



EDA CASE STUDY

MADE BY:

RAVI PRAKASH GUPTA

SANYASHREE SRIVASTAVA

INTRODUCTION

- ▶ The case study deals with credit risk analytics in the banking and financial sector
- ▶ The Exploratory Data Analysis(EDA) needs to be done to minimize the risk of losing money while lending to customers
- ▶ In this case study we will use EDA to understand how consumer attributes and loan attributes influence the tendency of default

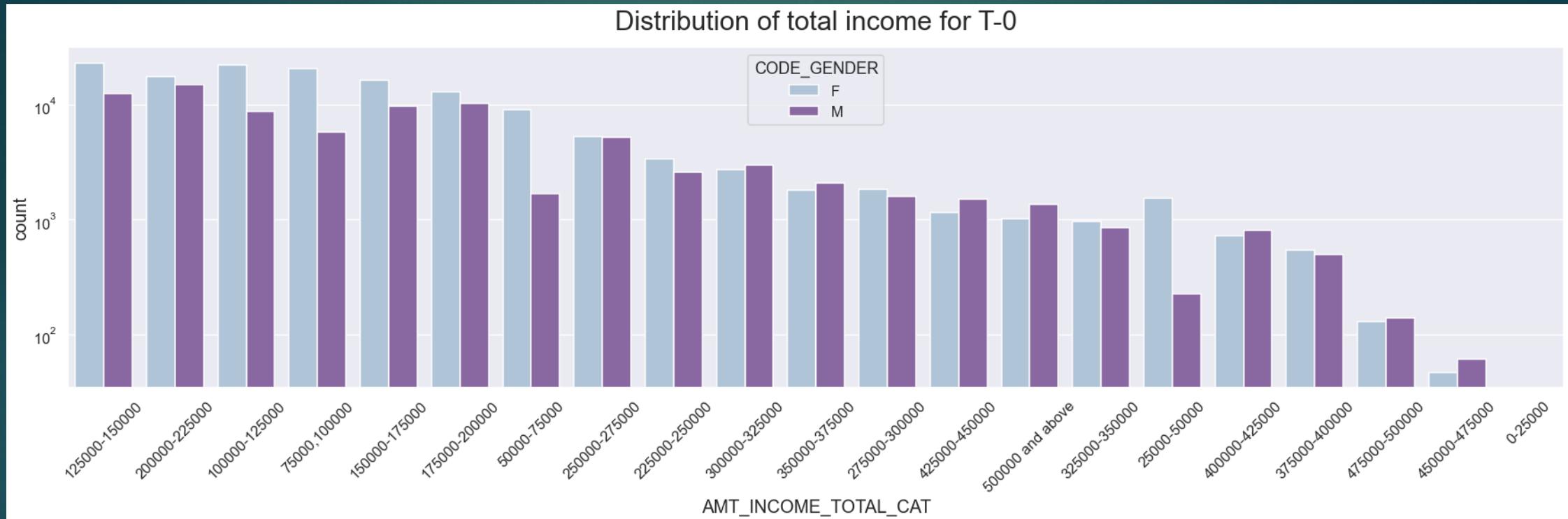
PROBLEM STATEMENT

- ▶ The company wants to understand the driving factors(driver variables) behind loan default and utilize this knowledge for its portfolio and risk assessment

ANALYSIS APPROACH

- Understanding the Risk analysis domain
- Loading the data
- Gathering meta data from the data frame
- Identifying the missing values & treating them
- Check for outliers
- Determining data imbalance
- Correlation Analysis
- Univariate & Bivariate Analysis

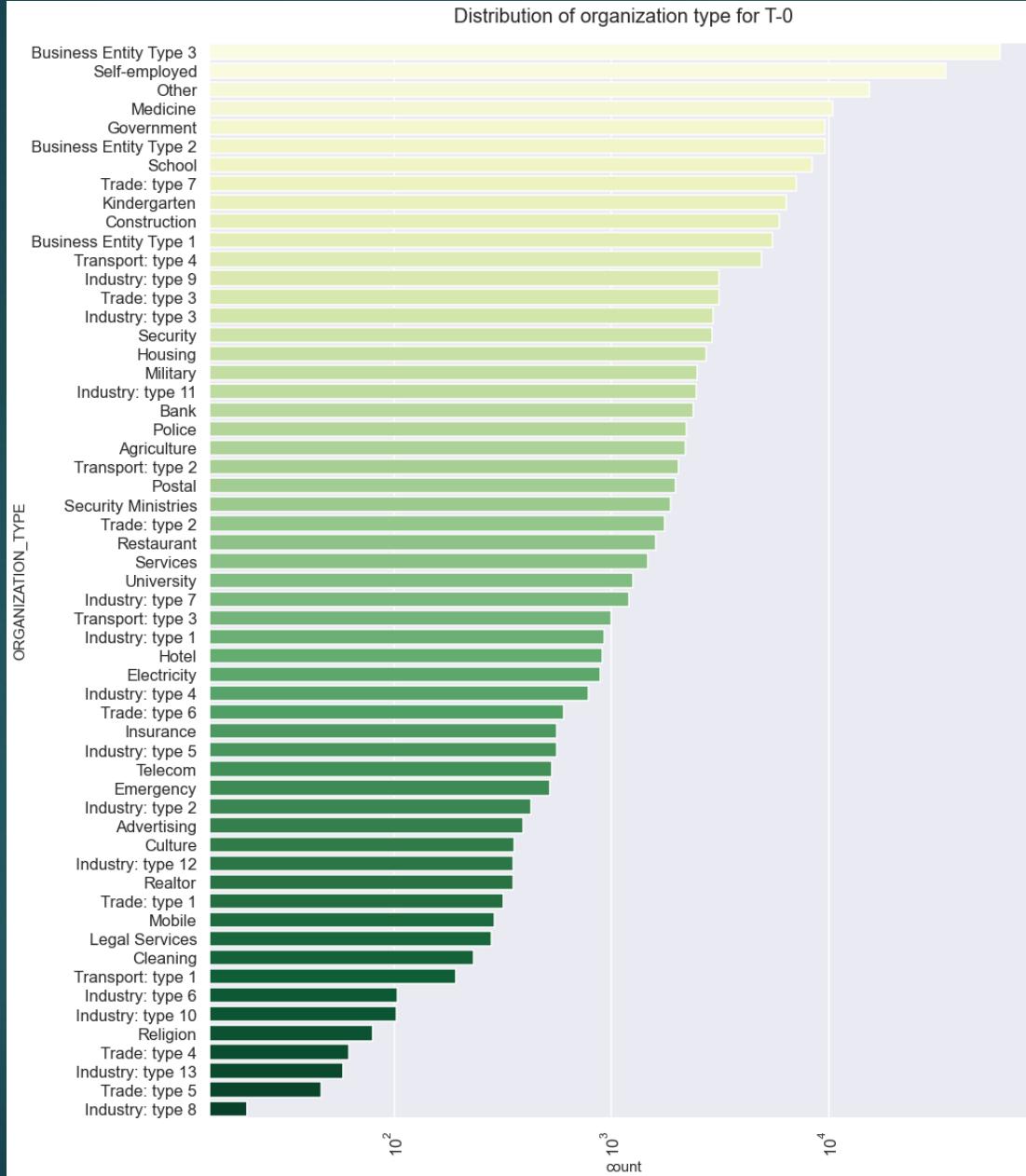
Distribution of total income for T-0



Inferences from the plot:

- ▶ Female counts are higher than male.
- ▶ Income range from 75000 to 125000 is having more number of credits.

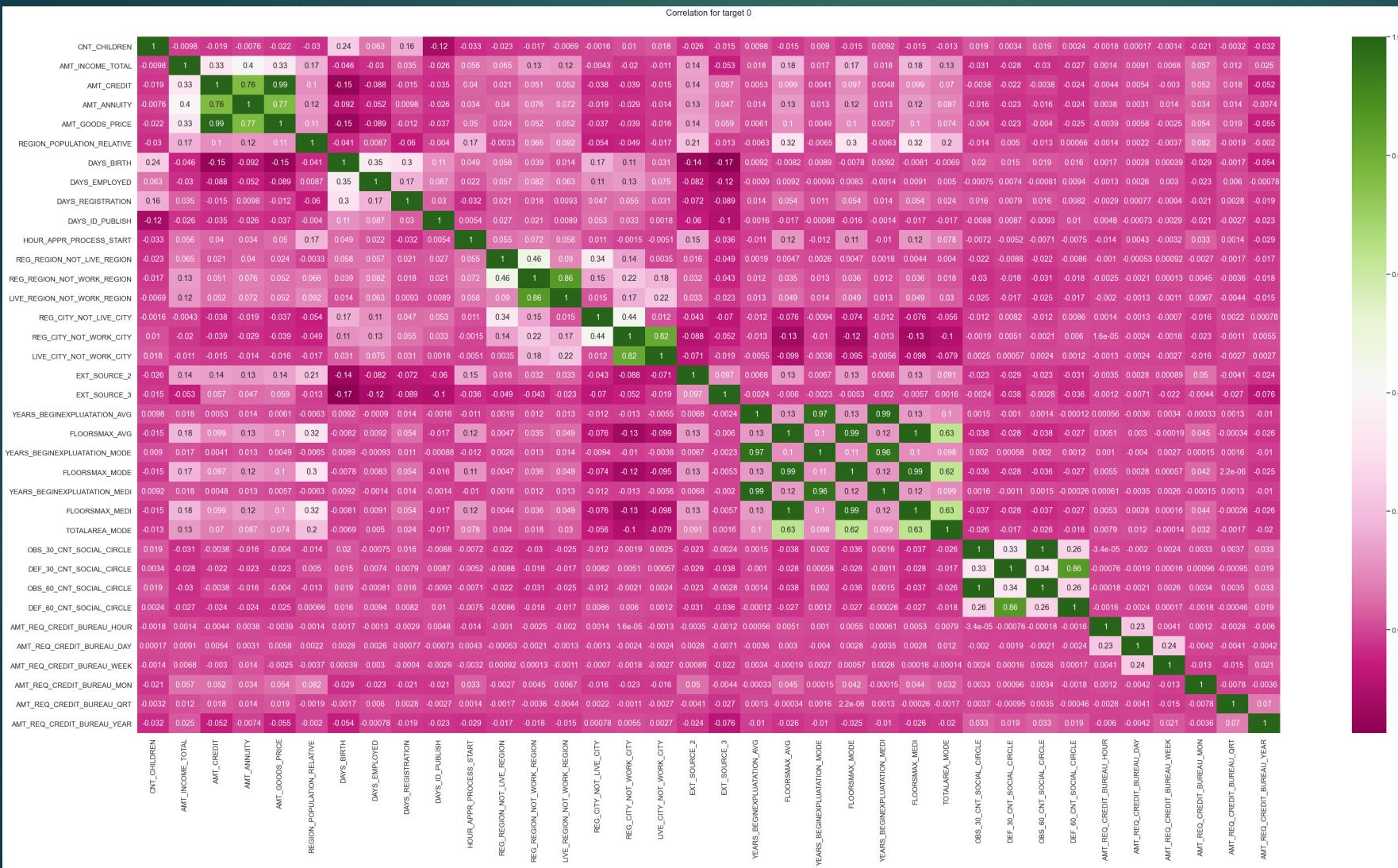
Distribution of organization type



Inferences from the plot:

- ▶ Business Entity Type 3 and Self-employed are the categories which have applied for the credits higher than any others.
- ▶ Industry : type 8 and Trade : type 5 are lowest in number of applications.

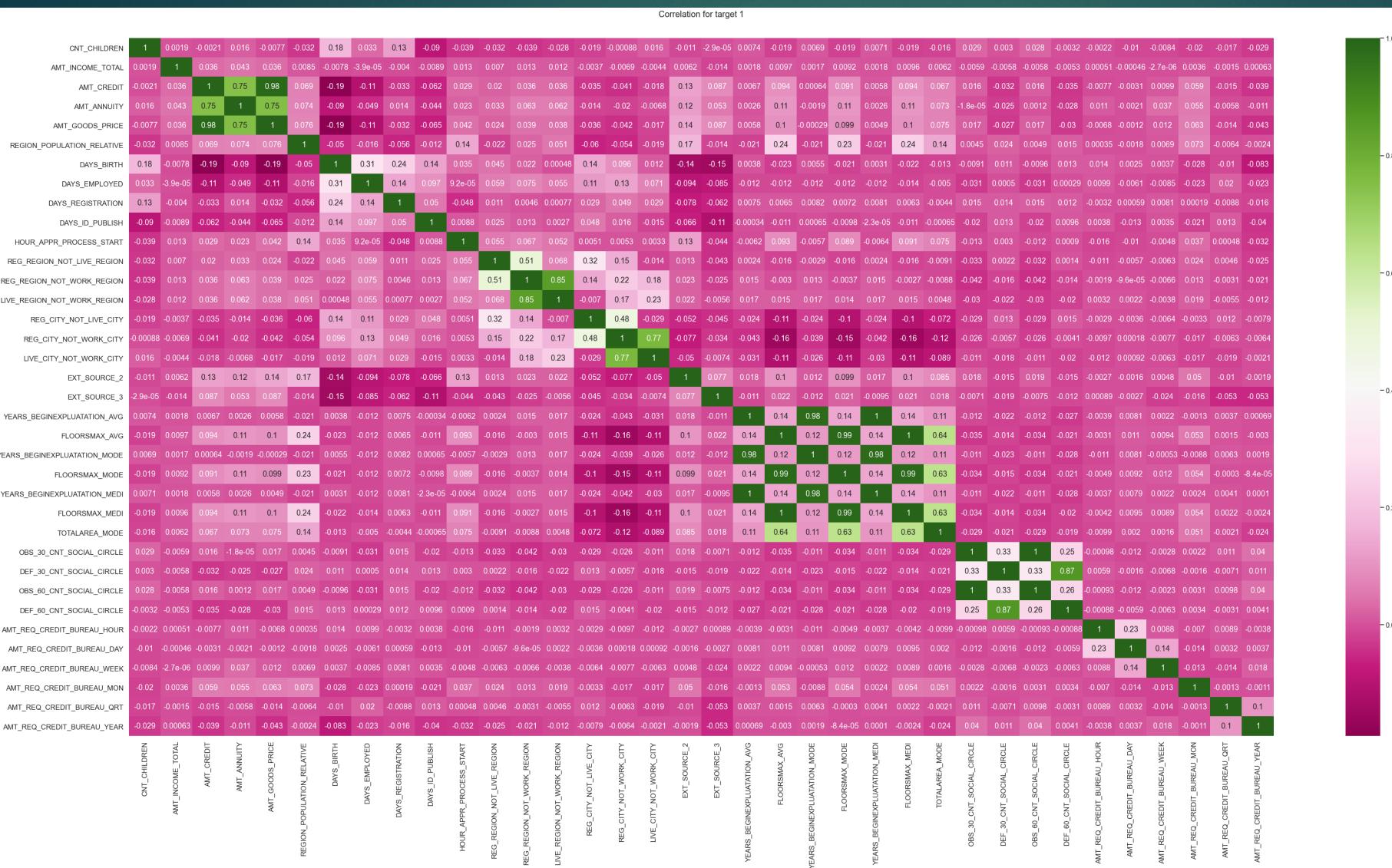
CORRELATION FOR TARGET 0



CORRELATION FOR TARGET 1

Inferences from the graph:

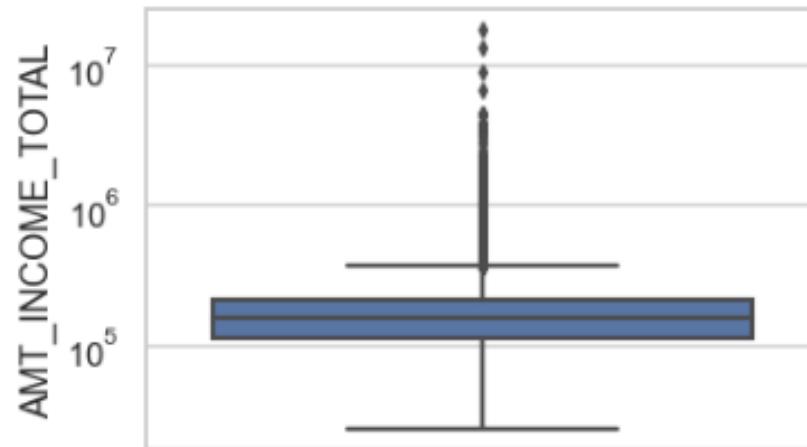
- ▶ This heat map for Target 1 is also having quite a same observation just like Target 0.
- ▶ If a client has his permanent address which is not same as work address, he/she is having less children and vice-versa
- ▶ Amount credited is less if client has his permanent address which is not same as contact address and same goes for work address



Univariate analysis for continuous variables

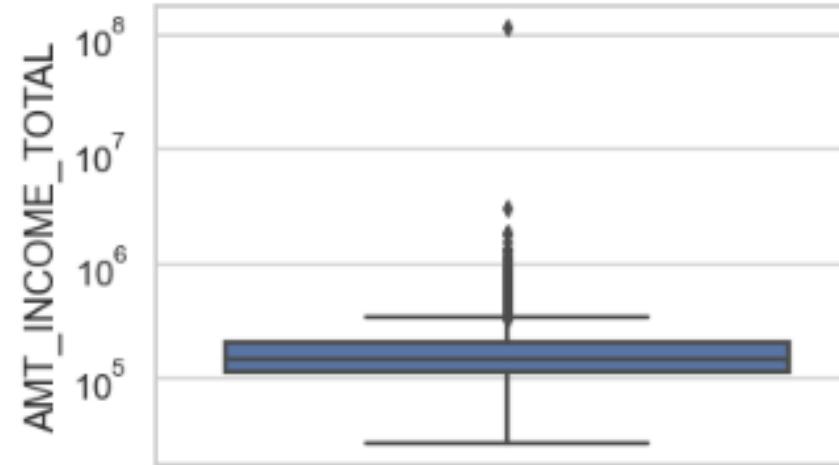
Distribution of AMT_INCOME_TOTAL

Distribution of total income amount for T-0



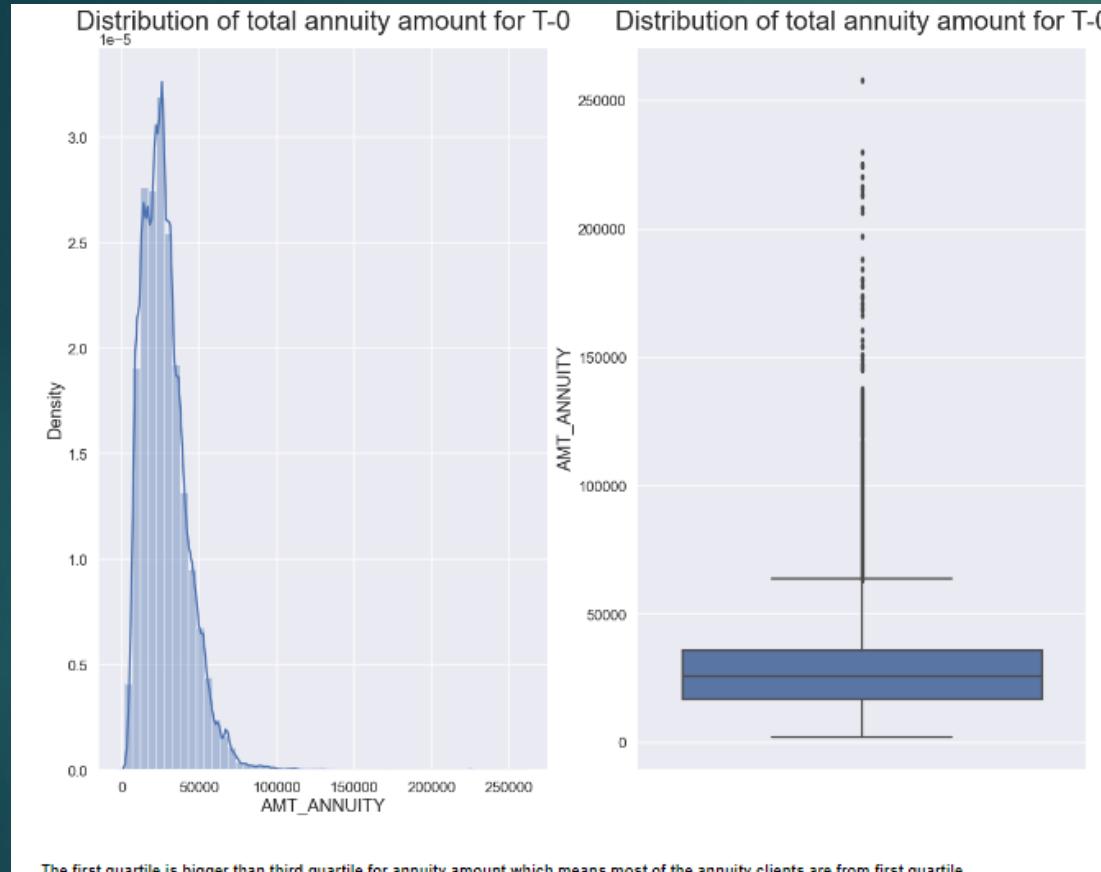
The third quartiles is very slim for income amount.

Distribution of total income amount for T-1

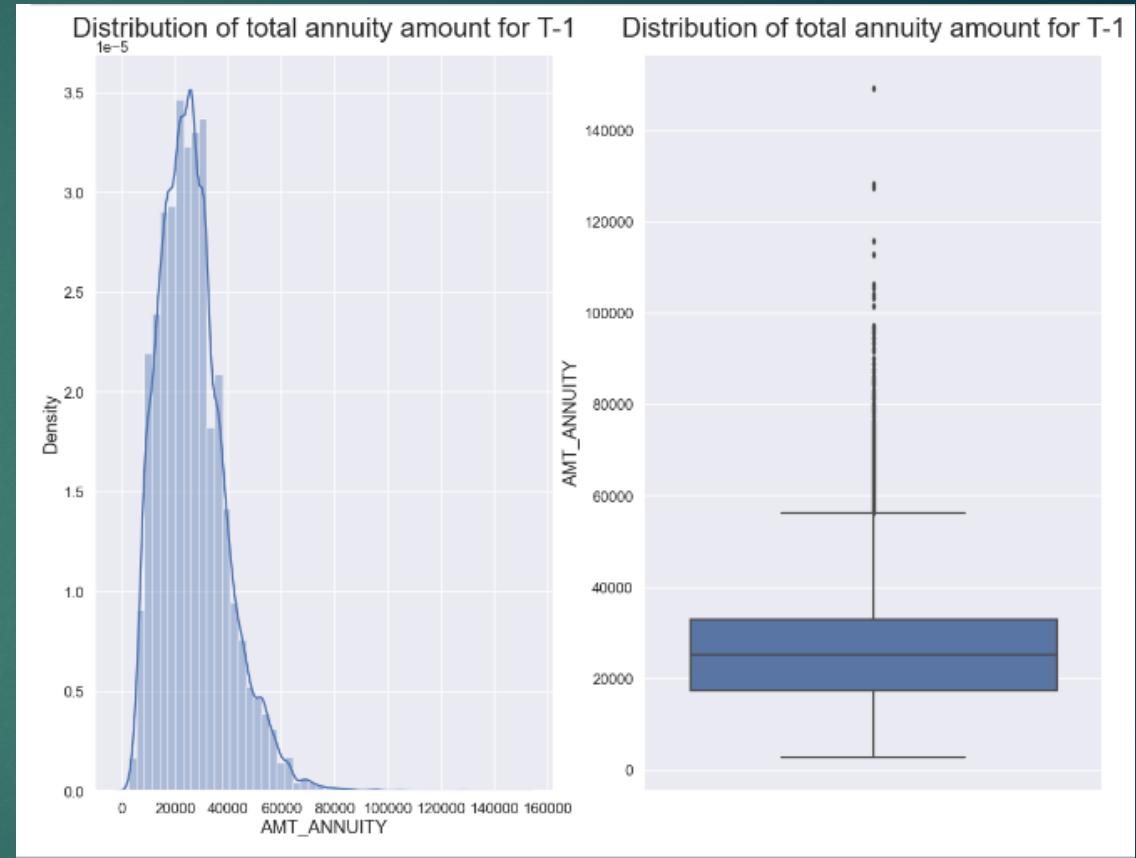


Most of the clients of income are present in first quartile.

Distribution of total annuity amount

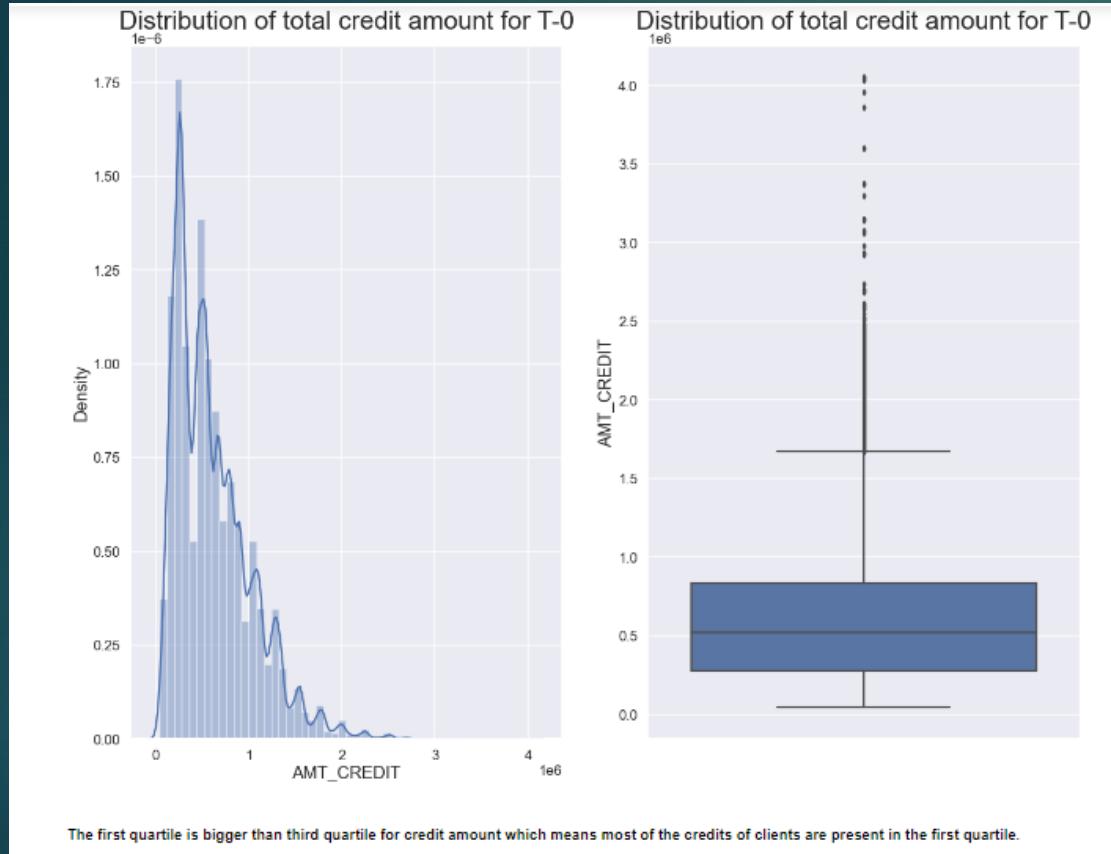


The first quartile is bigger than third quartile for annuity amount which means most of the annuity clients are from first quartile.

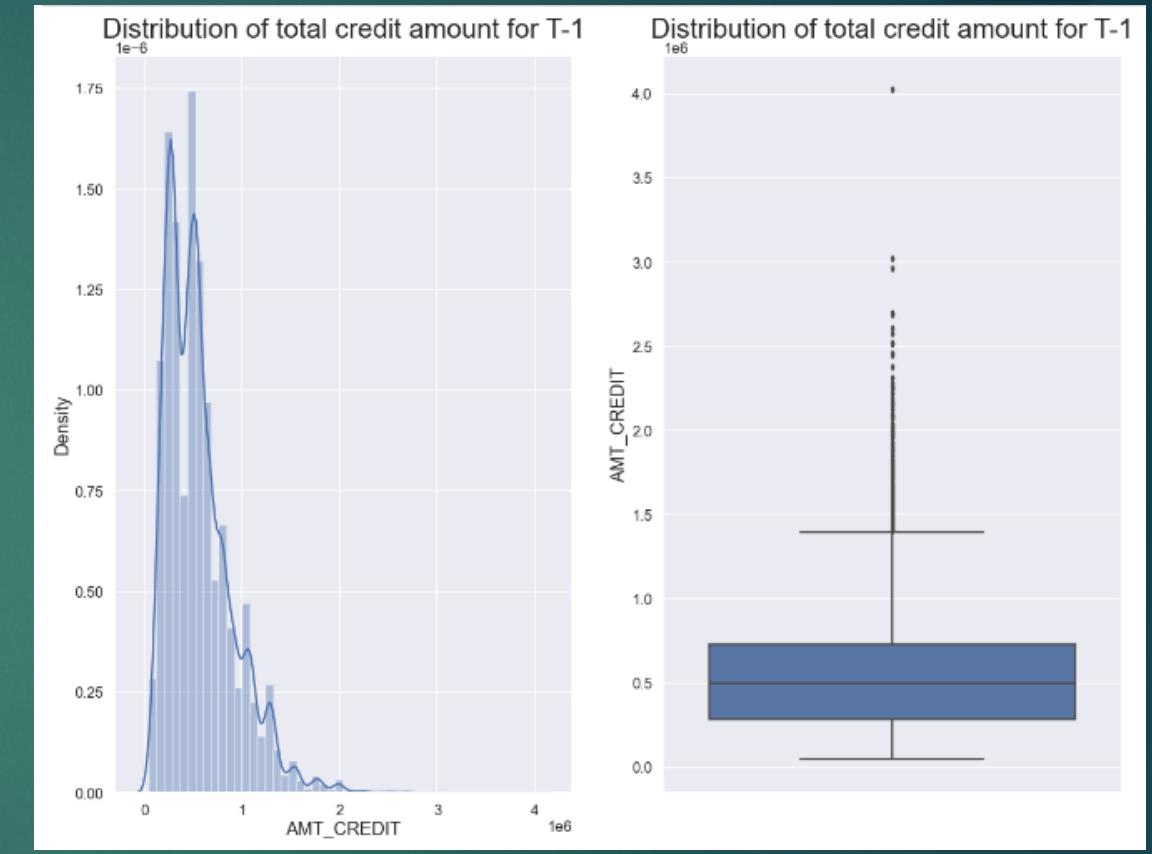


The first quartile is bigger than third quartile for annuity amount which means most of the annuity clients are from first quartile.

Distribution of total credit amount

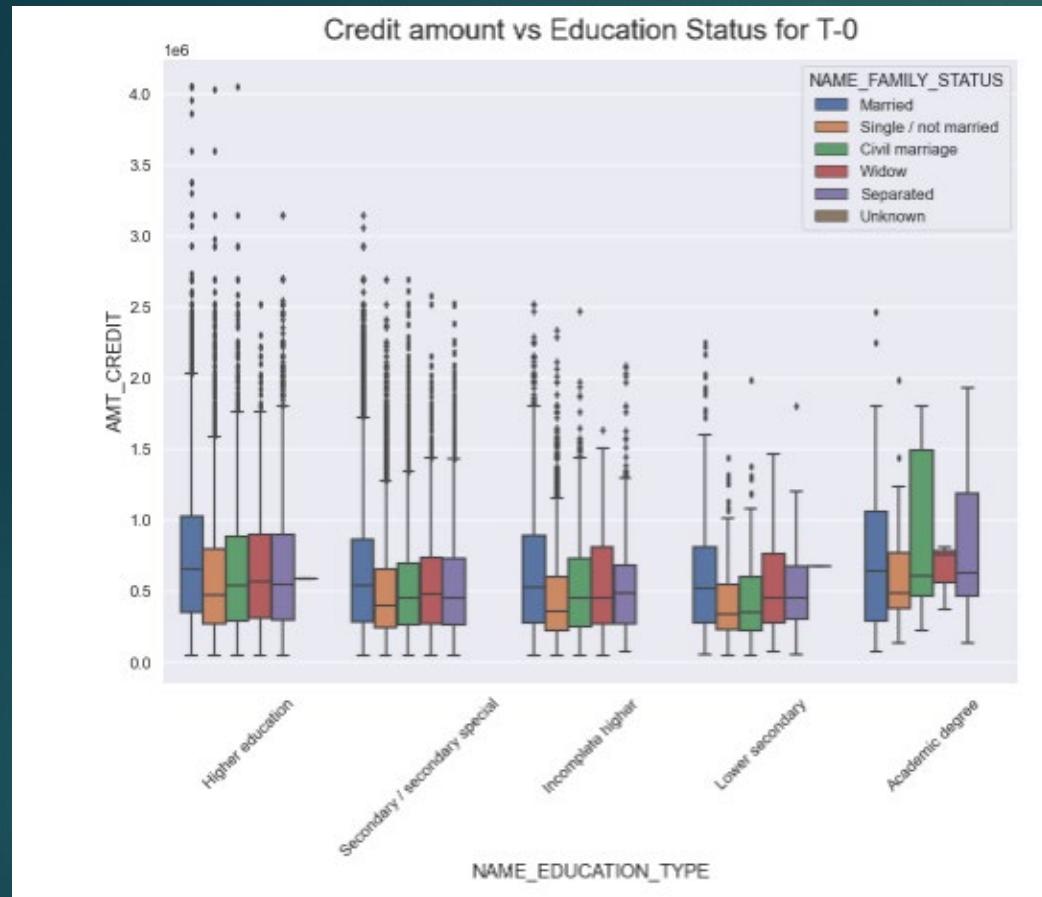


The first quartile is bigger than third quartile for credit amount which means most of the credits of clients are present in the first quartile.



The first quartile is bigger than third quartile for credit amount which means most of the credits of clients are present in the first quartile.

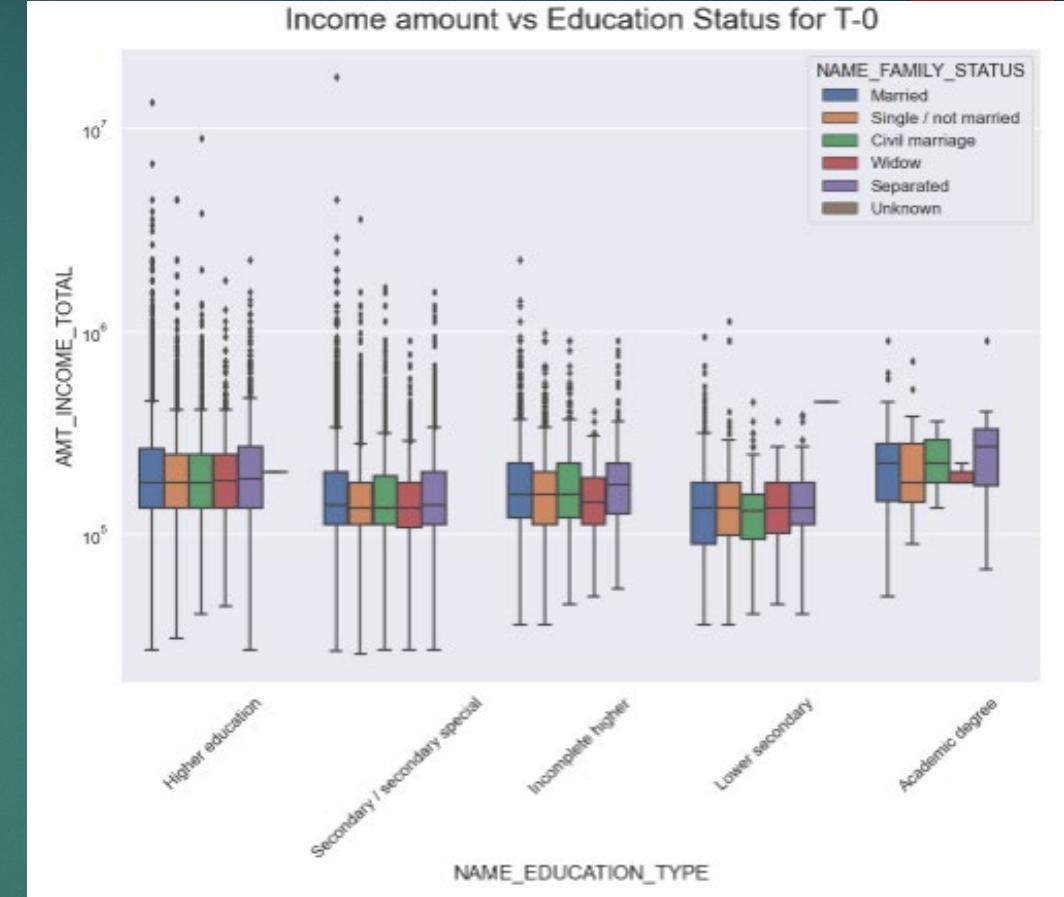
Bivariate analysis for variables- Target 0



Relationship between credit amount and education type

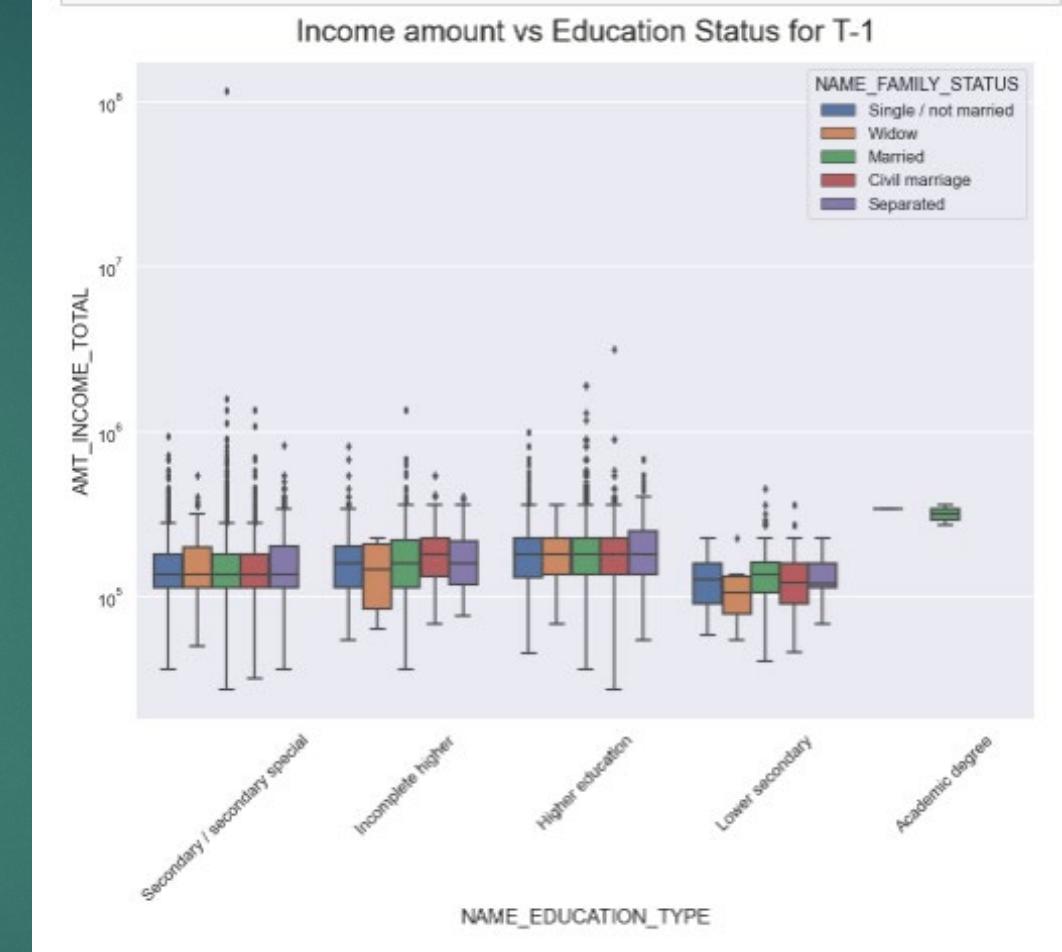
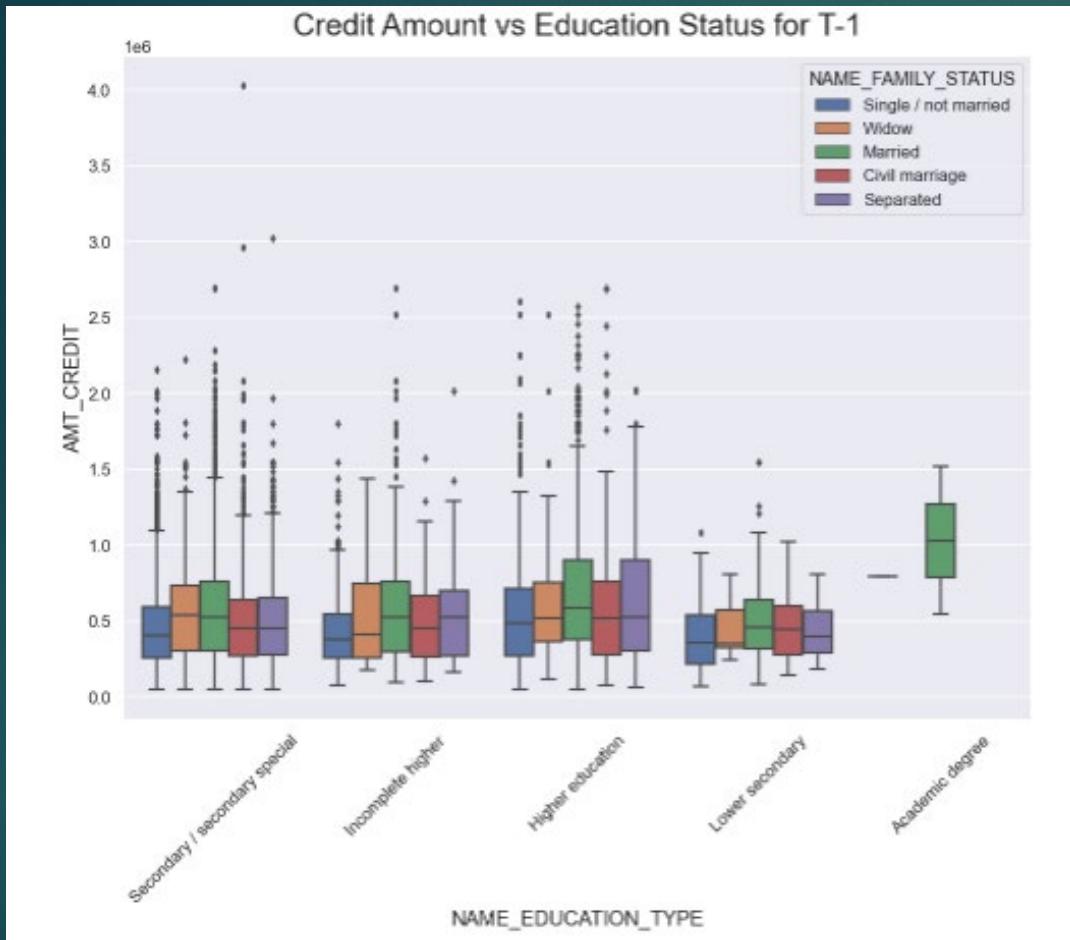
Inference:

- ▶ For Education type 'Higher education' the income amount is mostly equal with family status and has many outliers. Less number of outliers are present for Academic degree and the income amount is higher than that of Higher education.
- ▶ Lower secondary of civil marriage status have less income amount than others.



Relationship between income amount and education type

Bivariate analysis for variables- Target 1



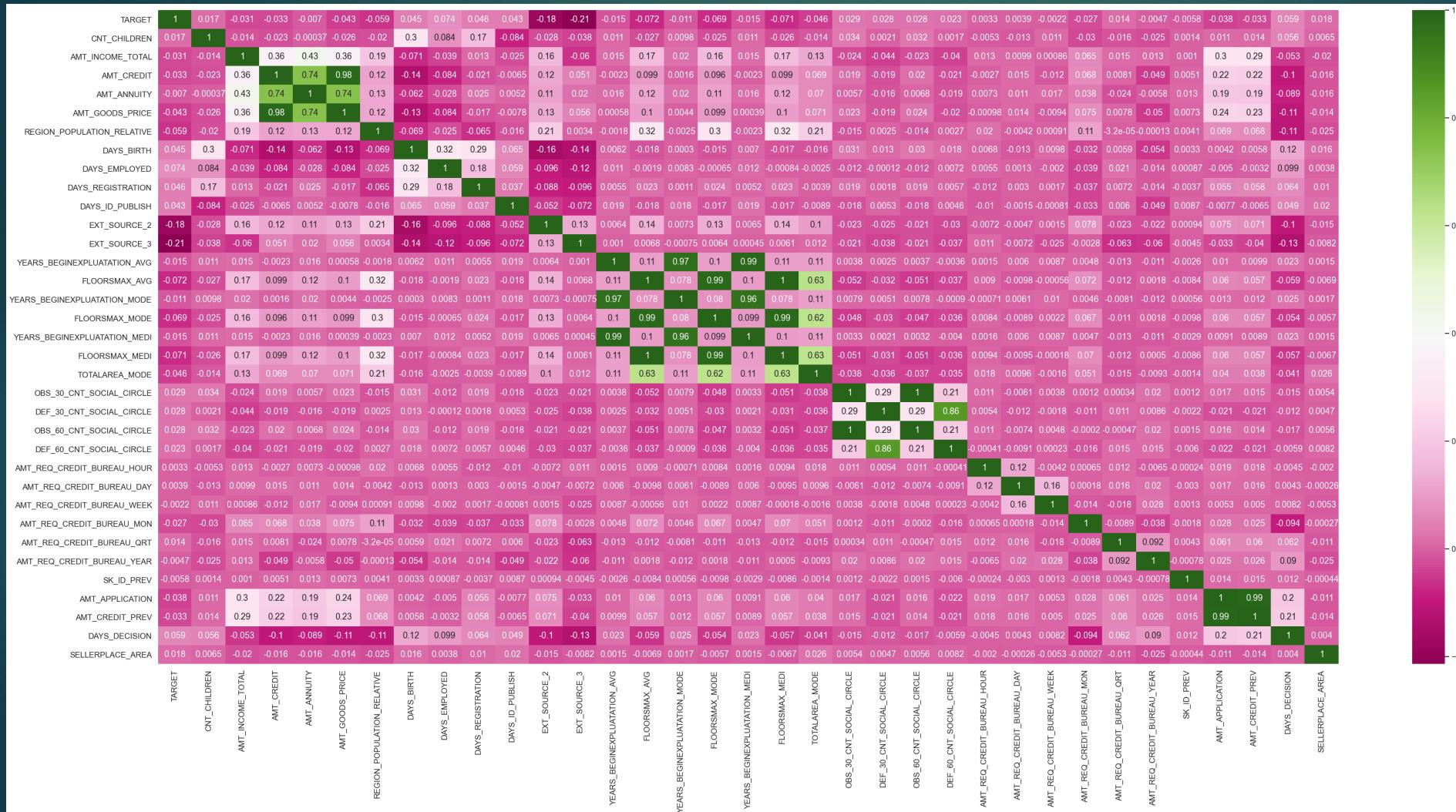
Relationship between credit amount and education type

Inference:

- ▶ T-1 plot has some similarity with T-0. For the 'Higher education' type, the income amount is mostly equal with family status. Less numbers of outliers are present for Academic degree and the income amount is higher than that of Higher education.
- ▶ Lower secondary are have less income amount than others. High number of outliers are present for Married status people who are secondary/ secondary special qualified.

Relationship between income amount and education type

Analysis with Previous loan application- Correlation





THANKYOU!!