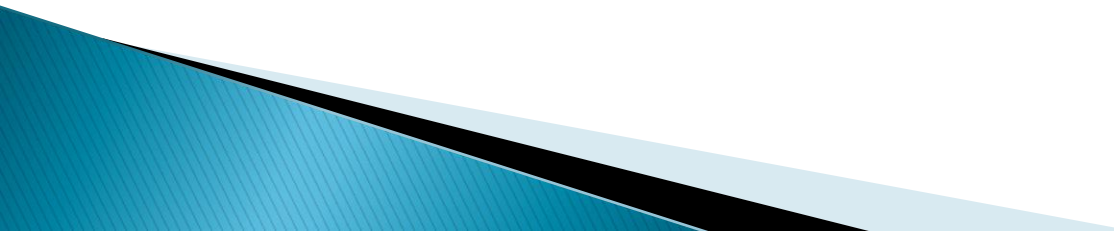




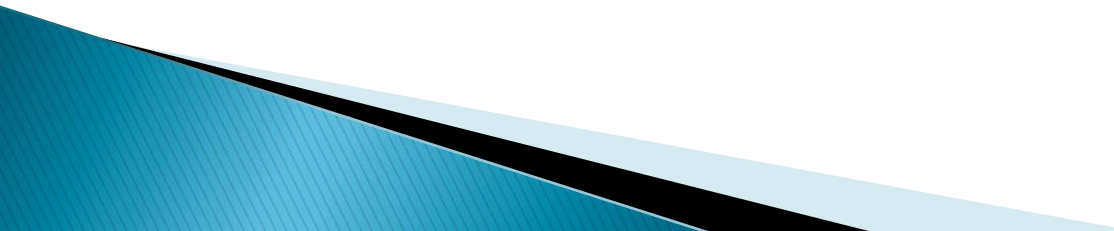
# CREDIT EDA CASE-STUDY

KHUPAT VISHNU PRIYA


# AIM

- ▶ Credit Risk Analysis aims at identifying the loan\_applicants who are capable of repaying the loan and also help the company to make a decision for loan Approval based on applicant profile. Which controls loss of business to the company and also avoid financial loss for the company.
- 

# APPROACH

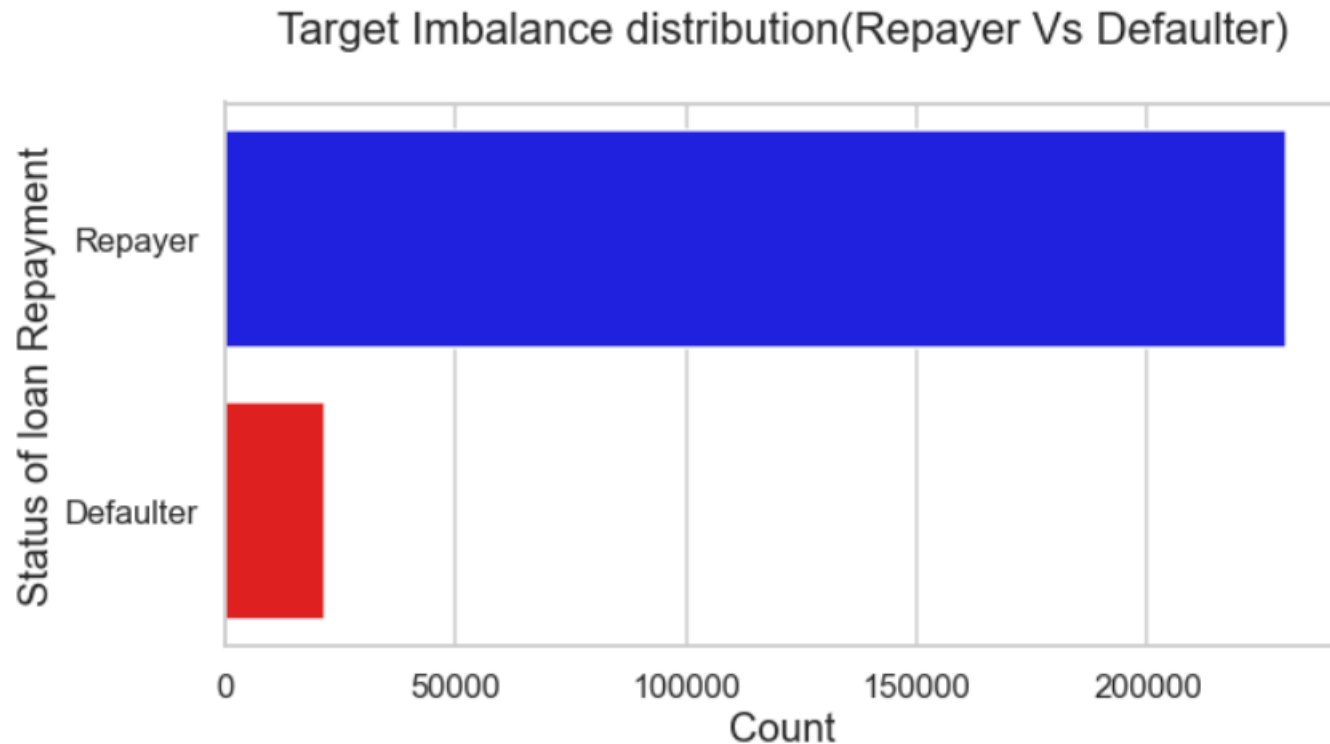
- ▶ Data Loading and Understanding
  - ▶ Data Cleaning and Manipulation
  - ▶ Data Imbalance
  - ▶ Performing Univariate, Segmented Univariate and Bivariate Analysis
  - ▶ Merging of application dataset with previous application
  - ▶ Performing data analysis on them to gain insights
  - ▶ Recommendations and Risks
- 

# DATA-CLEANING:

- ▶ For Data – Cleaning , we calculated the percentage of missing values and if the percentage is greater than 30 then we dropped those columns
  - ▶ For the columns whose missing values are less than 30, then depending upon the column we have imputed the values.
  - ▶ For example: If the column is of categorical type then we have imputed the values with mode. If the column is of numerical type then after checking the outliers we have imputed the values either with mean or mode, depending upon the outliers.
- 

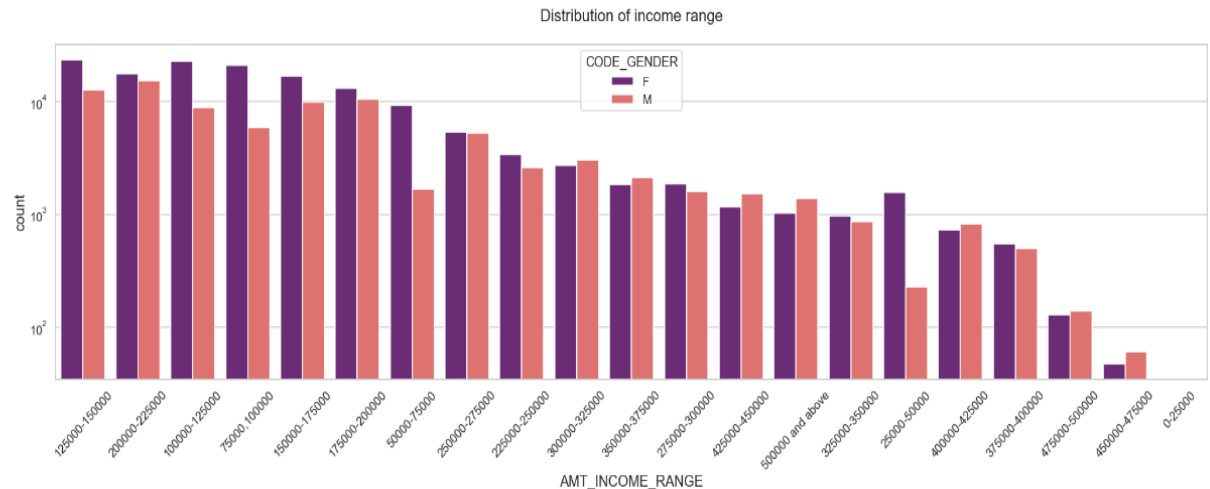
# DATA IMBALANCE

- ▶ The Imbalance percentage is 10.55



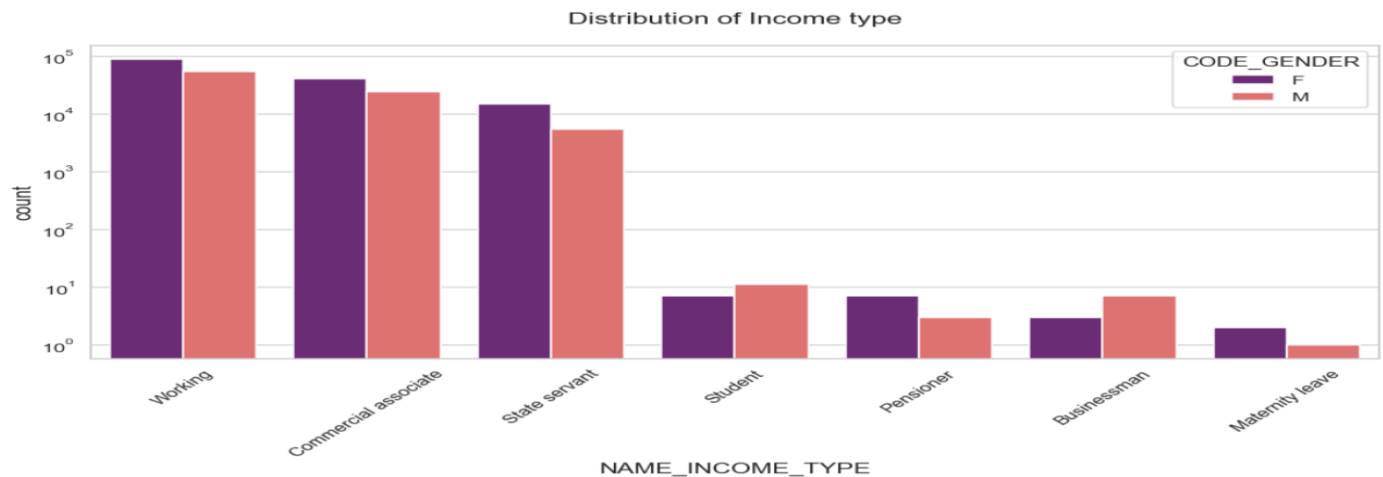
# UNIVARIATE ANALYSIS ON INCOME RANGE

- ▶ Female counts are higher than male.
- ▶ More number of credits are having for Income ranging from 100000 to 200000
- ▶ It displays that females are more than in having credits for that range.
- ▶ For the income range 400000 and above count is very less.



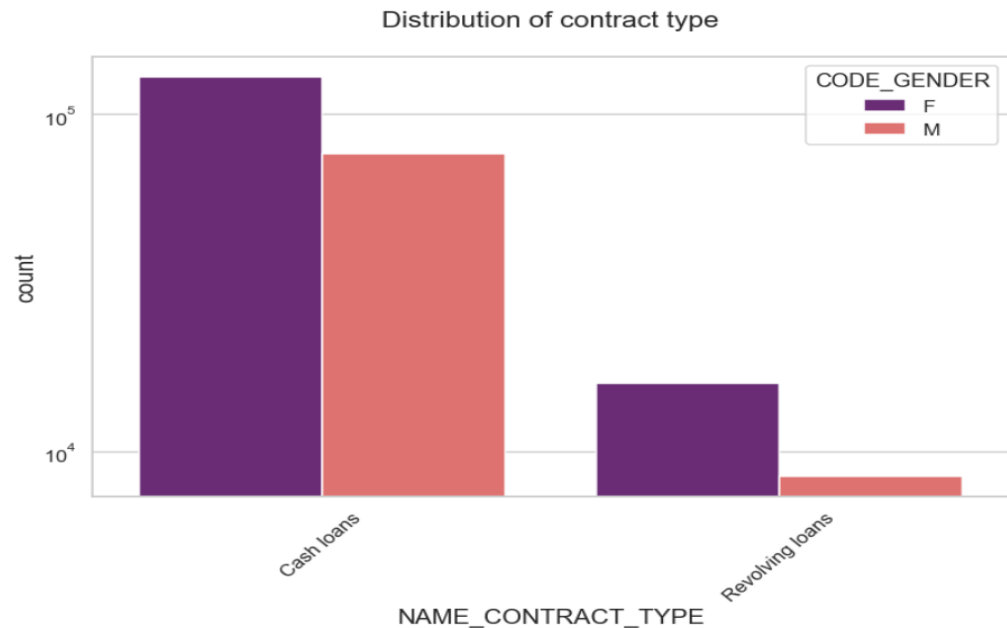
# UNIVARIATE ANALYSIS ON INCOME TYPE

- ▶ The number of credits are higher for the income type 'working', 'commercial associate' and 'State Servant' than the others.
- ▶ The females are having more number of credits than male.
- ▶ The number of credits are less for the income type 'student', 'pensioner', 'Businessman' and 'Maternity leave'.



# UNIVARIATE ANALYSIS ON CONTRACT TYPE

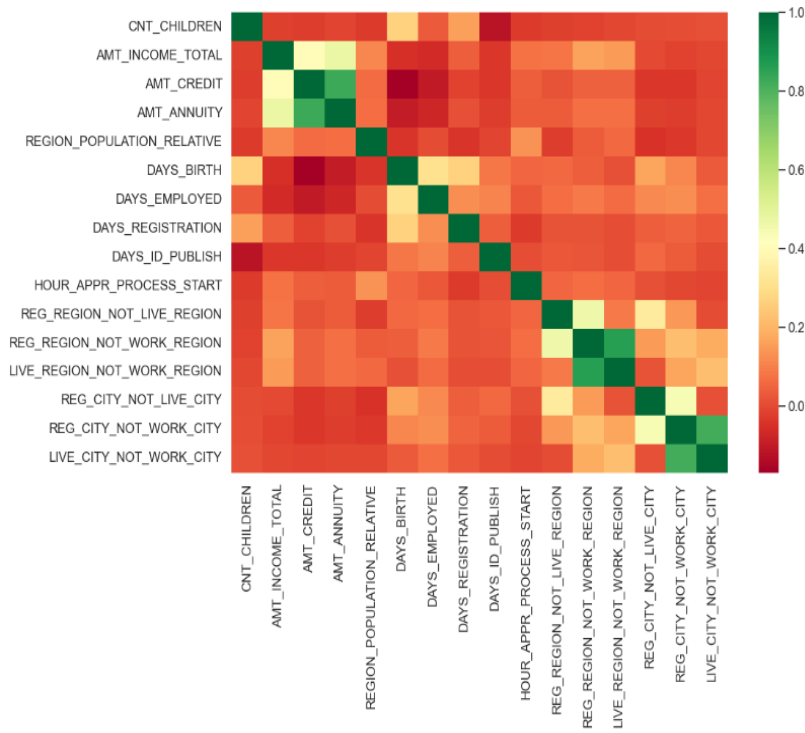
- ▶ The 'Cash loans' contract type is having higher number of credits than 'Revolving loans' contract type.
- ▶ Females are leading for applying credits.



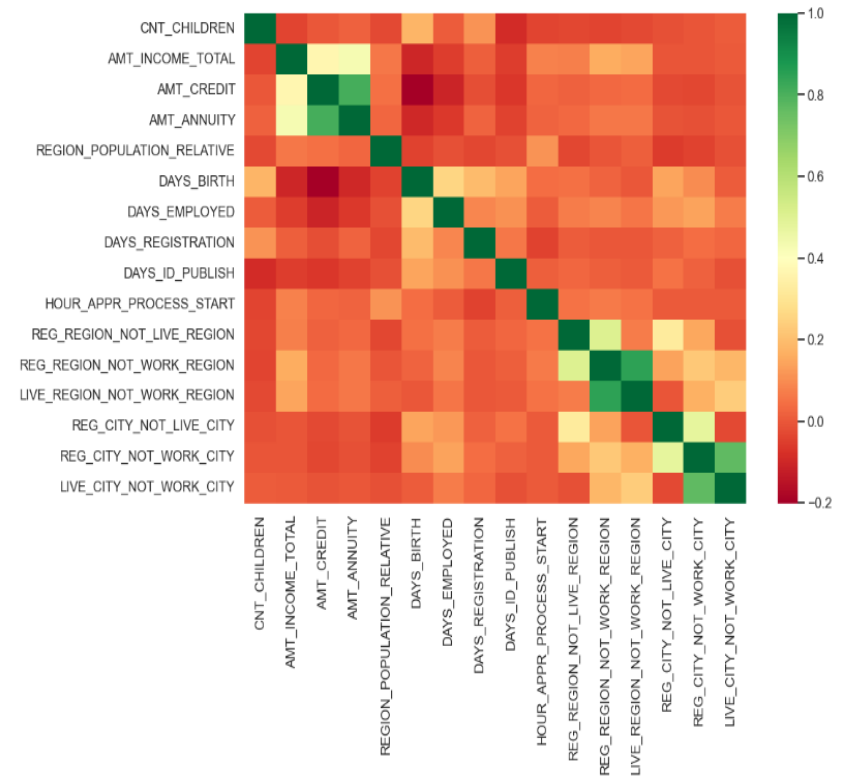


# CORRELATIONS

Correlation for target 0

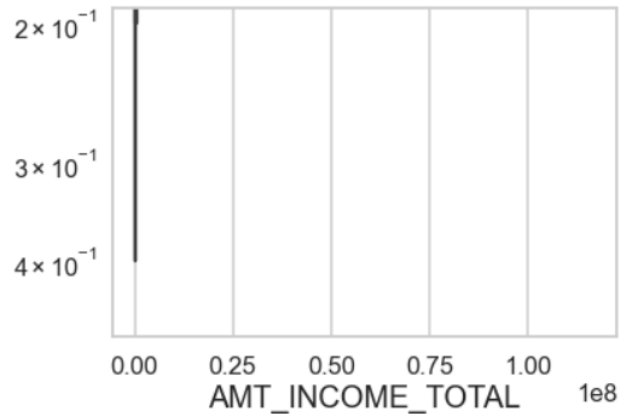


Correlation for target 1

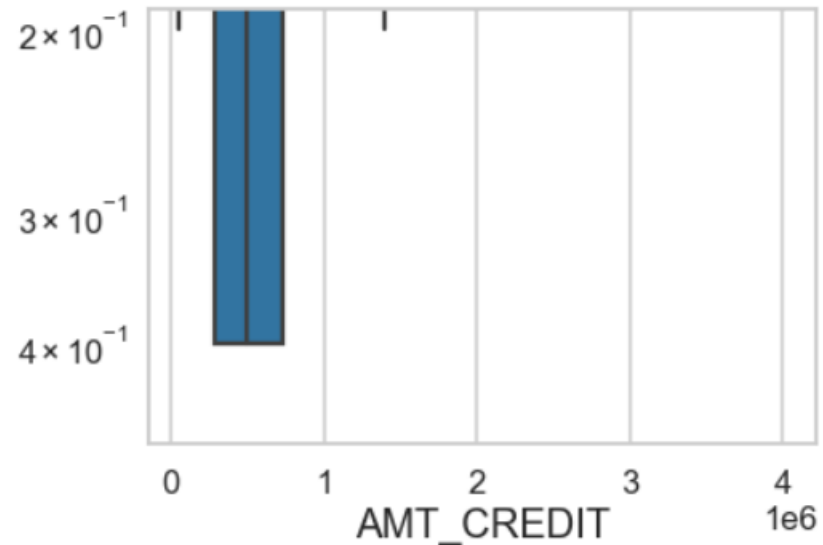


# UNIVARIATE ANALYSIS FOR VARIABLES

Distribution of income amount

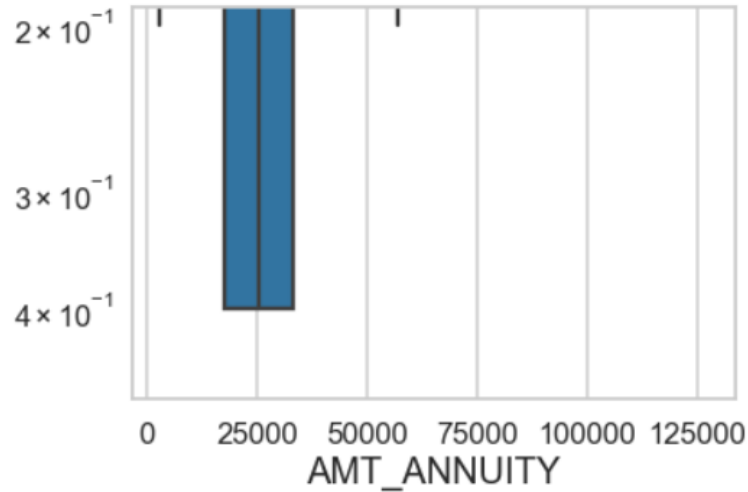


Distribution of credit amount

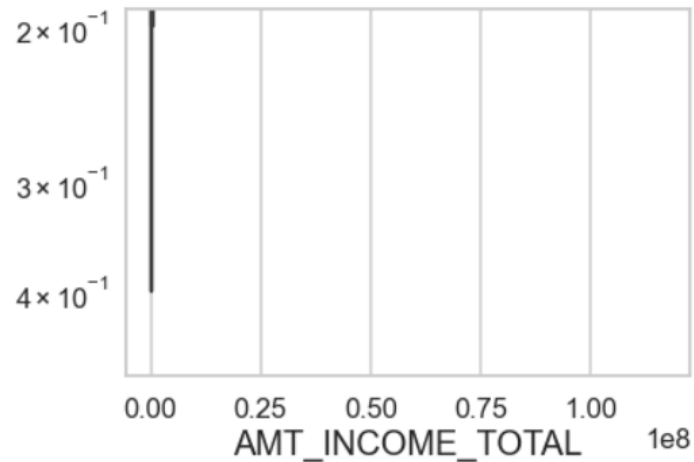


# UNIVARIATE ANALYSIS FOR VARIABLES

Distribution of Annuity amount

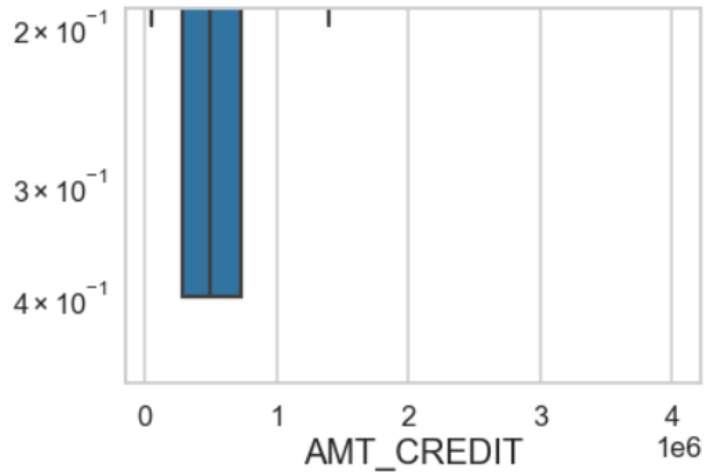


Distribution of income amount

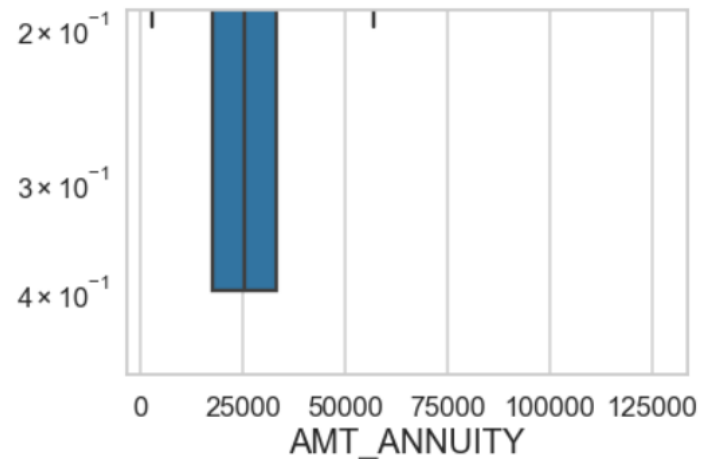


# UNIVARIATE ANALYSIS FOR VARIABLES

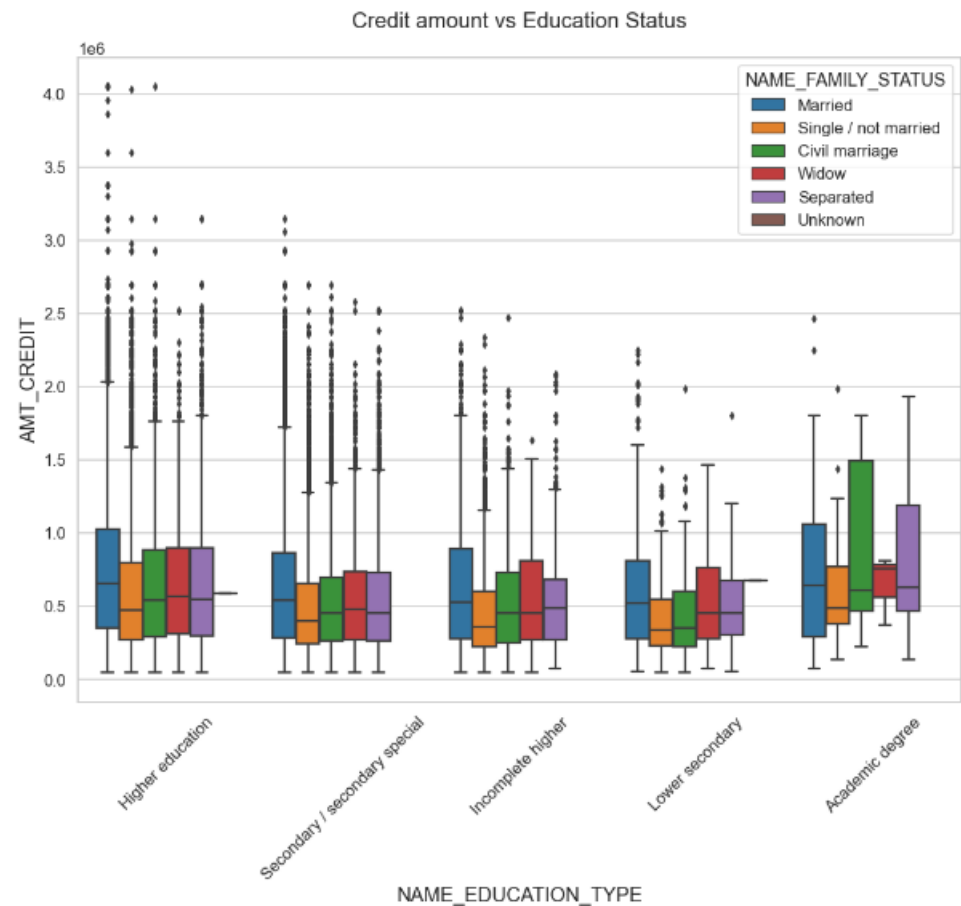
Distribution of credit amount



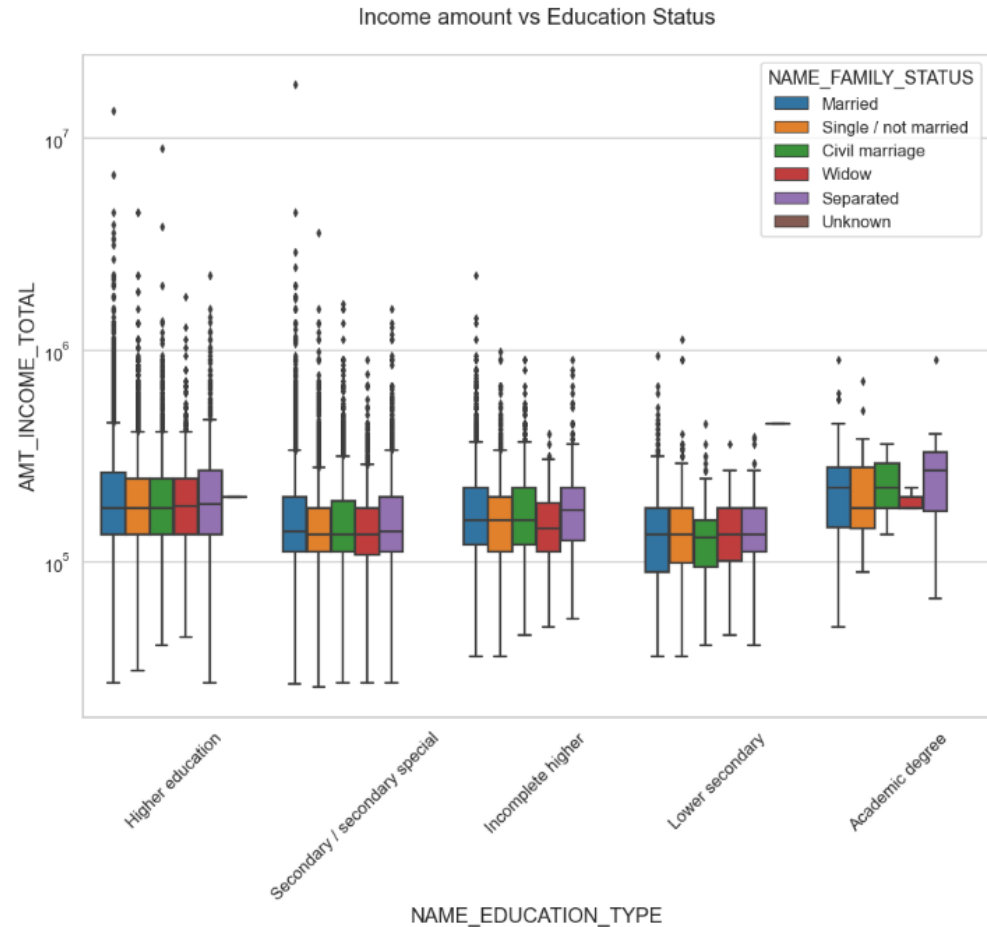
Distribution of Annuity amount



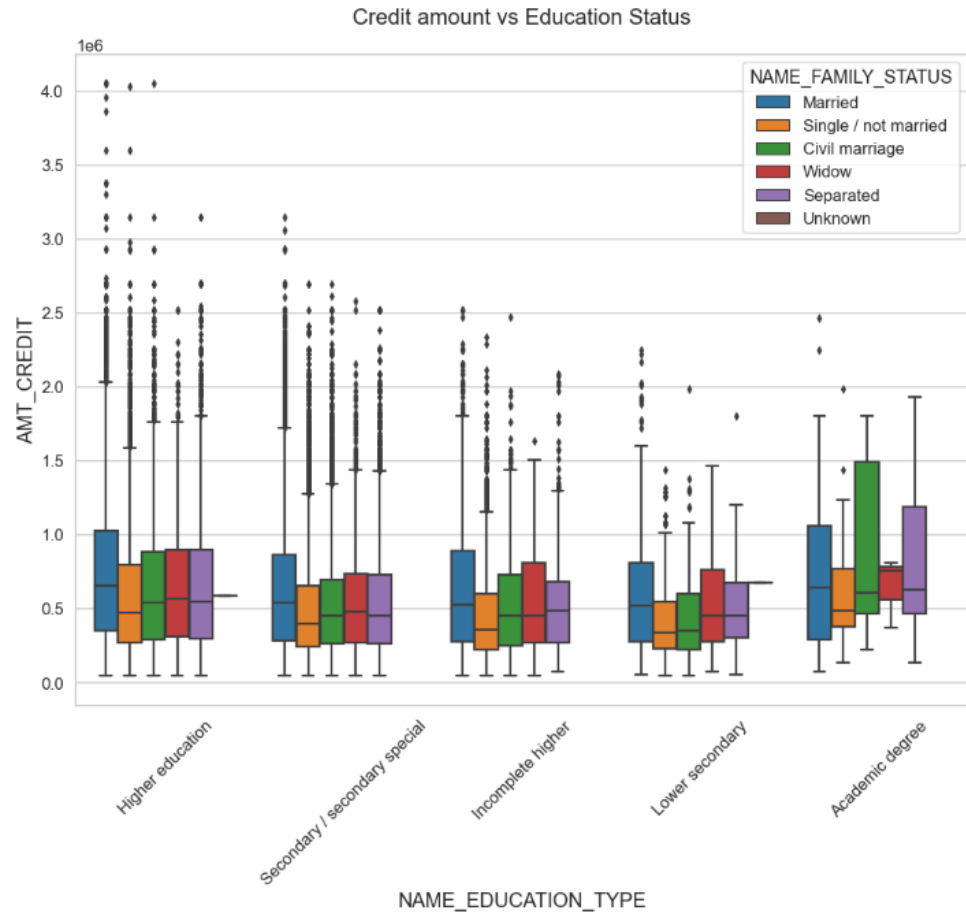
# Bivariate analysis for numerical variables (Credit Amount for target 0)



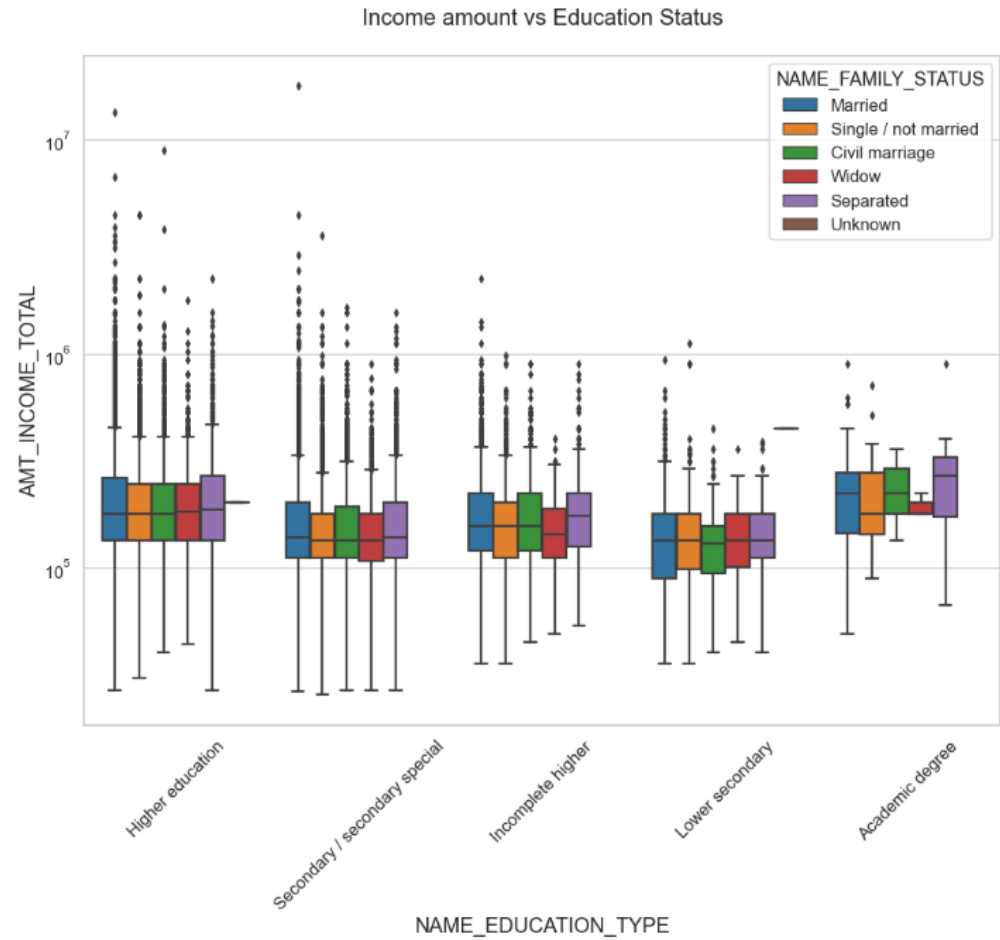
# Bivariate analysis for numerical variables (Income Amount for target 0)



# Bivariate analysis for numerical variables (Credit Amount for target 1)



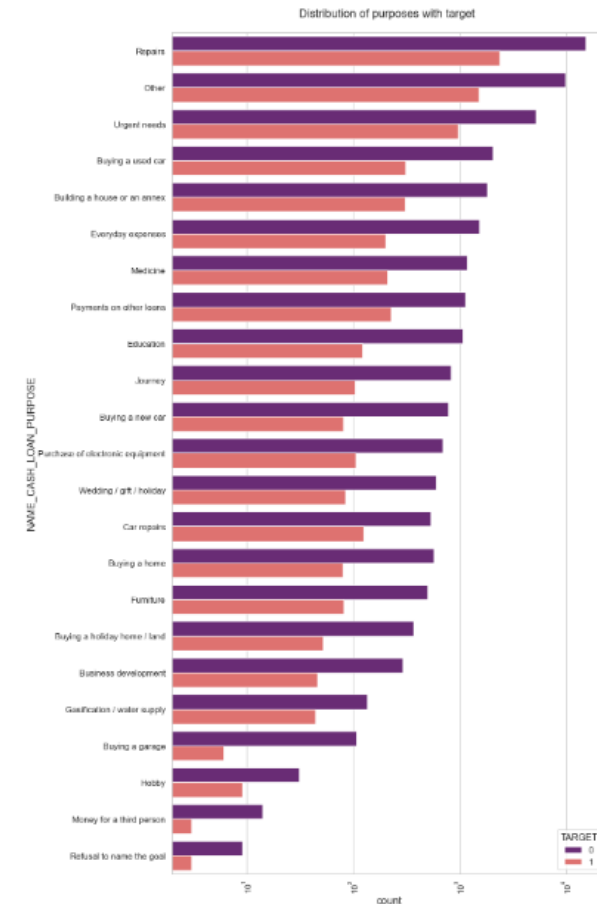
# Bivariate analysis for numerical variables (Income Amount for target 1)





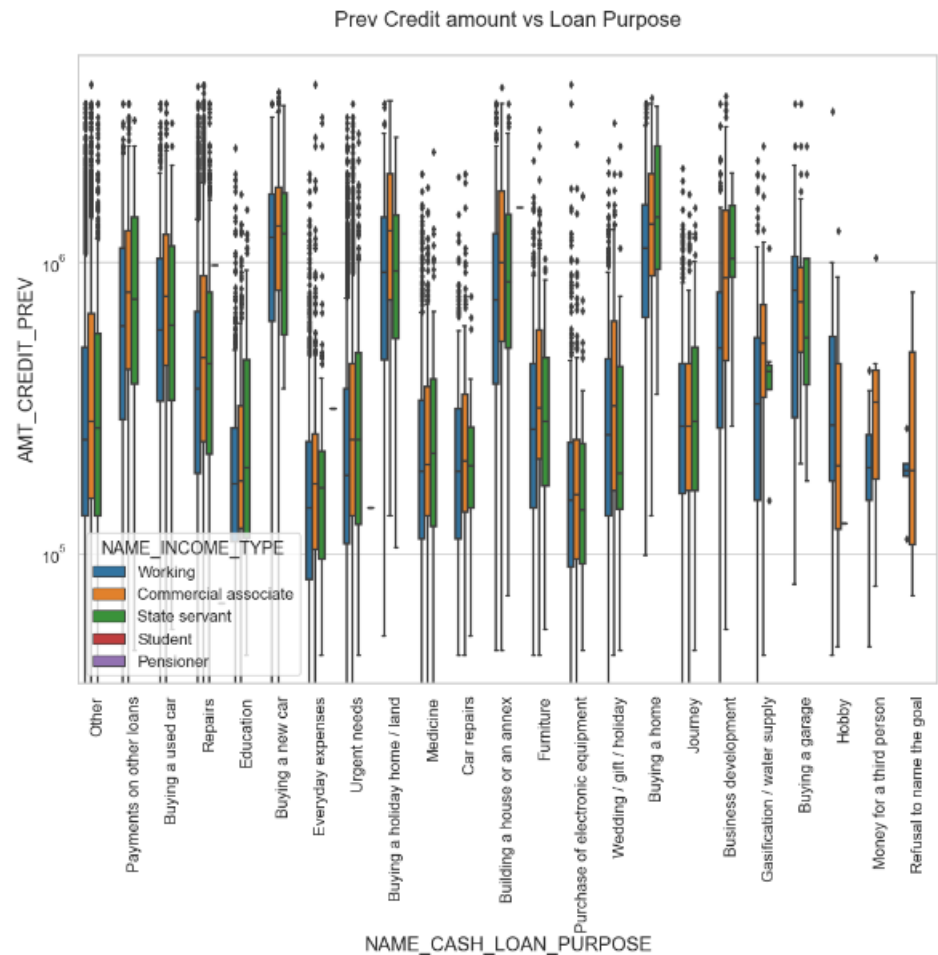
# ANALYSIS AFTER MERGING TWO DATASETS

## ► Univariate Analysis



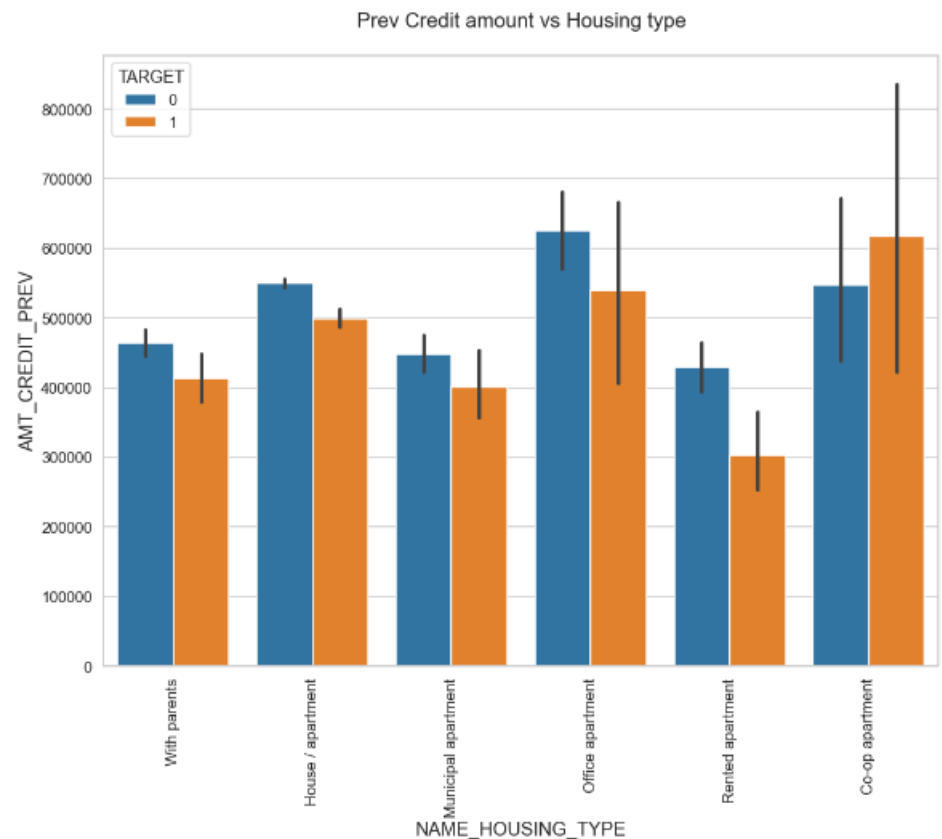
# ANALYSIS AFTER MERGING TWO DATASETS

## ► Bivariate Analysis

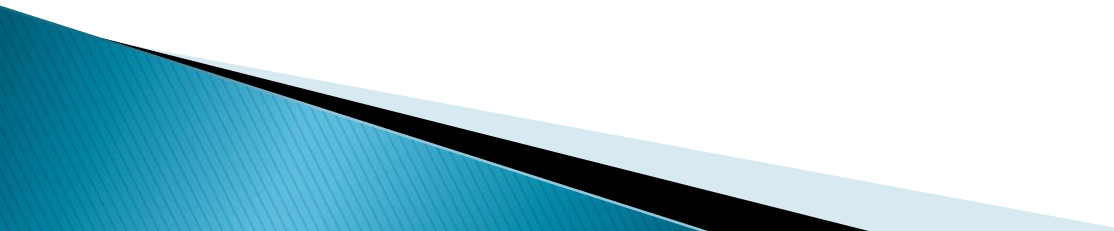


# ANALYSIS AFTER MERGING TWO DATASETS

- ▶ Credit amount prev vs Housing type



# CONCLUSION

- ▶ Banks should focus more on contract type 'Student', 'pensioner' and 'Businessman' with housing 'type other than 'Co-op apartment' for successful payments.
  - ▶ Banks should focus less on income type 'Working' as they are having most number of unsuccessful payments.
  - ▶ Also with loan purpose 'Repair' is having higher number of unsuccessful payments on time.
  - ▶ Get as much as clients from housing type 'With parents' as they are having least number of unsuccessful payments.
- 

**THANK YOU**