



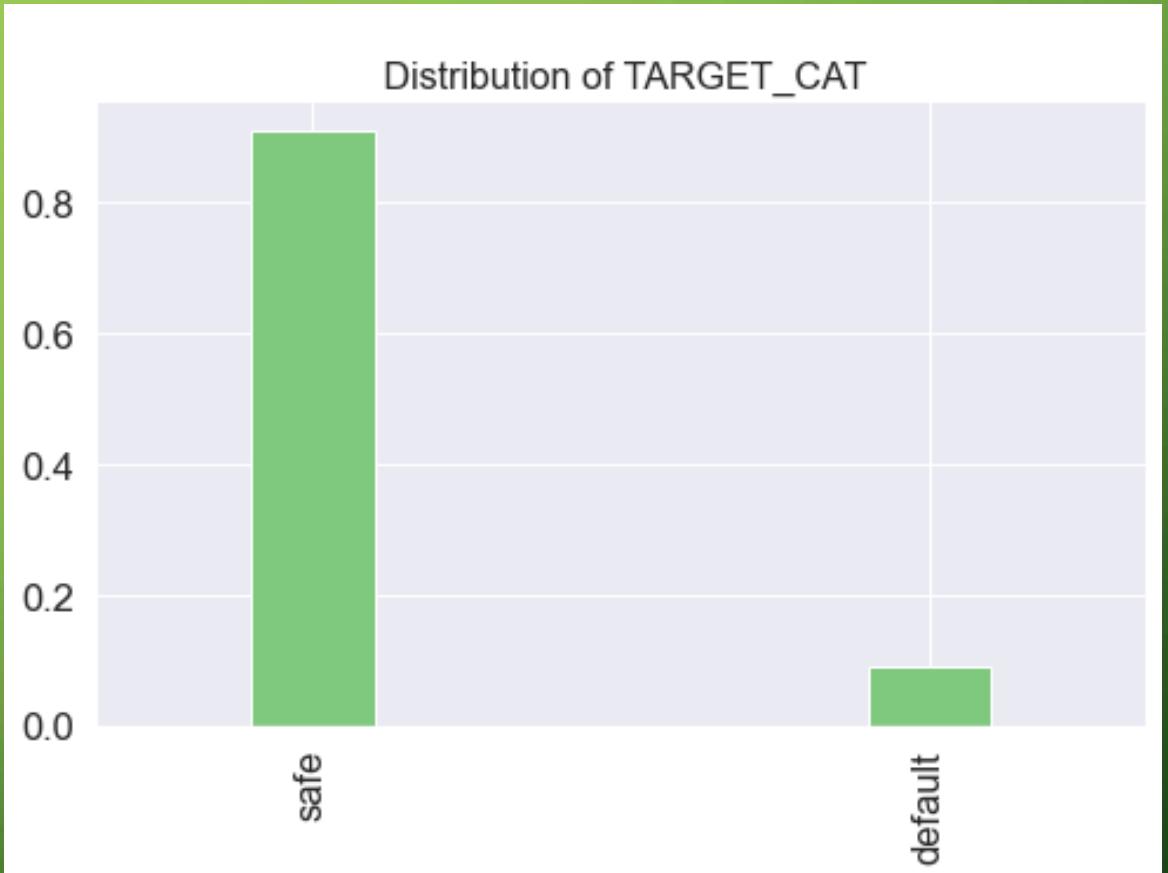
CREDIT RISK ANALYSIS

BY GLORIYA THOMAS & SUDEEP SHAH

UNIVARIATE ANALYSIS – CATEGORICAL VARIABLES

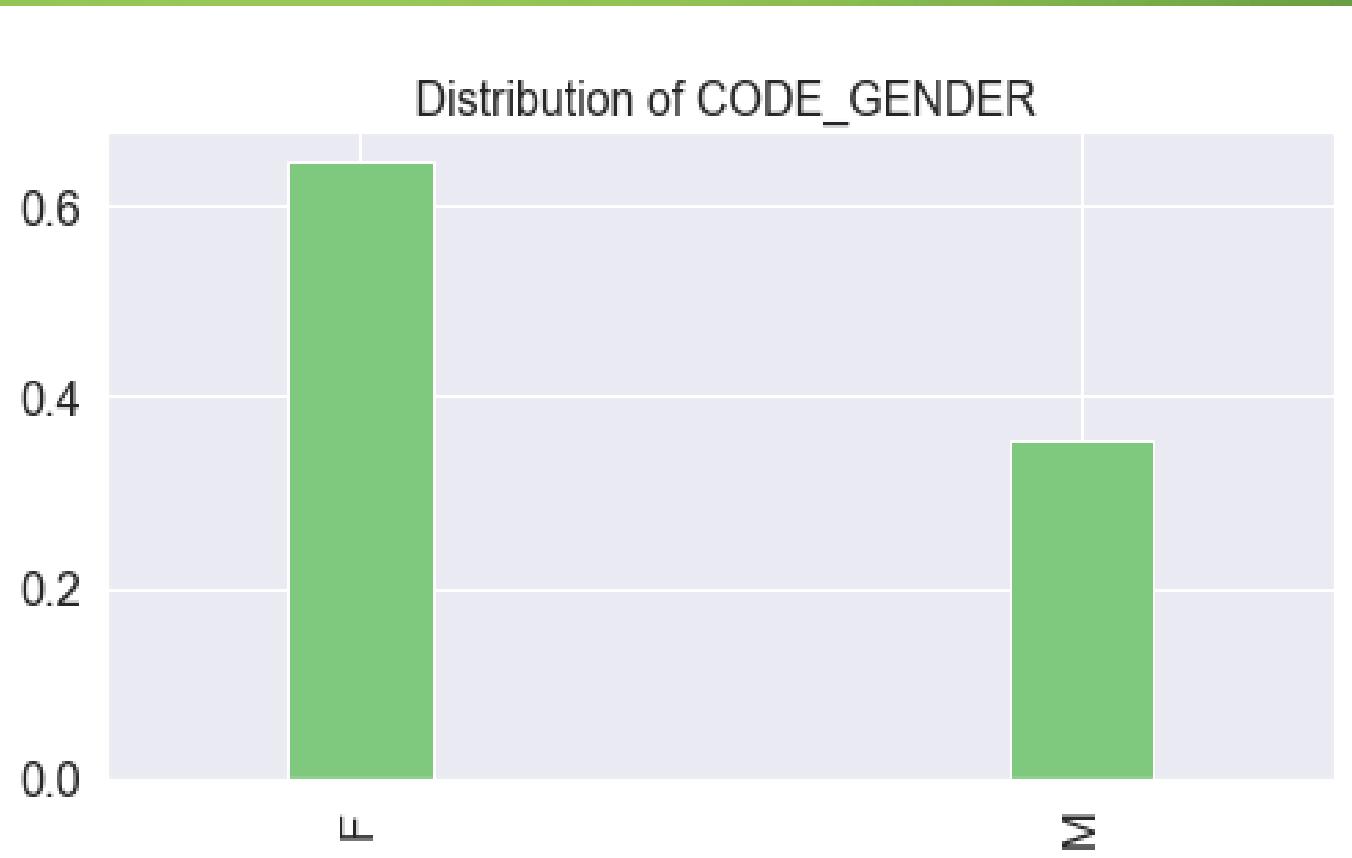
DISTRIBUTION OF TARGET

- Almost 91% of the customers paid back the loan, the remaining 9% turned default. Hence bank should keep an eye on the category who defaults more.



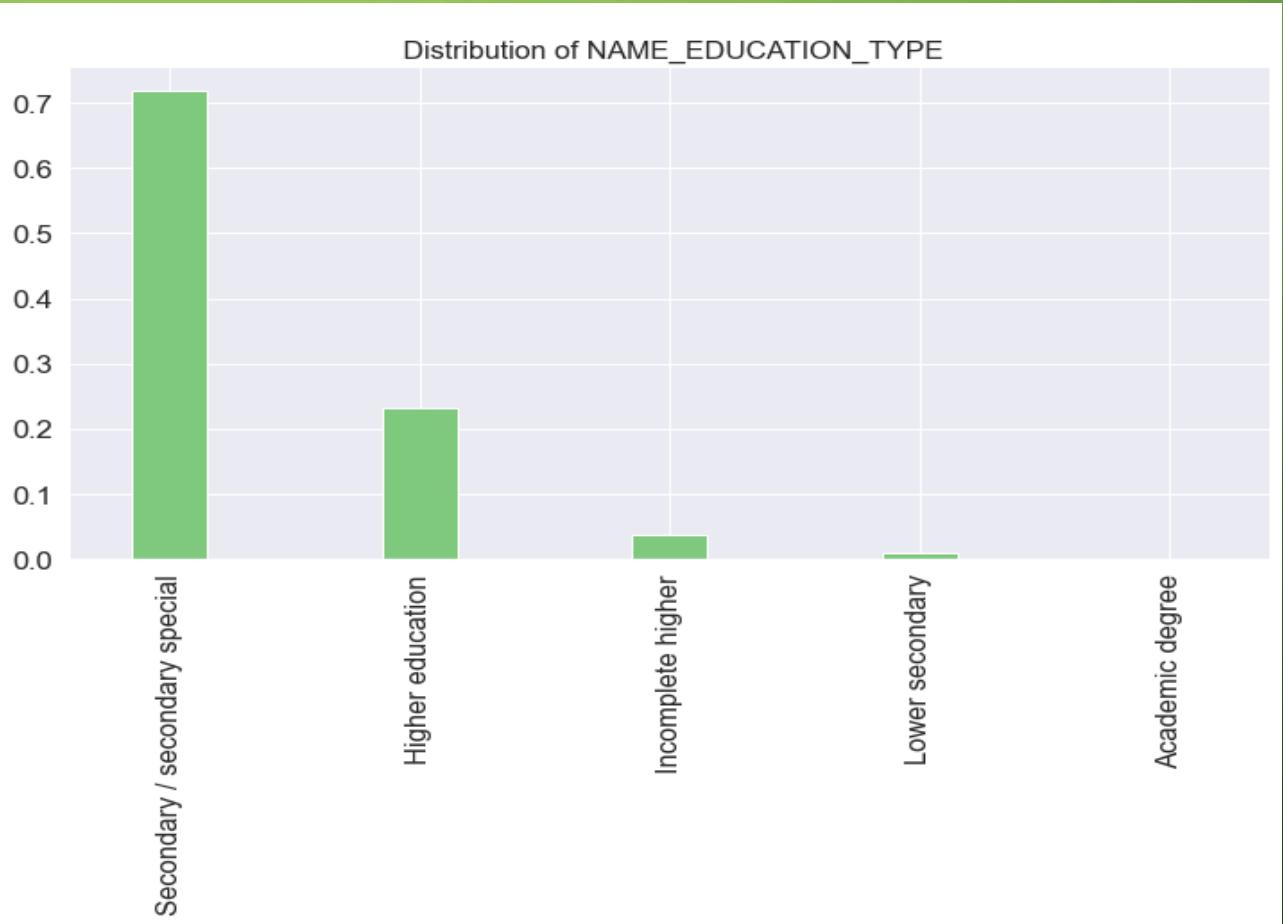
DISTRIBUTION OF GENDER

- Count of female customers are almost double as that of males.



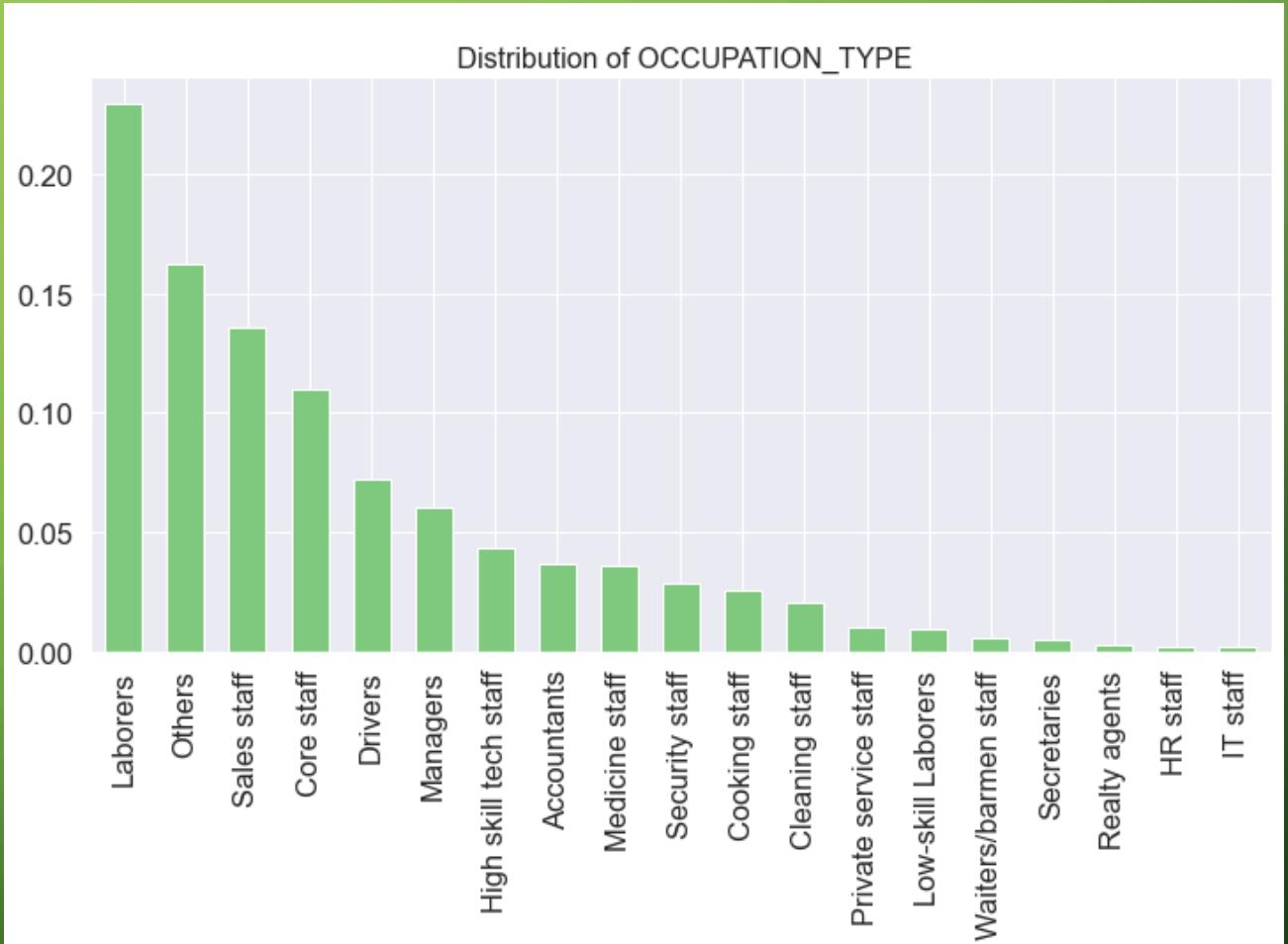
DISTRIBUTION OF EDUCATION

- The highest educational qualification of most of the loan applicants are higher education.
- Not many graduates have applied for loans.



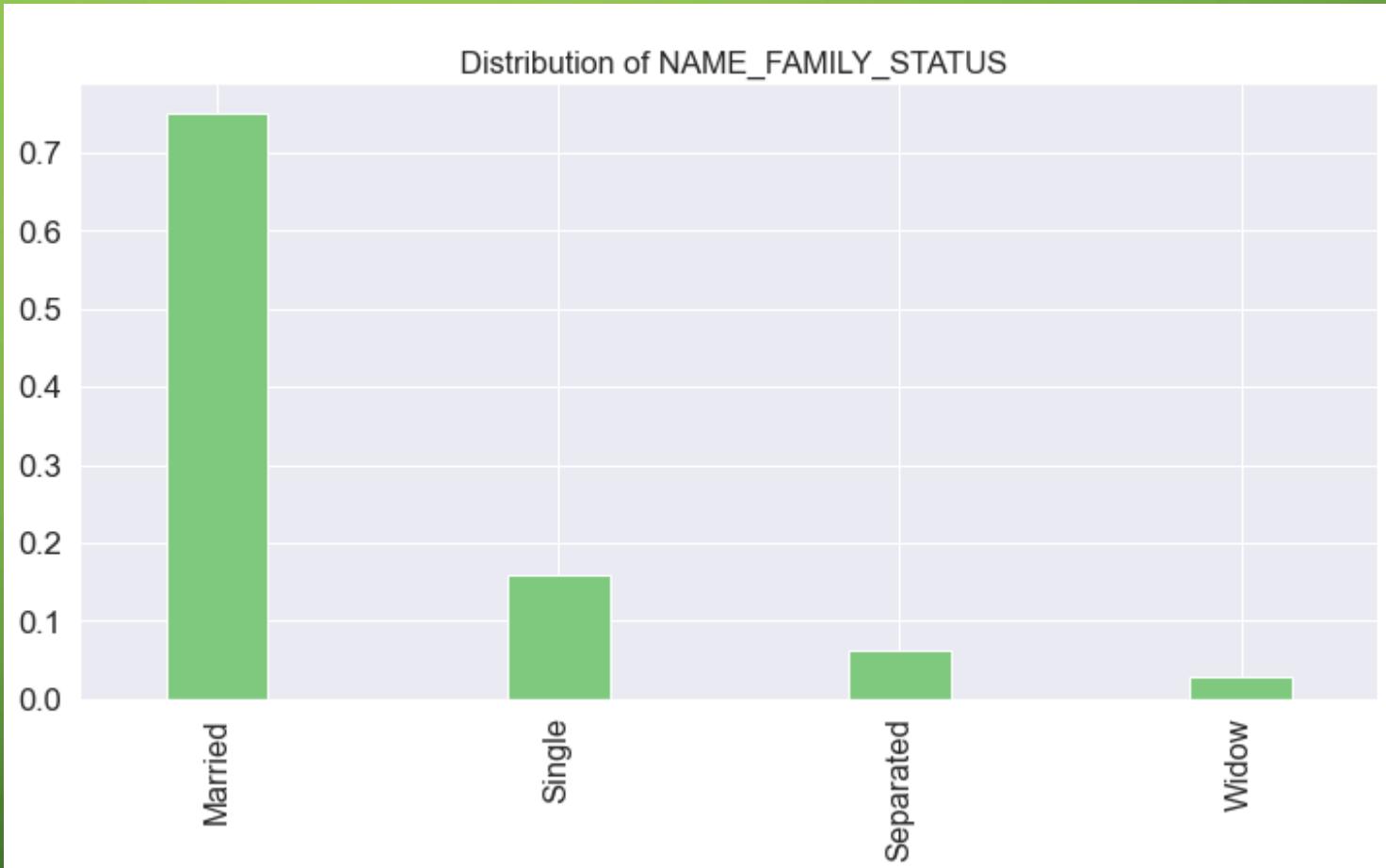
DISTRIBUTION OF OCCUPATION

- Most of the applicants are laborers followed by sales staff and core staff.
- The least categories are IT staff, HR staff and realty agents.



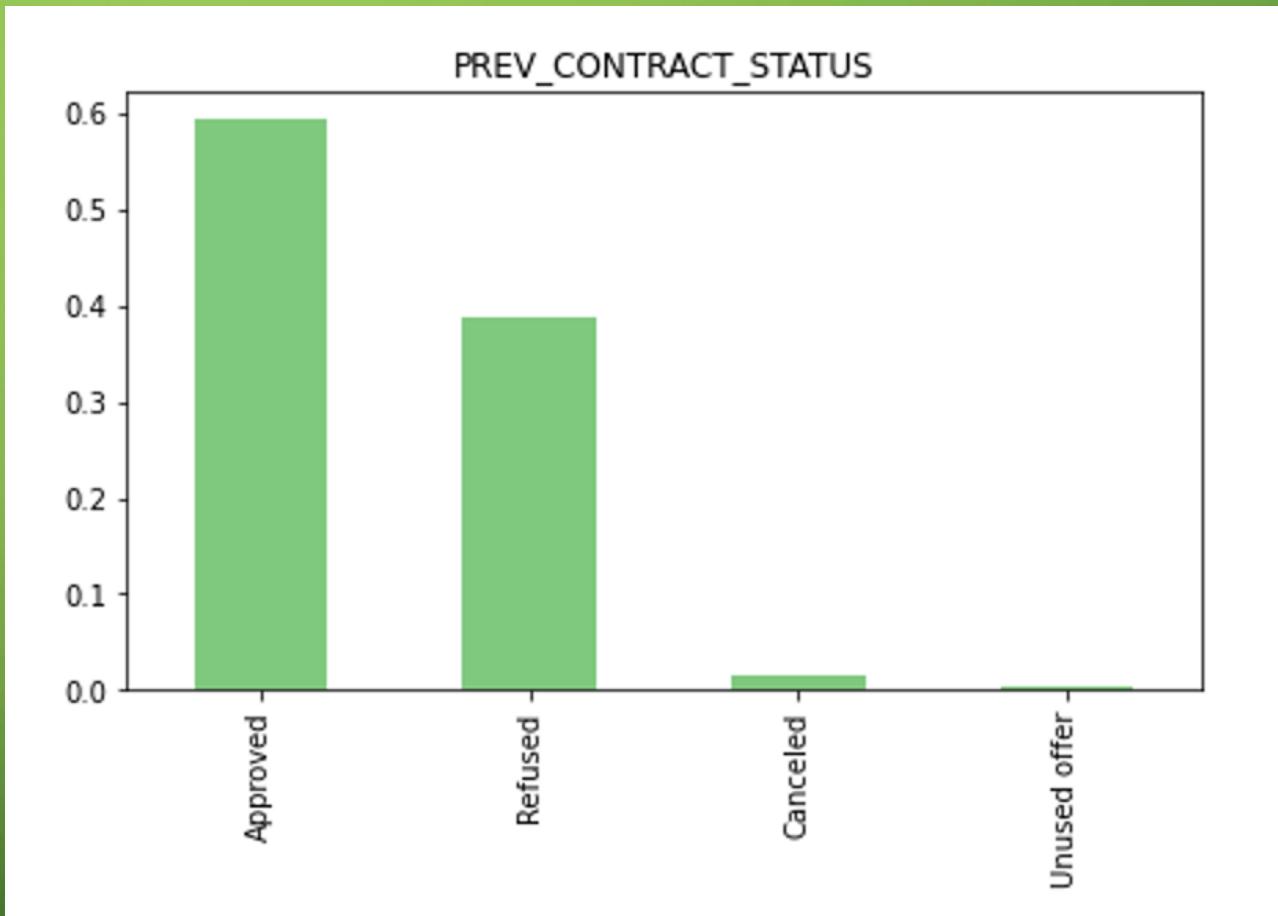
DISTRIBUTION OF FAMILY STATUS

- **Majority of the applicants are married people.**



DISTRIBUTION OF PREVIOUS CONTRACT STATUS

- Around 60% of the previous loan applications have been approved by the bank whereas around 38% has been refused.
- A small portion of the loans have been either cancelled or never used by the customer.



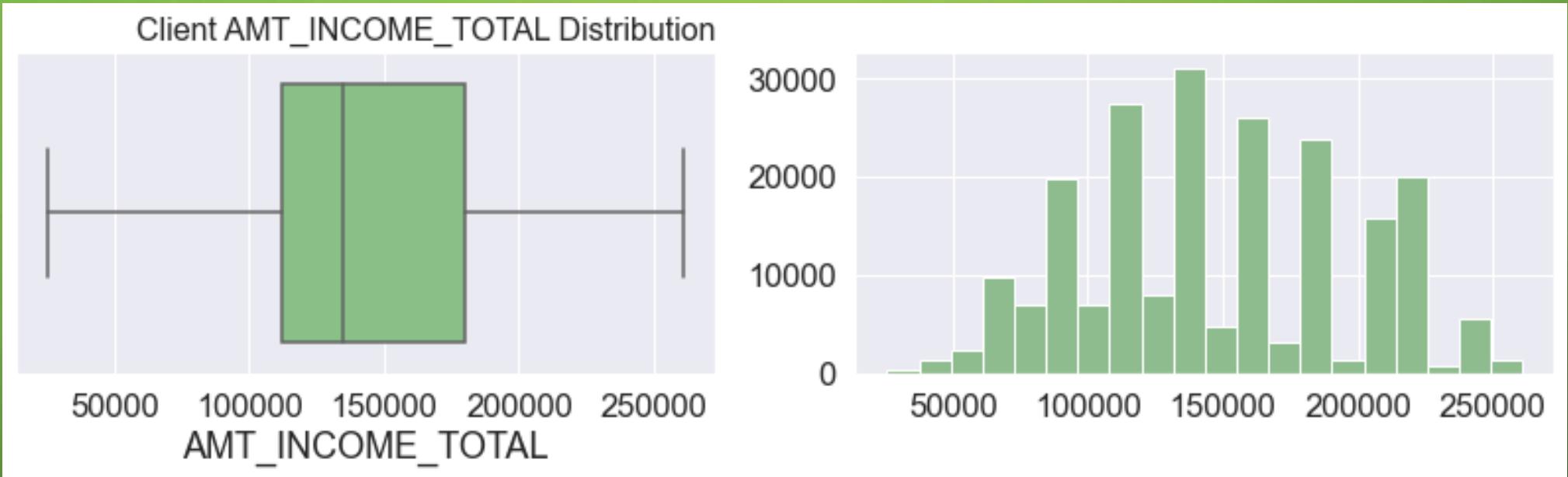
DISTRIBUTION OF AGE

- **Majority of the applicants are 30-40 years old.**



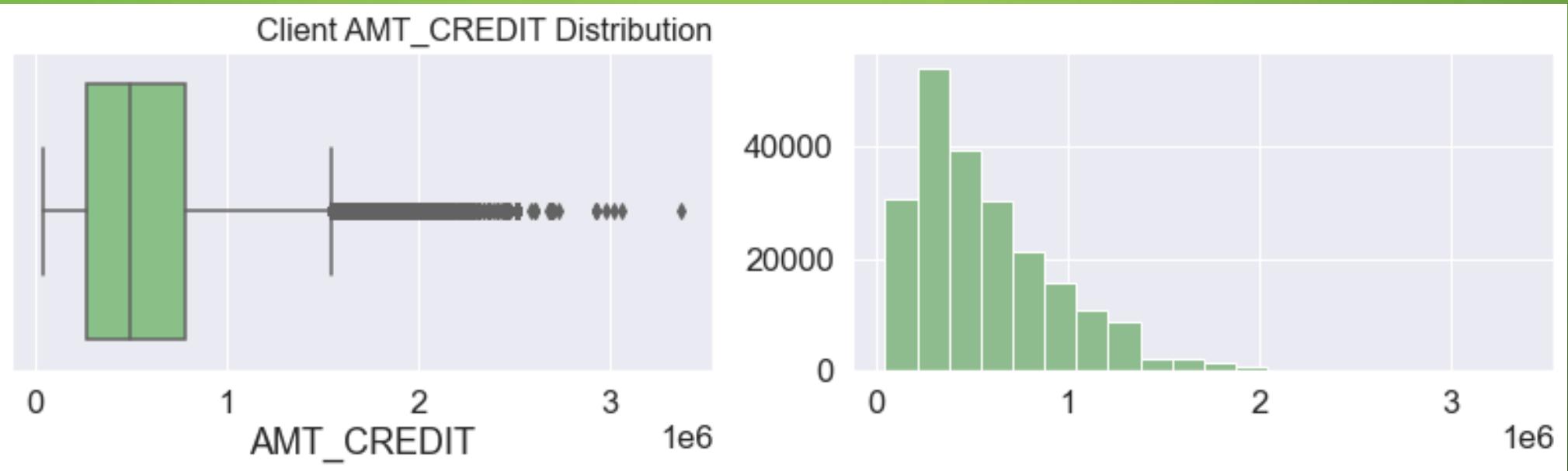
UNIVARIATE ANALYSIS – NUMERICAL VARIABLES

DISTRIBUTION OF INCOME



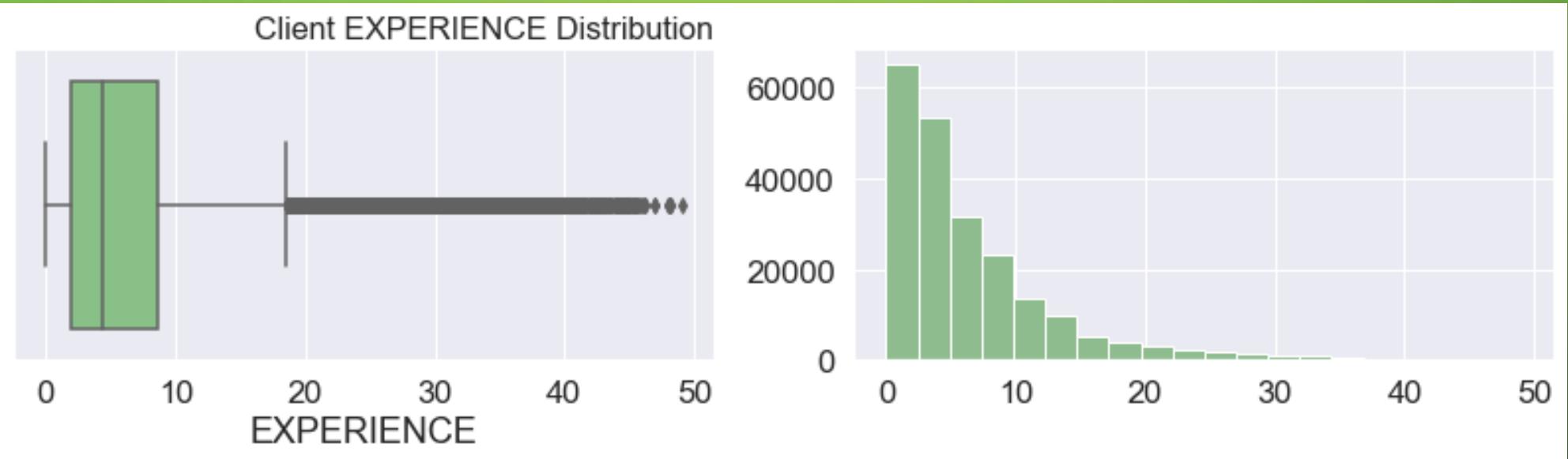
- The mean income of an applicant is around 1.4 Lakhs.
- However, most of the applicants are having a higher income.

DISTRIBUTION OF CREDIT



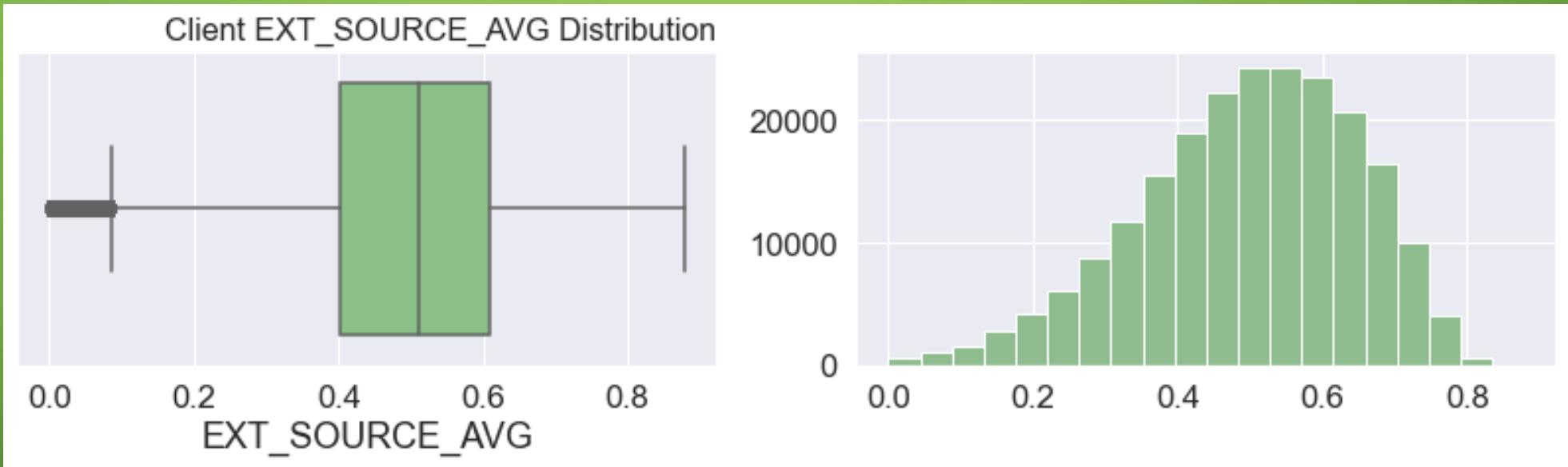
- The average credit for an applicant is around 5 Lakhs.
- However, majority of the applicants have applied for a higher credit.

DISTRIBUTION OF EXPERIENCE



- Majority of the applicants are in the early years of their employment with an average work experience of nearly 6 years.

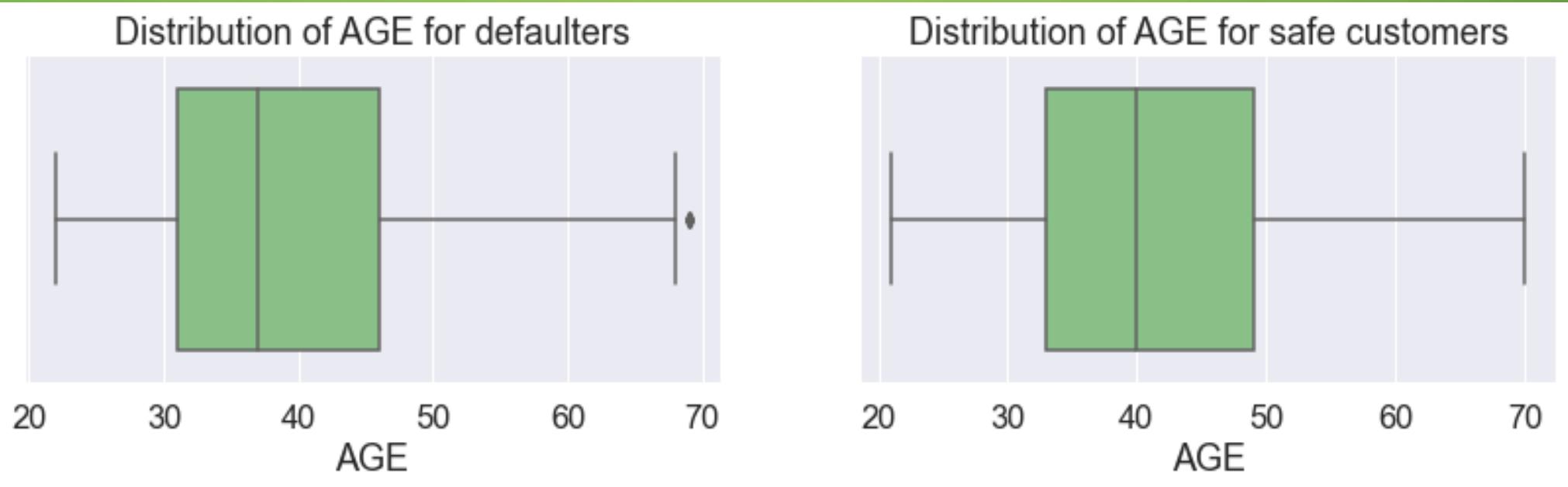
DISTRIBUTION OF EXTERNAL SOURCES AVERAGE



- The **average rating given by external agency is around 0.5**

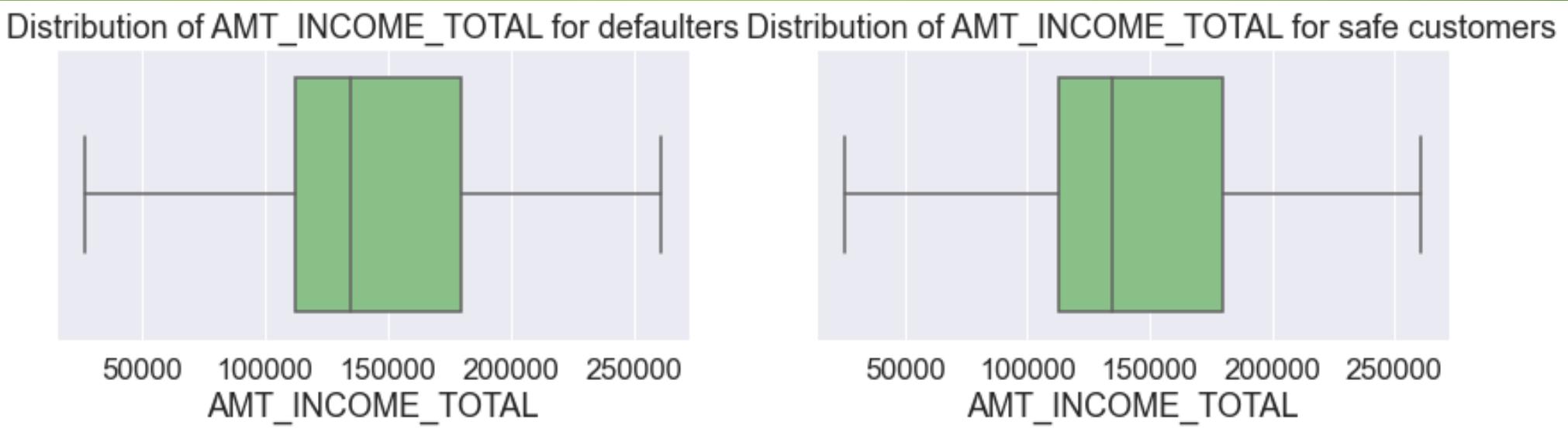
SEGMENTED UNIVARIATE ANALYSIS

DISTRIBUTION OF AGE - TARGET 0 VS TARGET 1



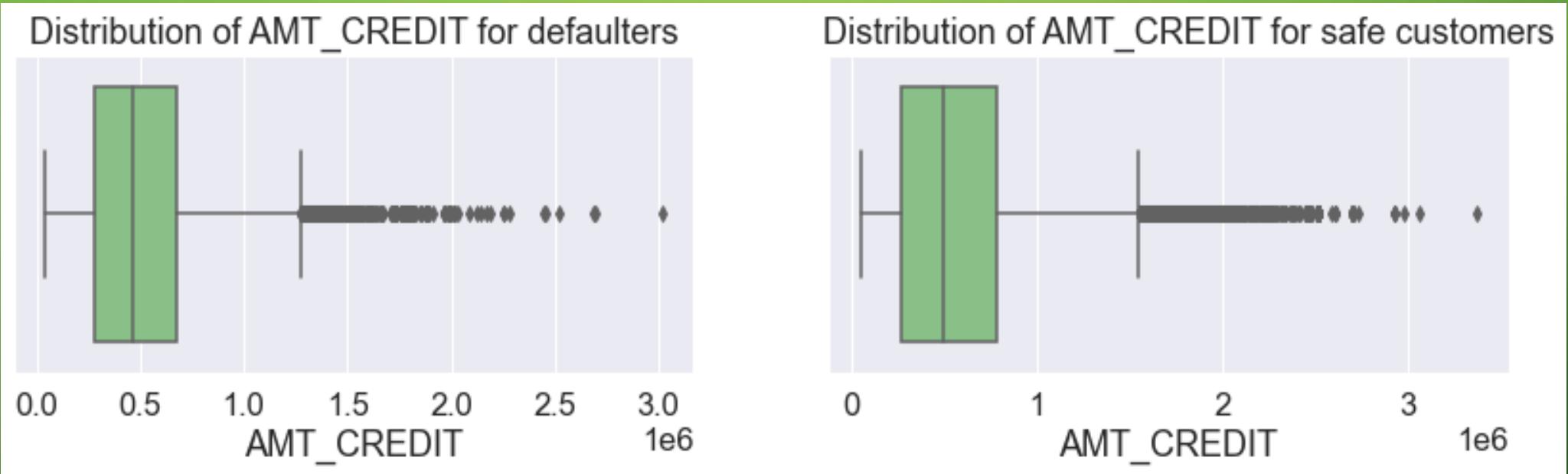
- Average age of default customers are around 38 years whereas the average age of safe customers are around 40, implies that younger applicants tend to default more.

DISTRIBUTION OF INCOME - TARGET 0 VS TARGET 1



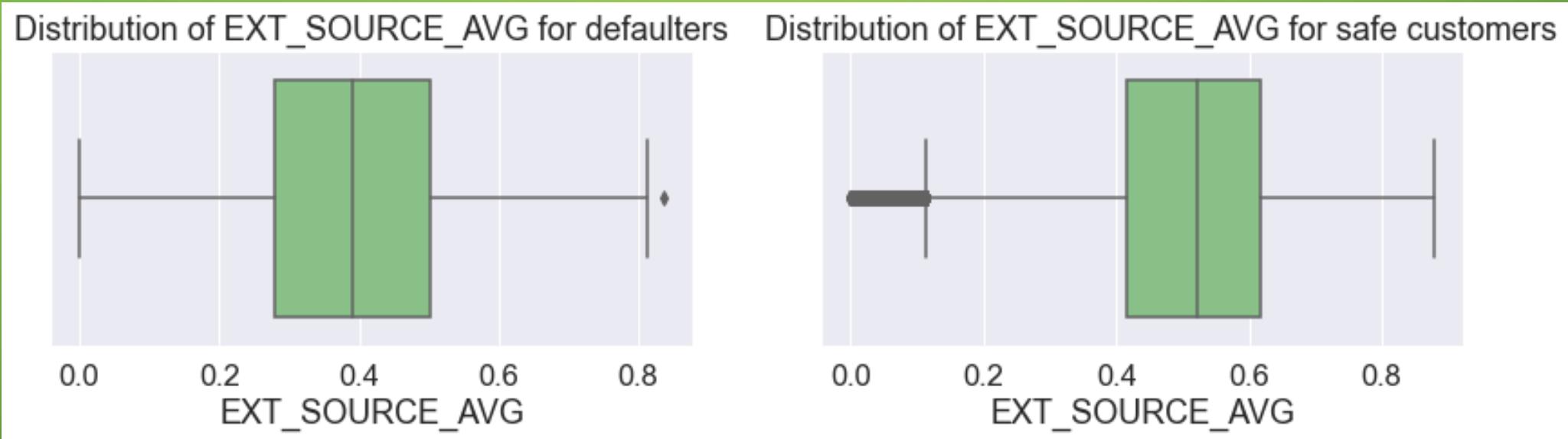
- The average income and its distribution for both the type of customers are almost the same.

DISTRIBUTION OF CREDIT - TARGET 0 VS TARGET 1



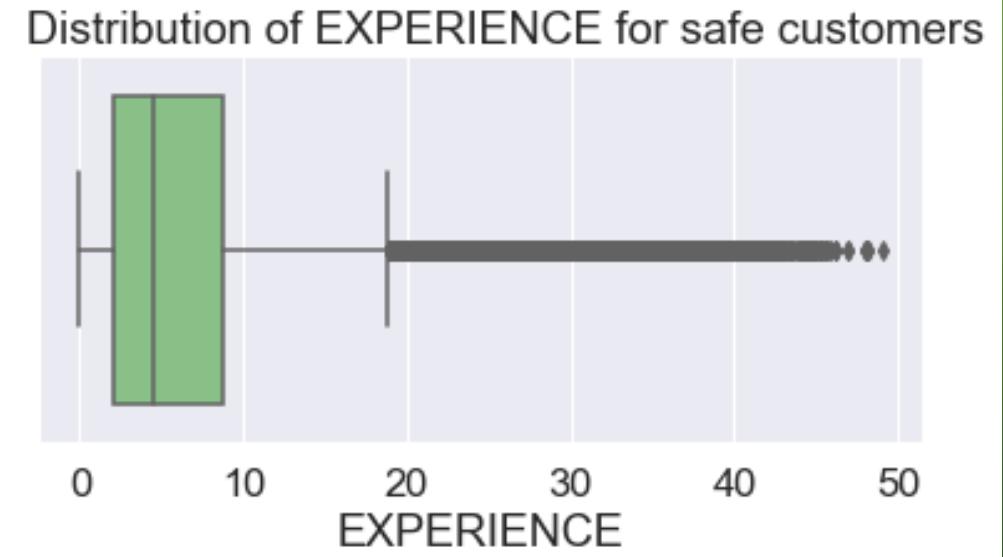
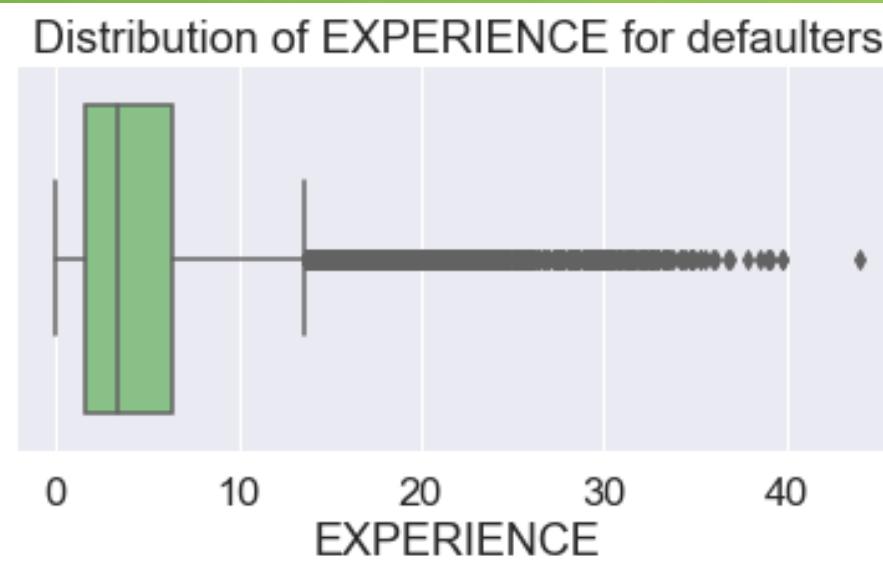
- The mean credit amount for both defaulters and safe customers are same however the safe customers have taken a slightly larger credit.

DISTRIBUTION OF EXTERNAL SOURCES AVERAGE - TARGET 0 VS TARGET 1



- The mean EXT_SOURCE_AVERAGE of defaulters are slightly less than 0.4 whereas for safe customers the rating is higher, slightly more than 0.5
- This shows that good external rating is a crucial key for banks while approving loans.

DISTRIBUTION OF WORK EXPERIENCE - TARGET 0 VS TARGET 1



- The mean work experience of both defaulters and safe customers are same however, the safe customers are slightly more experienced.

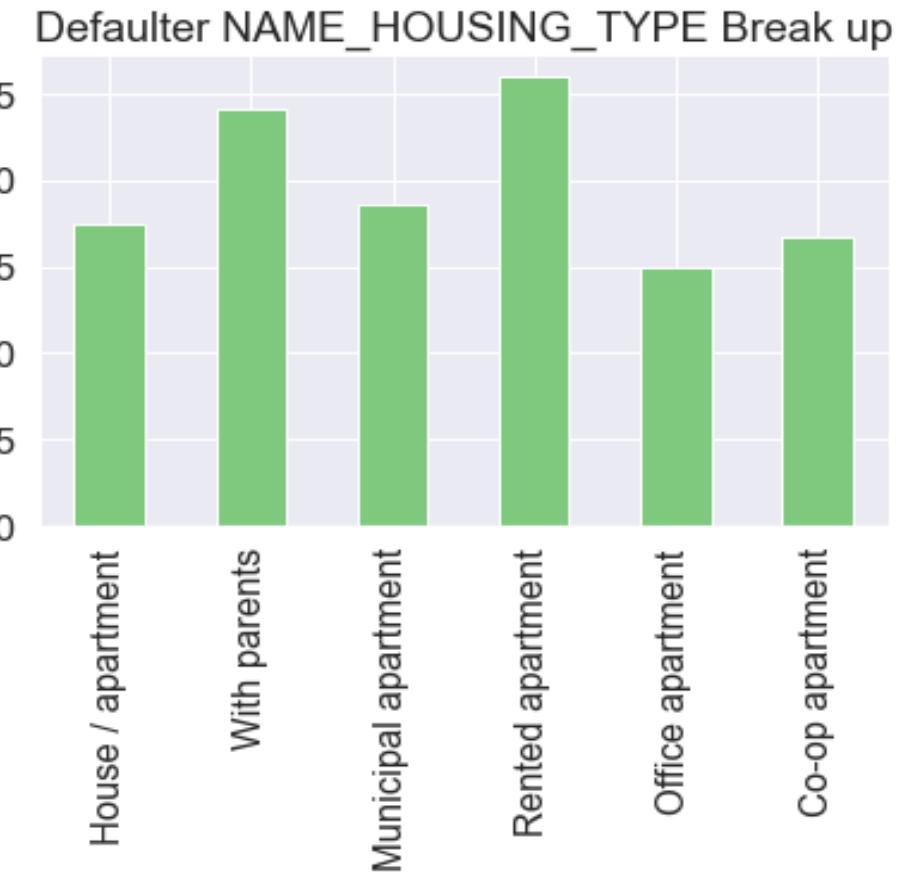
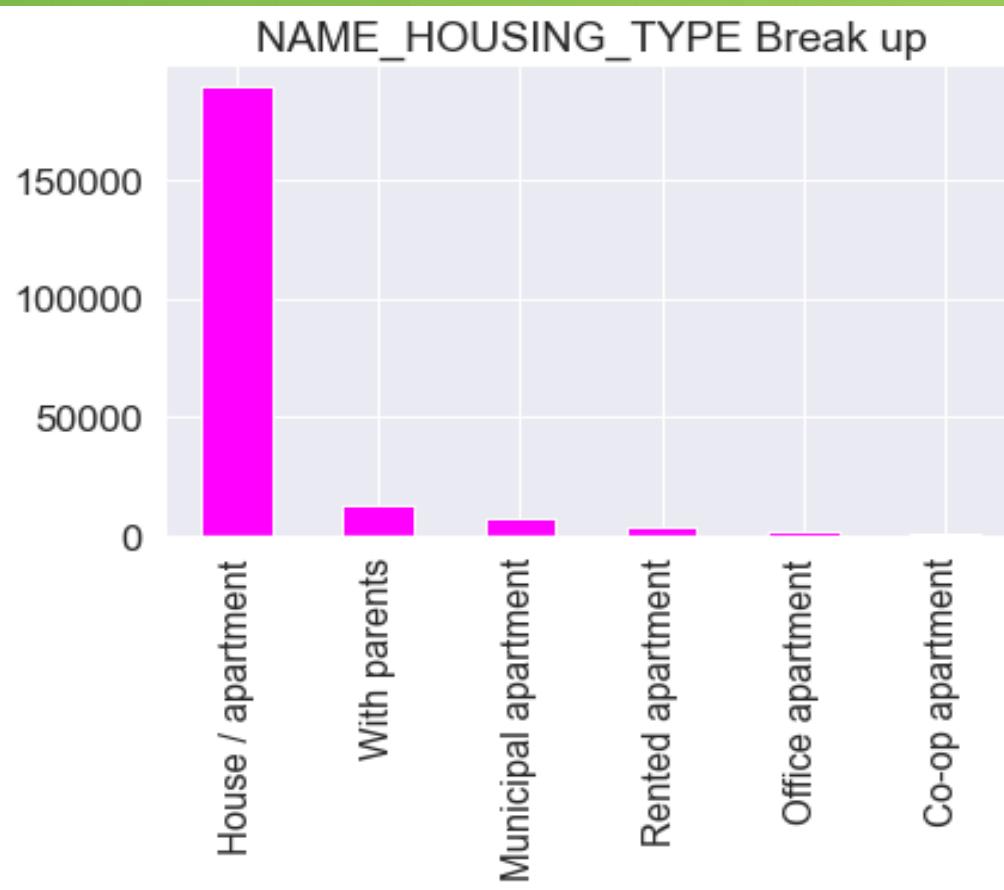
BIVARIATE ANALYSIS WITH TARGET VARIABLE

GENDER VS TARGET 1



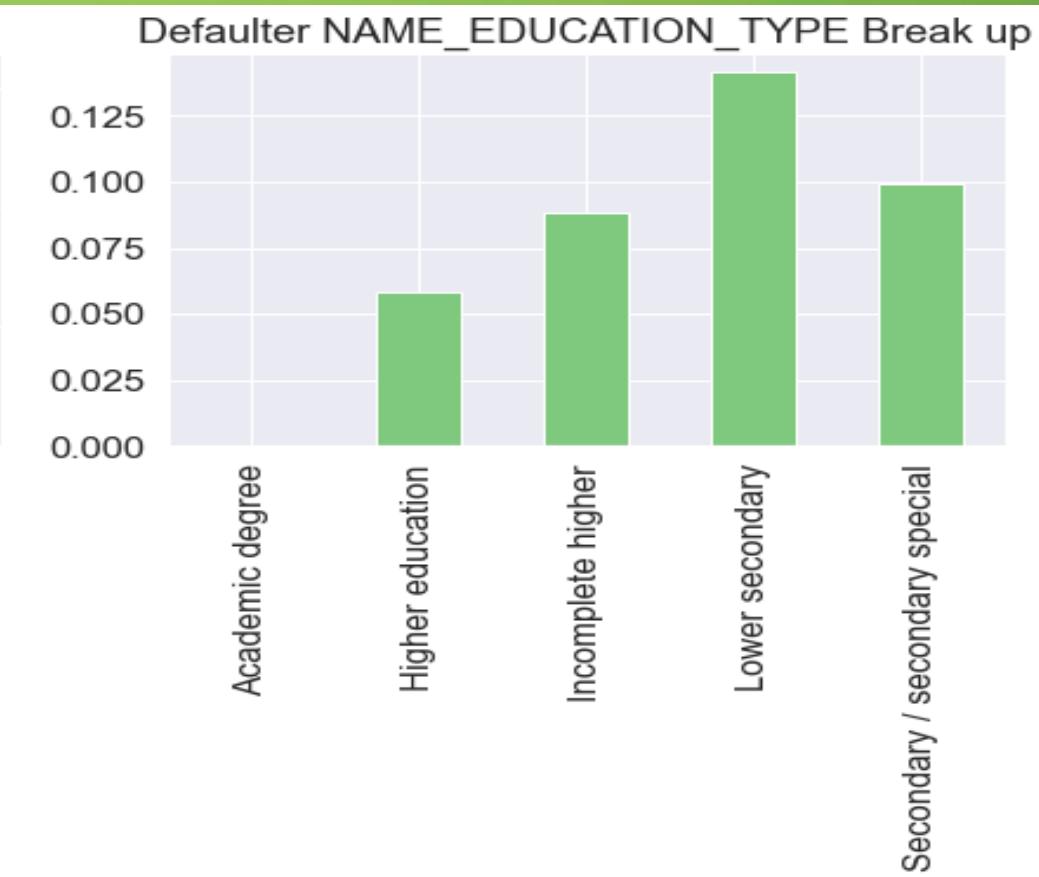
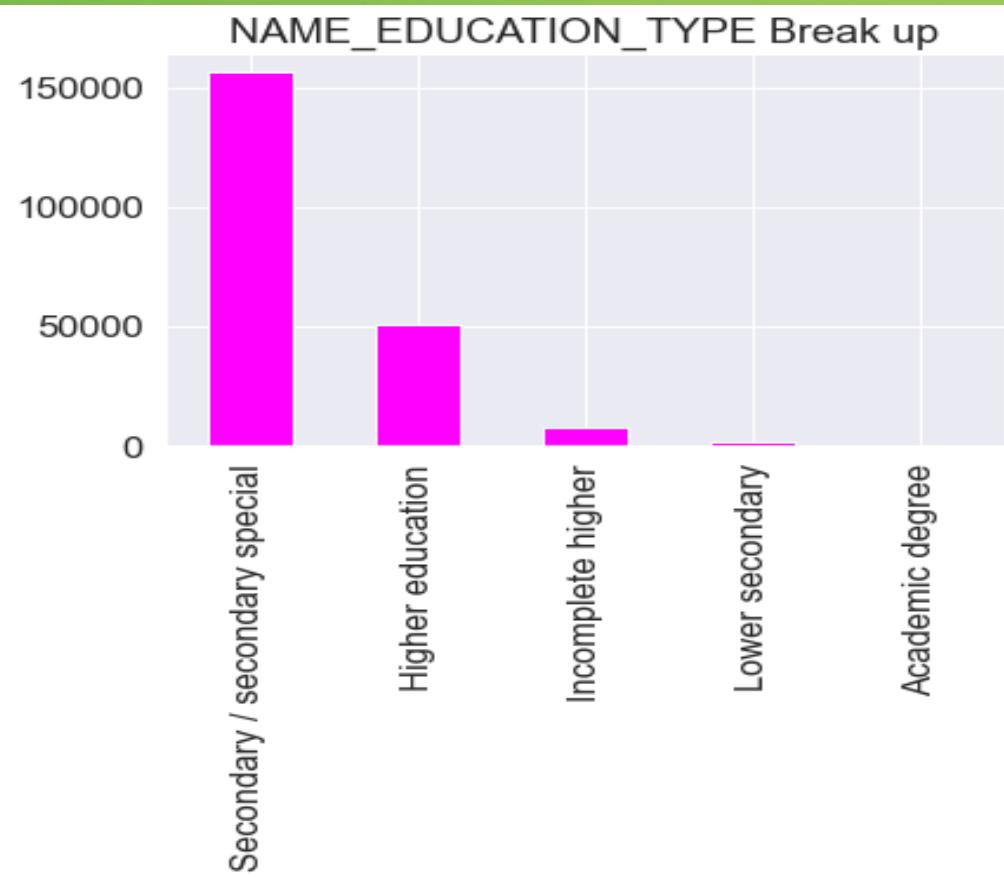
- Males have payment difficulties compared to females almost by ~3%

HOUSING TYPE BREAKUP VS TARGET 1



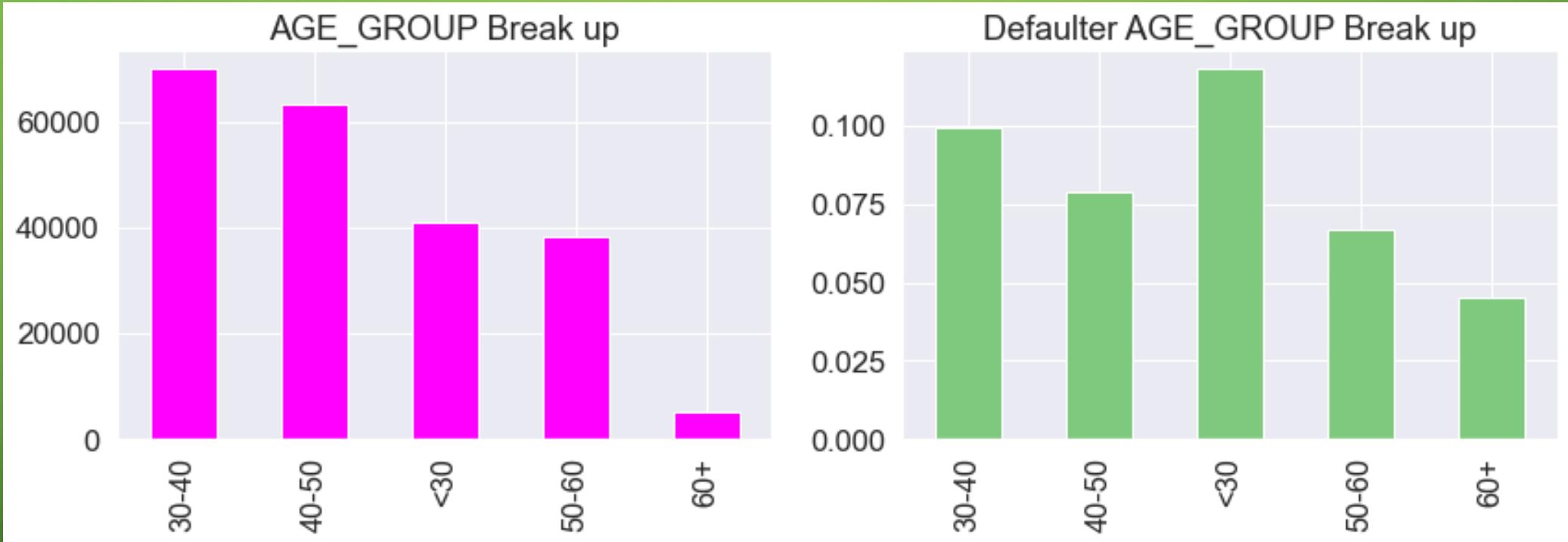
- Majority of loan applicants have own houses, followed by the ones living with parents.
- However, people living with parents, rented houses and in co-op apartments have payment difficulty.

EDUCATION VS TARGET 1



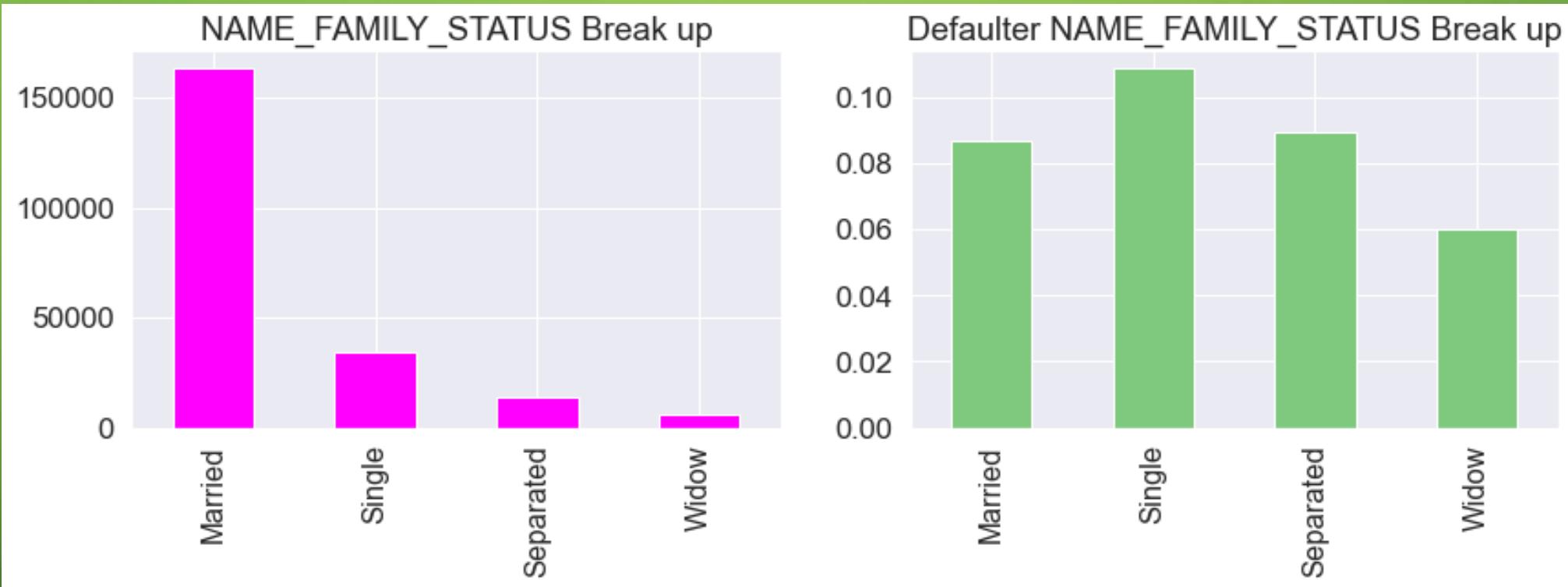
- **Majority of loan applicants are having secondary education followed by people with higher education.**
- **Applicants with lower secondary education has the highest percentage of payment related problems. Hence bank should focus on graduates and people with higher education.**

AGE GROUP VS TARGET 1



- Age group 30-40 are the majority who applied for loans.
- Below 30 years have high default rate.

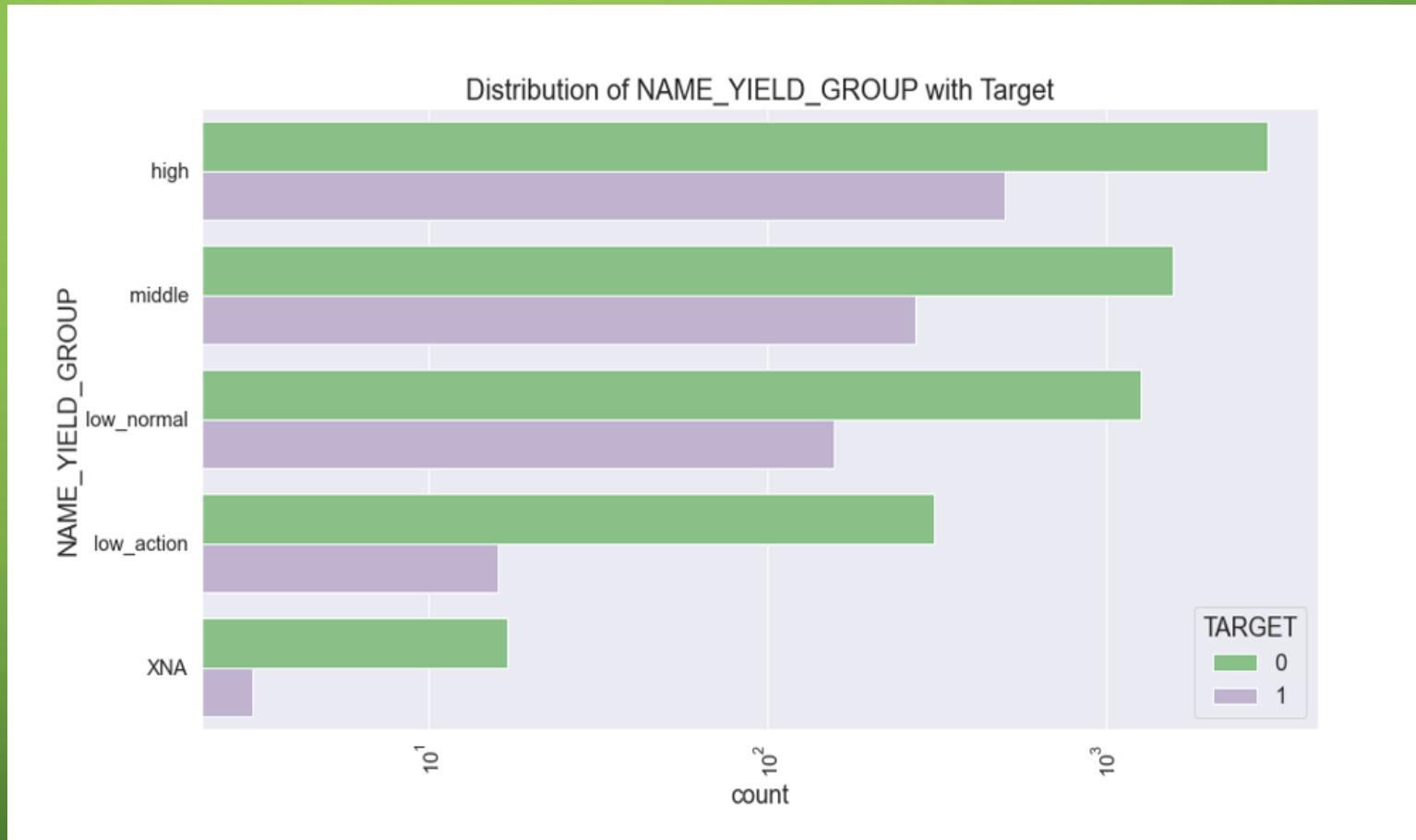
FAMILY STATUS VS TARGET 1



- Majority of the applicants are married people then singles. Whereas widows contributes the least.
- Singles are the major category of defaulters followed by separated people.

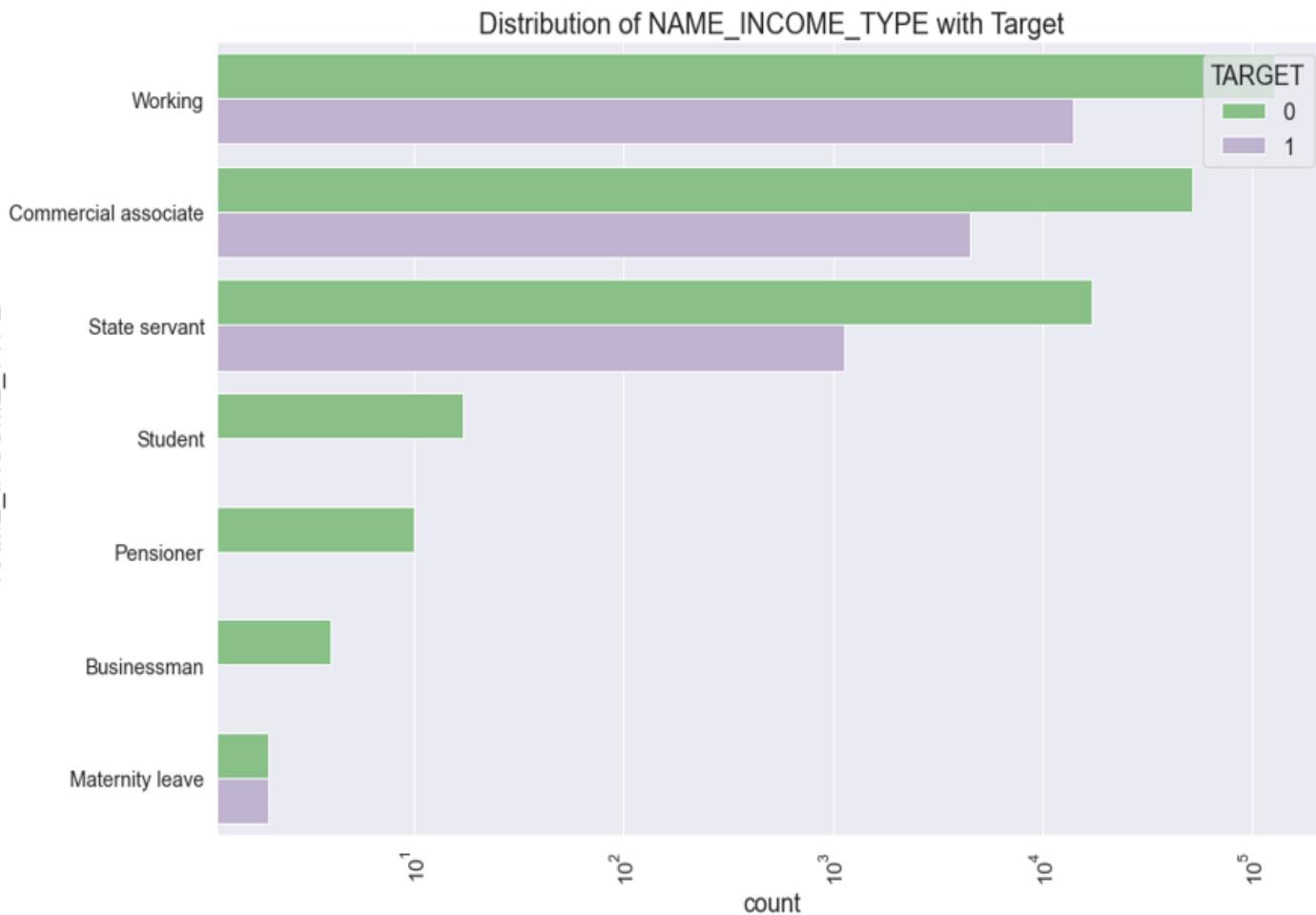
NAME YIELD GROUP VS TARGET 1

- Default rate is comparatively low for low Interest loans.



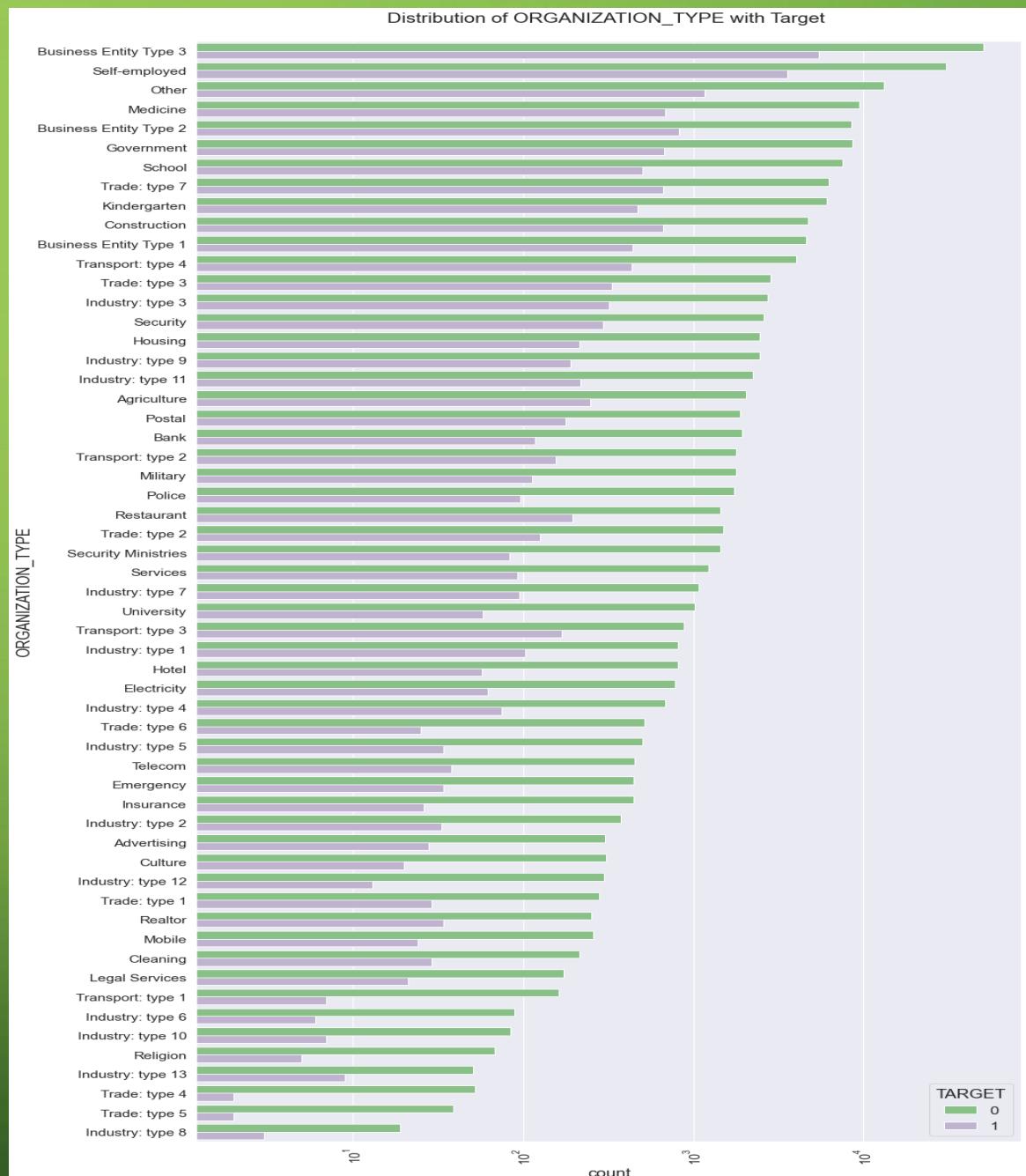
INCOME TYPE VS TARGET 1

Majority of the applicants are working, commercial associate and state servants, however their default rates are also high.



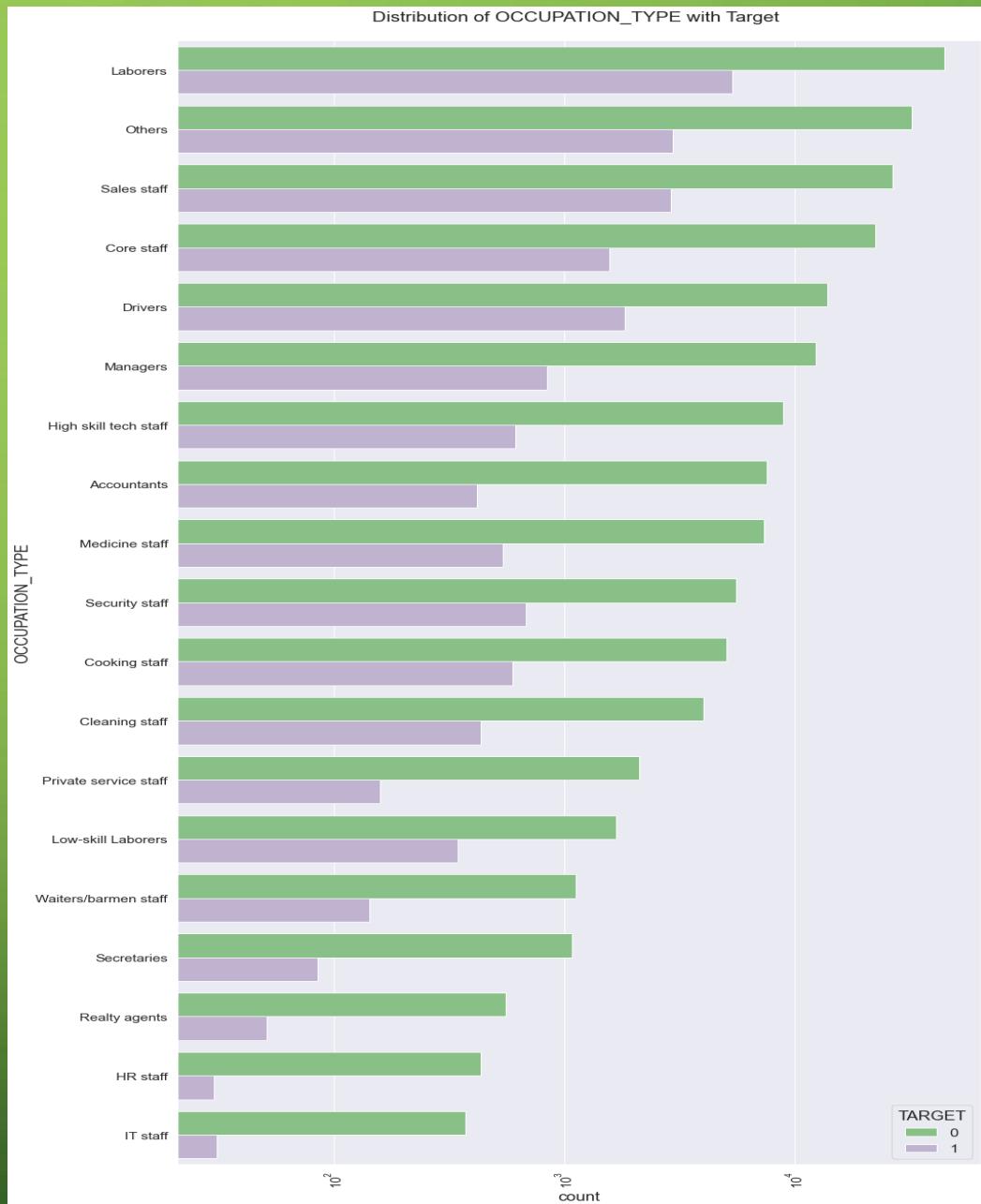
ORGANIZATION TYPE VS TARGET 1

- Most of the applicants are from the organization type 'Business entity Type 3' , 'Self employed' , 'Medicine' and 'Government'.
- Comparatively very less clients are from Industry type 8,type 6, type 10,type 12, Religion and trade type 5, type 4 and type 8.
- But their default rate is quite less, infact these categories are safe customers.



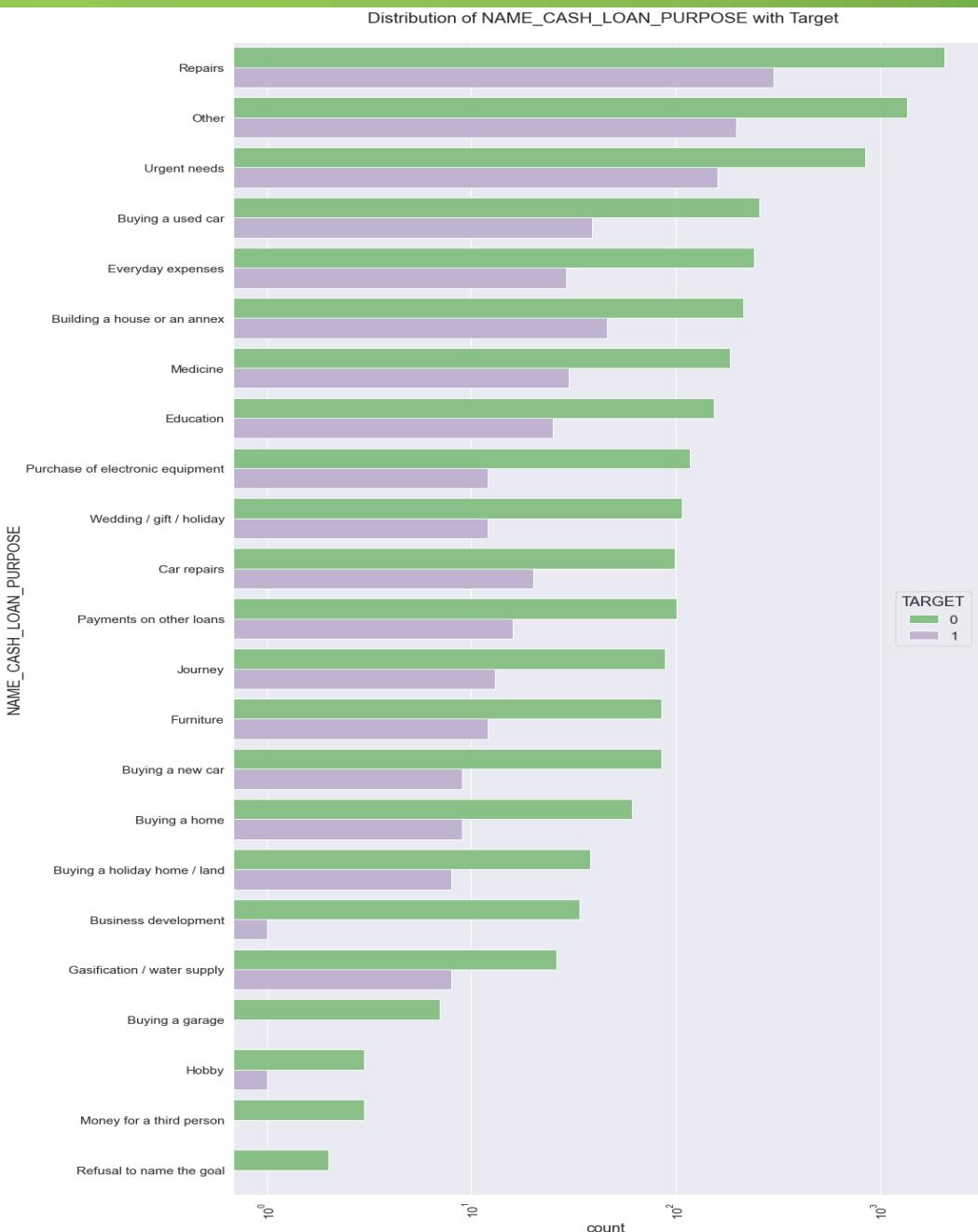
OCCUPATION TYPE VS TARGET 1

- We can see that laborers, Sales staff and core staff are the major category of loan applicants.
- Low-skill laborers, drivers, waiters, cooking staff etc. have a high default rate.

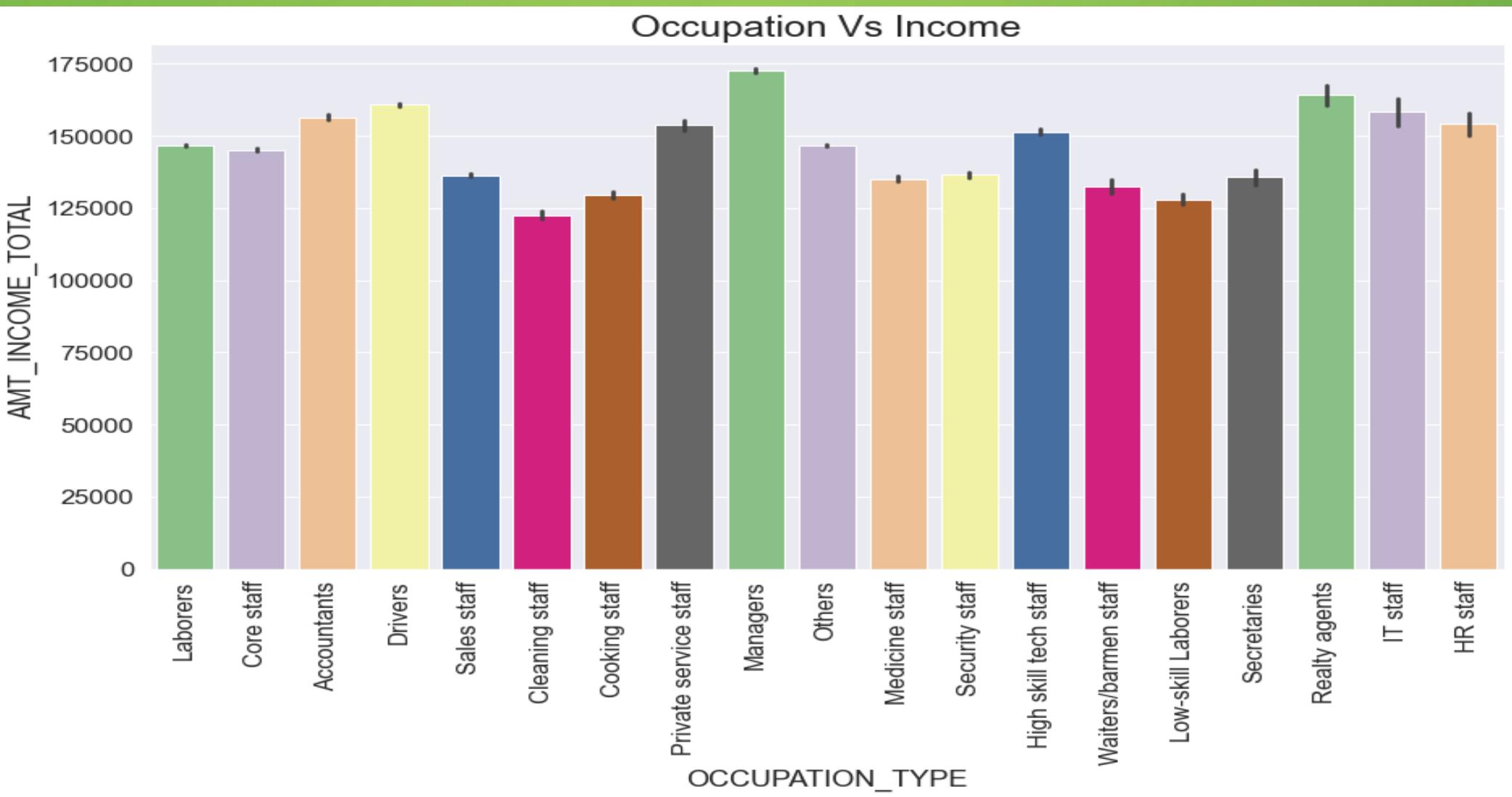


LOAN PURPOSE VS TARGET 1

- The major loan purposes are repairs, urgent needs, buying a used car.
- For business development, buying a new car, purchase of electronic equipment - the default rate is very less.
- Water supply, car repair, urgent needs, other repair have high chance of non-repayment.



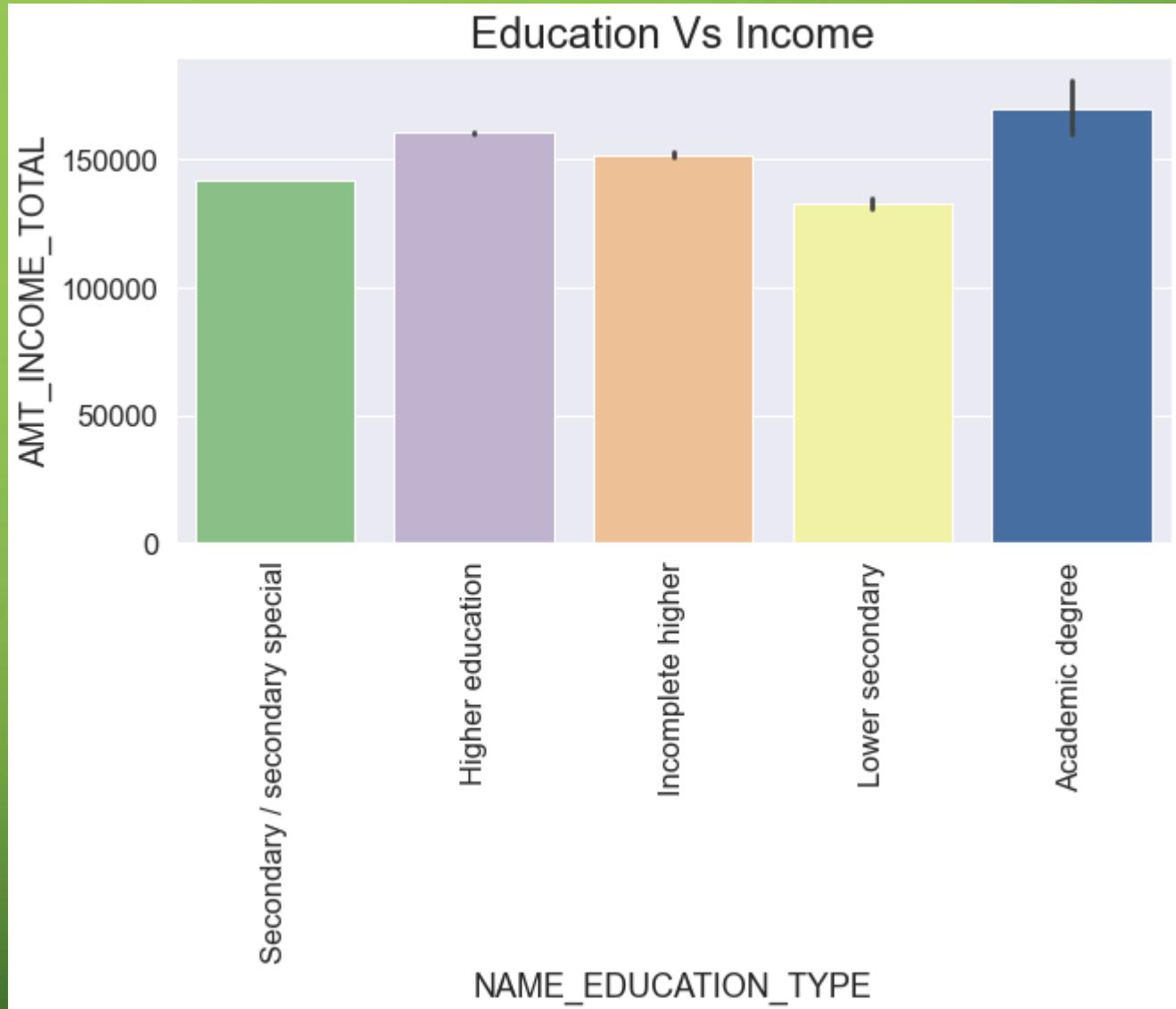
BIVARIATE ANALYSIS - CATEGORICAL VS NUMERIC



- Manager, realty agents and IT staff are at having a high-income level.
- Low skill laborers have a lesser income level and are the riskiest group with highest rate of default.

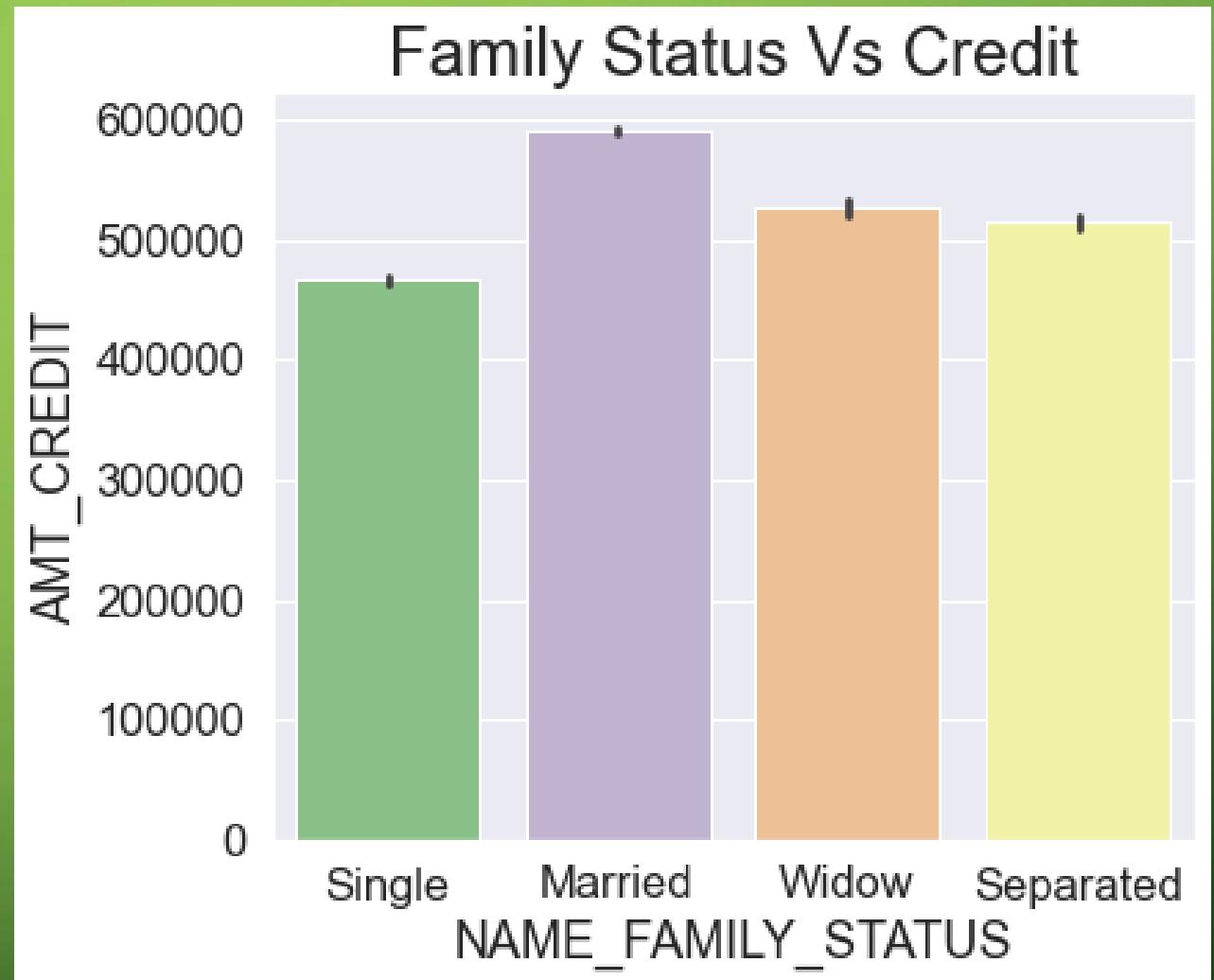
EDUCATION VS INCOME

Graduates are having highest income range whereas lower secondary applicants have a low income.



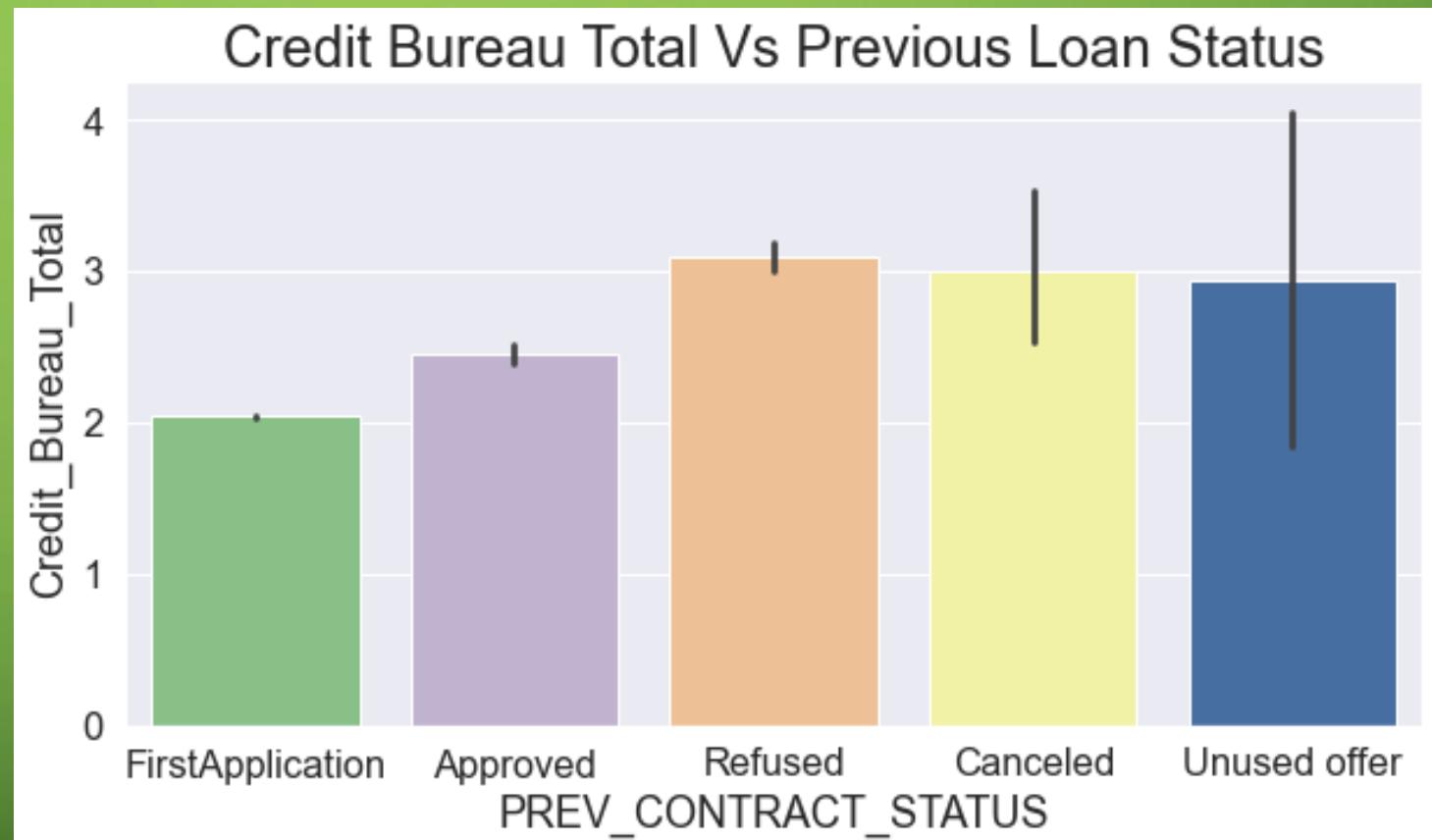
FAMILY STATUS VS CREDIT

It is implied that in general, married people usually go for more credit.



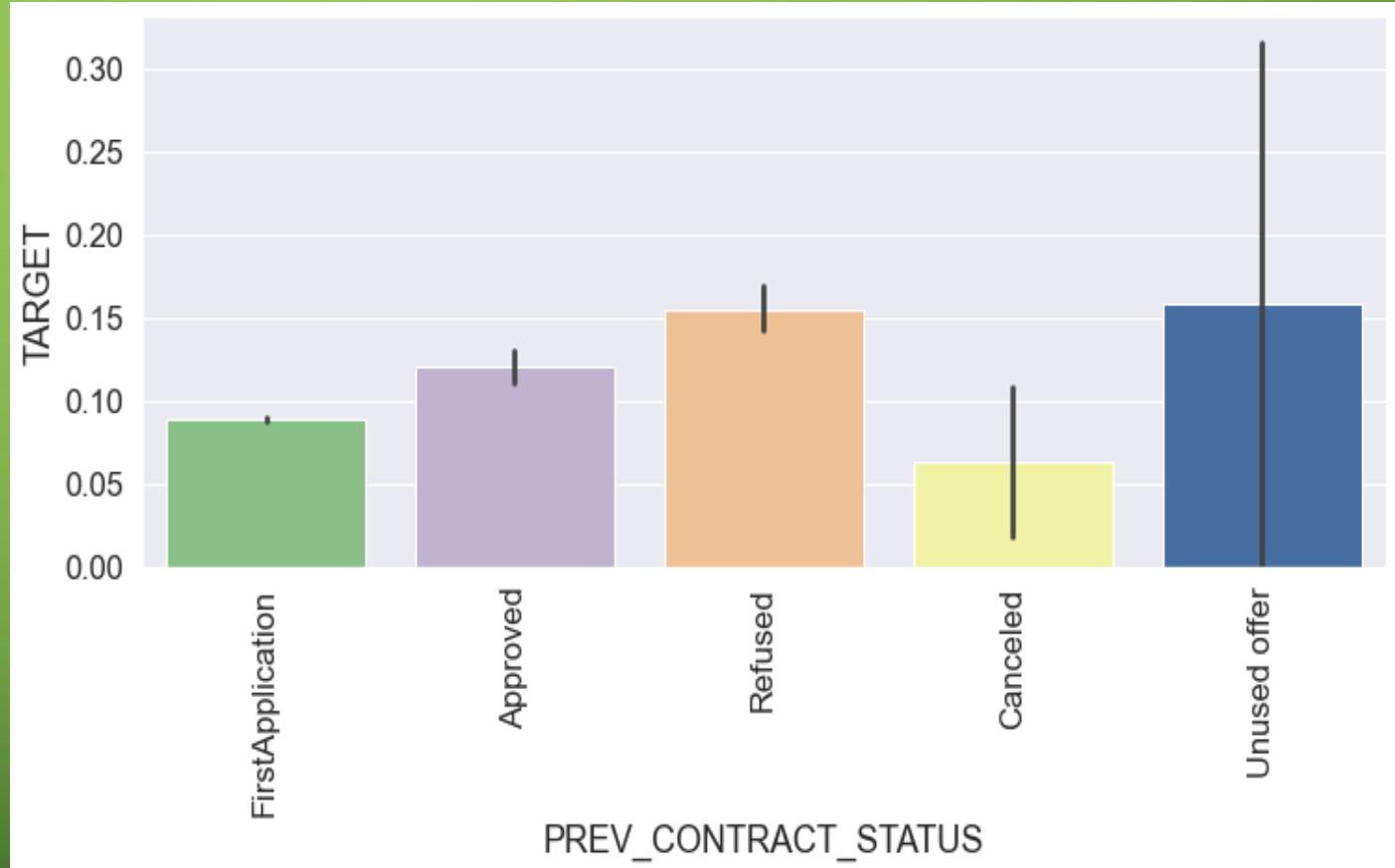
CREDIT BUREAU TOTAL VS PREVIOUS LOAN STATUS

- It is quite evident that people who have refused or unused offer in the past are likely to fall into default.
- First time applicants are more likely to repay.



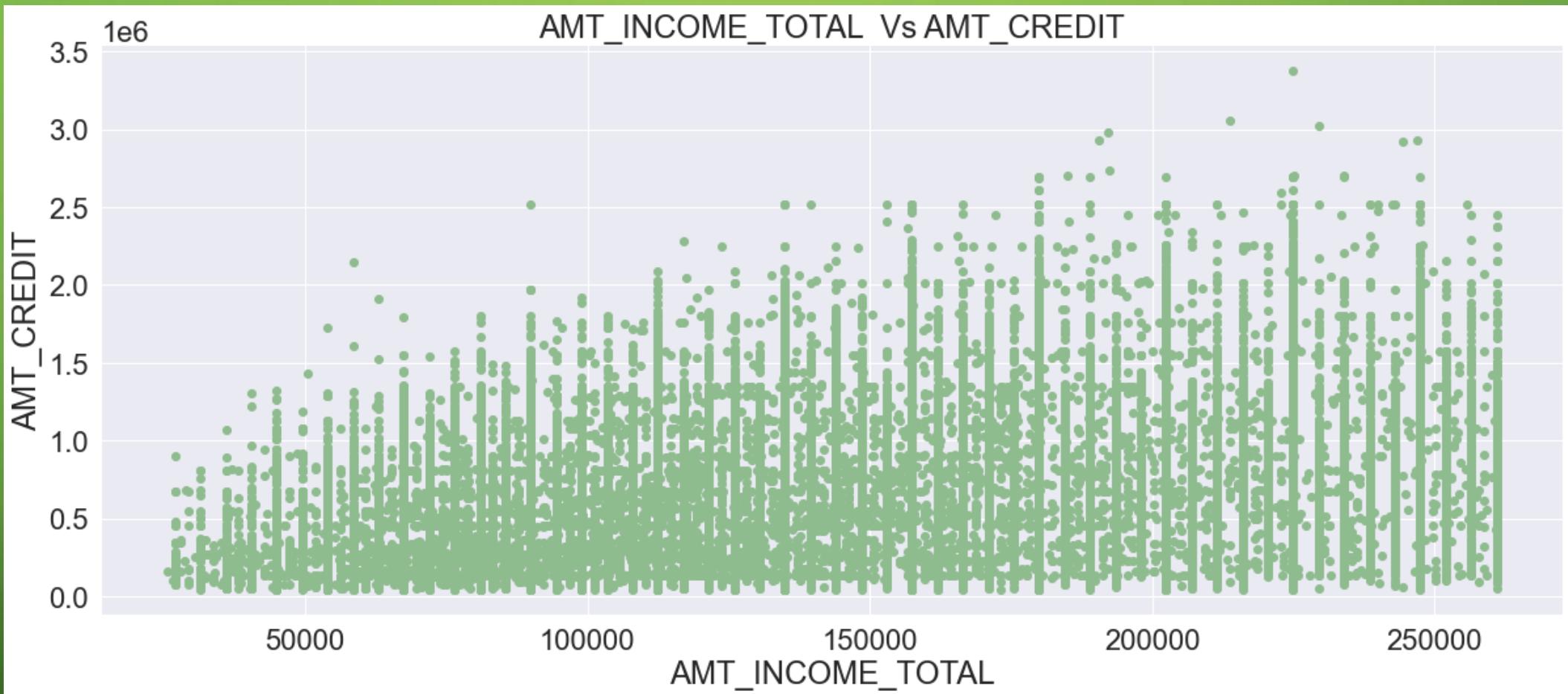
PREVIOUS LOAN STATUS VS TARGET

- It is quite evident that people who have refused or unused offer in the past are likely to fall into default.
- First time applicants are more likely to repay.



BIVARIATE ANALYSIS - NUMERIC VS NUMERIC

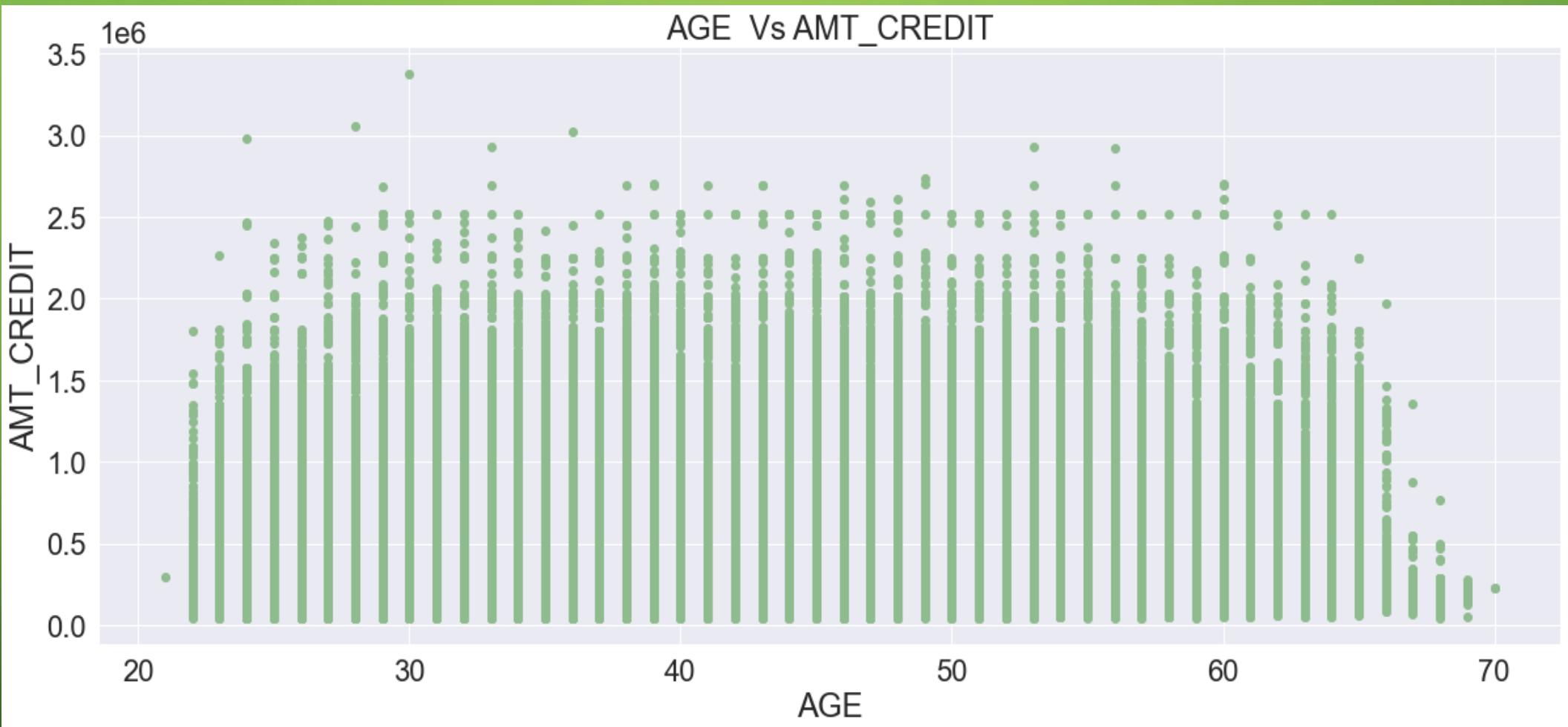
INCOME VS CREDIT



- Can't see any specific pattern but overall, as income increases the bank approves for more credit.

AGE VS CREDIT

AGE Vs AMT_CREDIT



- No specific pattern or correlation in general.

GOODS PRICE VS CREDIT

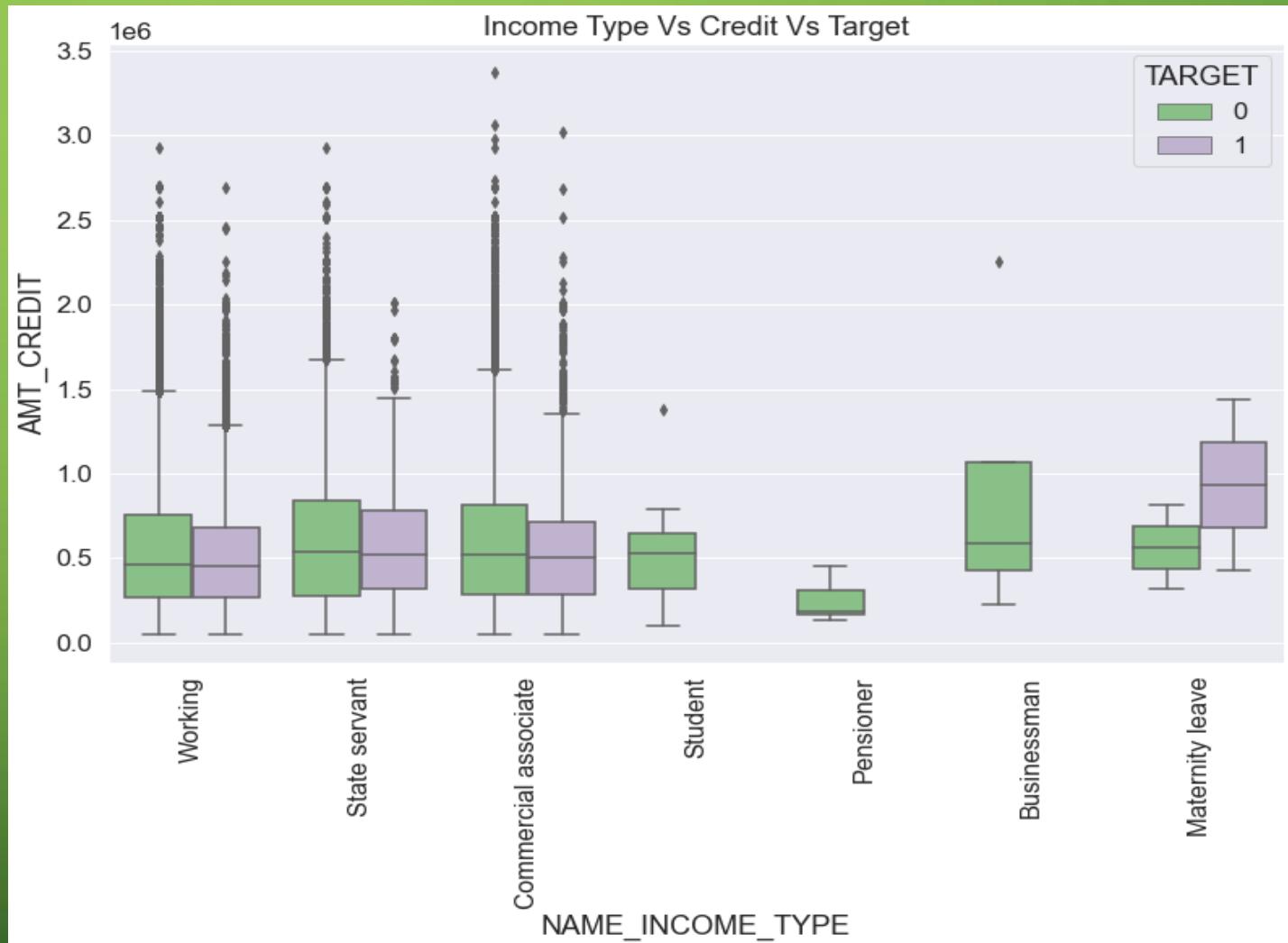


- There is a **strong positive correlation** between credit amount and the goods Price.

MULTIVARIATE ANALYSIS

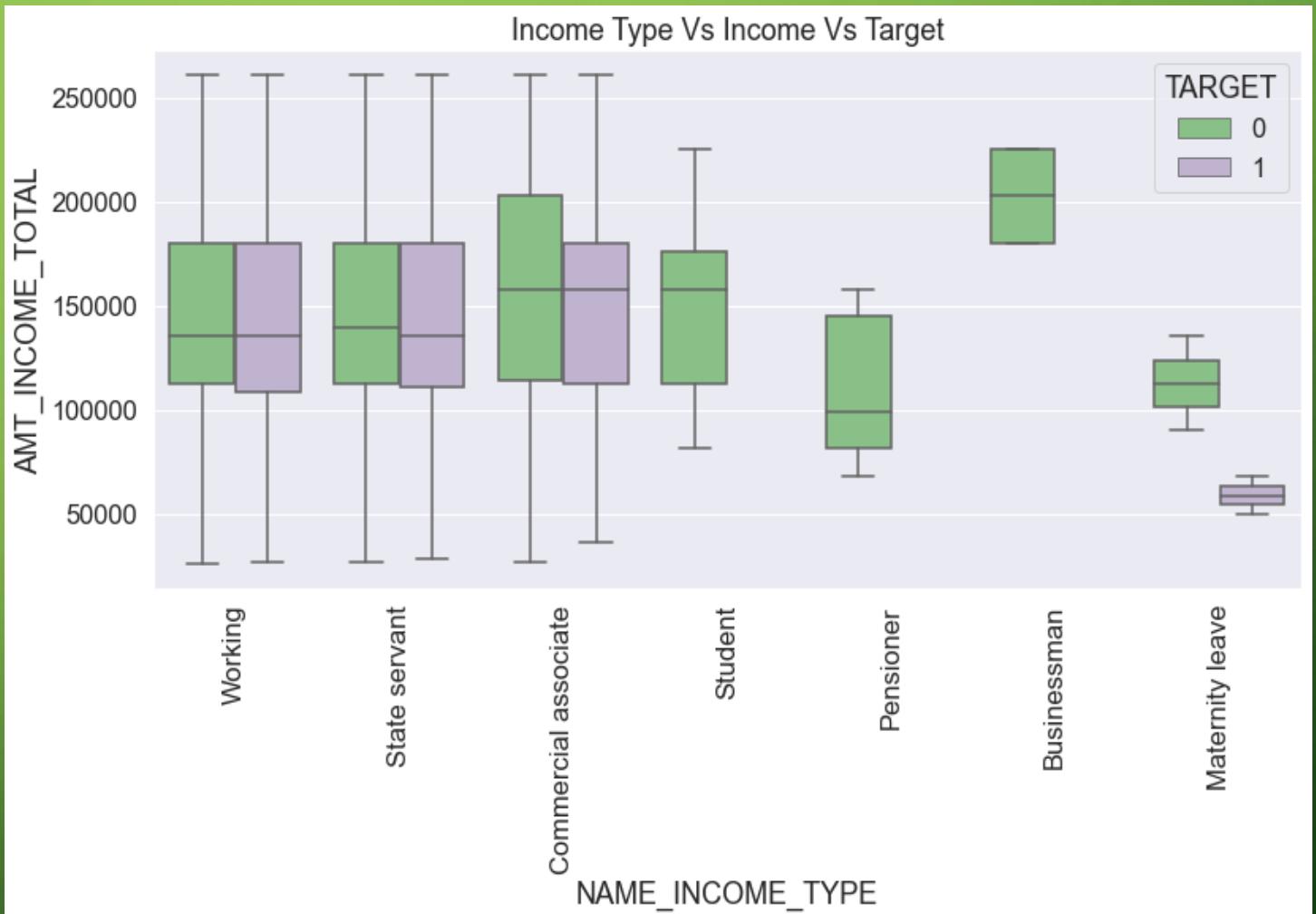
INCOME TYPE VS CREDIT VS TARGET

- The credit for businessmen and pregnant women are quite high.
- Pensioners, students and businessmen are safe customers; however, they are too less in count.
- Pregnant with higher credit tend to default hugely. Hence bank should reconsider their credit limit.
- Working, state servants and commercial associates who have taken lower credits tend to default.



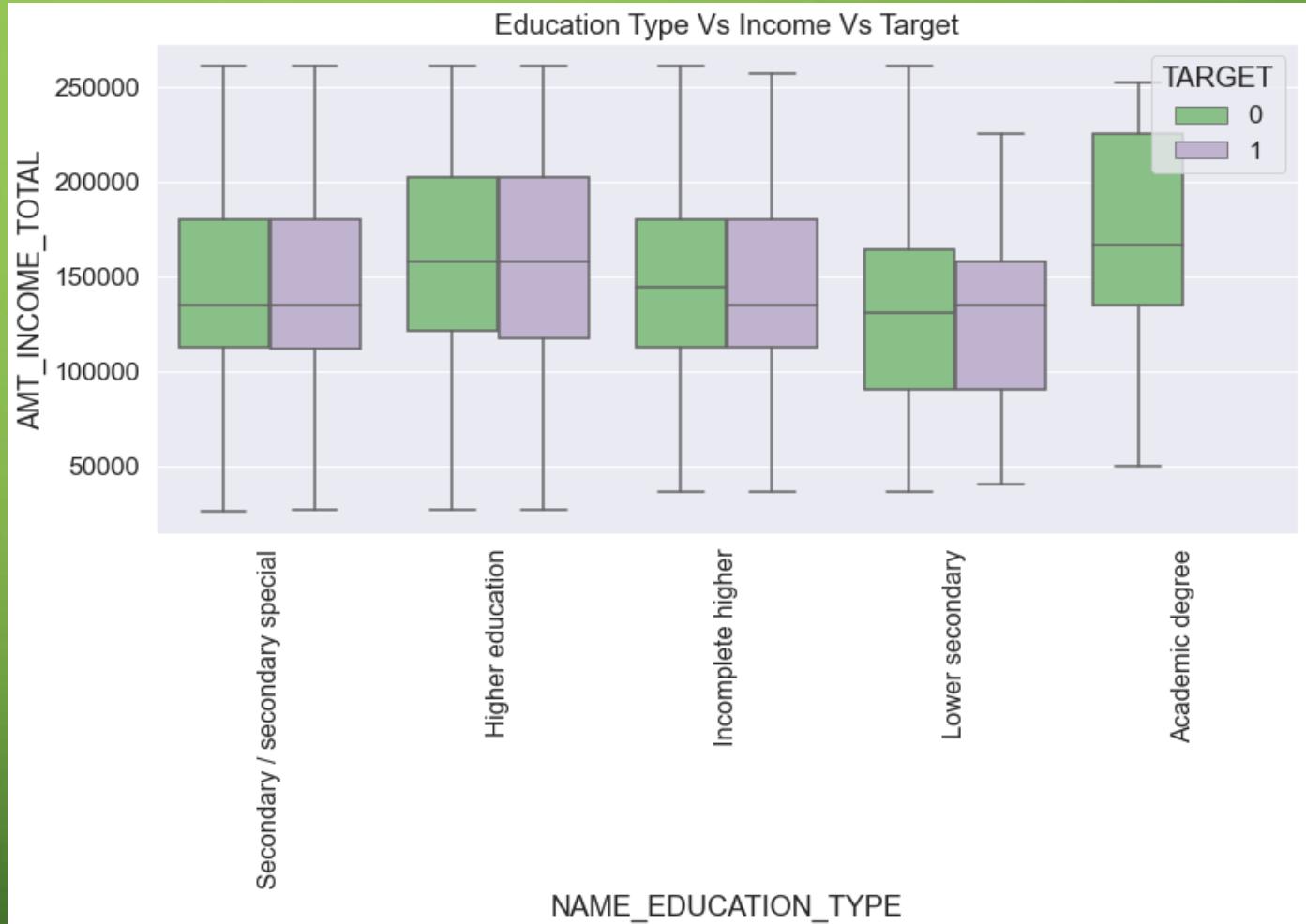
INCOME TYPE VS INCOME VS TARGET

- Commercial associates with lower income tend to default.
- Pensioner's income is minimal, but they are very consistent with repayment.
- Pregnant women with low-income level tend to default very badly.



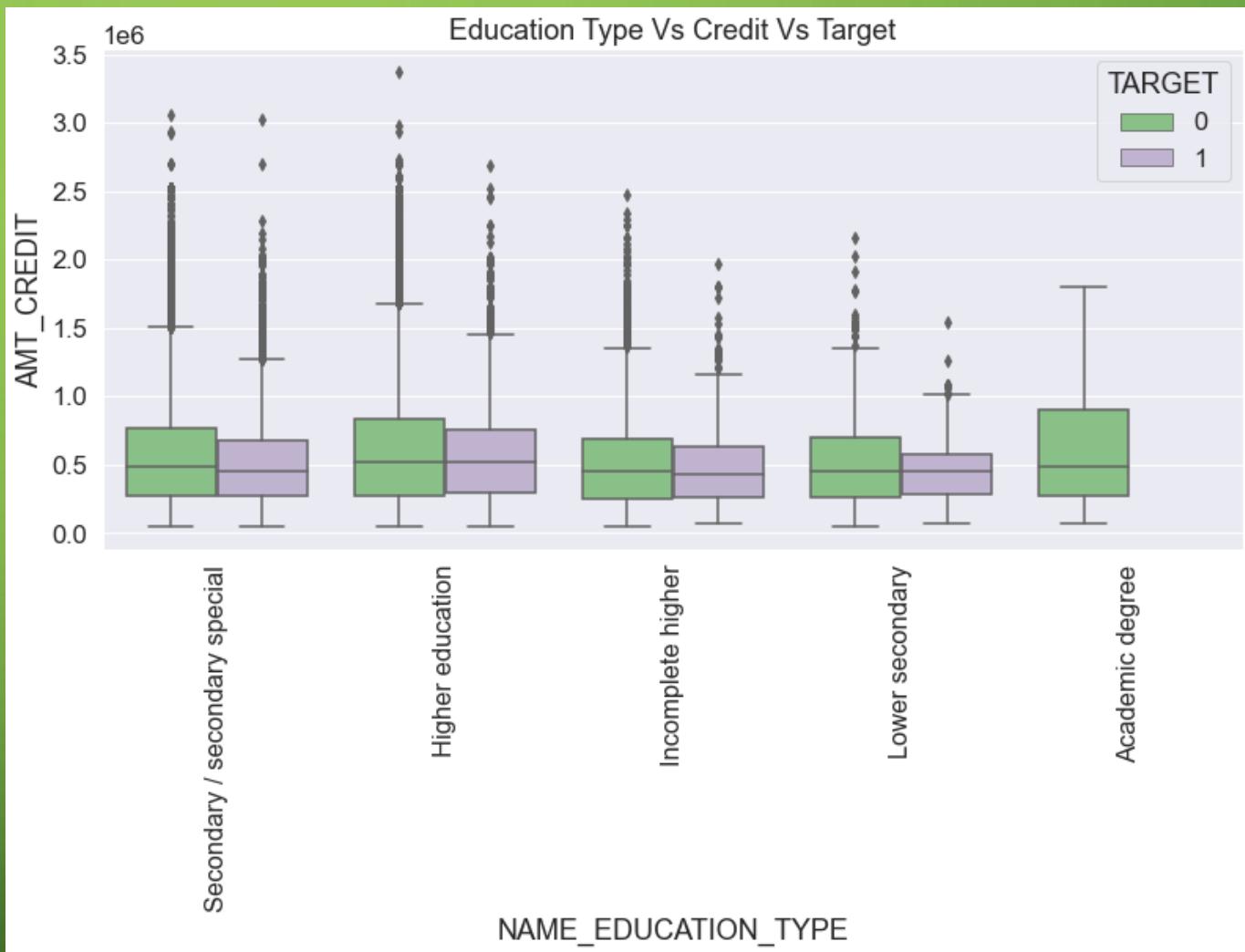
EDUCATION TYPE VS INCOME VS TARGET

- All the graduates are at high income slabs and are safe customers.
- People who have higher education also have a higher income but there is an equal chance to default.
- People having lower secondary education with more income, are rather safe customers.



EDUCATION TYPE VS CREDIT VS TARGET

- **Graduates in general have high income and have taken large credits are very consistent with repayment.**
- **People who have higher education also took a larger credit are tend to repay whereas lower credit tends to default.**



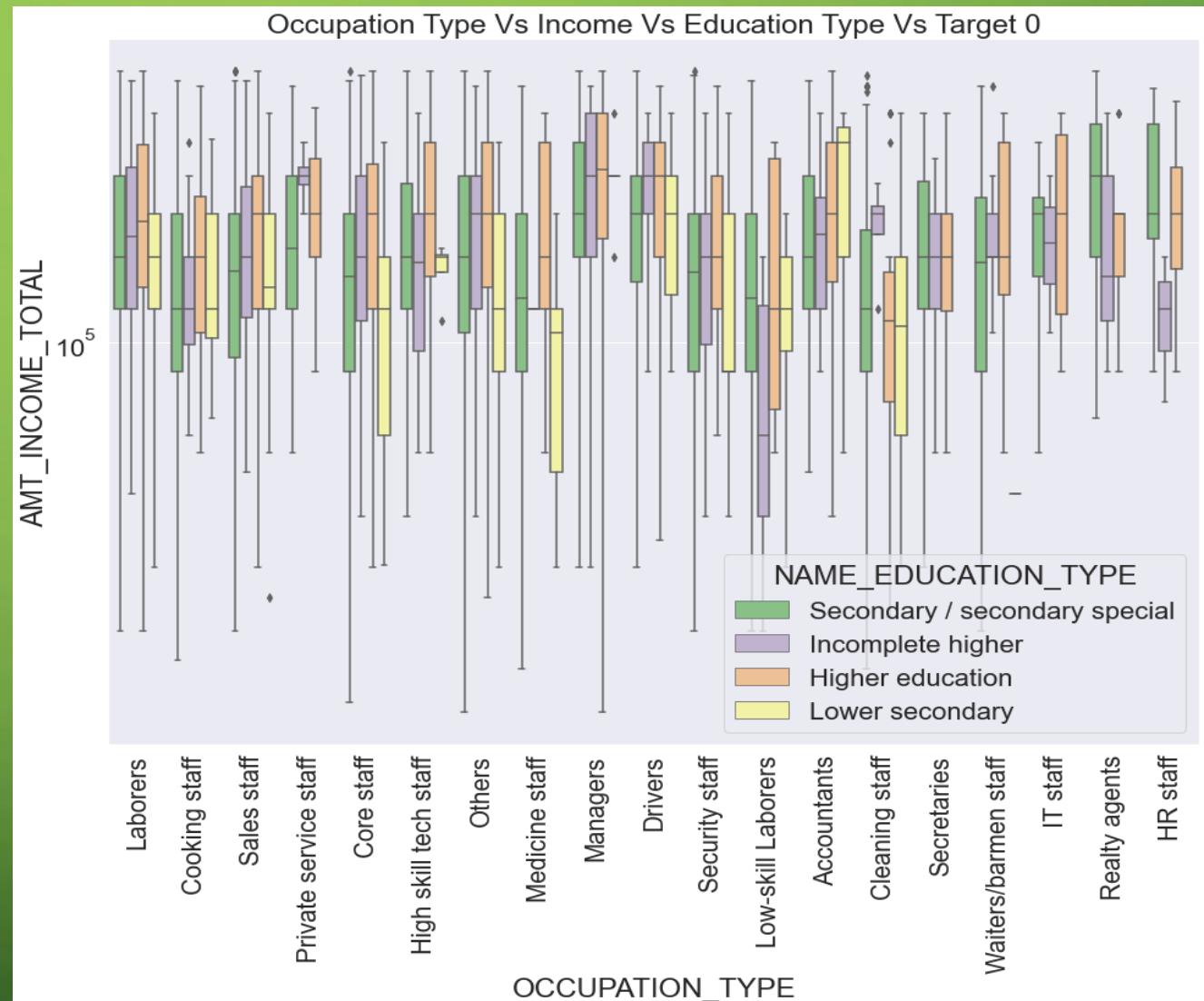
FAMILY STATUS TYPE VS CREDIT VS TARGET

- Singles who go for higher credits do repay in most of the cases.
- Widows and separated people tend to default.



OCCUPATION TYPE VS INCOME VS EDUCATION TYPE VS TARGET 0

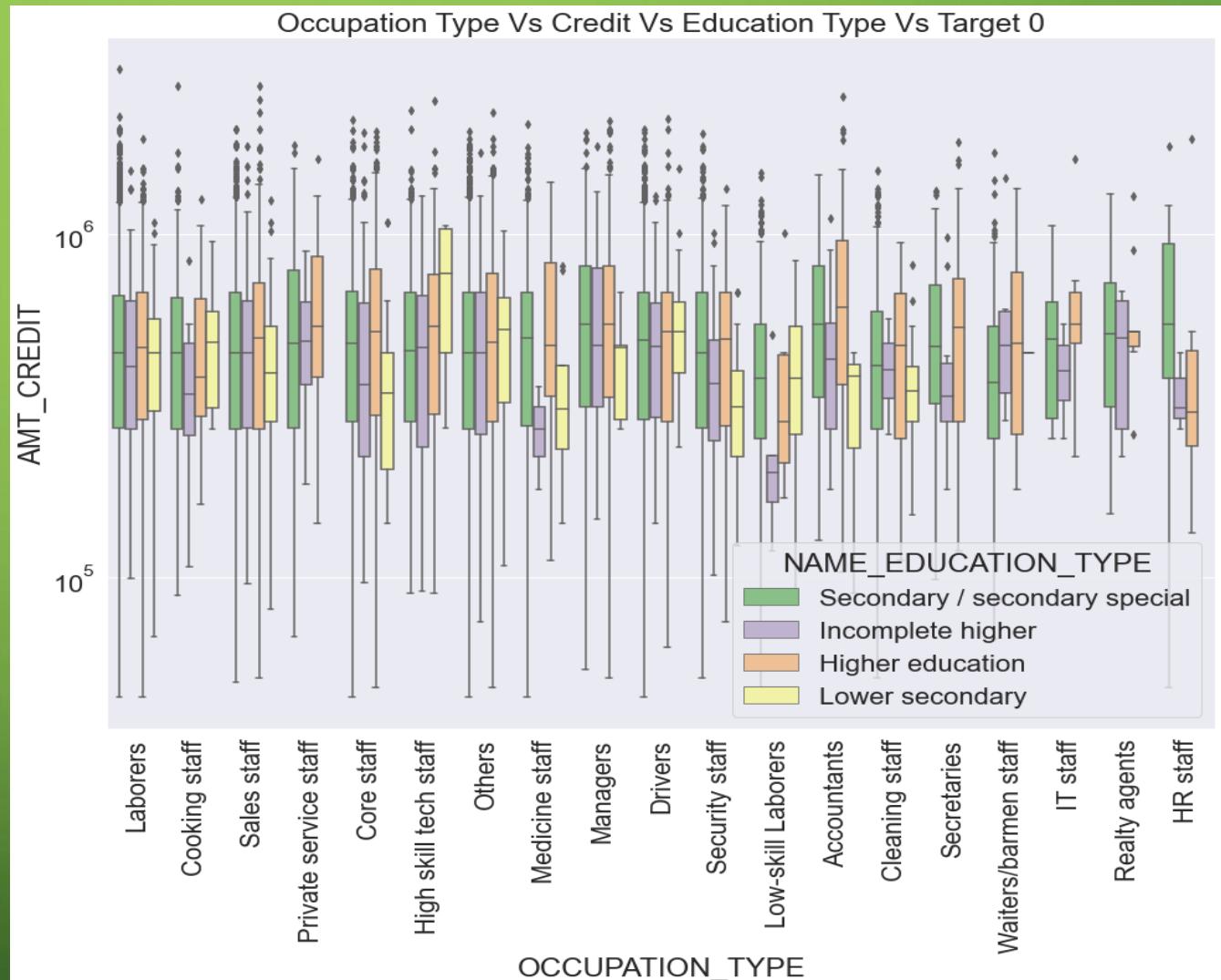
- People with higher education and drawing higher salary tend to default in general.
- People with lower secondary education and drawing lower salary tend to default in general.



OCCUPATION TYPE VS CREDIT VS EDUCATION TYPE VS TARGET 0

High Risk Categories:

- High skill tech staff with lower secondary qualification who has drawn higher credit limits.
- Accountants with a higher education or secondary who have drawn higher credit limits.
- HR staff with secondary who has drawn higher credit limits.
- Private service staff with higher education who has drawn higher credit limits.
- Bank should reconsider giving higher credit amount to the above category of applicants. May approve for lesser amounts.



HEATMAP FOR MAJOR NUMERIC VARIABLES



Findings from the heat map:

The correlation between various numeric variables remains the same regardless of the target variable.

1. AMT_CREDIT and AMT_GOODS_PRICE have a strong positive correlation with respect to the target Variable.
2. Similarly, PREV_AMT_CREDIT and PREV_AMT_GOODS_PRICE have a strong positive correlation with respect to the Target Variable.
3. CNT_CHILDREN and CNT_FAM_MEMBERS have a strong positive correlation with respect to the Target Variable.
4. Age and work experience are positively correlated.
5. Age and EXT_SOURCE_AVG are positively correlated.
6. Also, Age is positively correlated with AMT_INCOME_TOTAL, AMT_CREDIT and AMT_GOODS_PRICE.
7. PREV_AMT_ANNUITY is strongly and positively correlated with AMT_APPLICATION, PREV_AMT_CREDIT and PREV_AMOUNT_GOODS_PRICE.
8. AMT_INCOME_TOTAL is positively correlated with AMT_CREDIT, AMT_GOODS_PRICE, AMT_ANNUITY and PREV_AMT_ANNUITY.
9. REGION_POPULATION_RELATIVE and REGION_RATING_CLIENT are inversely correlated.
10. EXT_SOURCE_AVG and TARGET are inversely correlated, means less EXT_SOURCE_AVG means more likely to default.

HIGH RISK CATEGORIES:

- Low skill labourers having lesser income level regardless of their education are risky.
- People with lower secondary education, drawing lower salary tend to default in general. However lower secondary separated people are the riskiest.
- Widows who are managers, accountants or HR staff who have taken higher credits tend to default.
- Women on maternity leaves with low income and taken higher credit tend to default badly.
- Realty agents with incomplete higher education and high income usually takes higher credit but fail to replay. However, their count is too less.
- Commercial associates with lower income tend to default.
- In general, the default rate is more where user doesn't own any car.
- Singles are in general defaults followed by separated people.
- HR staff, waiters/barmen, drivers, cooking staff and core staff tend to default more.
- People living with parents, rented and in co-op apartments are having difficulty in repayment.
- Managers and high skill tech staff with lower secondary education who are taken higher credit tends to default.
- It is quite evident that people who have refused or unused offer are likely to fall default next time.
- When it comes to the purpose of loan water supply, car repair, urgent needs, repair have high chance of non-repayment.

RECOMMENDATIONS

- Bank should reconsider giving higher credit amount to the above category of applicants. May approve for lesser amounts or even reject their application after scrutiny.

POTENTIAL CUSTOMERS:

- In general graduates and people with higher education repay the loan on time especially graduates drawing high income and have taken large credits are very consistent with repayment. However only a very few graduates have applied for loan. People having Higher Education is the second largest category of applicants and in general make lesser defaults.
- Widows and married people with higher education are in general safe customers.
- Singles who are secretaries and HR staff are generally safe, however the count of such category are too less to reach any conclusion.
- IT staff, Accountants who are drawing high salary and have applied for higher credit amount seems to be safe customers.
- First time applicants are more likely to repay.
- Pensioners, students and businessmen are more likely to repay even though the percentage of such customers are very less.
- IT staff, realty agents, drivers and accountants having a higher income are in general safe customers.
- Loan purposes like business development, buying a new car, purchase of electronic equipment - the default rate is negligible.
- High income commercial associates who have taken bigger credits are seems to be safe customers.
- Comparatively very less clients are from Industry type 8,type 6, type 10,type 12, Religion and trade type 5, type 4, type 8 - but their default rate is quite less.
- The last but least - people accompanied by children are comparatively safer!

SUGGESTION

There are a few categories of very consistent applicants however, at present their overall count is less. Hence bank should try to target such customers for better sales through different sales promotions and customer campaigns.



THANK YOU

Gloriya Thomas & Sudeep Shah