

CREDIT EDA CASE STUDY

BANK LOAN APPROVAL

IFFATH FARAZ

PROBLEM STATEMENT

- The Company wants to understand driving factors behind Loan default.
- There are two types of risk associated with a loan request in a Bank:
- >>> If an applicant is capable of repaying the loan, but their loan request is rejected this results in the loss of the bank.
- >>> If the applicant is likely to be defaulter and the bank approves their loan, the bank faces financial loss.

ANALYSIS STEPS

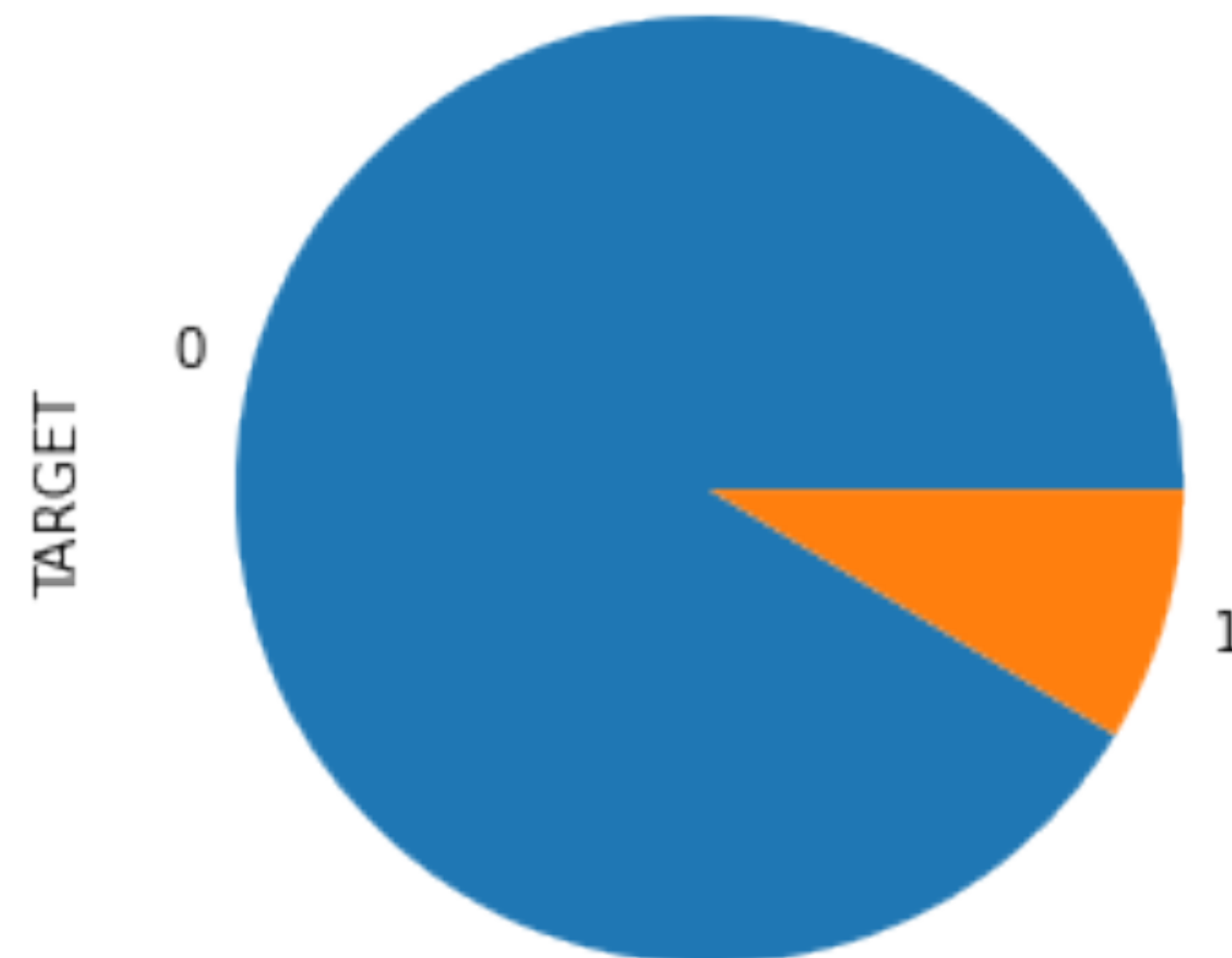
APPLICATION DATAFRAME,PRE APPLICATION DATAFRAME AND MERGED DATAFRAME

- Loaded the datafile, checked for formatting errors.
- Handled Missing values by either dropping/imputation
- Handled data imbalance and data irregularities
- Univariate Analysis
- Bivariate Analysis
- Segmented Univariate Analysis
- Top 10 correlation

PROPORTION OF DEFAULTERS IN THE DATAFRAME

8.7% of the Applicants are defaulters

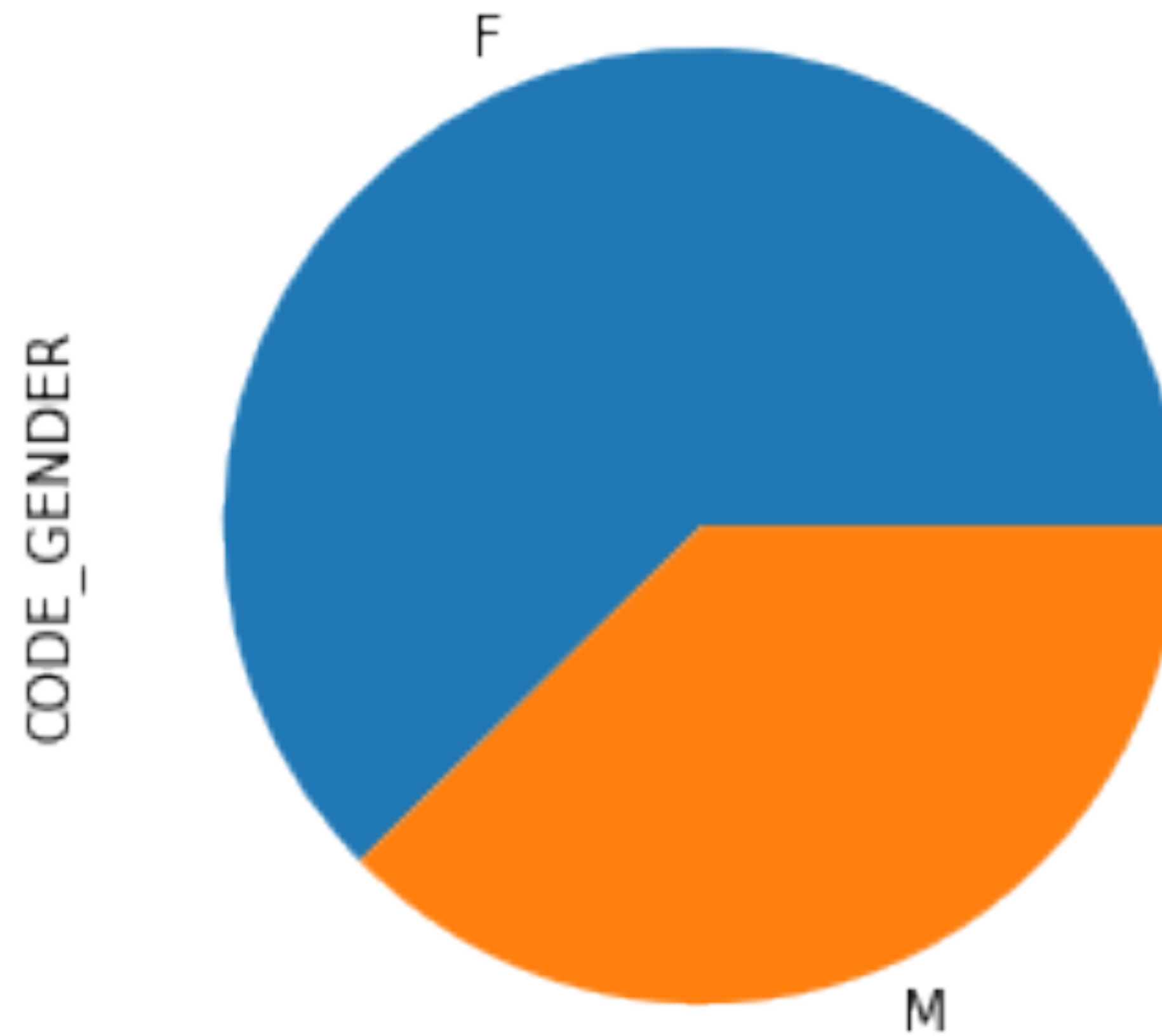
Univariate Analysis on Target Variable where 0 stands for Non defaulters and 1 stand for Defaulters



UNIVARIATE ANALYSIS ON GENDER

Univariate Analysis on Gender

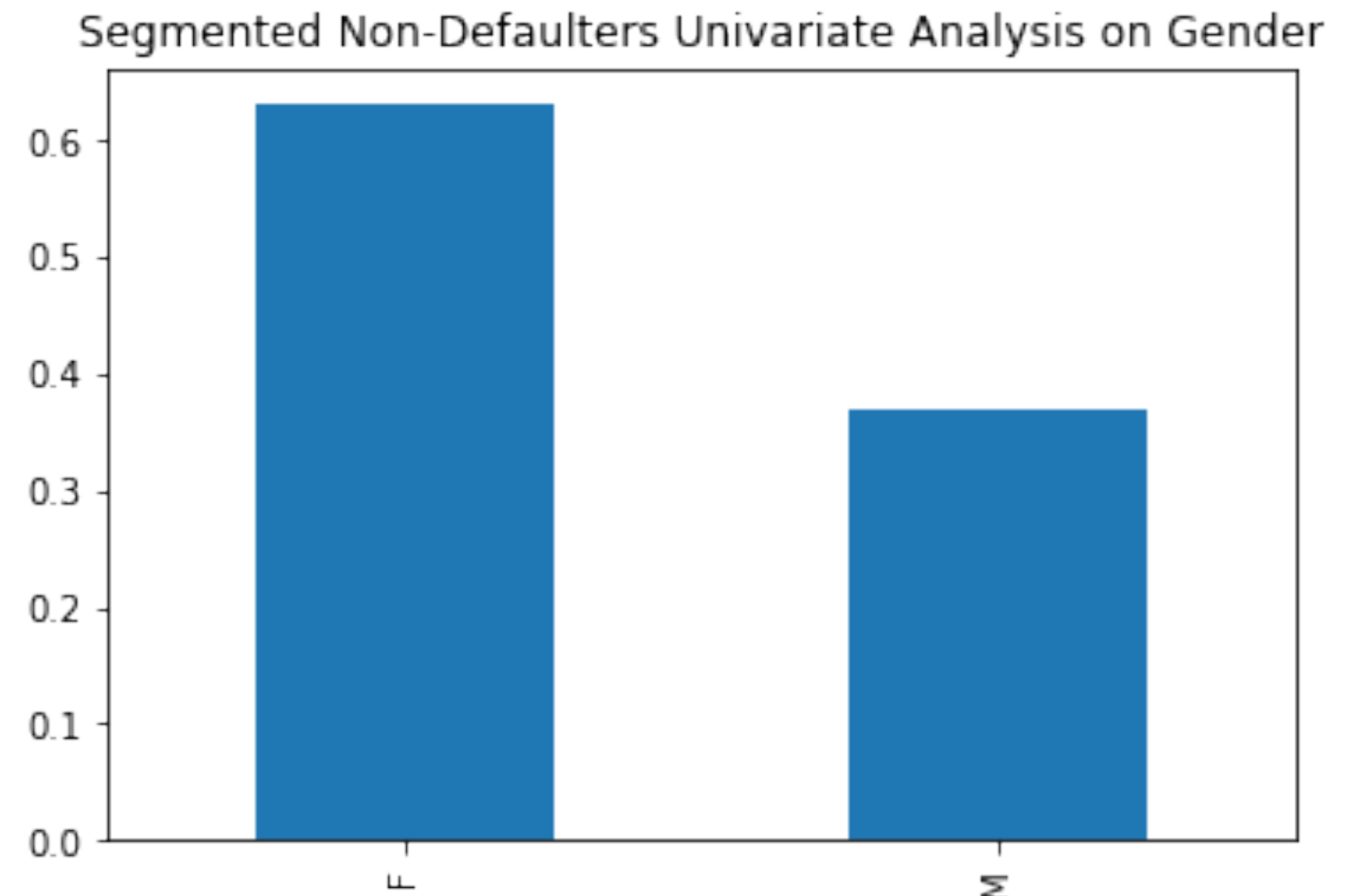
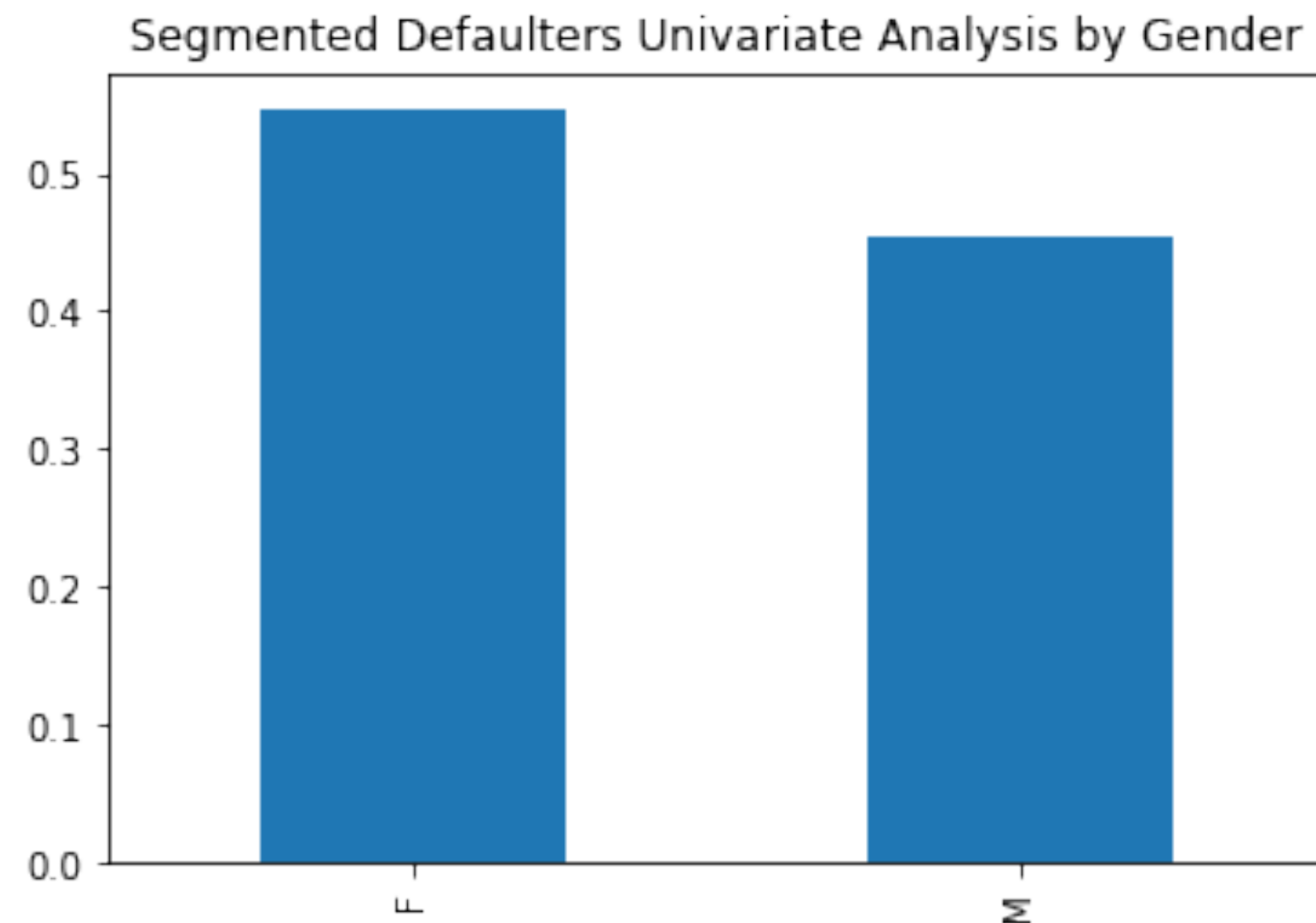
- FEMALE—62.3%
- MALE—37.7



DEFAULTERS AND NON DEFAULTERS

SEGMENTED UNIVARIATE ANALYSIS by GENDER

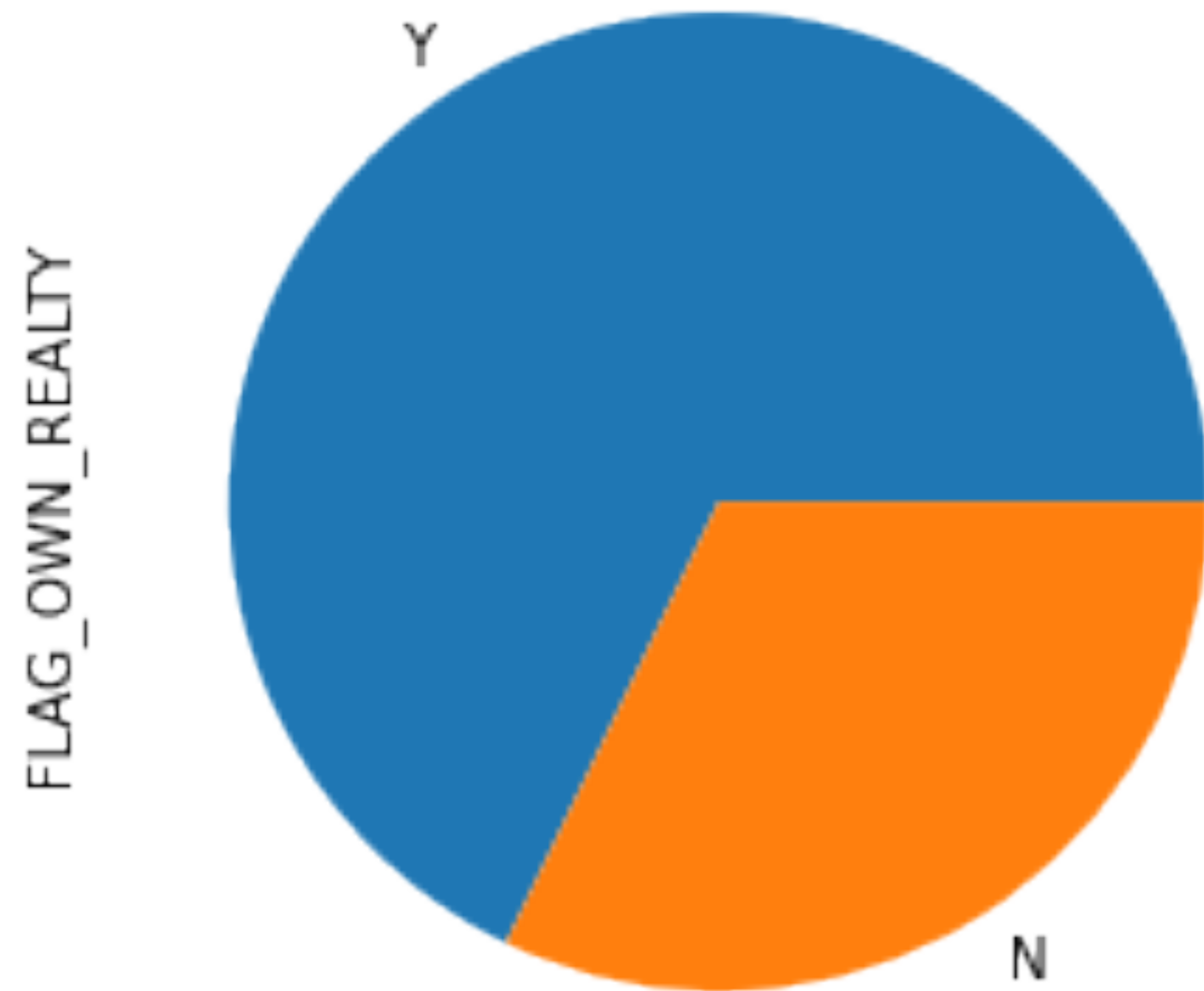
- Female applicants are in the higher ration and they are higher in the Defaulter section as well.



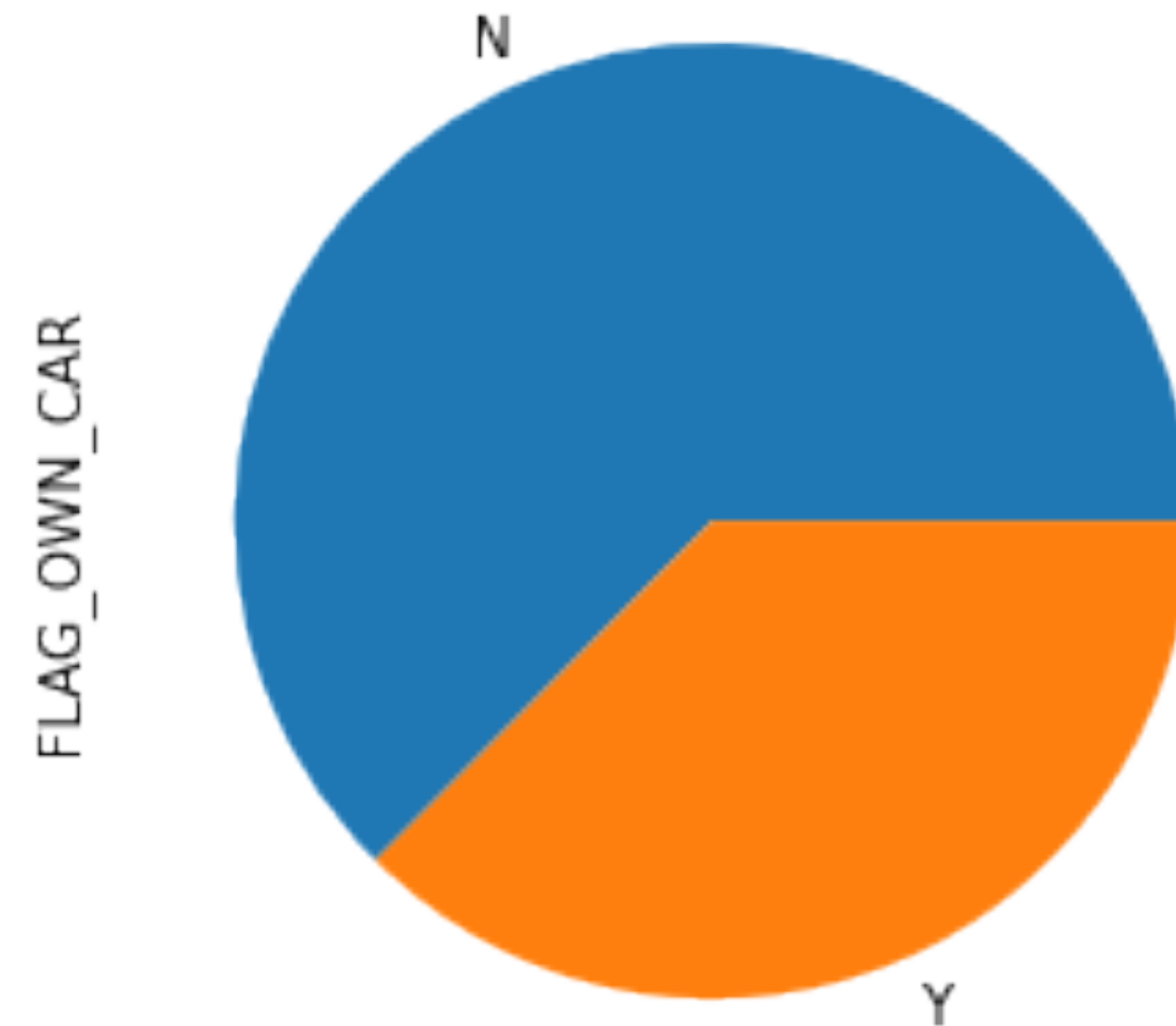
HOUSE AND CAR OWNERSHIP

- Majority of applicants own a house but not a car

Univariate Analysis on Owning a house



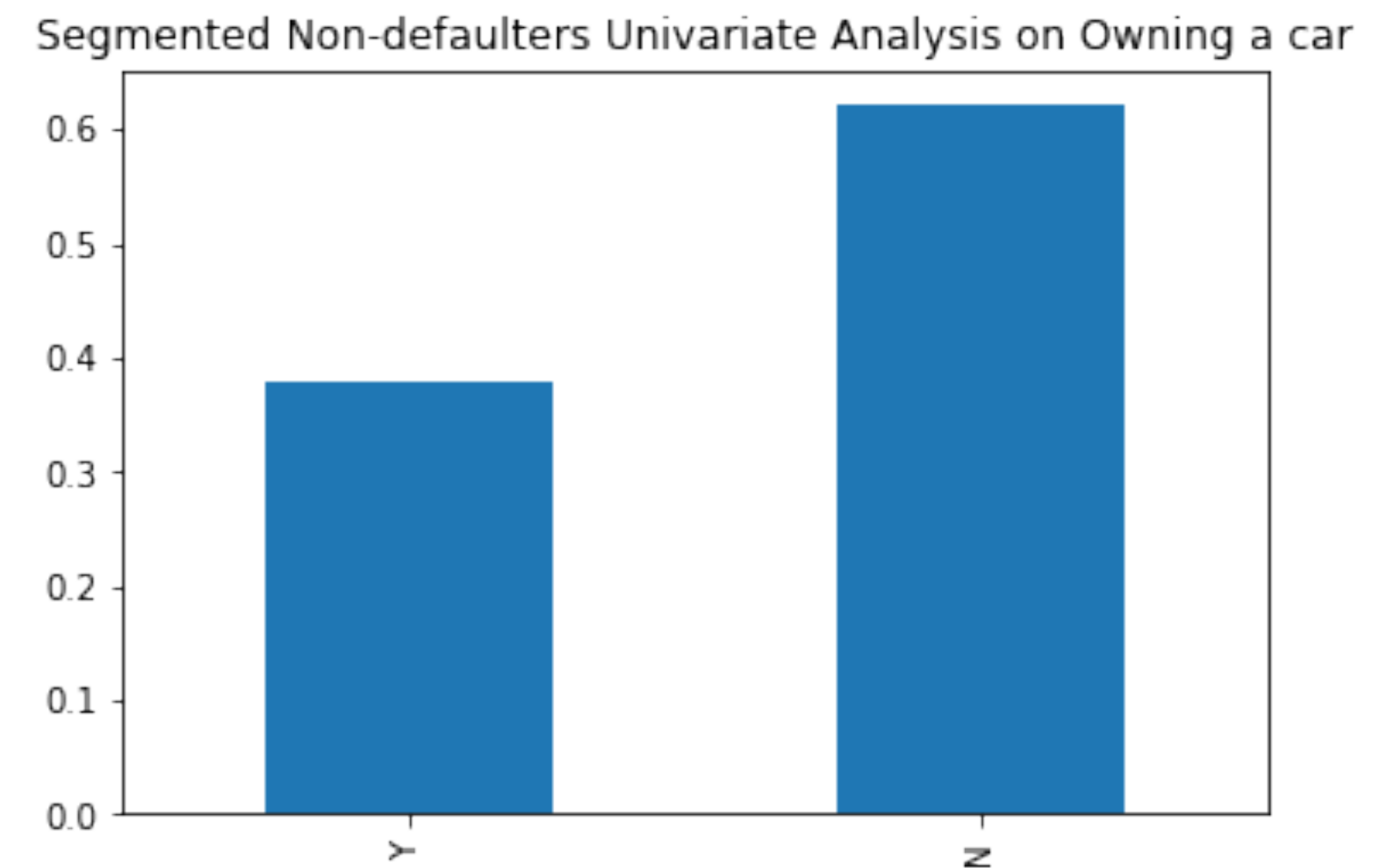
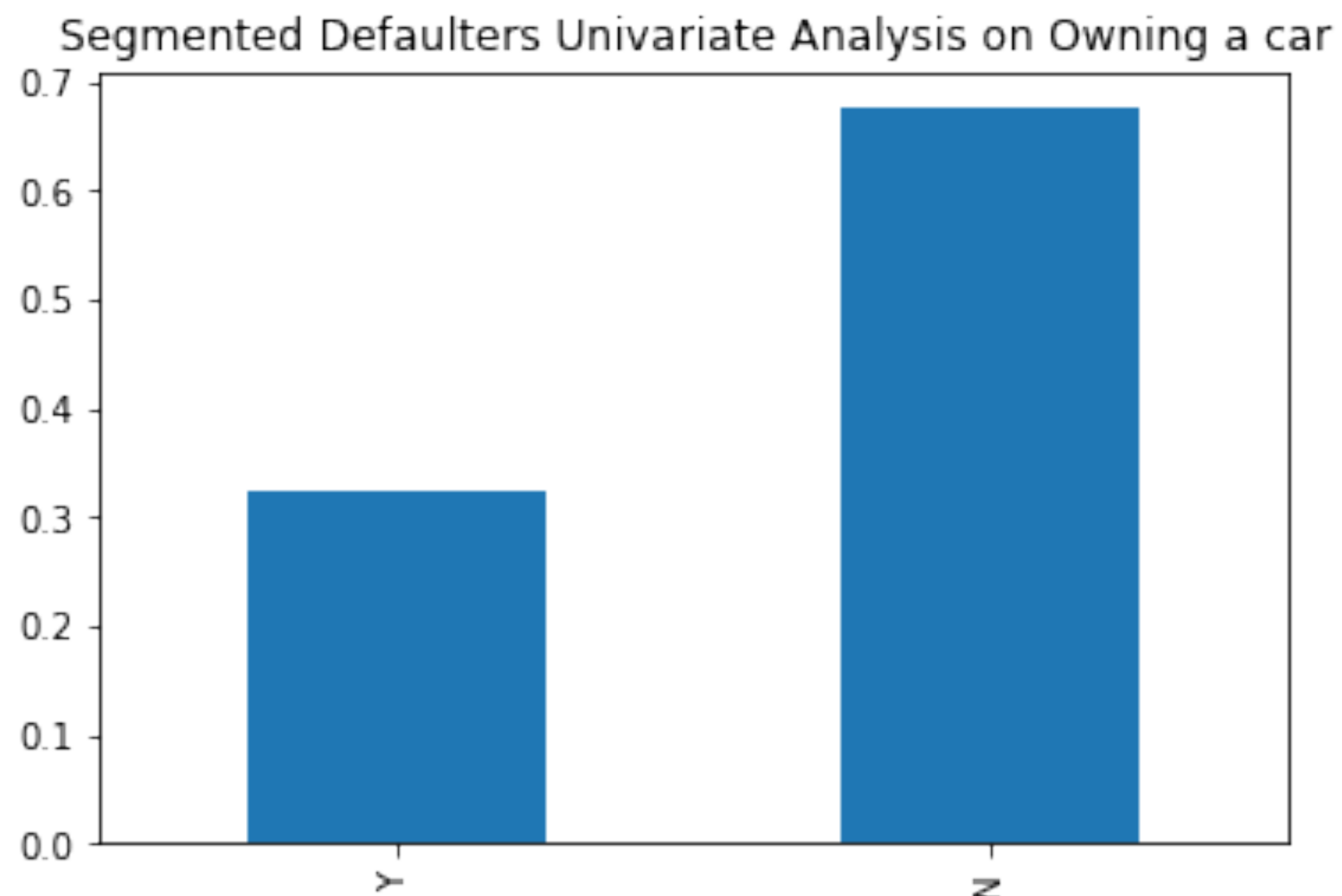
Univariate Analysis on Owning a Car



DEFAULTERS AND NON DEFAULTERS

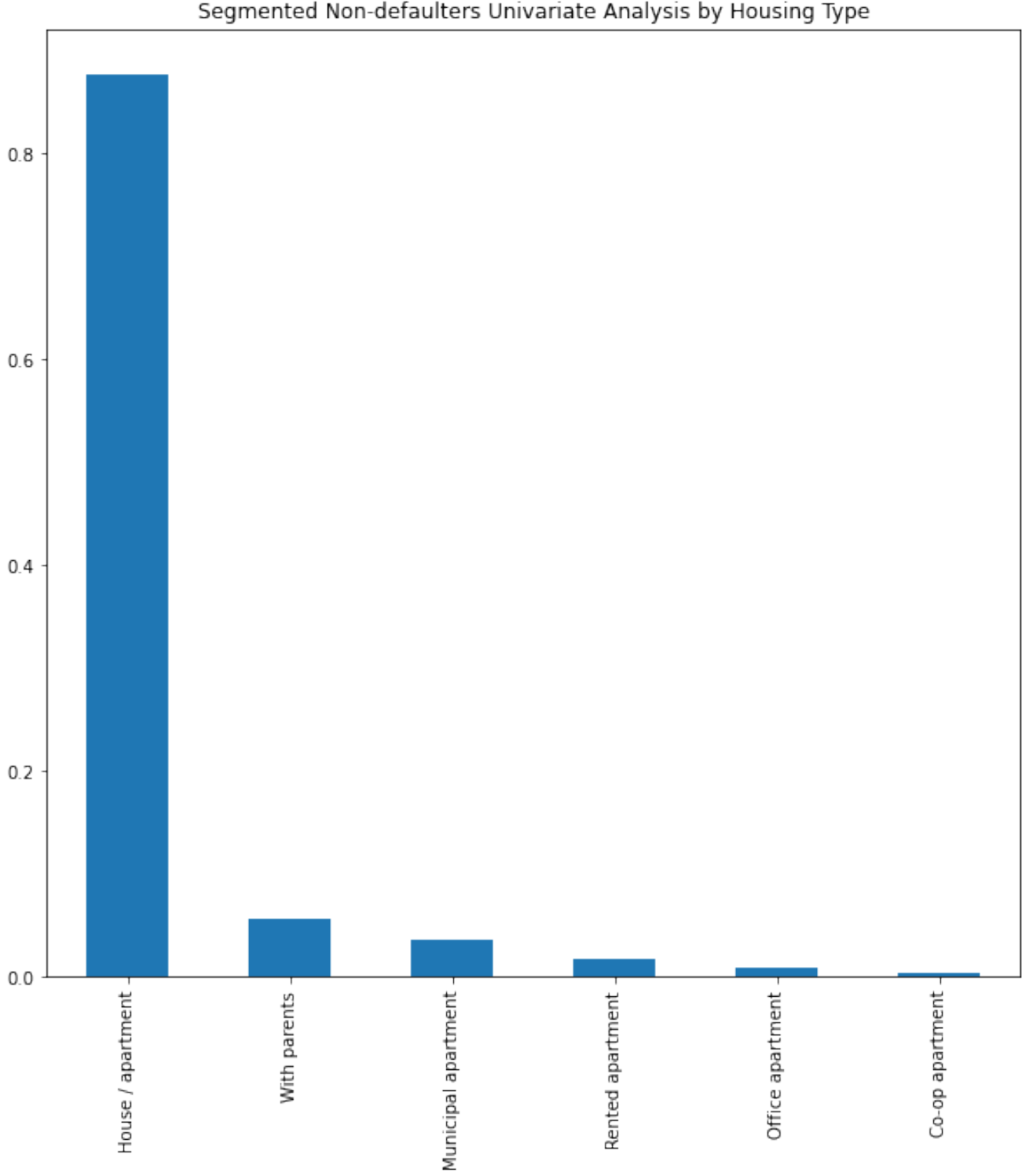
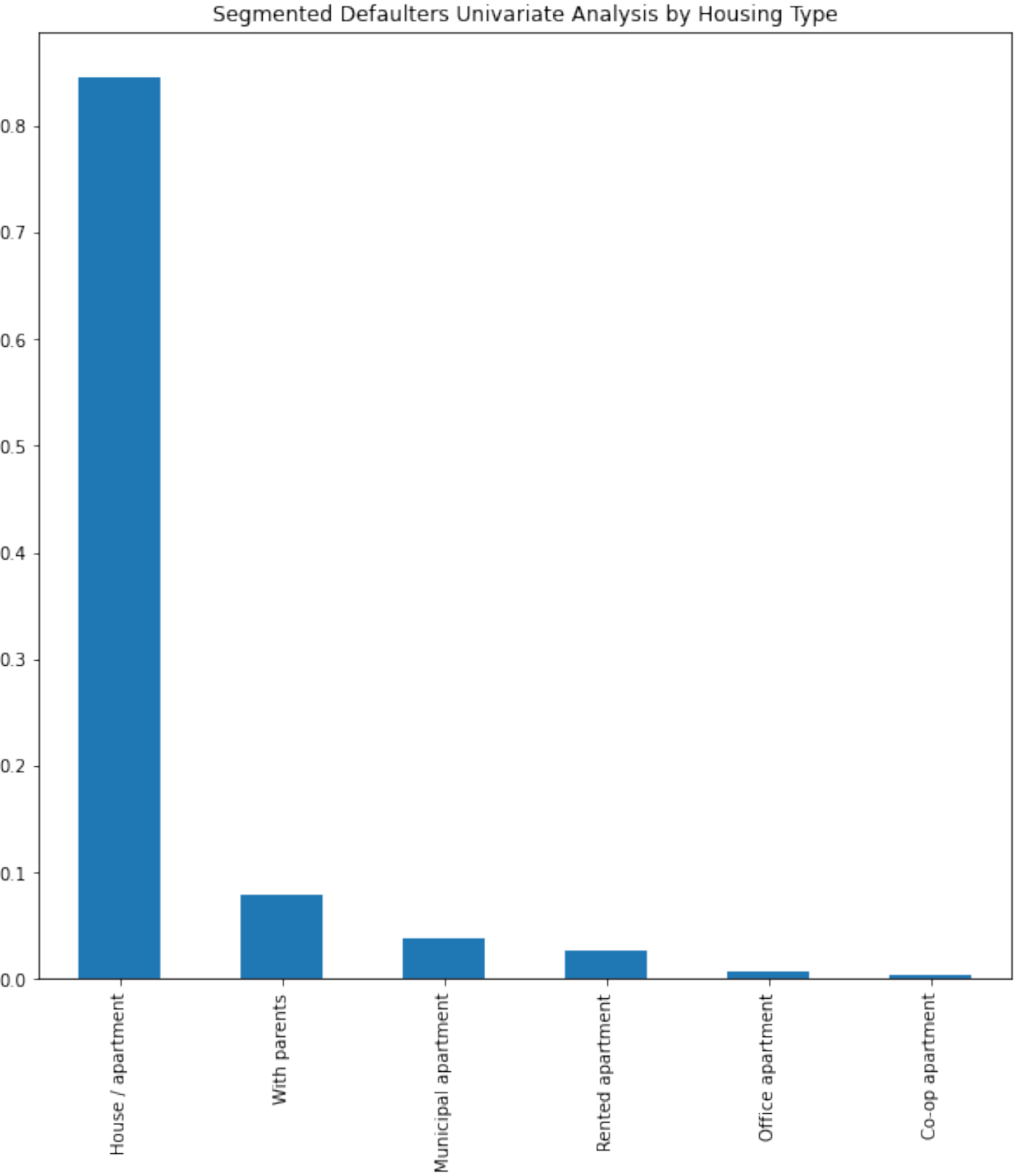
SEGMENTED UNIVARIATE ANALYSIS BY OWN CAR

- We can say that the rate of default of people owning car is less than the applicants who don't own a car.



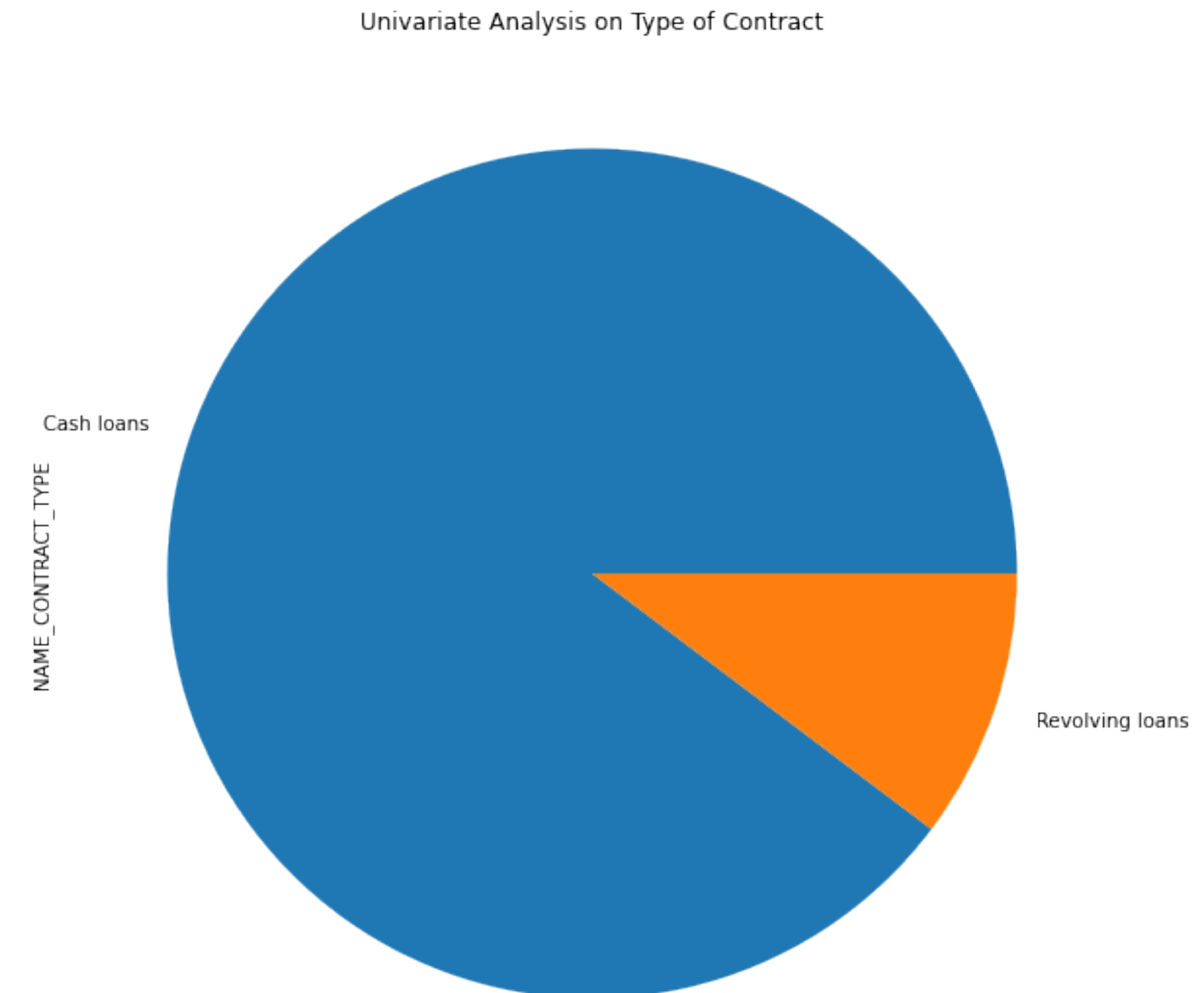
DEFAULTERS AND NON DEFAULTERS

SEGMENTED UNIVARIATE ANALYSIS BY OWN HOUSE



CONTRACT TYPE UNIVARIATE ANALYSIS

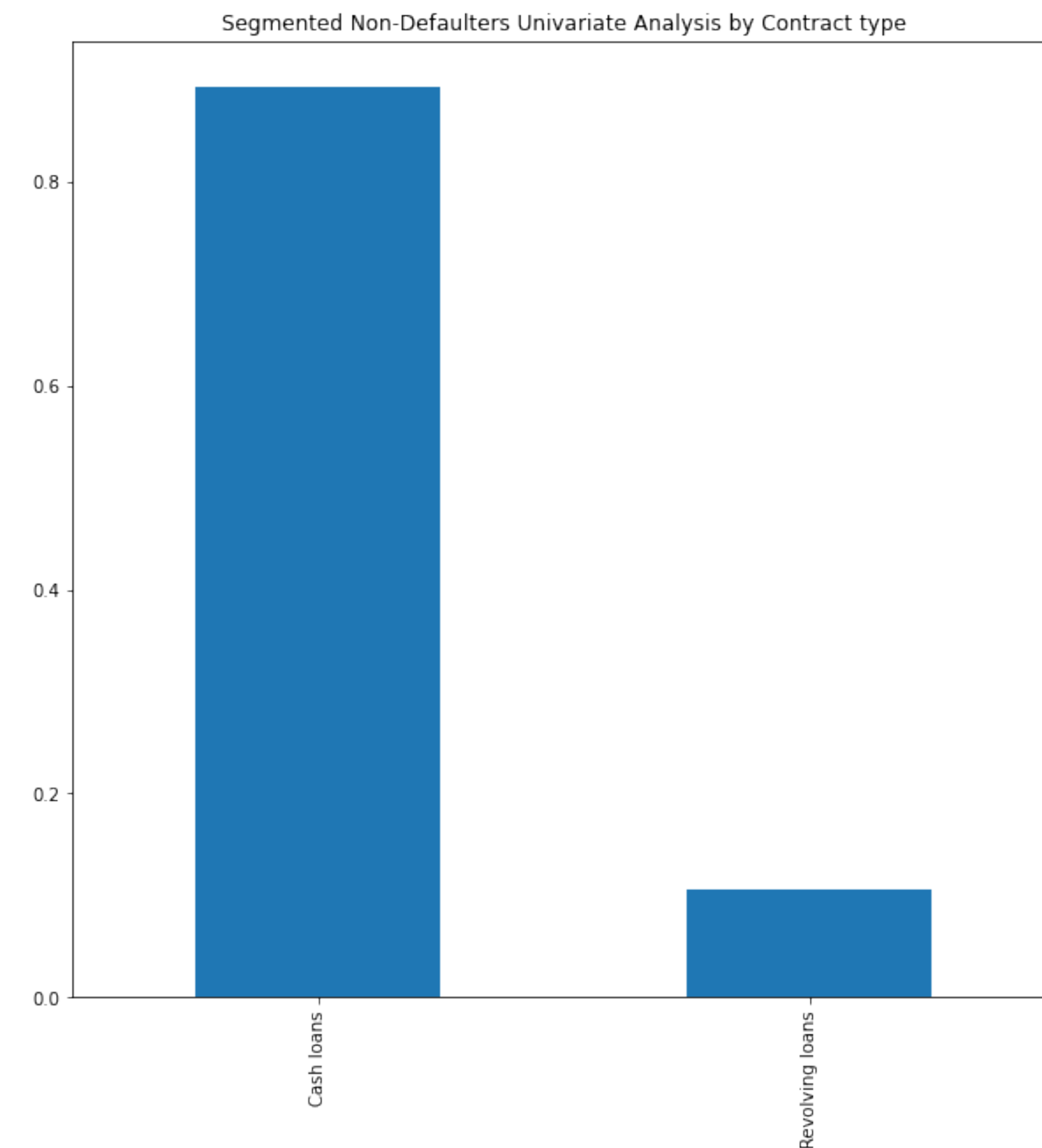
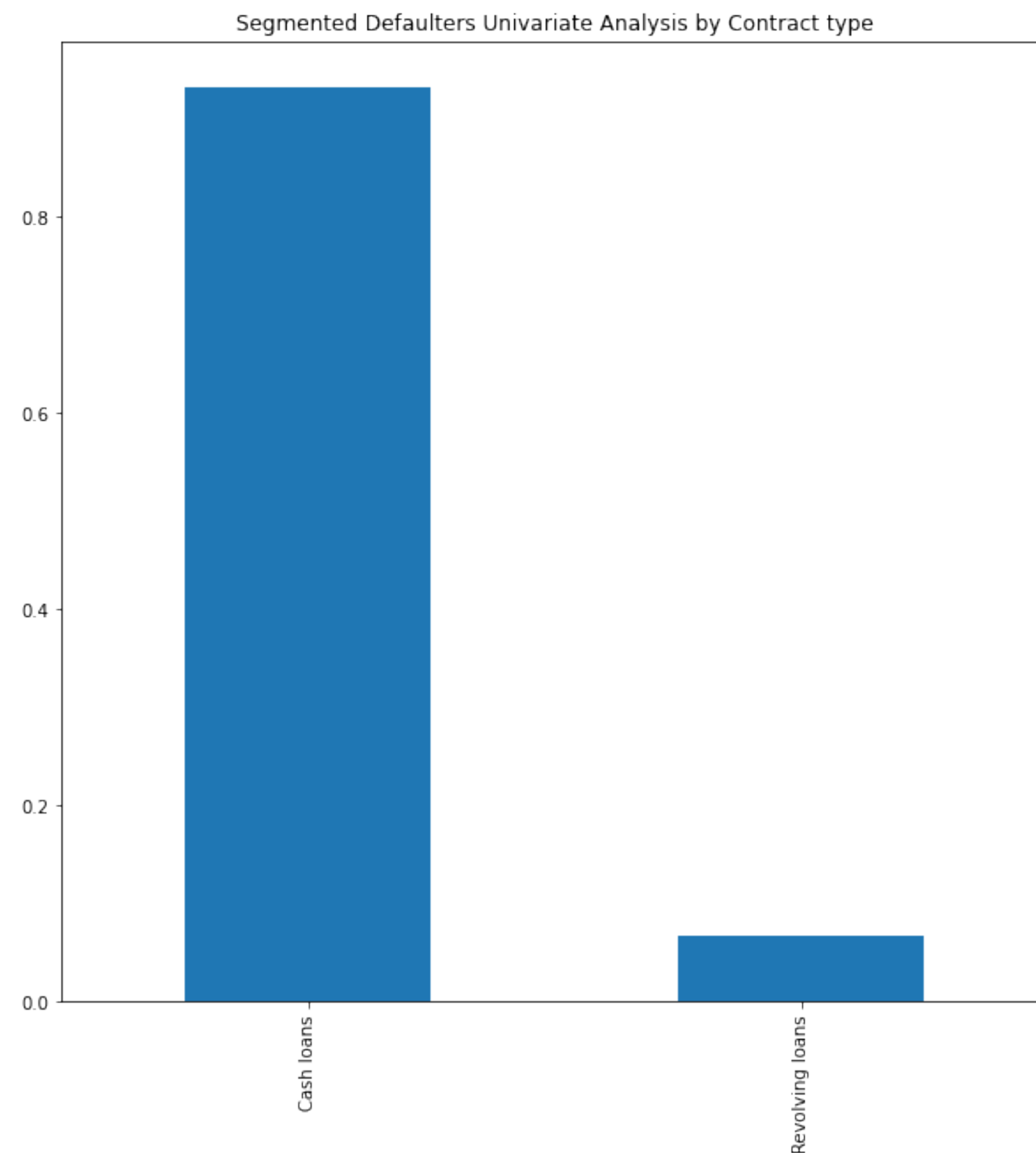
- There are two types of loans:
- >Cash loan
- >Resolving Loan
- 89.7% of loans are Cash loan.



DEFAULTERS AND NON DEFAULTERS

SEGMENTED UNIVARIATE ANALYSIS BY CONTRACT TYPE

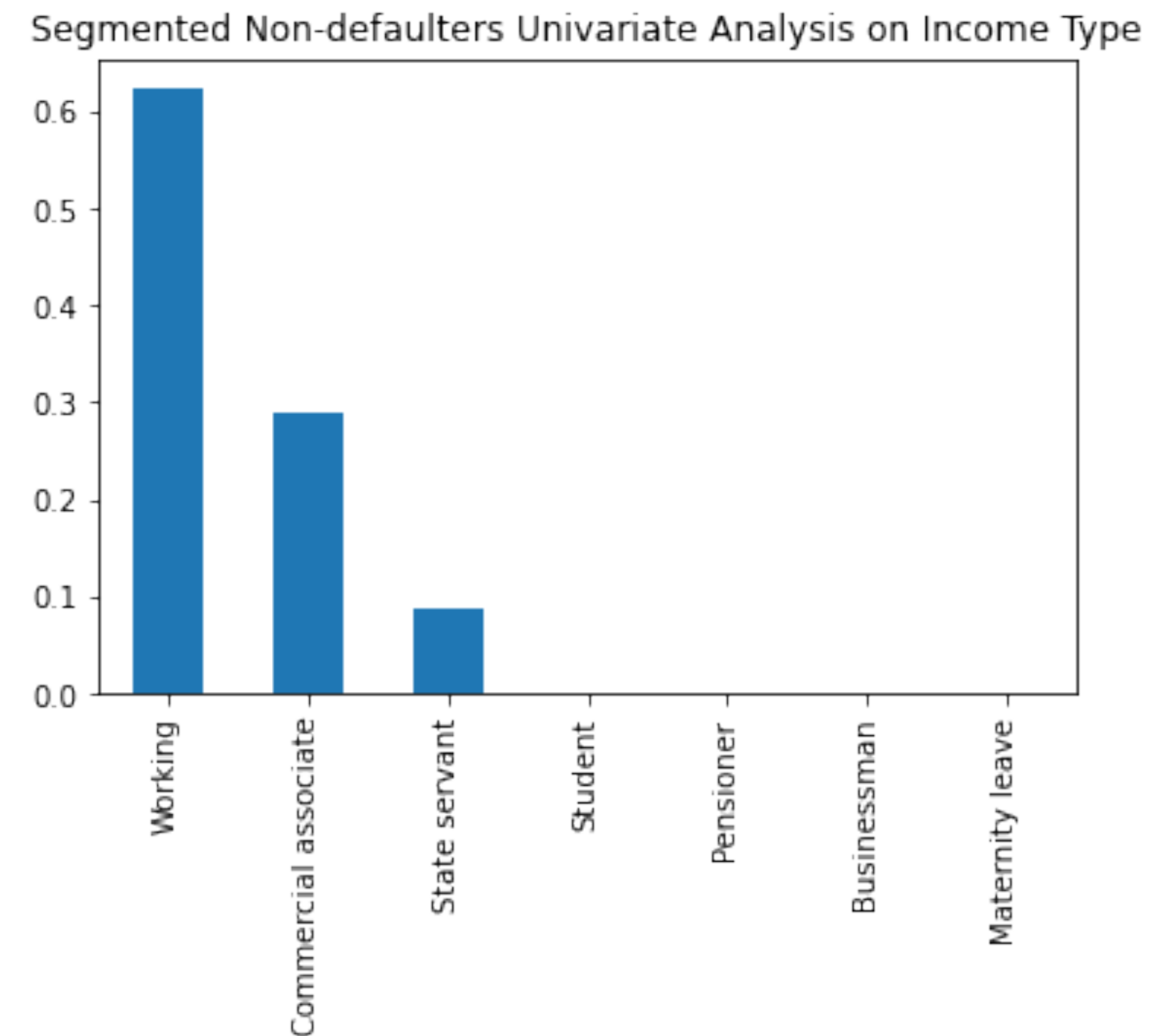
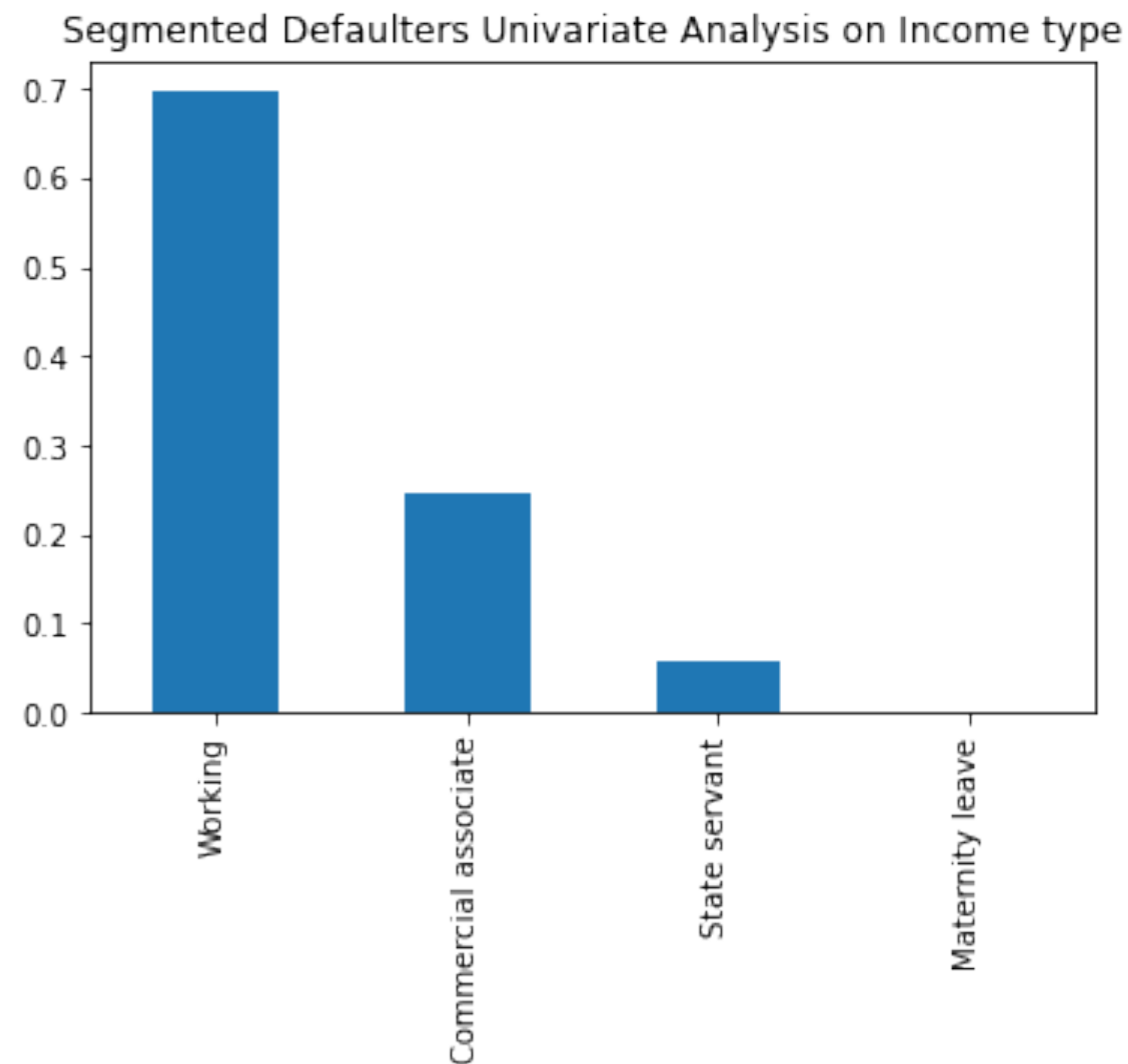
- Cash loans contribute majority of defaulters



DEFAULTERS AND NON DEFAULTERS

SEGMENTED UNIVARIATE ANALYSIS BY INCOME TYPE

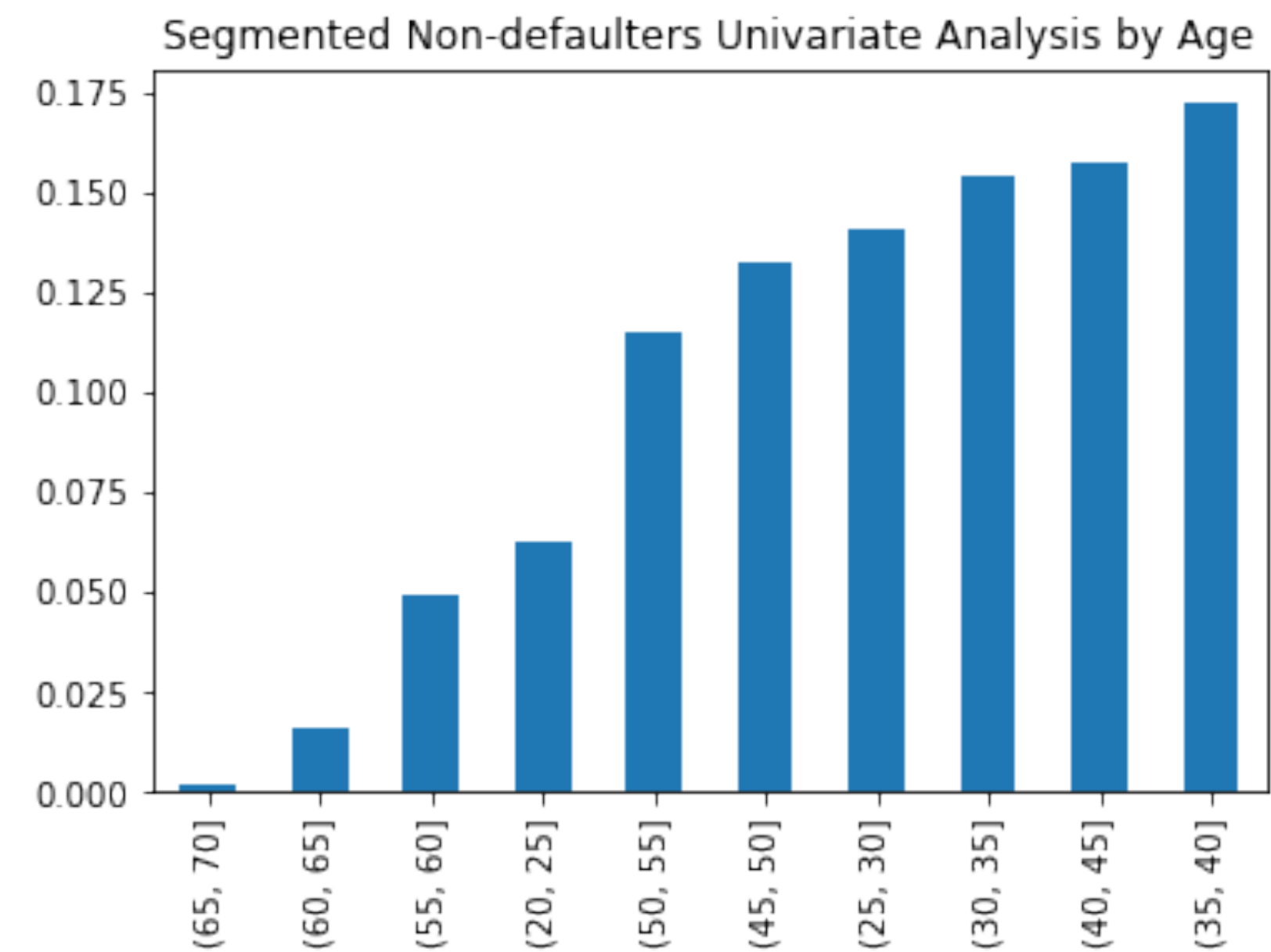
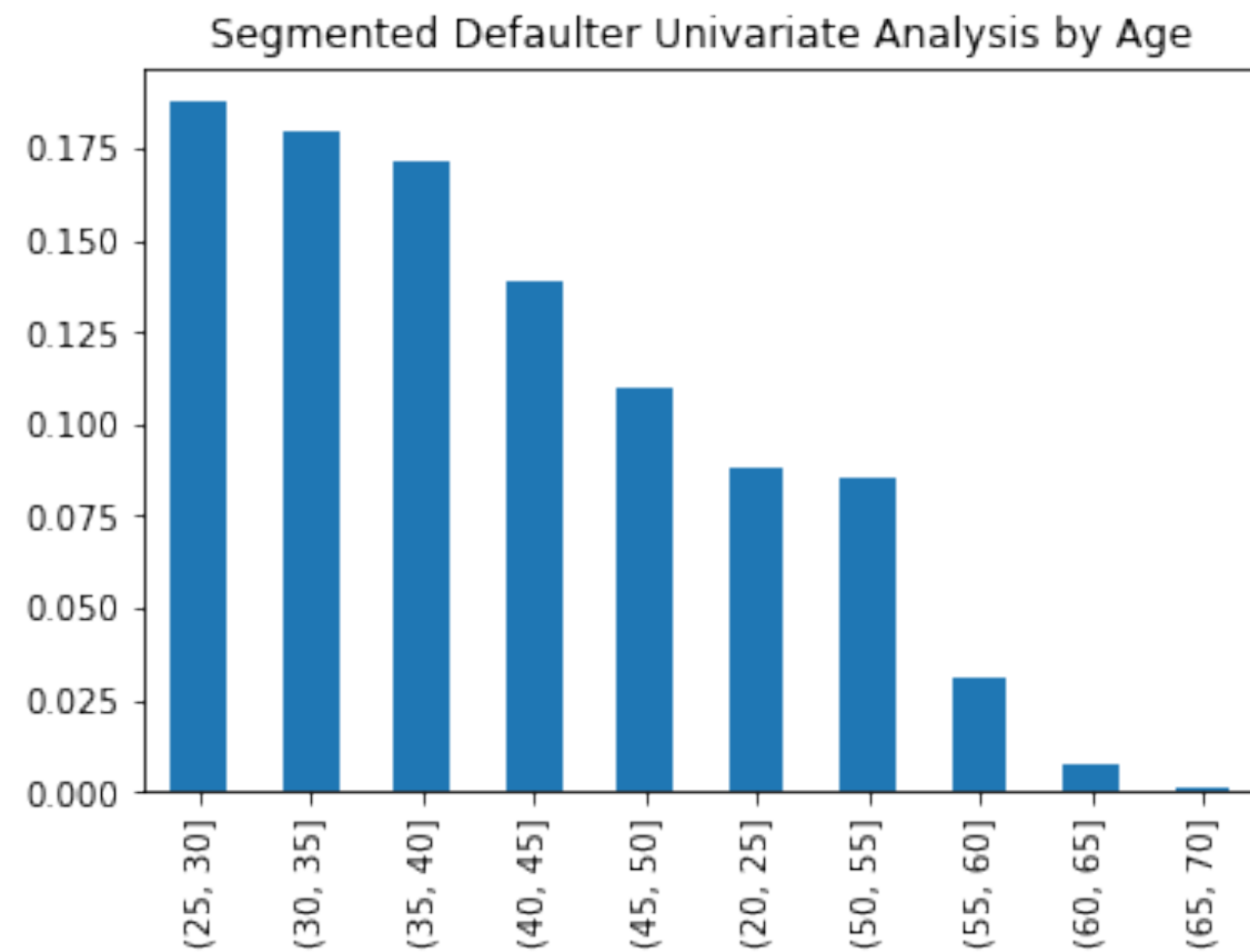
- Working class contribute to maximum number of loan applicants, Hence Working class have more chances to default.



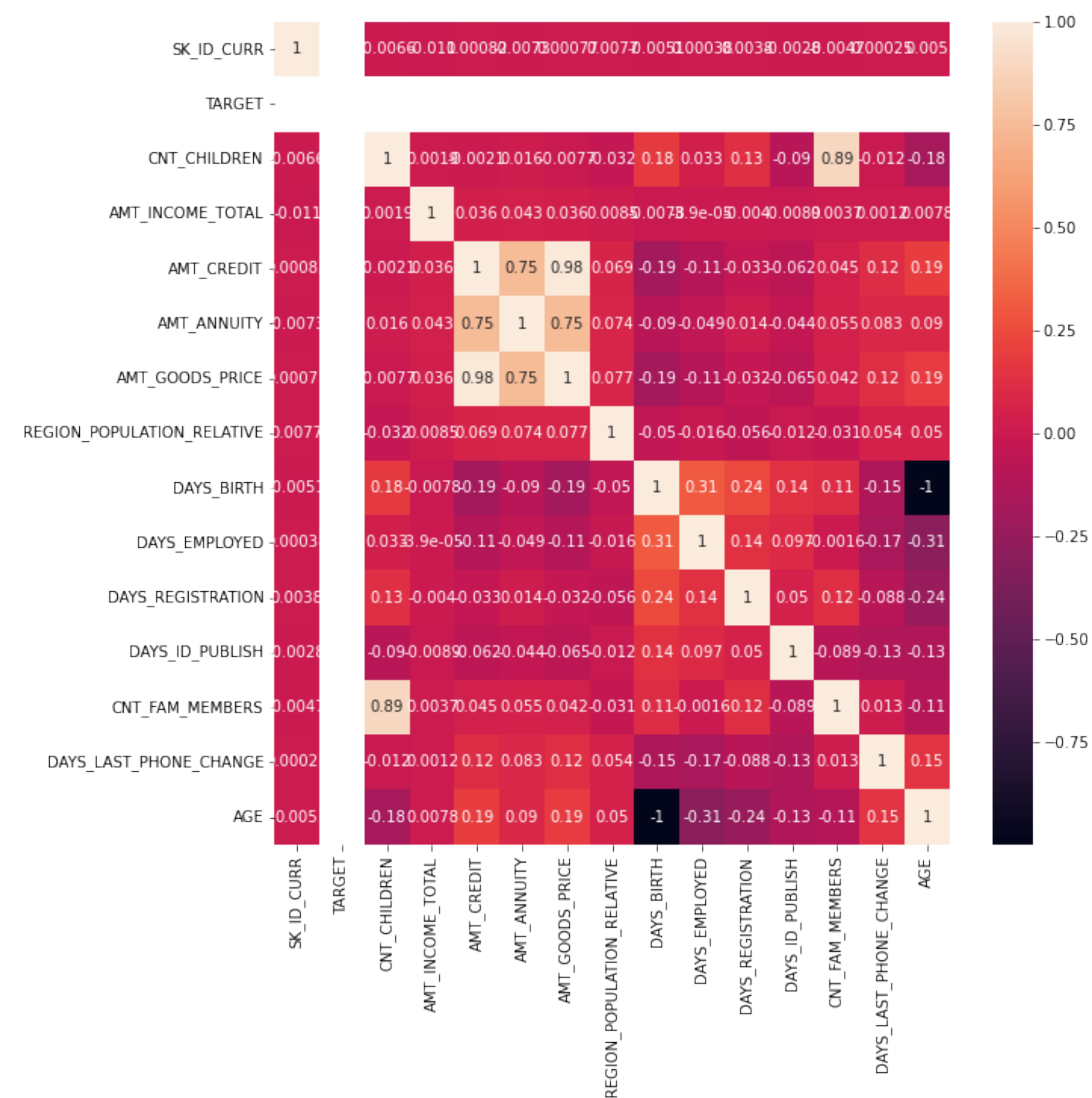
DEFAULTERS AND NON DEFAULTERS

SEGMENTED UNIVARIATE ANALYSIS BY AGE GROUP

- Majority of the defaulters are in the Age range(25-30)



CORRELATION METRIX (DEFAULTERS)



TOP 10 CORRELATED VARIABLES

	Column1	Column2	Correlation
94	AMT_GOODS_PRICE	AMT_CREDIT	0.982536
182	CNT_FAM_MEMBERS	CNT_CHILDREN	0.893829
95	AMT_GOODS_PRICE	AMT_ANNUITY	0.749591
79	AMT_ANNUITY	AMT_CREDIT	0.748708
143	DAYS_EMPLOYED	DAYS_BIRTH	0.307018
158	DAYS_REGISTRATION	DAYS_BIRTH	0.241202
214	AGE	AMT_CREDIT	0.189389
216	AGE	AMT_GOODS_PRICE	0.185020
122	DAYS_BIRTH	CNT_CHILDREN	0.176563
223	AGE	DAYS_LAST_PHONE_CHANGE	0.147100

CONCLUSION

CONCLUSION

- Banks must issue Loans to Business men(Both Male and Female) and students as they do not find it difficult to pay.
- Banks can issue loans to those applicants who live in House/Apartments, Rented Apartments.
- Banks should focus on promoting more cash loans and consumer loans as its the most common contract type amongst all applicants.
- Banks should focus on applicants with higher education and academic degrees who have high income.
- Bank must avoid giving loans to applicants with more number of children