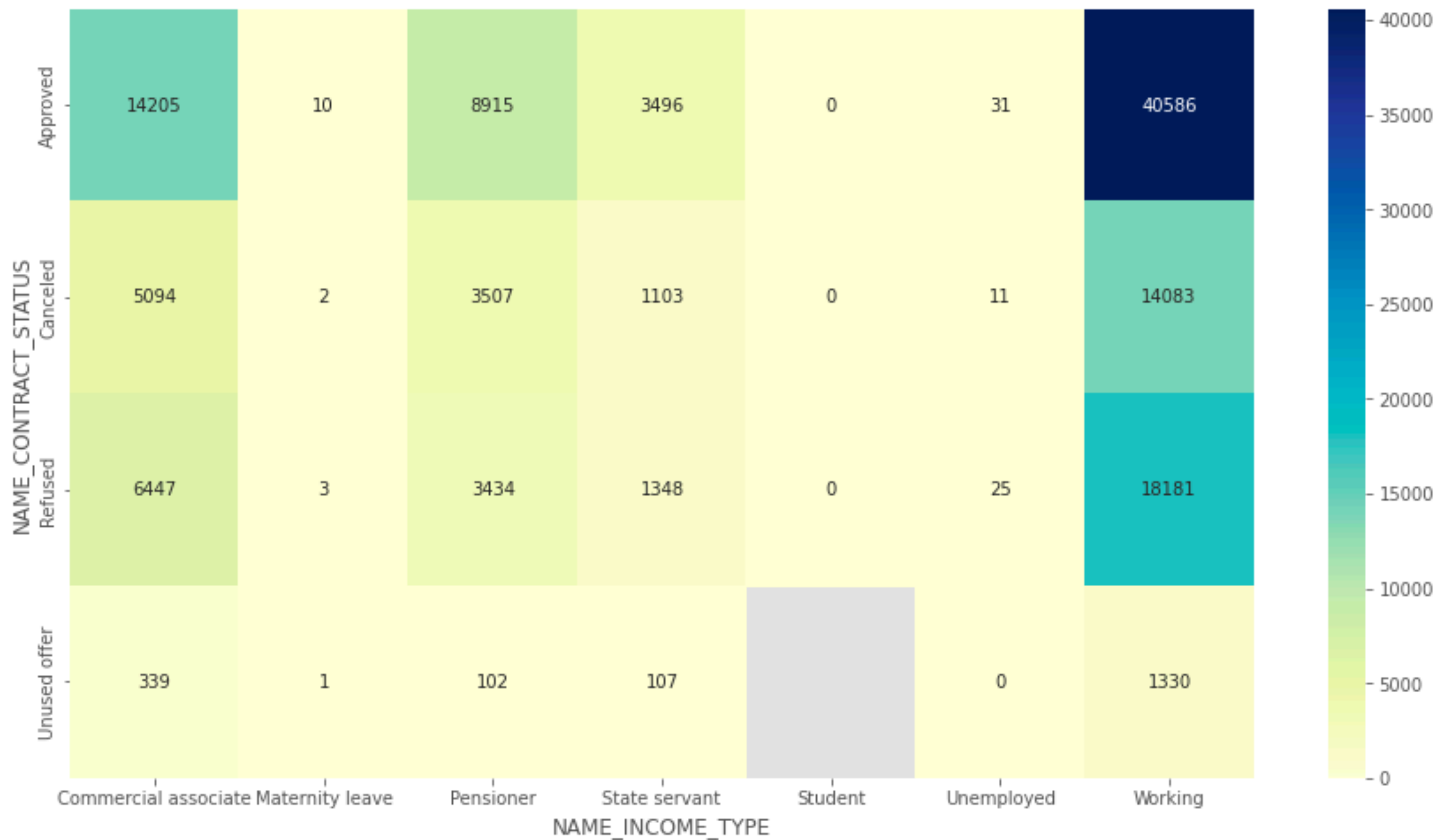
OBSERVATION:

1. Working applicant with Approved status have defaulted in highest numbers
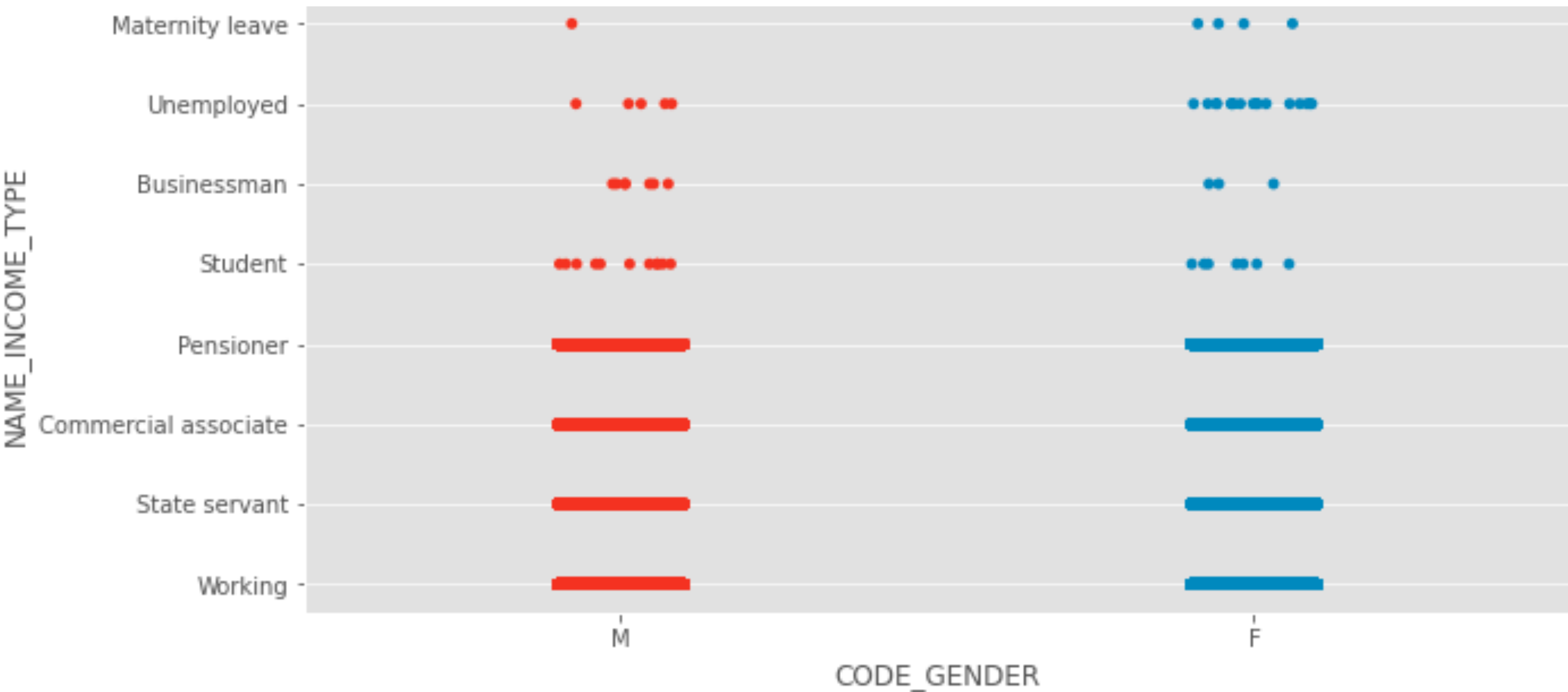2. Previous applications with Refused, Cancelled, Unused loans also have default which is a matter of concern. This indicates that the financial company had Refused/cancelled previous application, but has approved the current and is facing default on these loans.
3. 18,181 applicanst of working class were REFUSED earlier and now have defaulted.

Age Group Vs. Contact Status
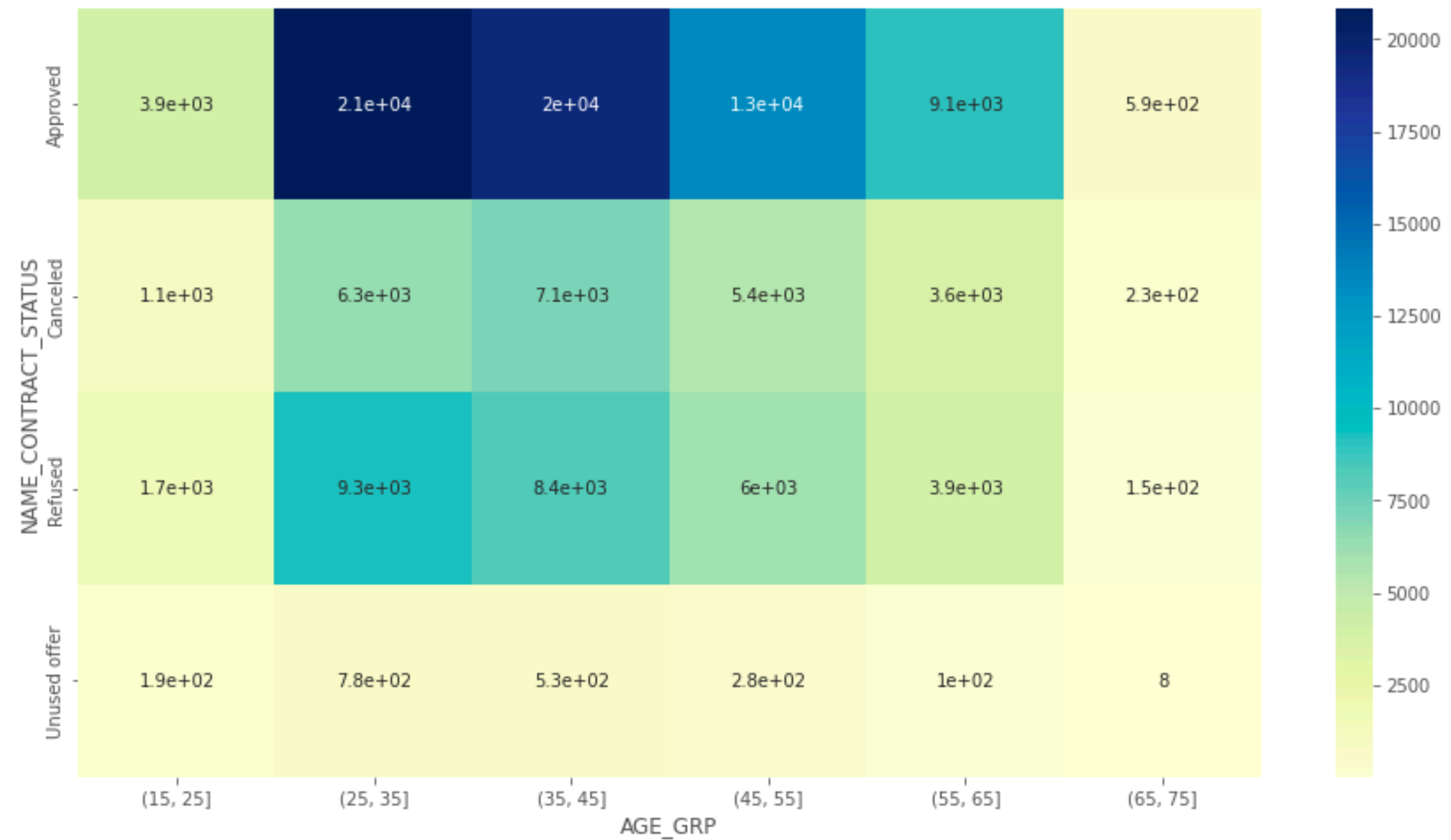
OBSERVATION:

1. Approved loans of age group 25-45 have higher defaults.
2. Refused, cancelled, loans in previous application have defaulted in current.

CONCLUSION:

* The count of Maternity Leave in *NAME_INCOME_TYPE* is very less and it also has maximum % of payment difficulties i.e. around 40%. Hence, client with income type as Maternity leave are the driving factors for Loan Defaulters.

* The count of Low skilled Laborers is comparatively very less and it also has maximum % of payment difficulties i.e. around 17%. Hence, applicant with occupation type as Low skilled Laborers are the driving factors for Loan Defaulters.

* Most Defaulters are coming under the age group of 25-35 years olds , followed by 35-45 years.

* Refused and cancelled loans in previous application were approved later and the same variables can can be seen defaulting loans.

* In overall data Working Class can be observed as applying for highr number of loans hence, loan defaulting numbers are higher.

* Major data of the occupation is missing which is effecting the analysis.

* The Cash Loan are sanctioned in a large number bank should give Revolving loan.

* Applicant who are single Defaults more.

* Business persons Doesn't default. Lending loan to them is safer.

THANK YOU