# Credit EDA Case Study

BY : SOUKHYA H S,
MUKESH RANA

# Introduction

- This case study aims to identify patterns which indicate if a client has difficulty paying their installments
- This will ensure that the consumers capable of repaying the loan are not rejected.
- In other words, the company wants to understand the driving factors (or driver variables) behind loan default, i.e. the variables which are strong indicators of default.

# Title and Content Layout with List

- Univariate Analysis
  - A. Categorial Data
  - B. Continuous Data

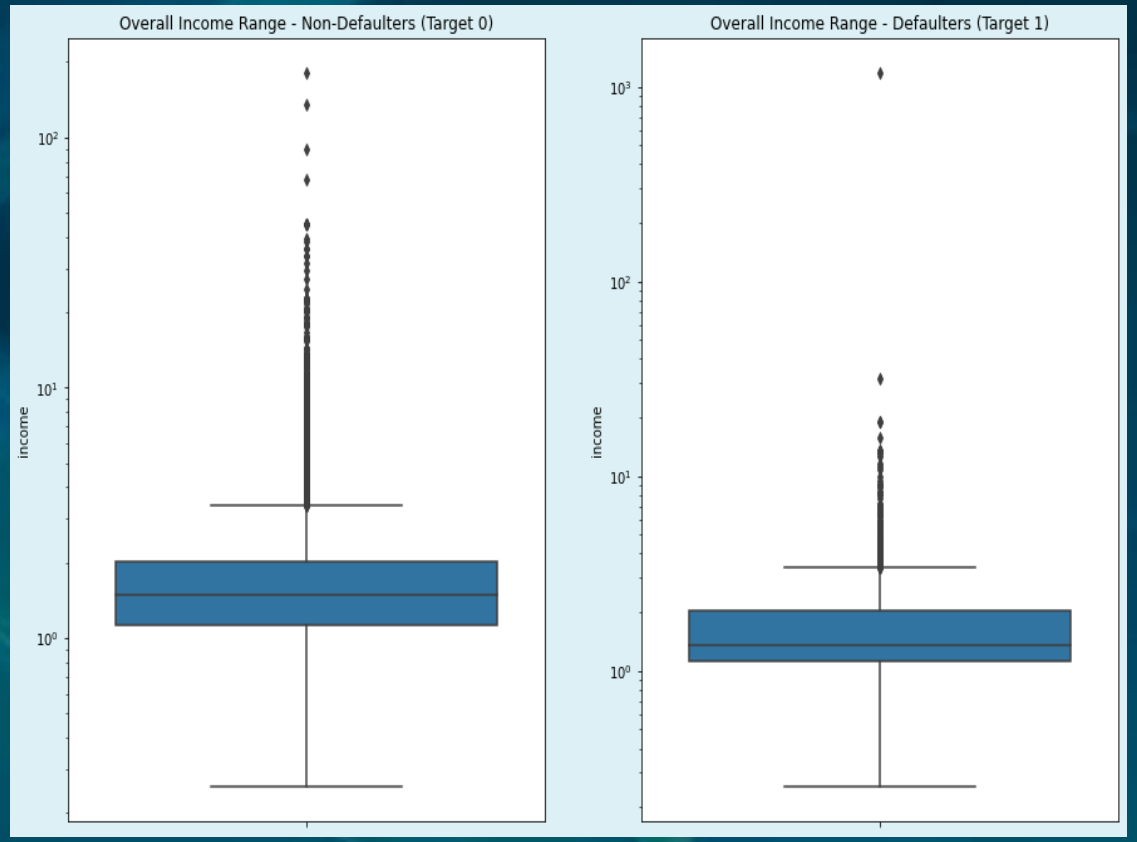
- Bivariate Analysis
  - A. Categorial Data
  - B. Continuous Data

- Multivariate Analysis

# 1. Univariate Analysis

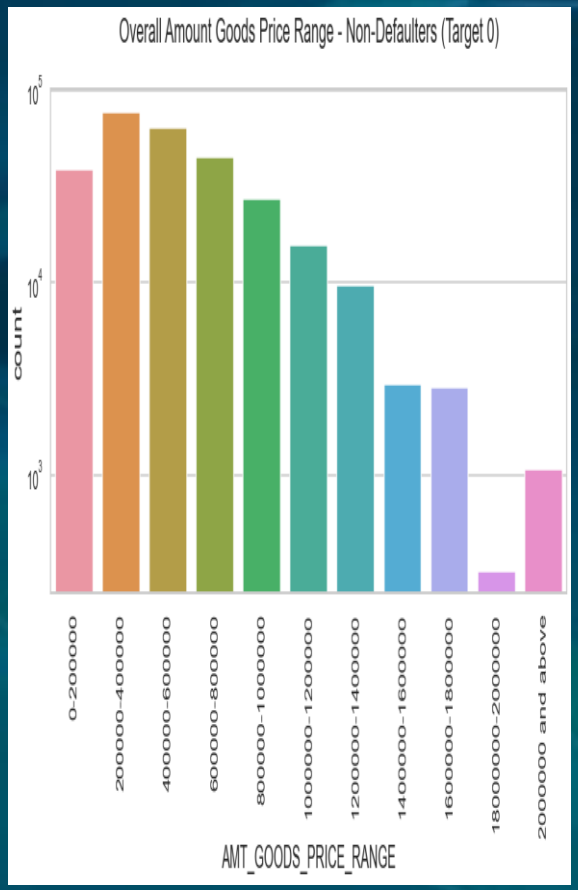
# ➢Univariate Analysis of Numerical Data (Income)

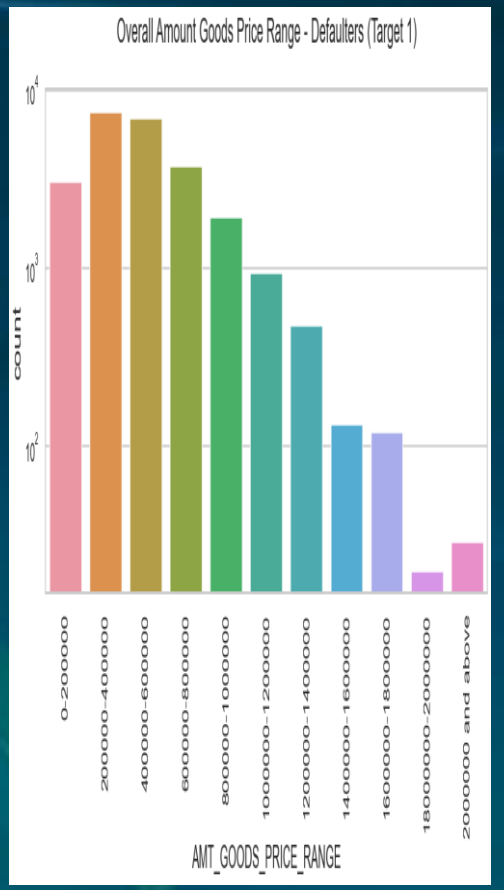
- We can see that the income range for Non-Defaulter is wider than Defaulter applicants.
- The average Income of Non-Defaulter is greater than Defaulter.
- So in this case Its better to give the loan to Higher income holder.



# ➢Univariate Analysis of Numerical Data (AMT_GOODS_PRICE_RANGE)

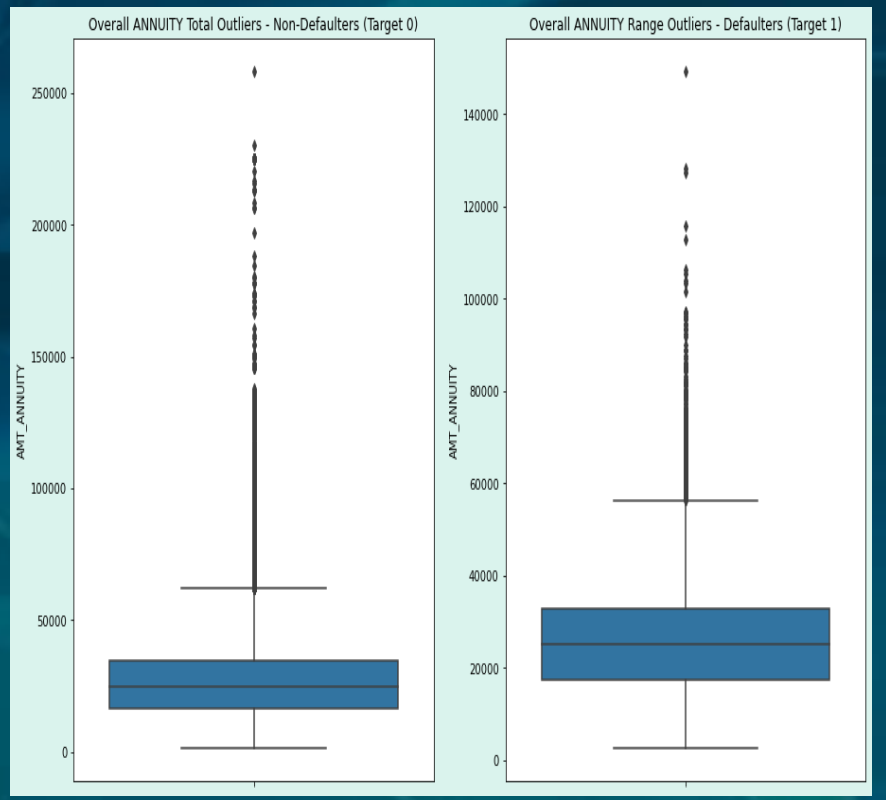
- From these charts Its clear that applicants having more Goods amount are more inclined towards Non-Defaulter.
- So it is safer to lend the loans to applicants owning more goods.

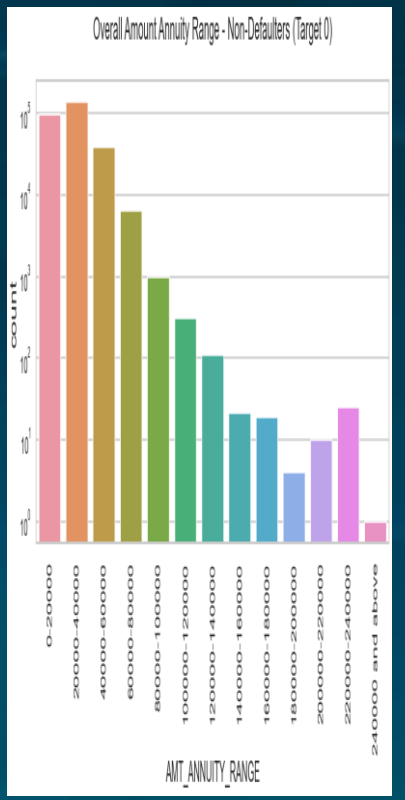


# ➢Univariate Analysis of Numerical Data (Annuity Amount)

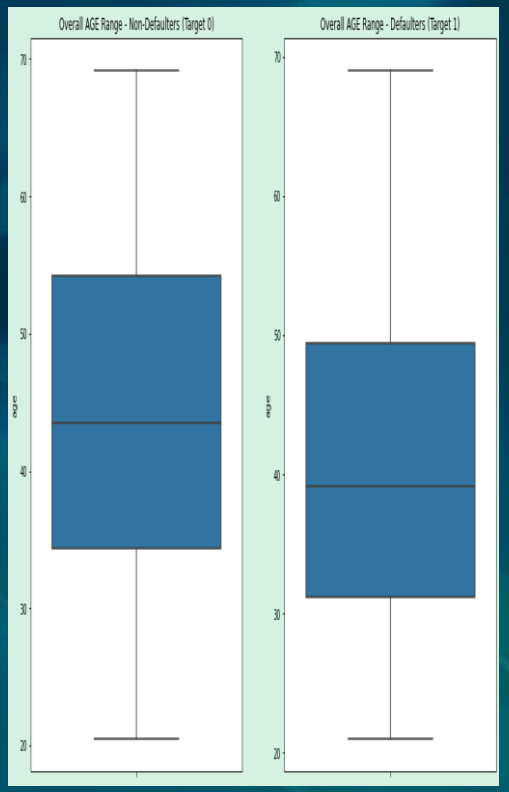
- It clearly showing that Non-Defaulters has less annuity amount.
- Hence, if the annuity amount is higher then applicants faces difficulties in paying back.
- So in this case Its better provide the loan at lesser or at the amount which applicants are capable to pay.
- From 3er chart we can say that, It is safer to keep the annuity amount lesser than 80000.

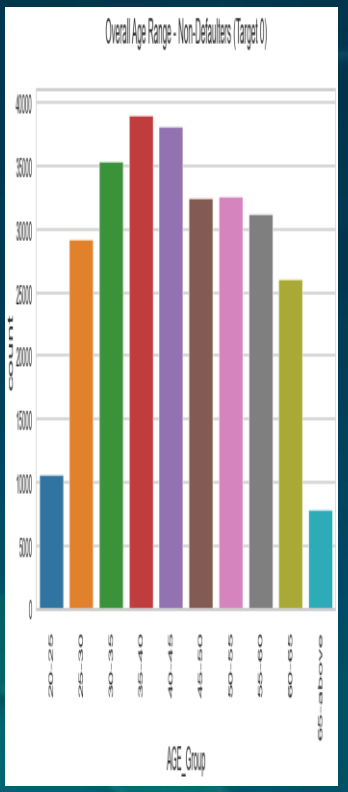


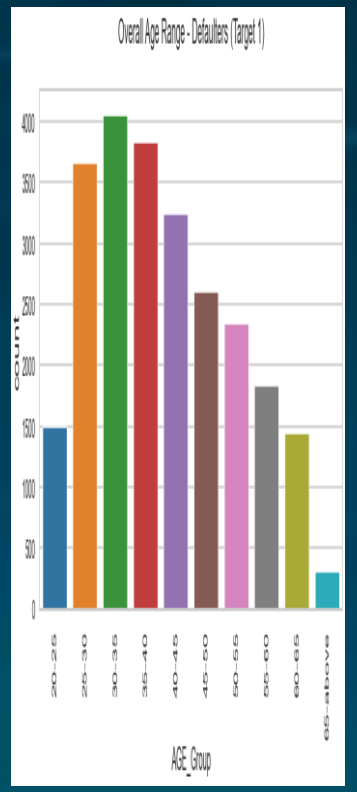



# ➢Univariate Analysis of Numerical Data (Age)

- If we talk about average age of applicants, that Defaulters are younger than Non-Defaulter
- The average age of Non-Defaulter is nearly 45 and for defaulter it is 39.
- Most of the defaulters comes under 25-40 age group, whereas most of the Non-Defaulters comes under 30-45age group.
- Here, It seem that it is safer to give loans to elders(more than 45 year old applicants) as they has less chances to be defaulter.

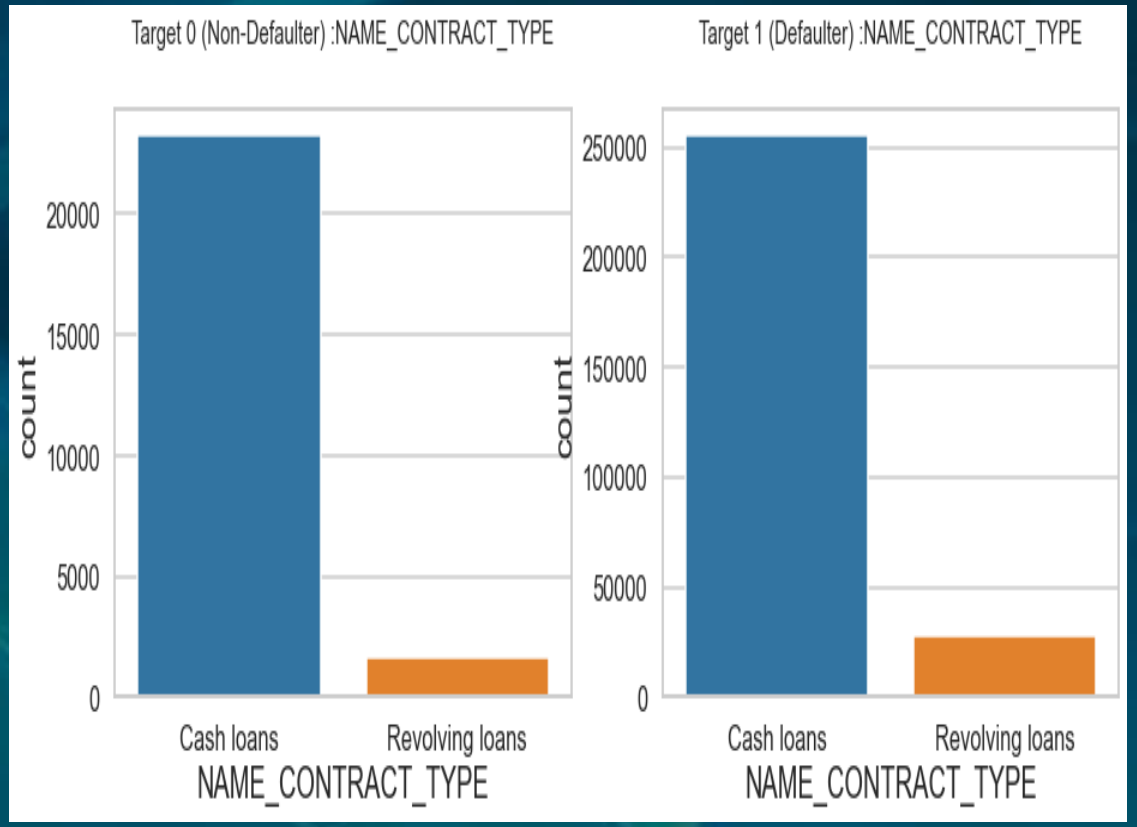


# ➢Univariate Analysis of Categorical Data (NAME_CONTRACT_TYPE)

- We can see that there is minor difference between Defaulter and non-Defaulter taking cash and Revolving Loans.
- Here we can say that Defaulter takes more loans  for cash and Revolving both.
- So here, because of minor difference we cannot say anything.

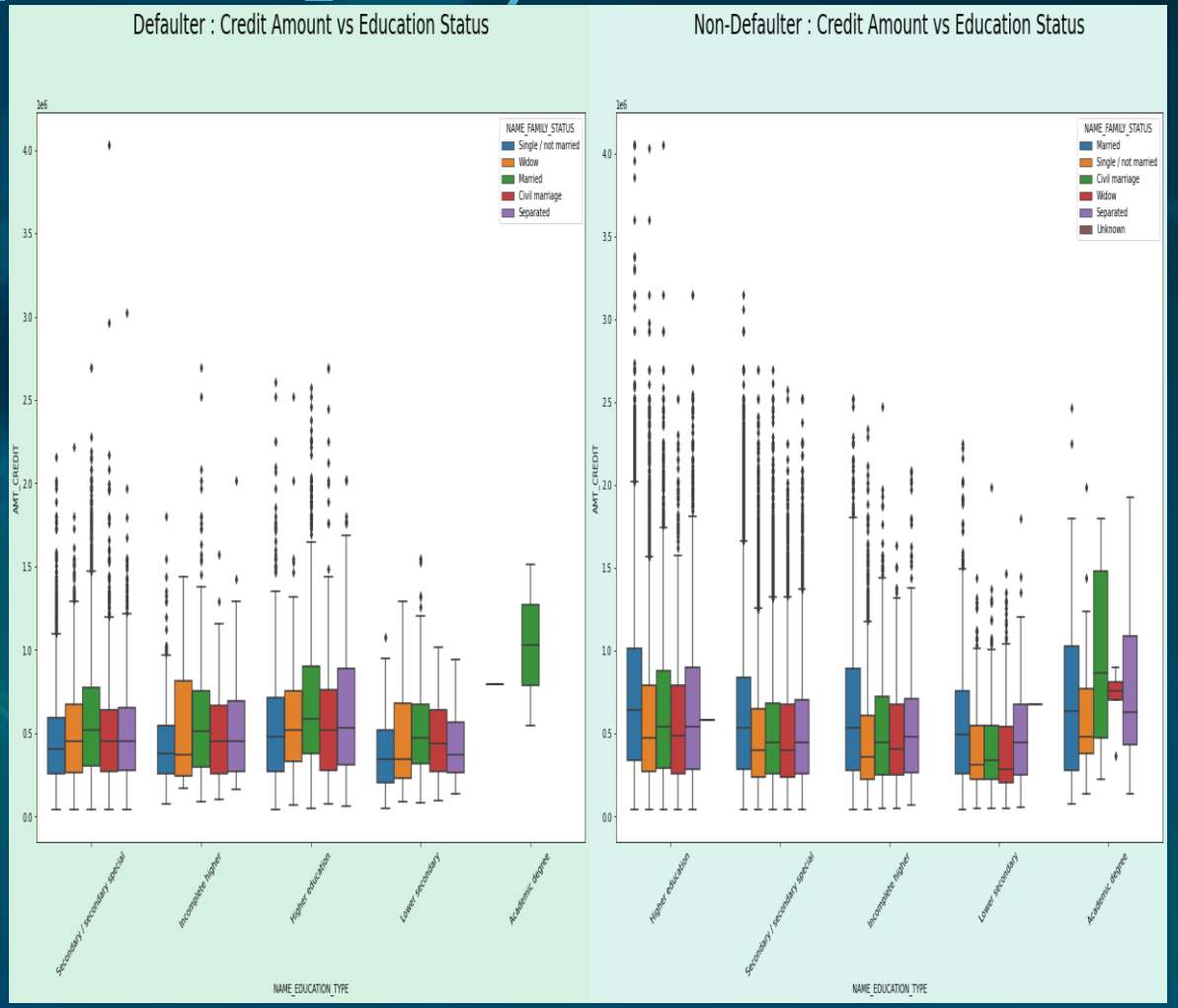


# 2. Bivariate Analysis

# ➢Bivariate Analysis of Categorical Data

## (NAME_EDUCATION_TYPE vs AMT_CREDIT)

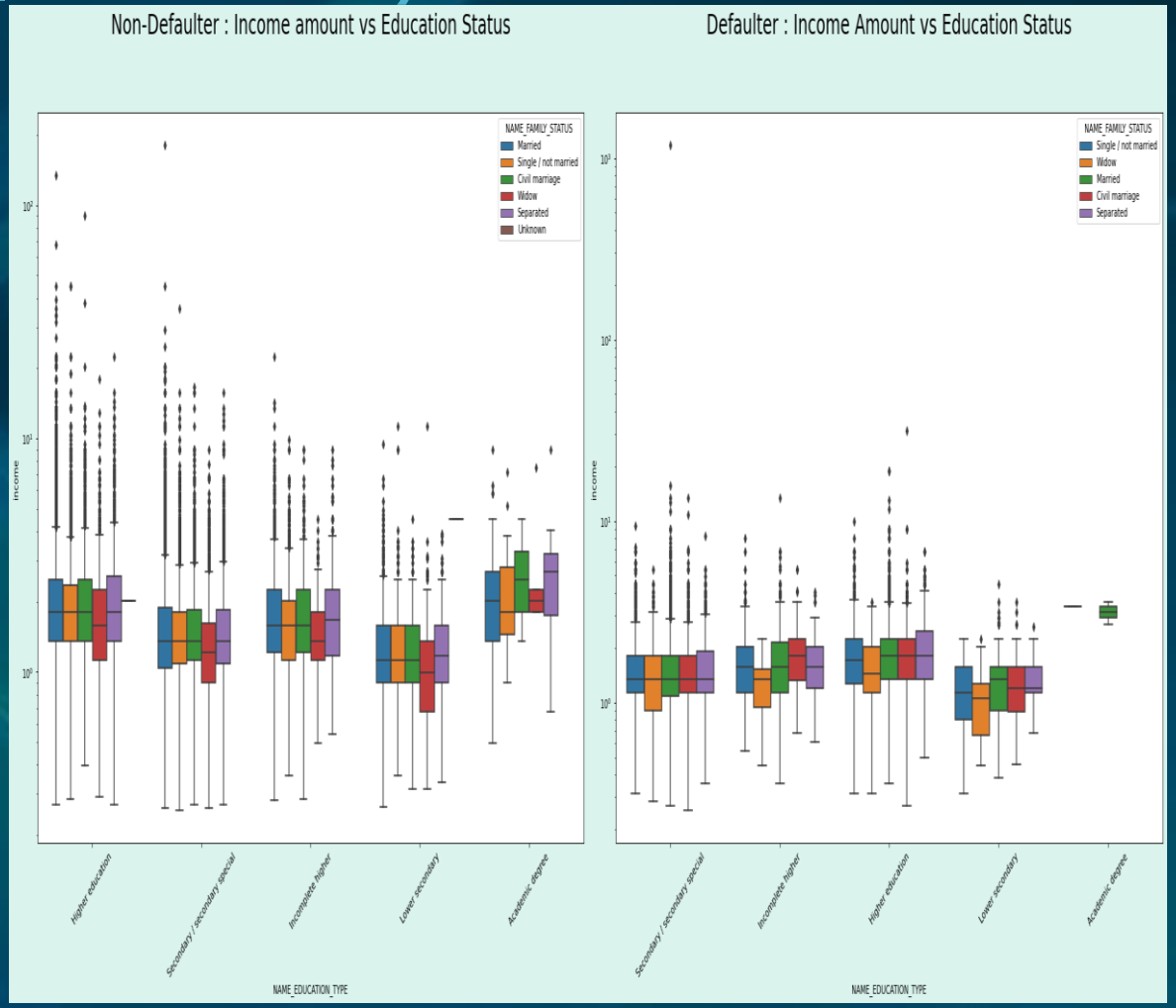
- For every level of education non-Defaulter has much spread-out in terms of credit.
- At the same time Defaulter has narrower rang of credit.
- It is safer to lend the money to Academic Degree holder as they are less tends to become Defaulter.



# ➢Bivariate Analysis of Categorical Data

## (NAME_EDUCATION_TYPE vs Income)

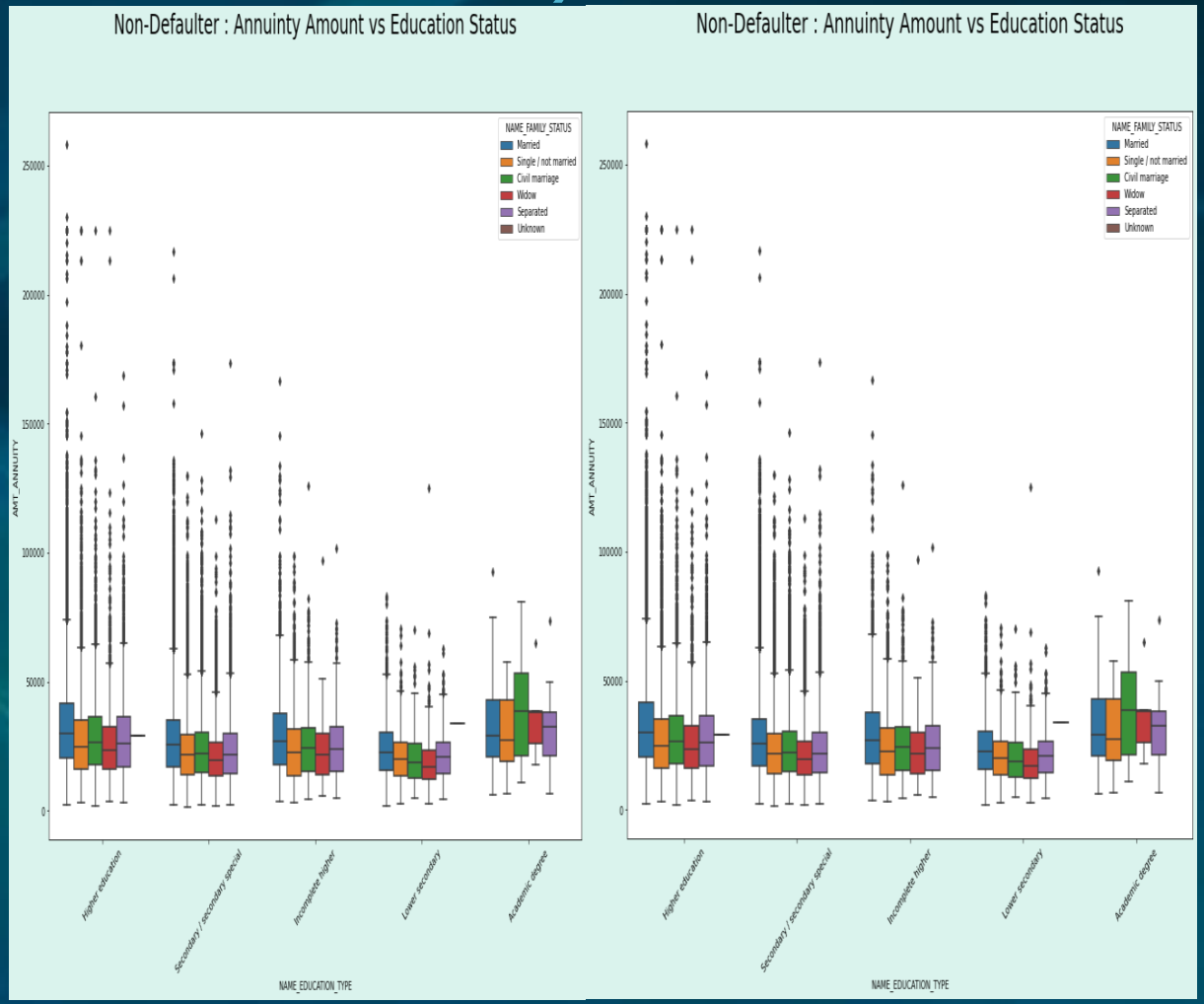
- As we have already seen that it is safer to lend money to higher income holder.
- Here, Highly Educated applicants has more spread of their income, hence they are safer applicants.
- Almost all of the Academic Degree holder are Non-Defaulter.



# ➢Bivariate Analysis of Categorical Data

## (NAME_EDUCATION_TYPE vs AMT_ANNUITY)

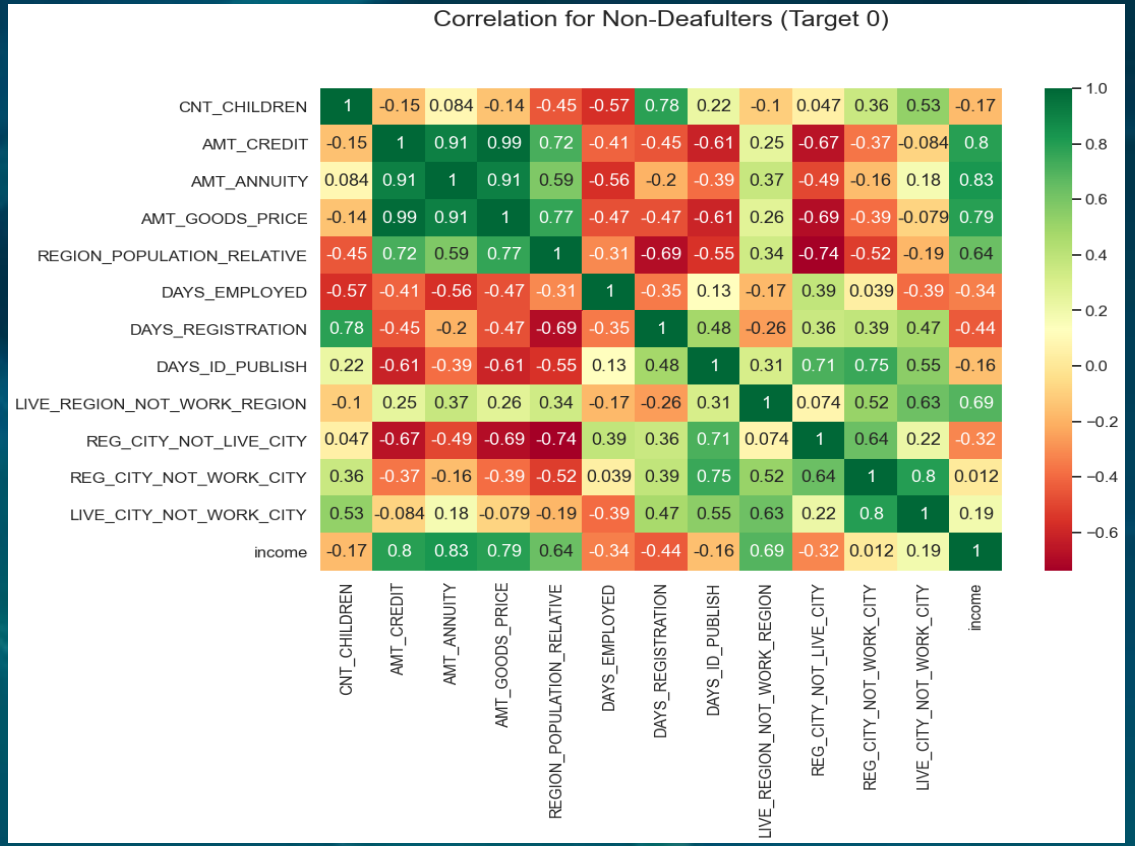
- We have already seen that, lesser the annuity amount lesser chances of default.
- Here, married applicants tends to have more annuity amount in all levels of education.



# 3. Multivariate Analysis

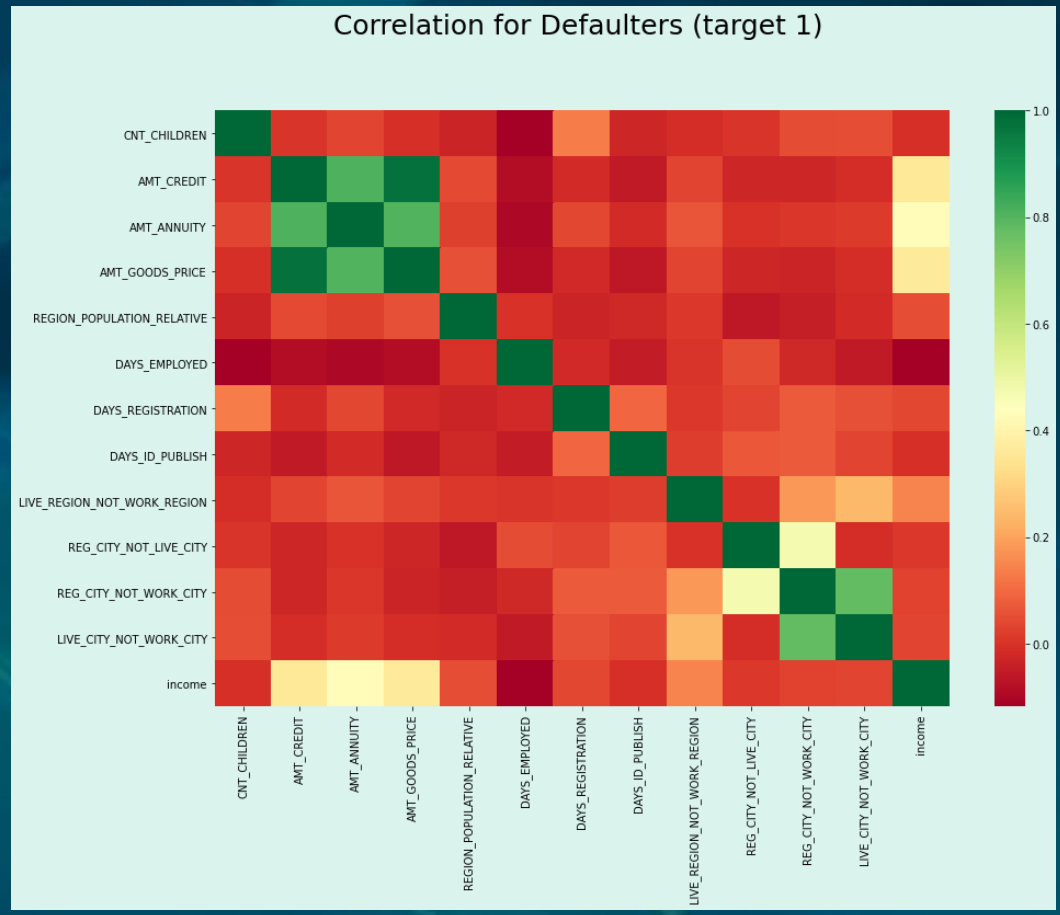
# Multivariate Analysis
## (Correlation for Non-Defaulter)

- Most of the variables has higher correlation coefficient.
- For example annuity amount and credit amount are highly correlated.
- Similarly for the applicants owning more goods has very high credit amounts and so on.



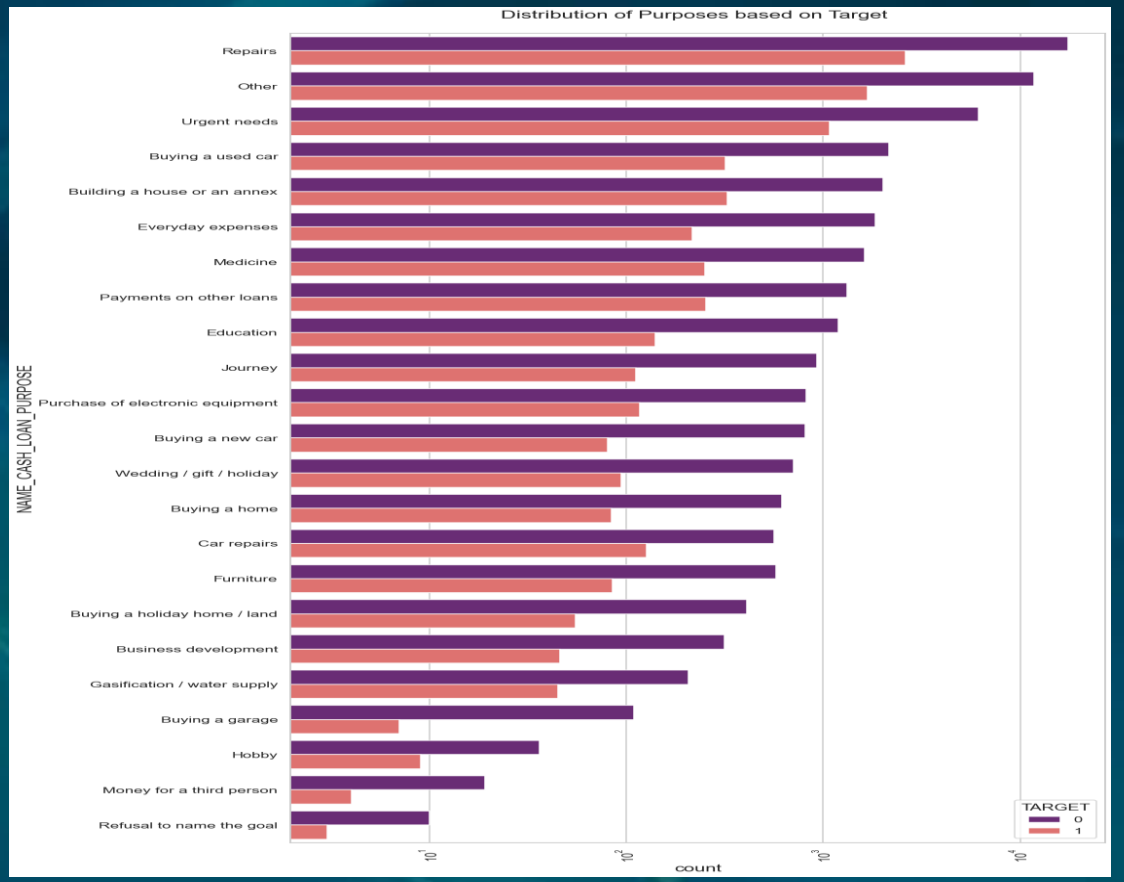
# Multivariate Analysis
## (Correlation for Defaulter)

- For Defaulters, Credit amount, annuity amount, Goods amount are highly correlated to each other.
- Whereas most of the variable/factors has negative correlation coefficient.



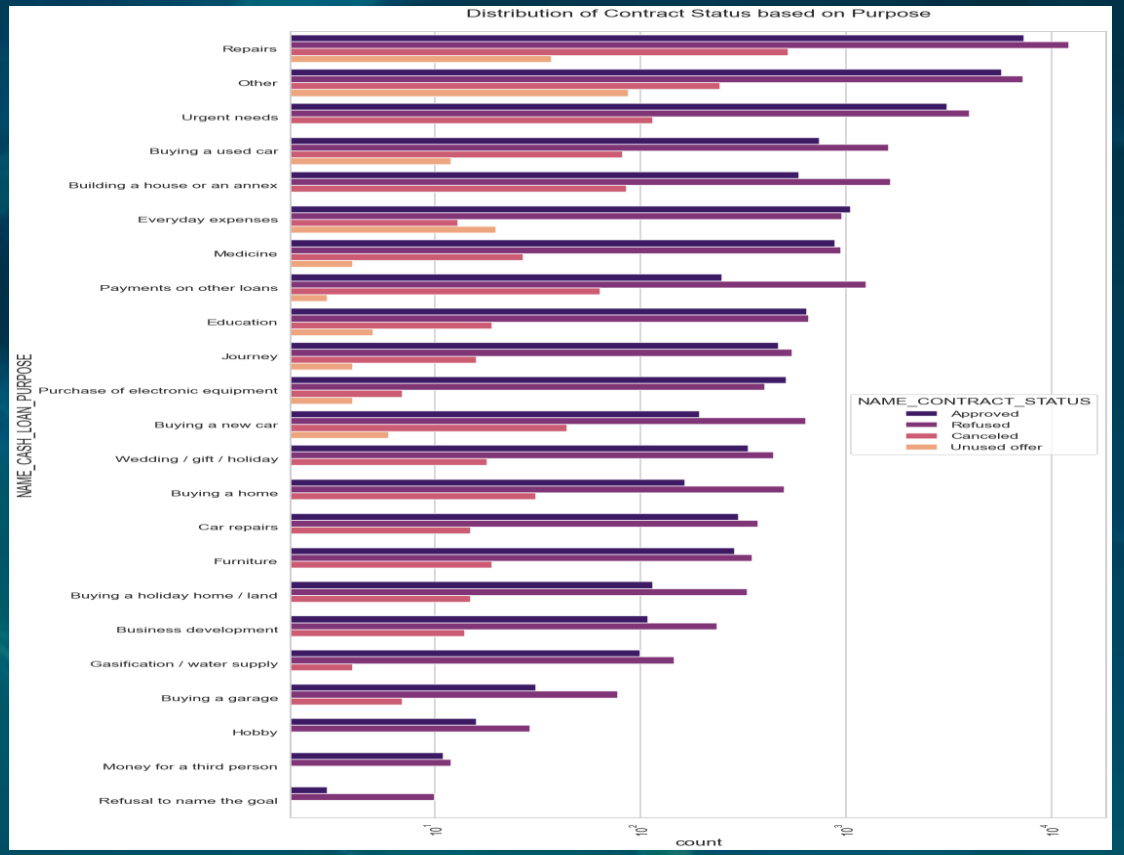
# Multivariate Analysis
## (Purposes based on Target)

- From this Non-Defaulter chart we can conclude that Repairs, Others, and Urgent need are very popular purpose of lending loans.
- All purposes has chances of default as well as chances of non-default. But, chances of non-default is always more in all purposes.



# Multivariate Analysis
## (Contract Status based on Purpose)

- This shows the loan status for different purpose.
- Repairs, Others, Urgent needs are the purpose which are approved and rejected maximum time.
- If the purpose is Purchasing electronic equipment then there is high chances of getting approved, otherwise all purposes has high chances of Rejection of application.



# CONCLUSION

- It is safe to give loan to Applicants who are holding "Higher Degree" or holding "Academic Degree".
- Among "Academic Degree" holders people having Family status other then "Married" are the most safest category to give loan.
- It is very unsafe to give loan to "Widow" Applicants who have studied only till "Lower Secondary".
- Loan purposes with "Repairs" are facing more difficulties in payment on time.
- Its not safe to give loans to people living in Co-op apartment.
- It is safe to give loan to the people living with there parents as there is very less chance of Defaults.
- Banks should focus more on contract type "Student" ,"pensioner" and "Businessman" with housing "type other than "Co-op apartment" for higher rate of successful payments.

# Thank You