



# EDA BANK LOAN RISK ANALYSES CASE STUDY

Chandana Konduru  
Shilpa

# PROBLEM STATEMENT

- When the company receives a loan application, the company has to decide for loan approval based on the applicant's profile. Two types of risks are associated with the bank's decision:
- If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company
- If the applicant is not likely to repay the loan, i.e. he/she is likely to default, then approving the loan may lead to a financial loss for the company.
- 
- The data given below contains the information about the loan application at the time of applying for the loan. It contains two types of scenarios:
- **The client with payment difficulties:** he/she had late payment more than X days on at least one of the first Y instalments of the loan in our sample,
- **All other cases:** All other cases when the payment is paid on time.

# DATA AVAILABLE

- This dataset has 3 files as explained below:

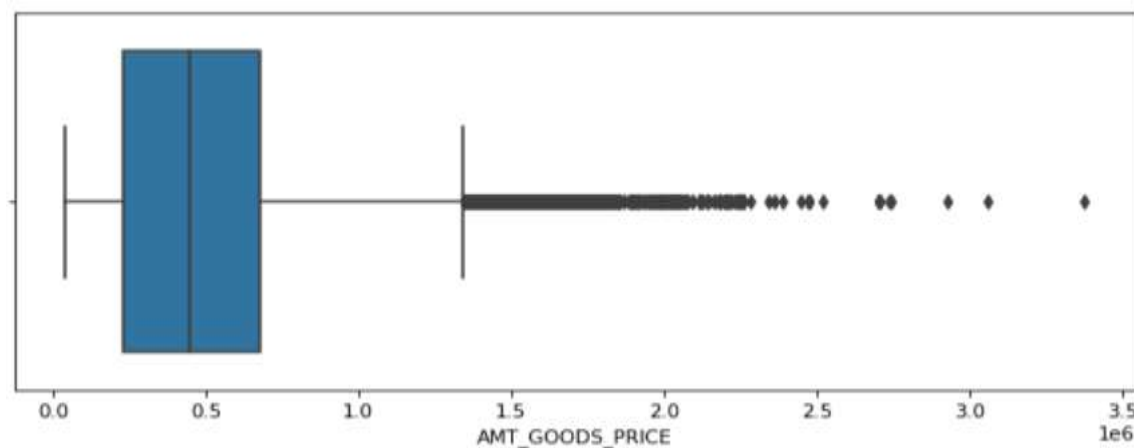
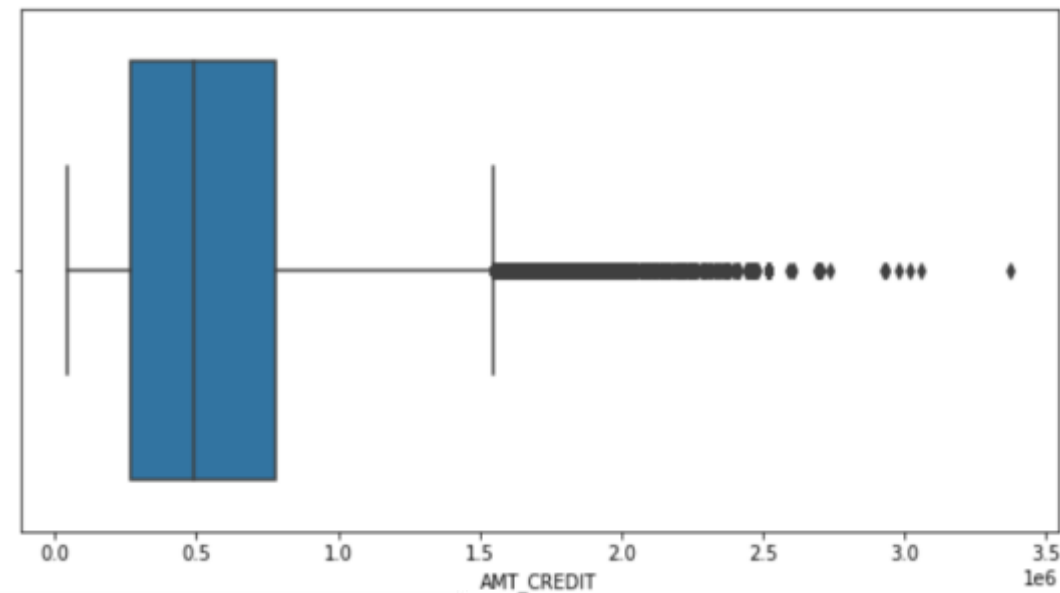
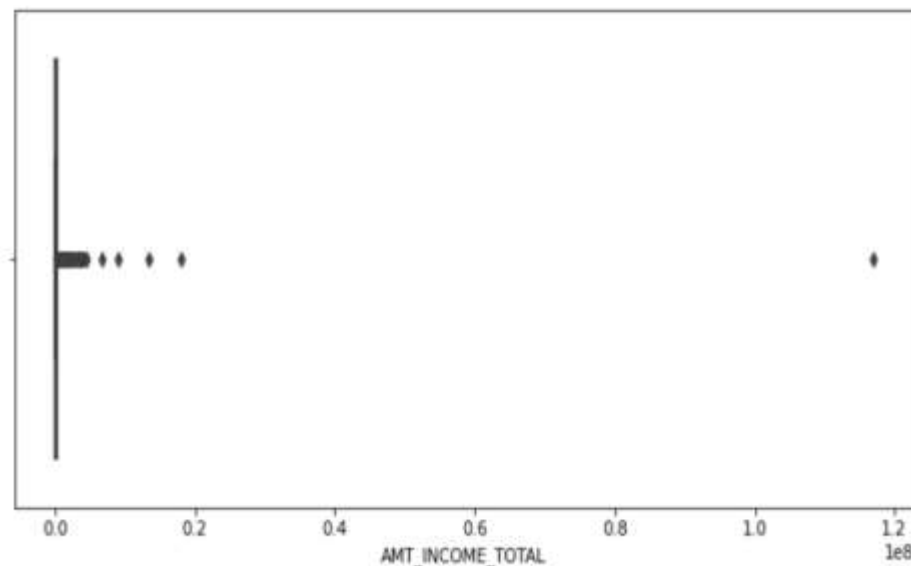
1. '*application\_data.csv*' contains all the information of the client at the time of application.

The data is about whether a **client has payment difficulties**.

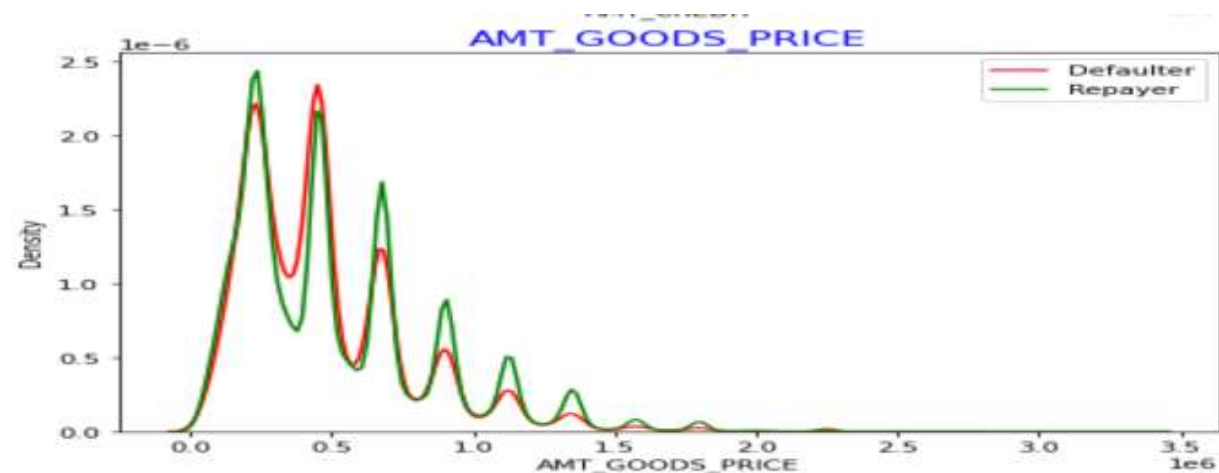
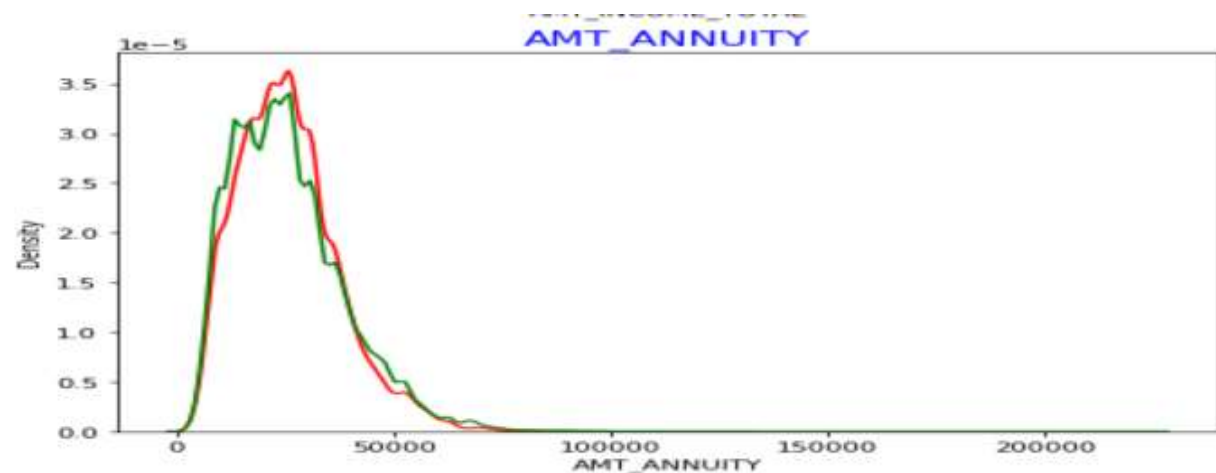
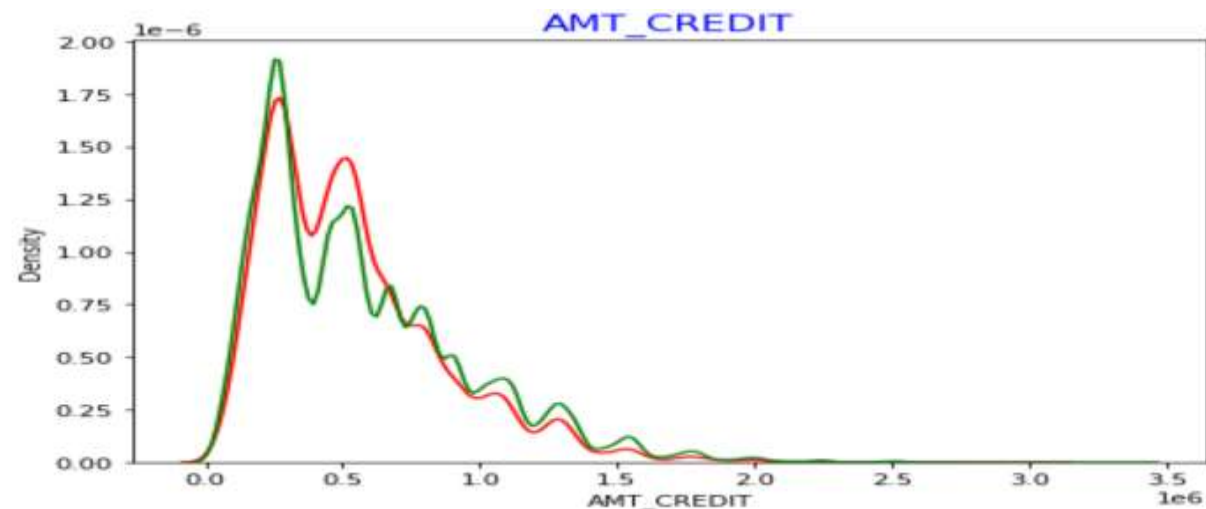
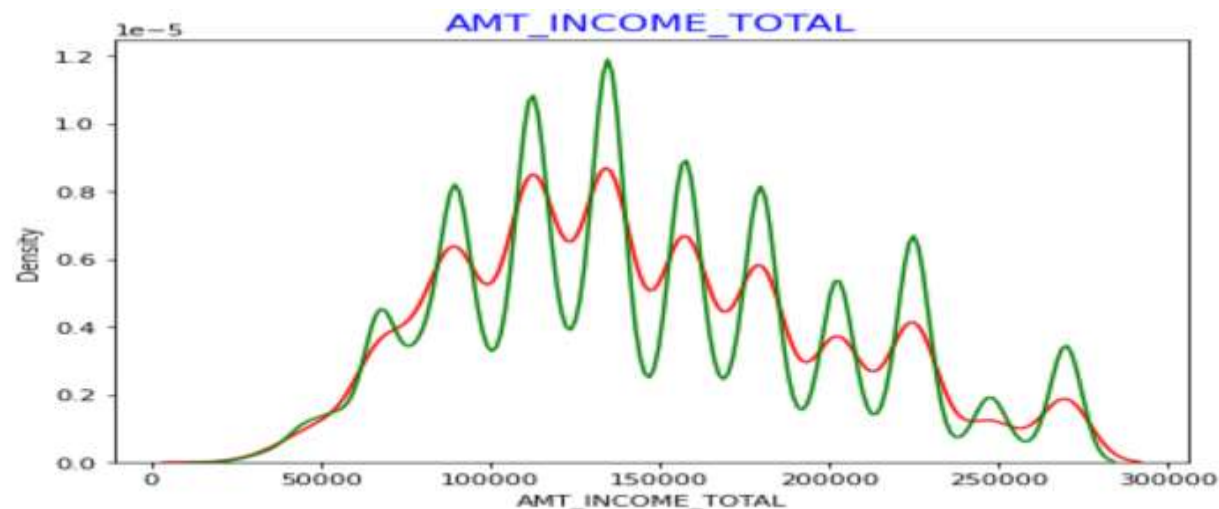
2. '*previous\_application.csv*' contains information about the client's previous loan data. It contains the data whether the previous application had been **Approved, Cancelled, Refused or Unused offer**.

3. '*columns\_description.csv*' is data dictionary which describes the meaning of the variables

# OUTLITES ANALYSIS



# NUMERICAL DATA ANALYSIS



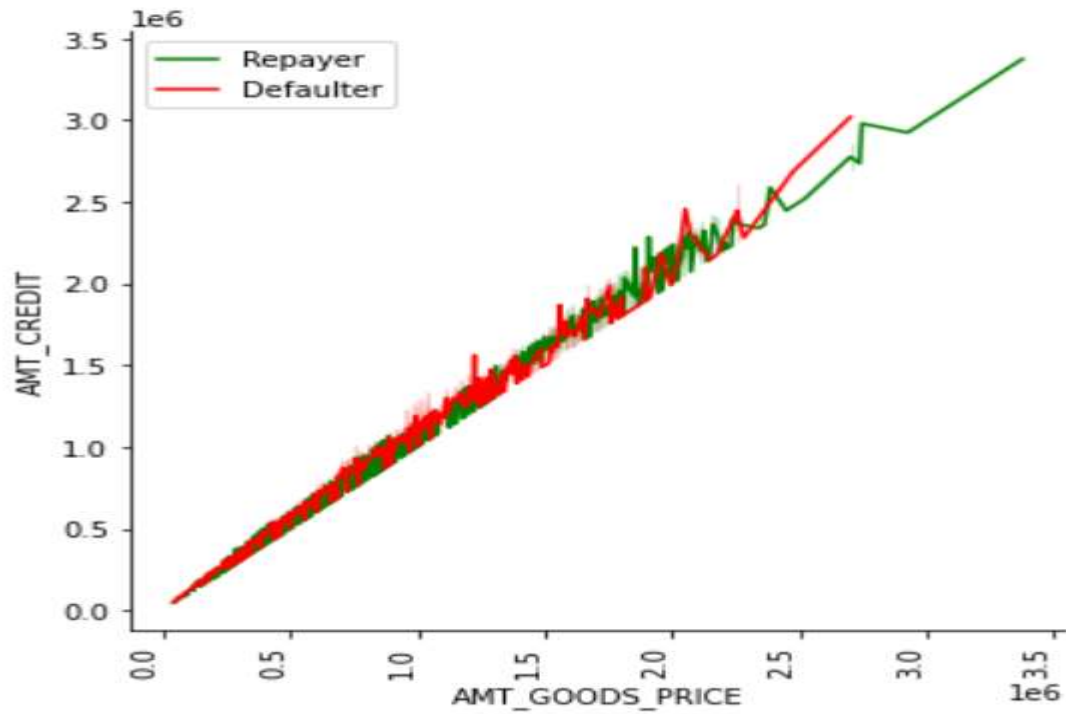




# NUMERICAL DATA ANALYSIS

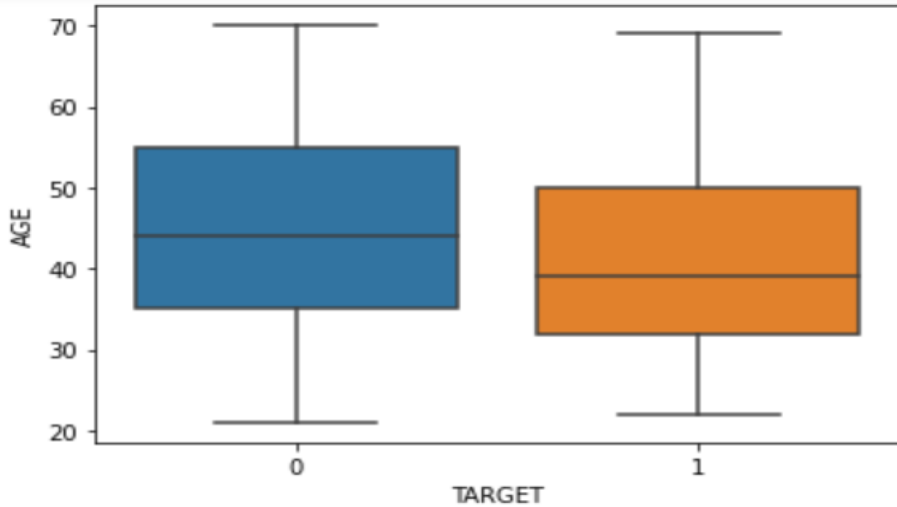
- Most no of loans are given for goods price below 10 lacs
- Most people pay annuity below 50000 for the credit loan
- Credit amount of the loan is mostly less then 10 lacs
- The re-payers and defaulters distribution overlap in all the plots and hence we cannot use any of these variables in isolation to make a decision

# NUMERICAL BIVARIATE ANALYSIS



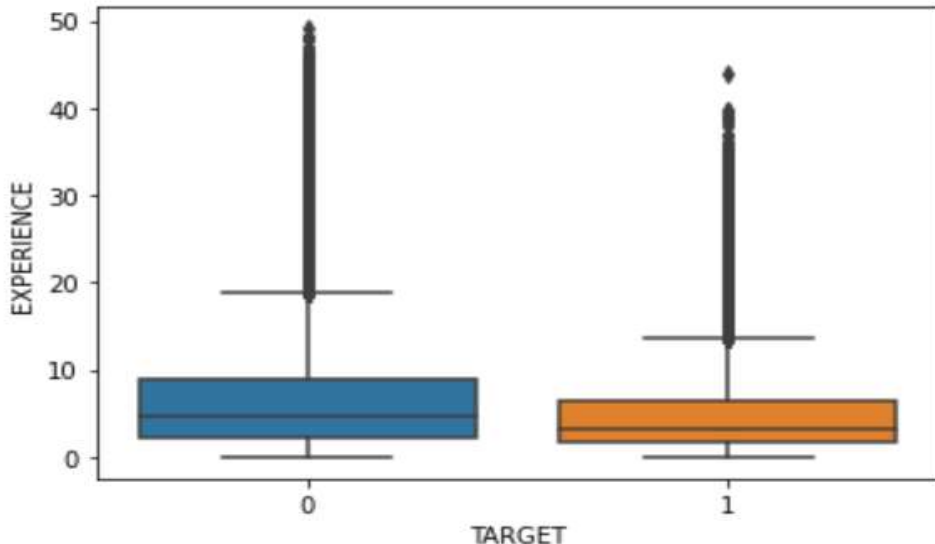
- When the credit amount goes beyond 3M, there is an increase in defaulters

# DATA NUMERICAL ANALYSIS



Comparing age and loan payments, we can say that on average client who has less age (<40) are more likely to make delay in making lone payment

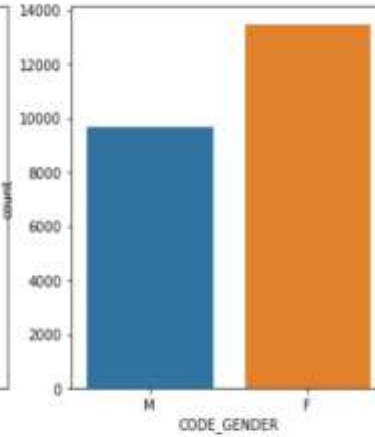
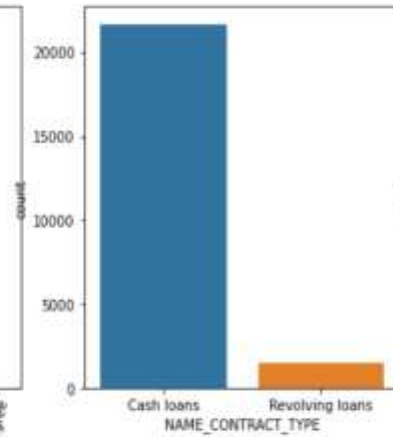
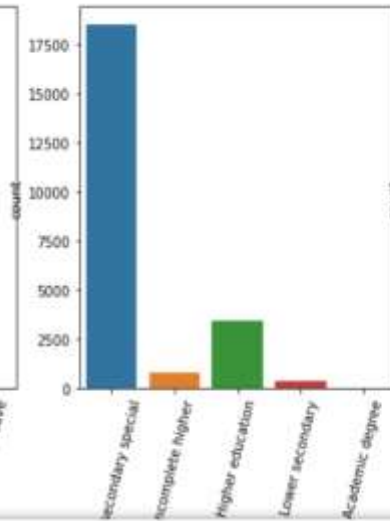
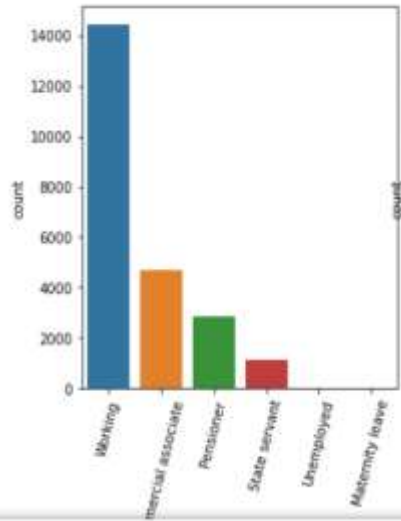
still if we compare lone payment with EXPERIENCE, average of payments is mostly common but the client who has less experience are more likely to make delay in payment if we check for upper extrema and 75% of payments



- As the client gets older are get more experienced, they are more likely to repay their loans on time more often



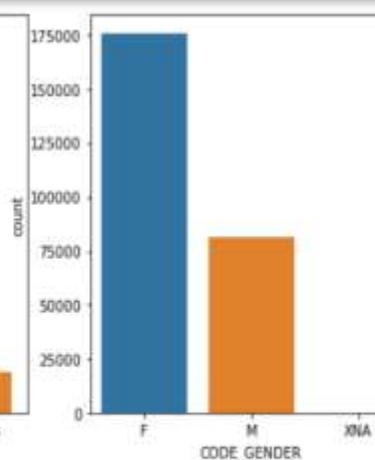
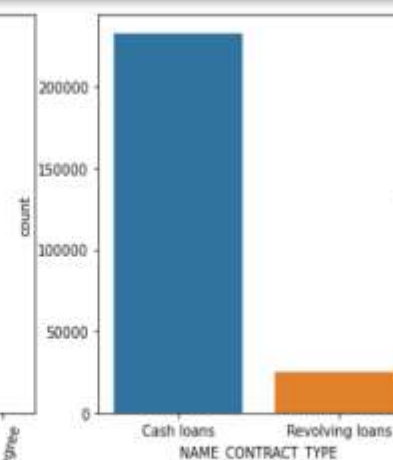
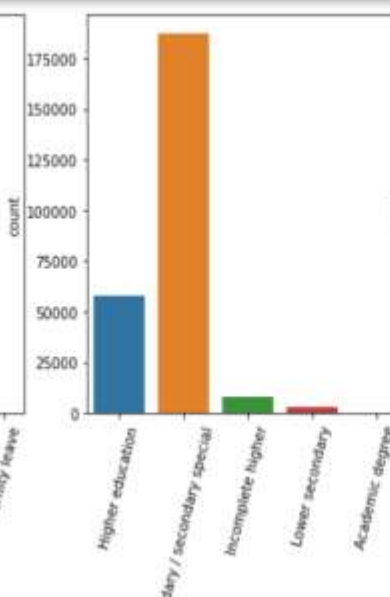
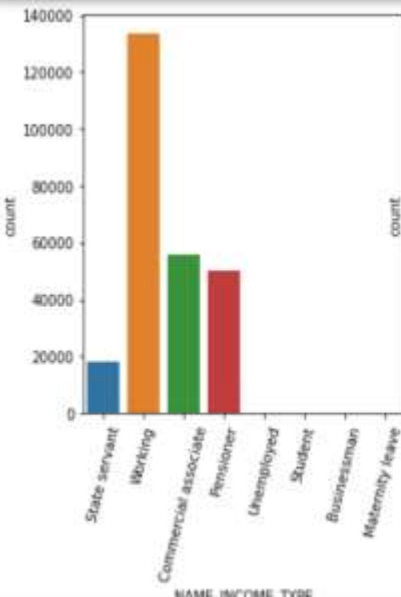
# CATEGORICAL DATA ANALYSIS

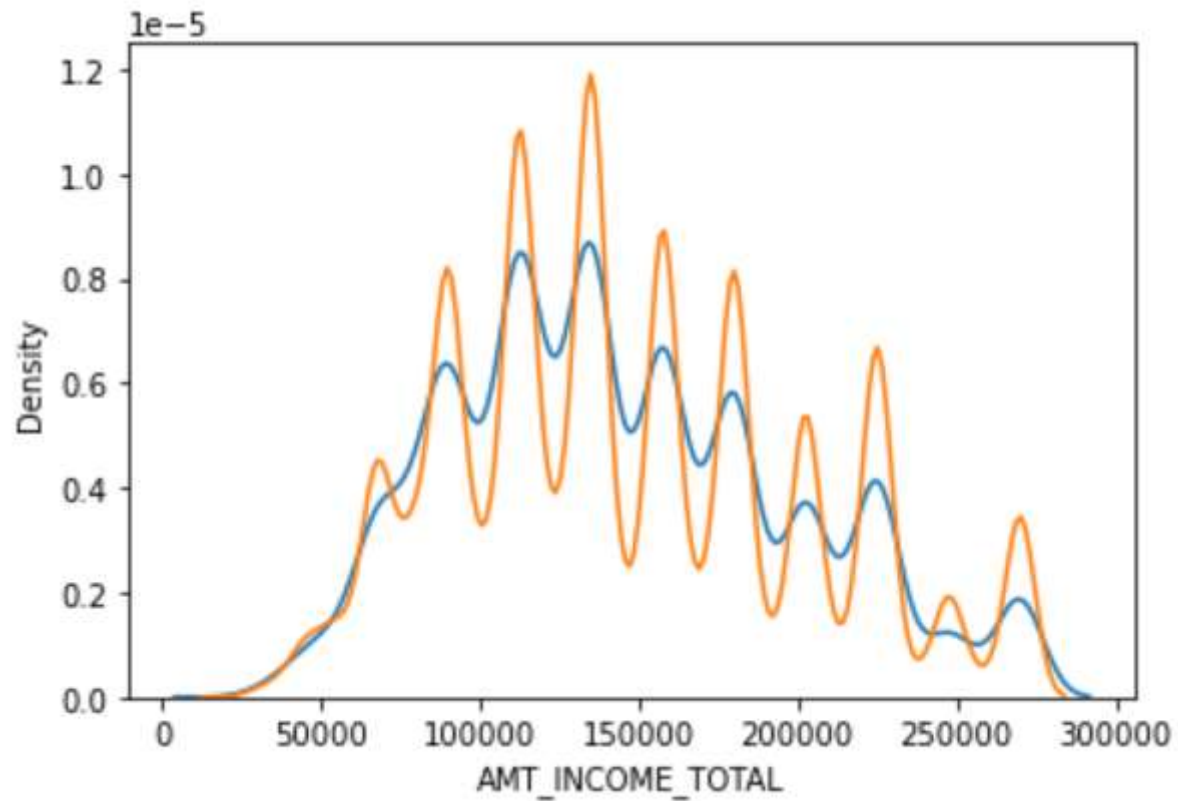


1.Type of NAME\_CONTENT\_TYPE doesn't affect much while coming to lone pay, in both case there are not many differences

2.for NAME\_INCOME\_TYPE with working is less likely to effect, but Pensioner are mostly paying on time

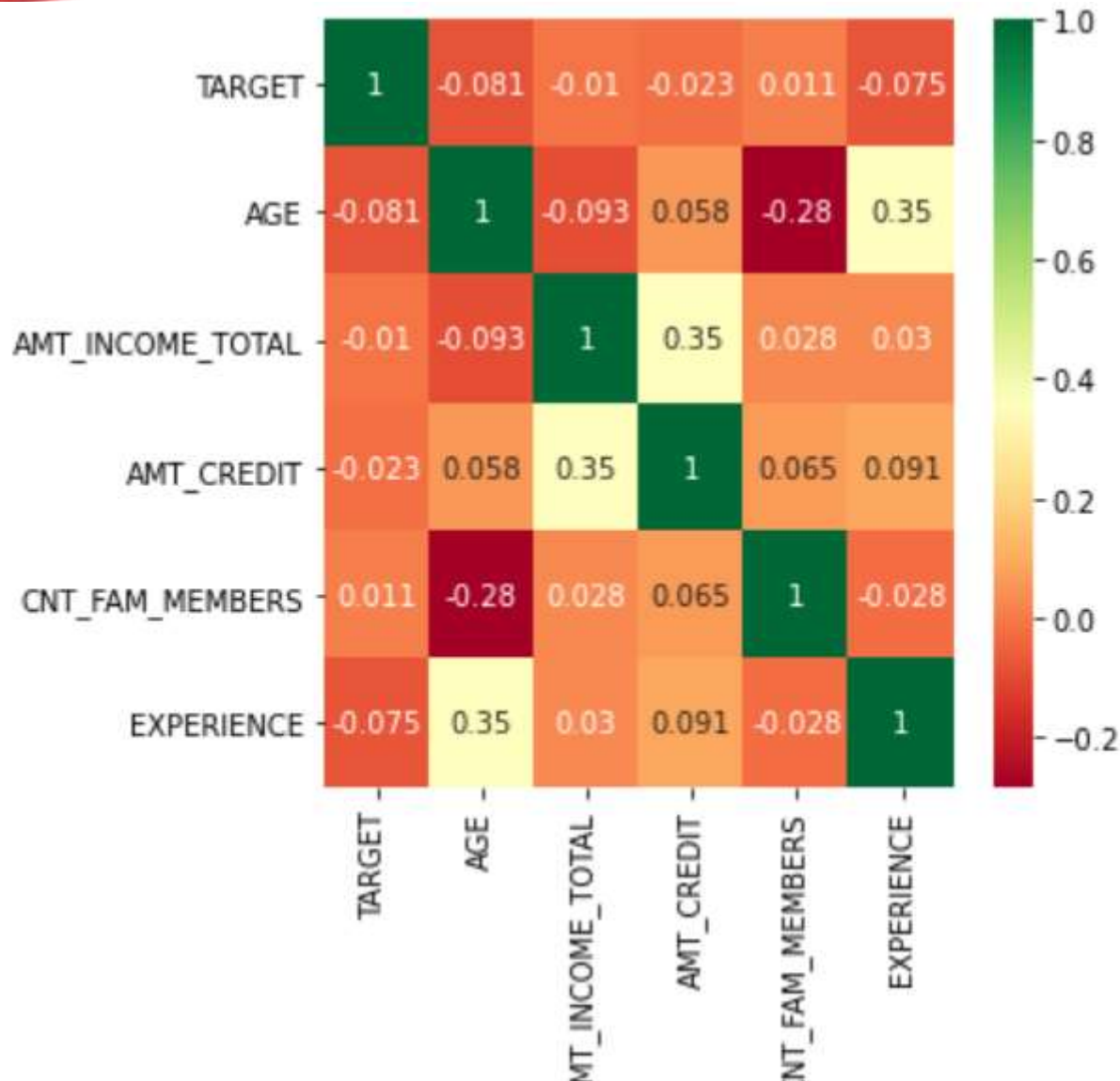
3.checking with GENDER Female are more likely to take lone and ratio of paying on time and delaying are most like same



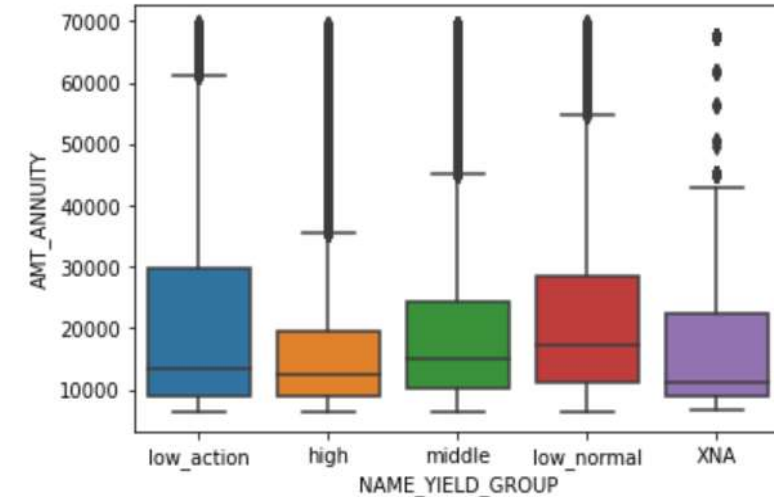
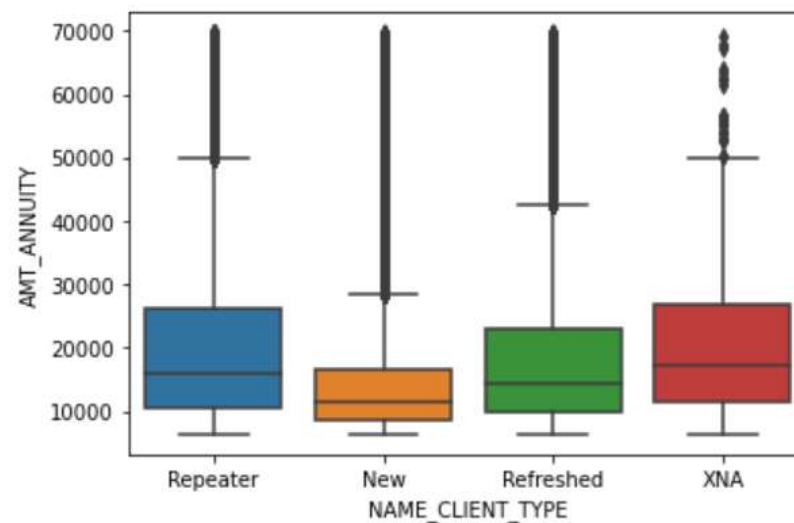
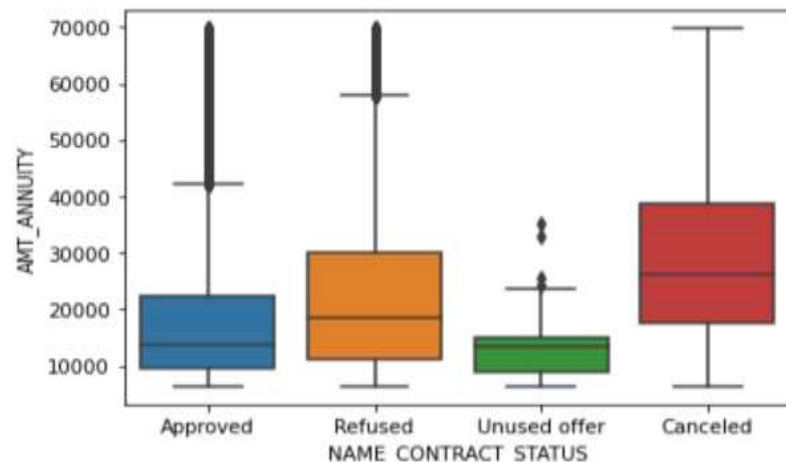
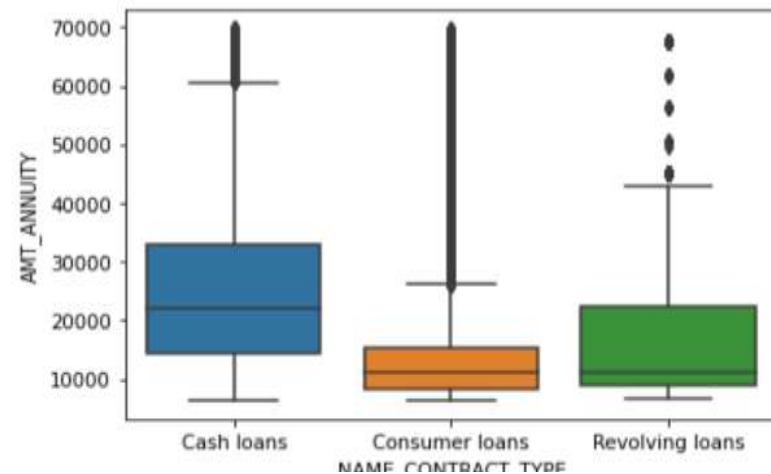
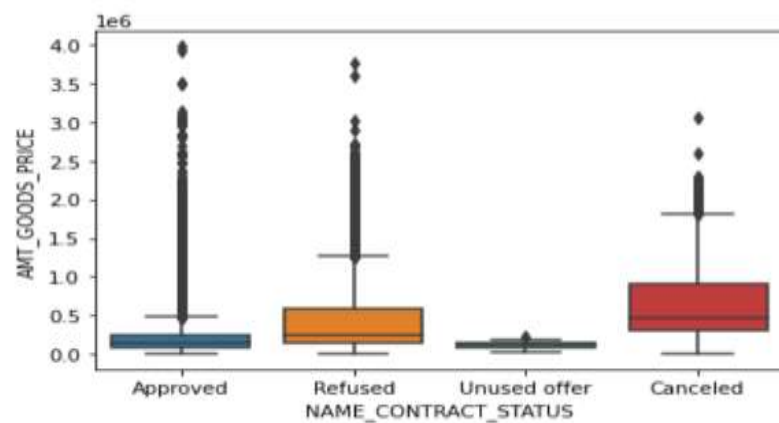


Client with `AMT_INCOME_TOTAL` 100000 and 150000 are more likely to delay payment

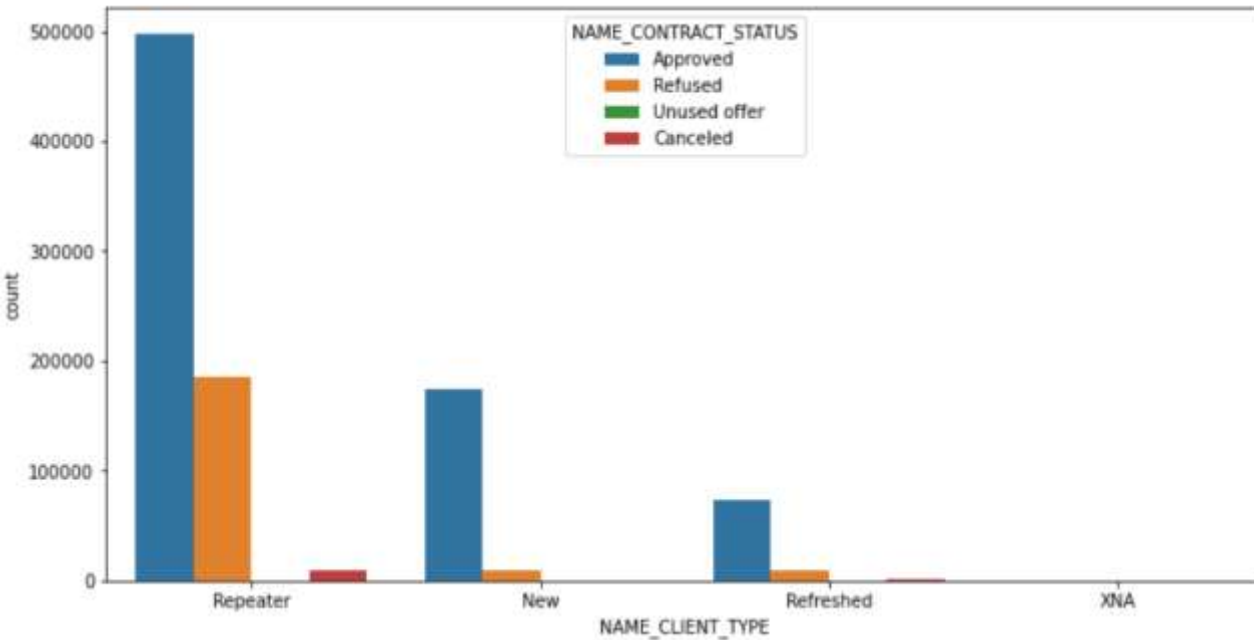
# Correlation



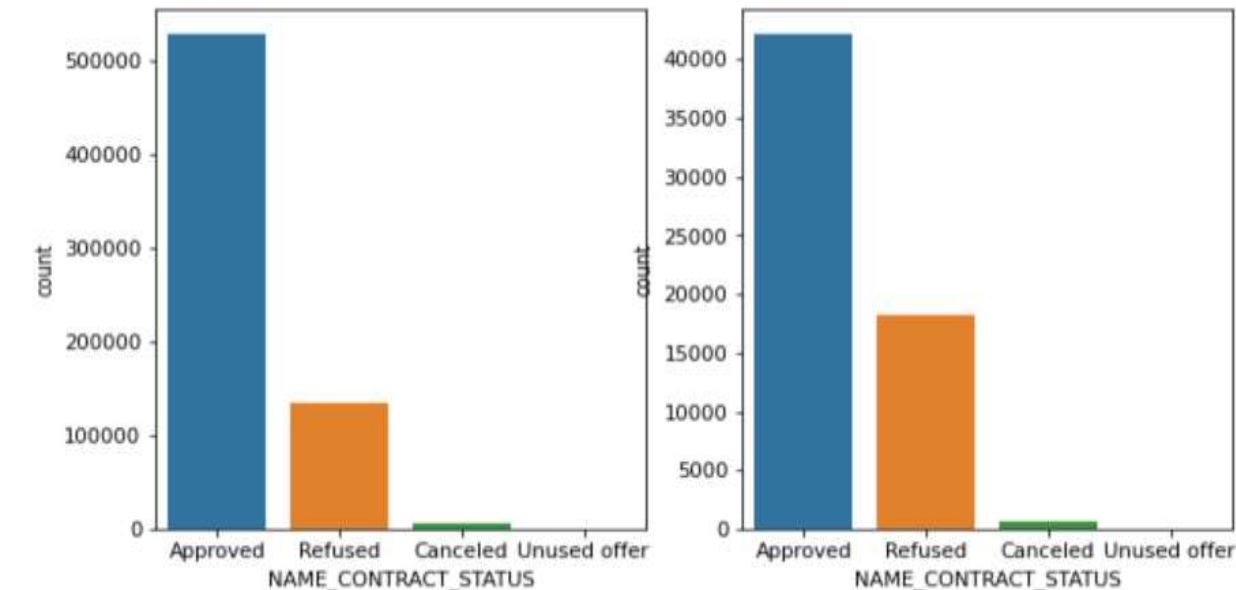
- Client with older age and heigh family count are more likely to delay payment as we are seeing high negative correlation
- Applicants who have higher family members have higher default rate.



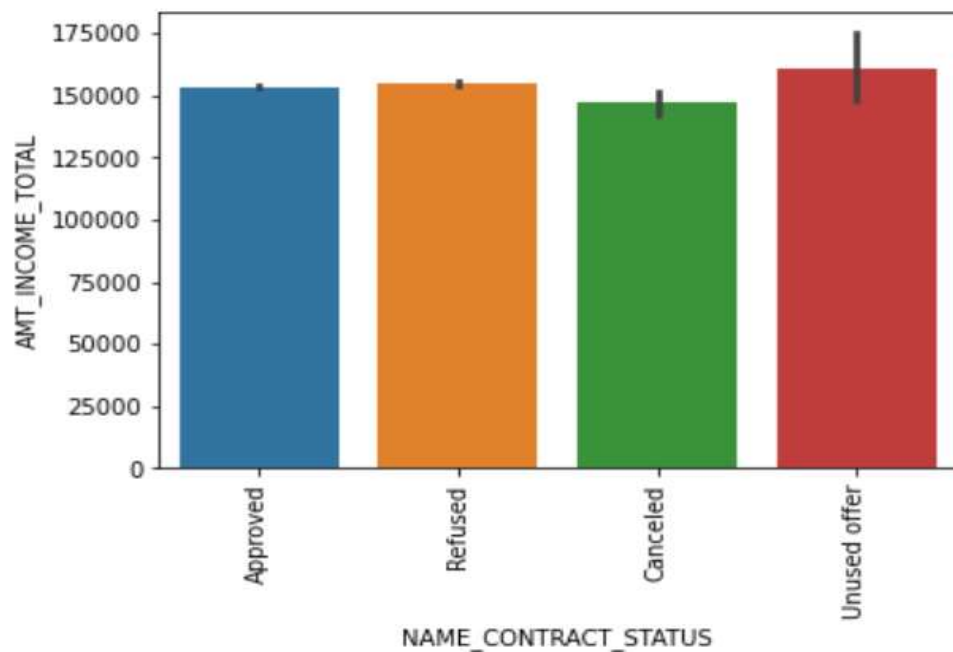
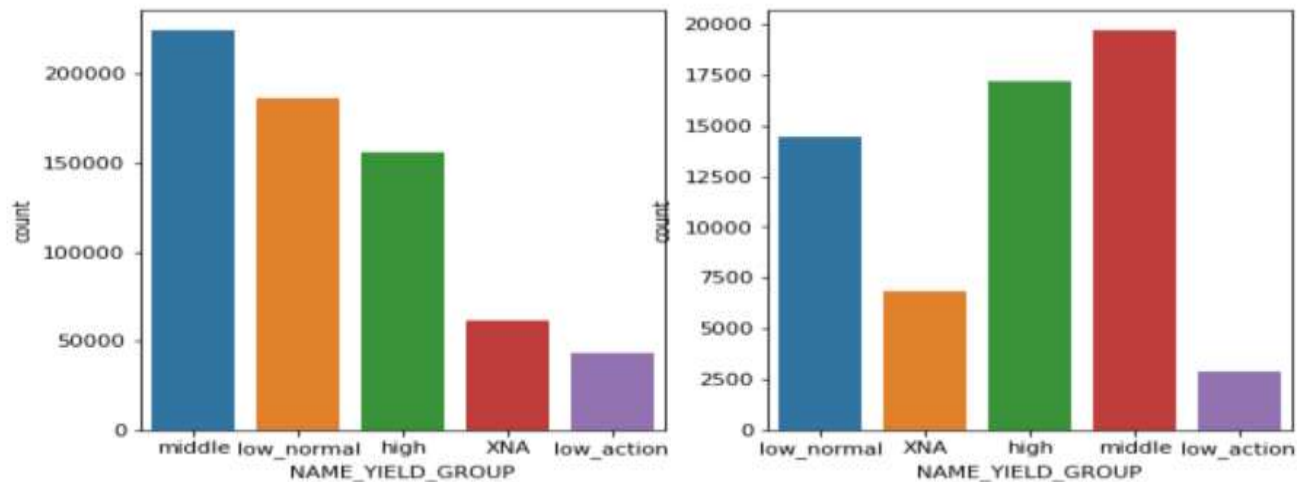
- Goods price of good that client has higher rate of approval
- on annual scale there are cash lone as main but customer lone has higher amount
- cancel rate of lone are comparatively higher than others



- Repeater clients has higher approval rate than refused but even when compared with over all type of clients repeater clients has better approved and Refused rate when compared to other(new, refreshed, XNA)
- applications rejected had same percentage for paying on time and defected
- defective clients with approval has lower







- clints who have not used offer earlier have defaulted even when there average income is higher than others
- Interest with low normal are little high being defect



*Thank you*