



# CREDIT EDA objective or problem statement

1. LOAN providing NBFCs and banks finds to decide whose type of client should be suitable for loan.
2. Bank want to do analysis on client to give loan who are eligible to repay the loan amount.
3. bank don't have their credit history so some clients take advantage of becoming defaulter.
4. If bank not give loan then there is a loss in business of bank and if banks gives loan then there is a credit risk for bank
5. Bank want to expand their business and interesting to avoid non defaulters.
6. Do analysis and finds trend to give loans those who are eligible to repay loan

# ASSUMPTIONS

1.n code gender there is an XNA category whose gender is unknown.

2.Day\_Birth columns have negative values which should be in positive.

# APPROACH

1. First start with application data file.
2. Fixing irregular rows and columns live aberrations etc.
3. Fixing gender code column and convert M to MALE.
4. Fixing gender code column and convert F to Female.
5. Do metadata of application data file
6. Find no of object data type columns
7. Find no of float and integer data type columns
8. Find missing values and percentage also
9. Set threshold that columns have missing value over 40% drop them.

# APPROACH

10. Drop the columns who have missing values above threshold
11. Fill the missing values with mean , median, mode.
12. Fill the missing value for categorical data type columns with Mode.
13. Do univariate and bivariate analysis on the columns of application data file.
14. Now review the previous application file for loans
15. Do some metadata on previous file
16. Check for missing values and outliers.
17. Remove the missing values according to the threshold of 40%.

# APPROACH

18. Do the univariate and bivariate analysis of previous application data

19. Merge both the file , application data and previous data file

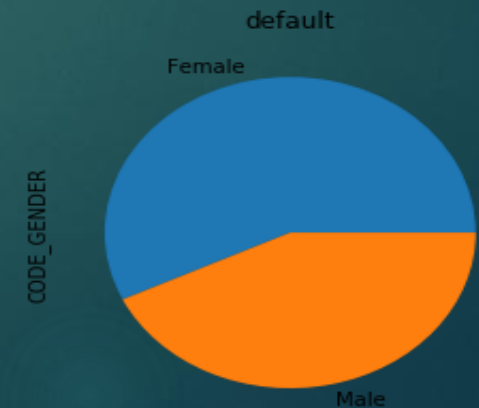
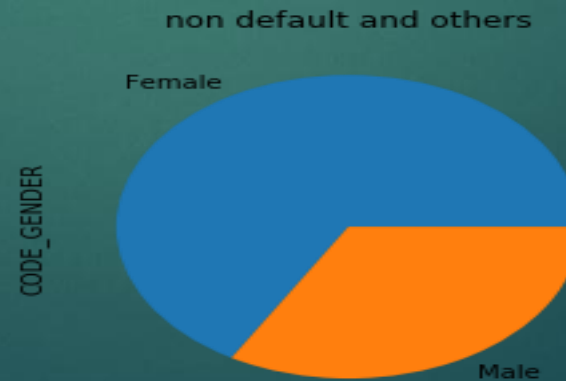
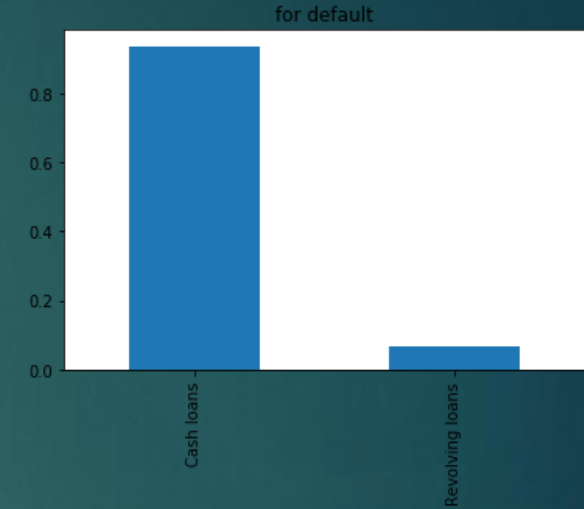
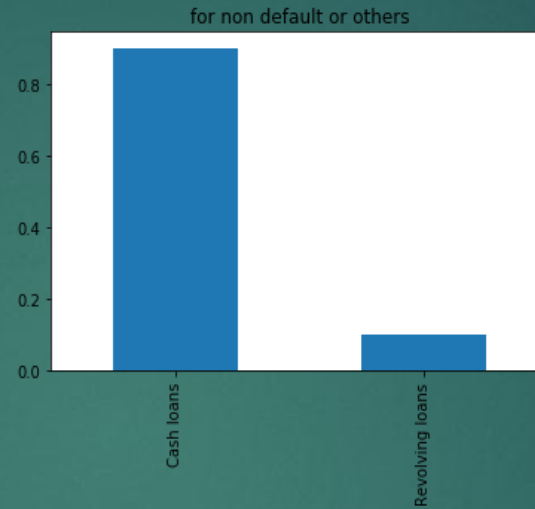
20. Go for univariate and bivariate analysis for merge file and tried to find out the trend and pattern.

# Analysis results

## UNIVARIATE ANALYSIS :-

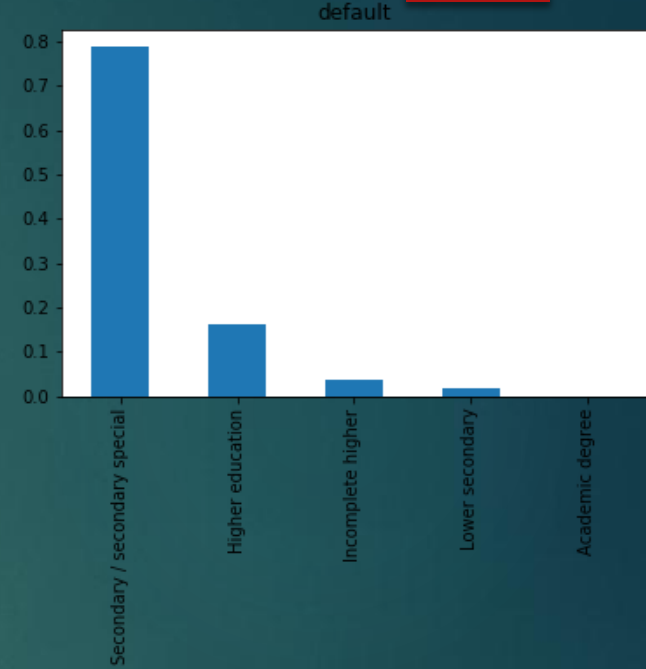
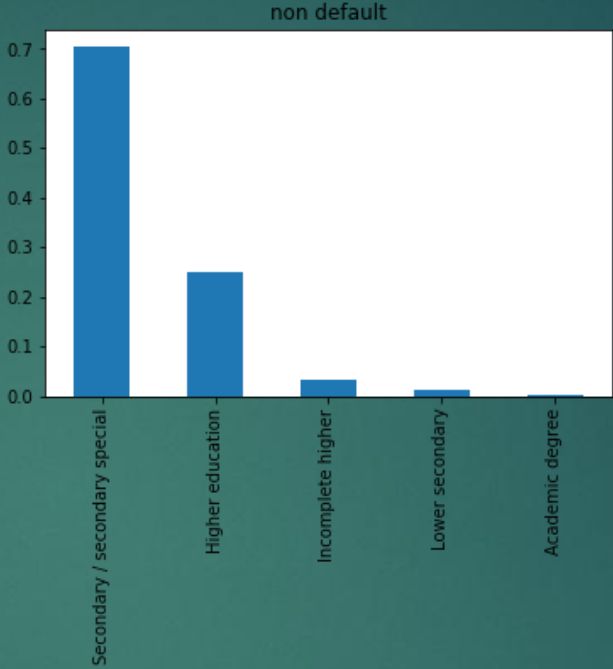
I OBSERVE FROM THE PLOT THAT PEOPLE ARE MORE INTERESTED IN CASH LOAN IN COMPARE TO RESOLVING LOAN.

I FOUND IN ANALYSIS ON GENDER THAT WOMEN MORE APPLY FOR LONE AND MORE WOMEN ARE DEFAULT IN LOAN

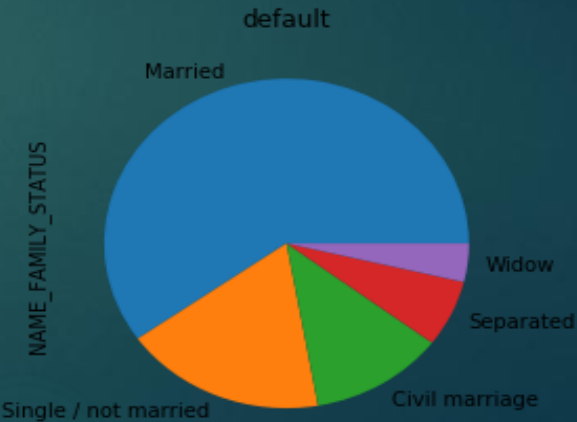
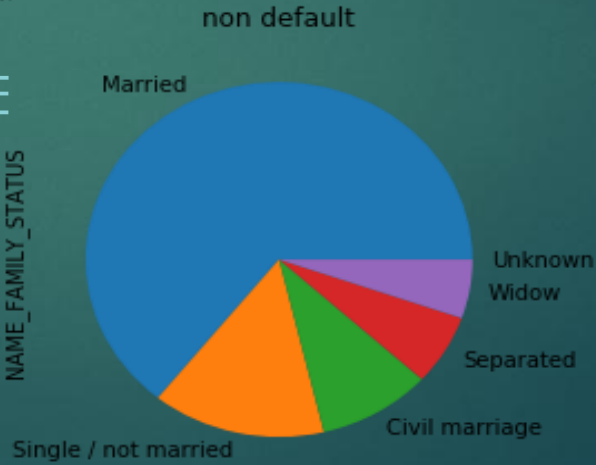


# Univariate analysis

I FIND THAT MOST OF LOAN APPLICANT ARE SECONDARY AND SECONDARY SPECIAL.

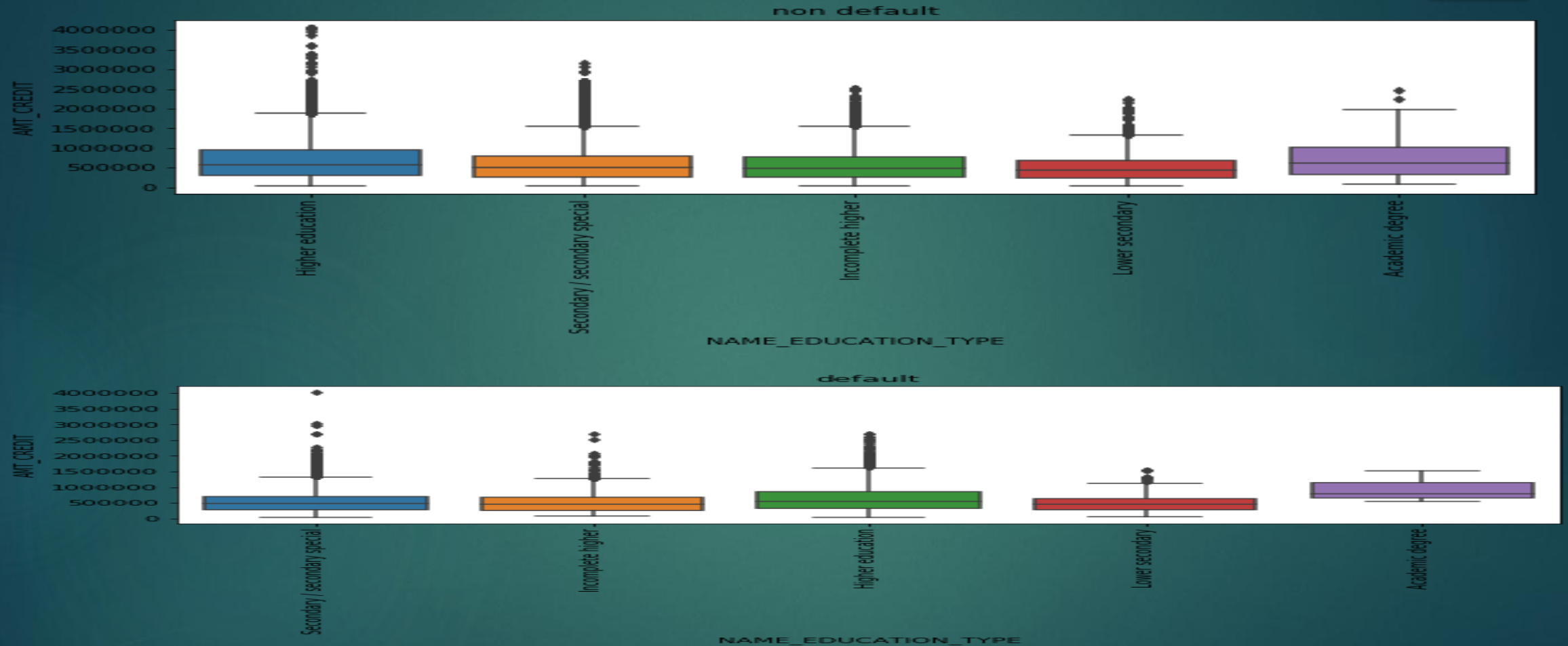


OBSERVE THAT MARRIED APPLICANTS MORE IN NUMBERS IN DEFAULT CASE





# Bivariate analysis



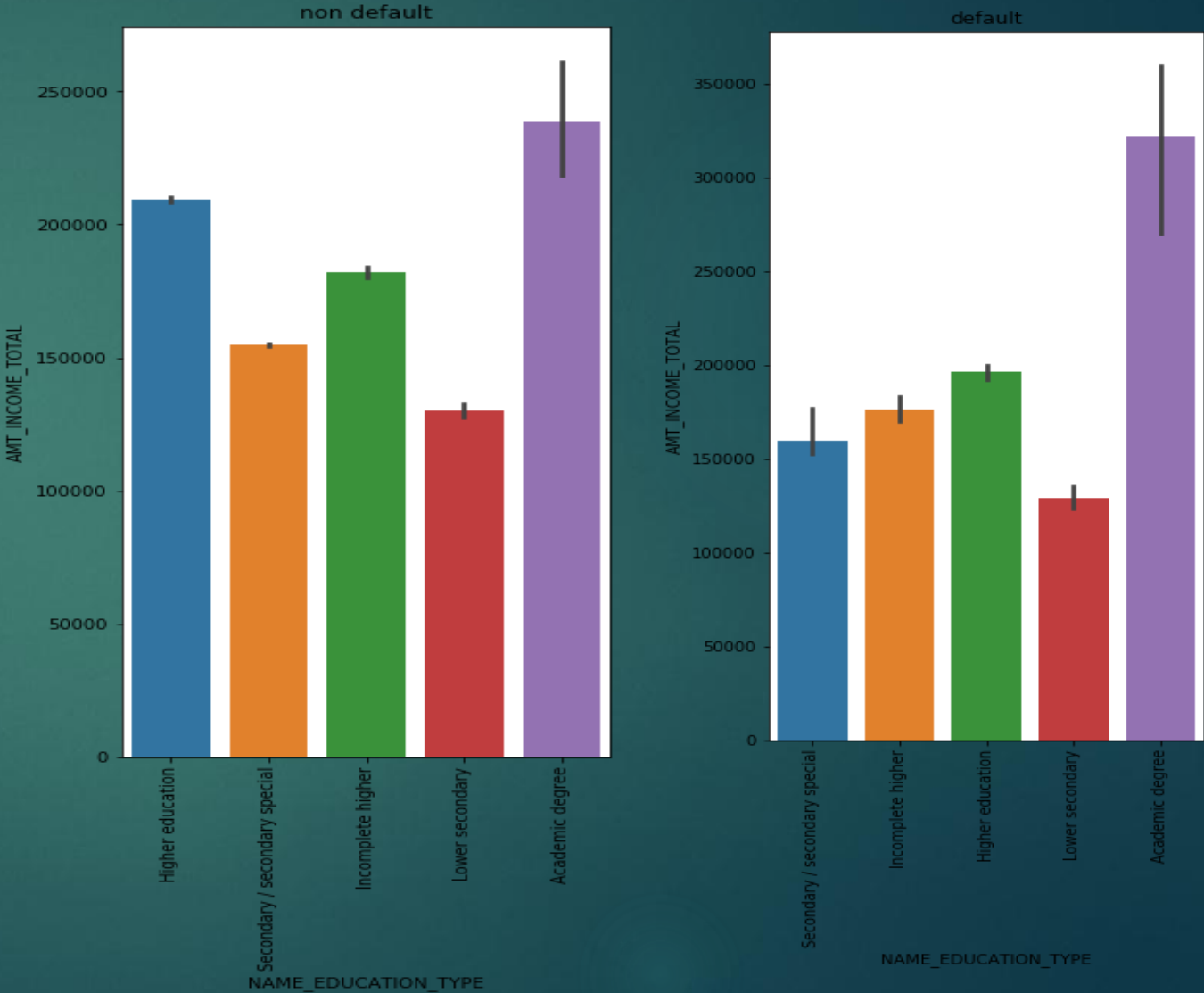
WE CAN CHECK FROM PLOT THAT CLIENT WITH DEFAULT OF ACADEMIC DEGREE IS HIGHER THAN NON DEFAULT



# Bivariate analysis

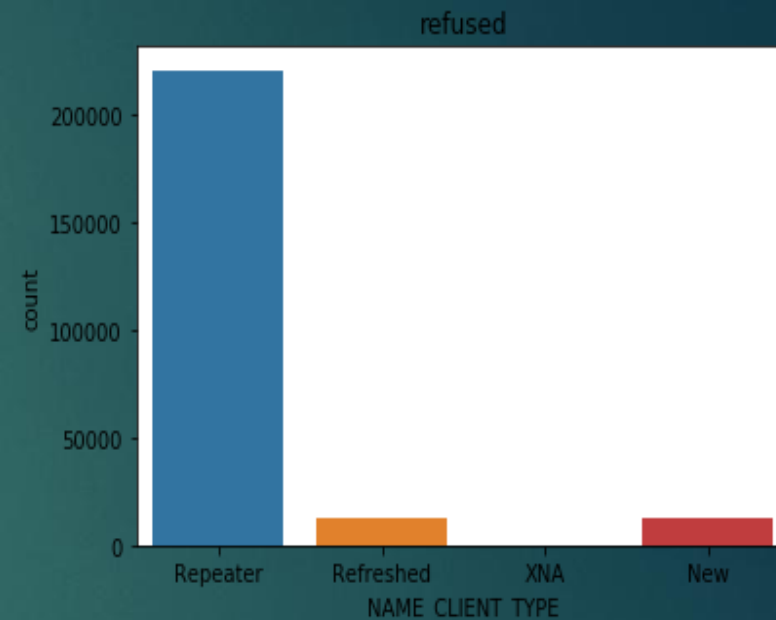
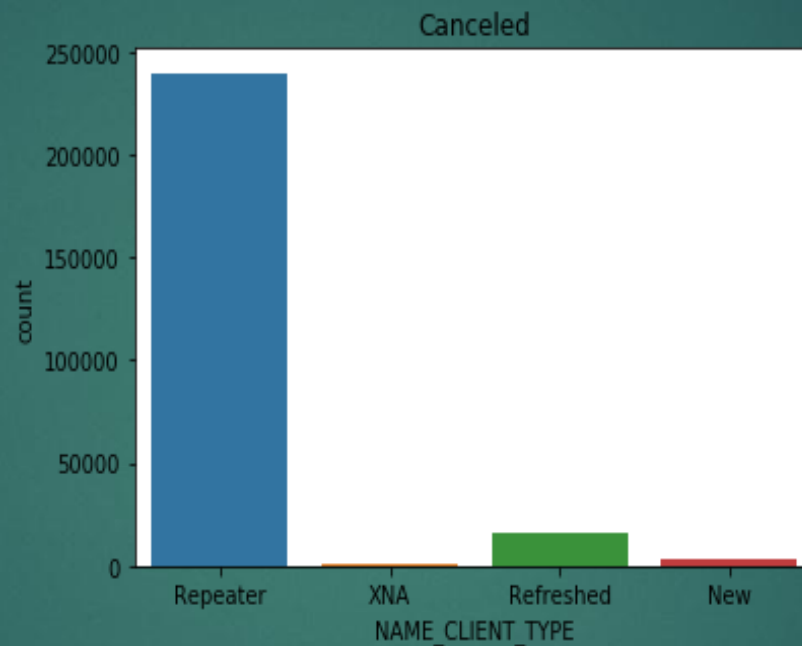
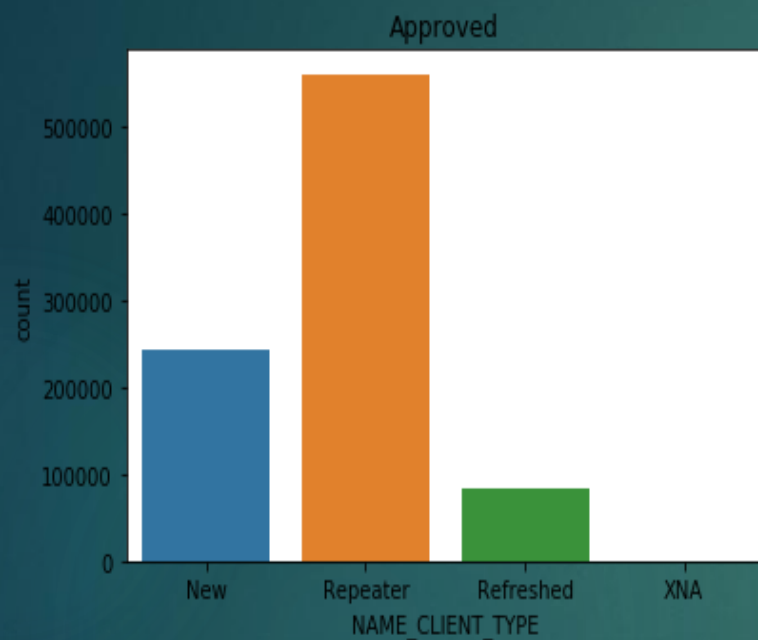
## AMT\_INCOME\_TOTAL AND NAME\_EDUCATION\_TYPE BIVARIATE ANALYSIS

WE CAN SEE THAT ACADEMIC DEGREE IS  
DEPENDS ON THE CLIENT TOTAL INCOME



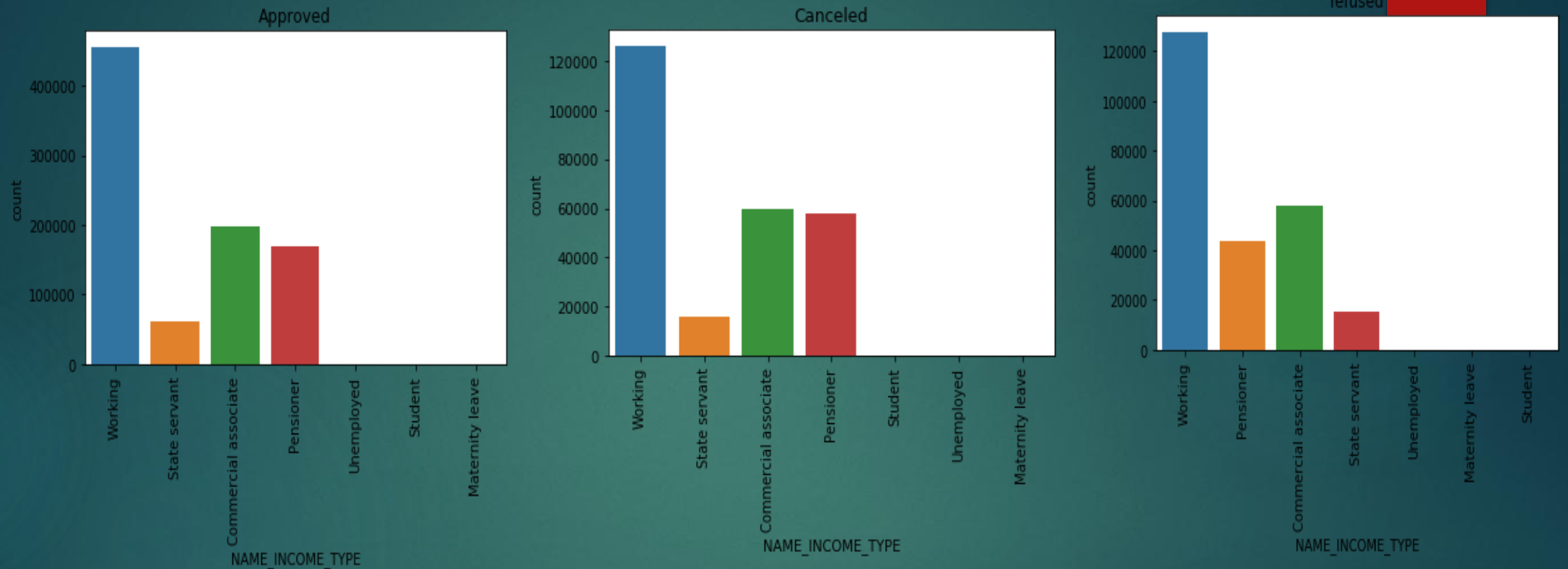
# Segmented univariate analysis

## CLIENT TYPE



1. WE CAN SEE HERE THAT REPEATERS GET MORE APPROVED
2. REPEATERS GET MORE LOAN CANCELATION
3. REPEATERS GET MORE REFUSED FOR LOANS

# NAME\_INCOME\_TYPE segmented univariate



WE CAN OBSERVE THAT WORKING CLASS GET MORE LOAN APPROVAL, REFUSAL, AND CANCELATIONS FOR LOAN

AND COMMERCIAL ASSOCIATES GET IN AVERAGE AMOUNT OF LOAN REFUSAL, APPROVAL, CANCELATION

# Conclusions

1. WE CAN SEE THAT WORKING CLASS MORE APPLIED FOR LOANS AND NO OF DEFAULTERS ALSO MORE IN WORKING.

2. BANK SHOULD FOCUS ON STUDENTS, PENSIONERS FOR BUSINESS.

3. WE OBSERVE THAT FEMALES APPLIED FOR LOANS IN LARGE NUMBERS.

BANK SHOULD APPROACH TO MALES APPLICANTS.

4. MORE APPLICANTS IN REPEATERS BANK ALSO NEED TO FOCUSED ON OTHER CONTRACTS TYPES ALSO