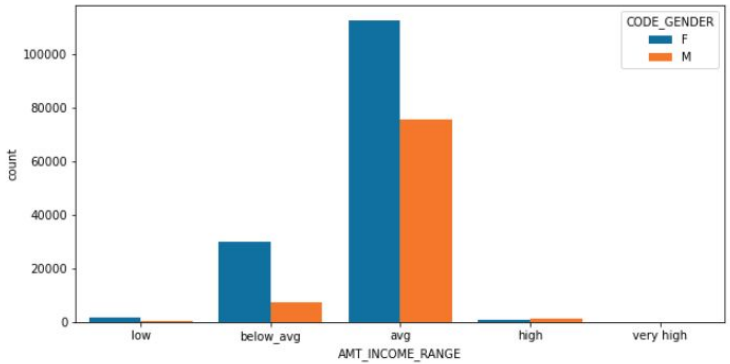
Credit EDA Case Study

By:
Manoj Kumar
Vidya Sagar

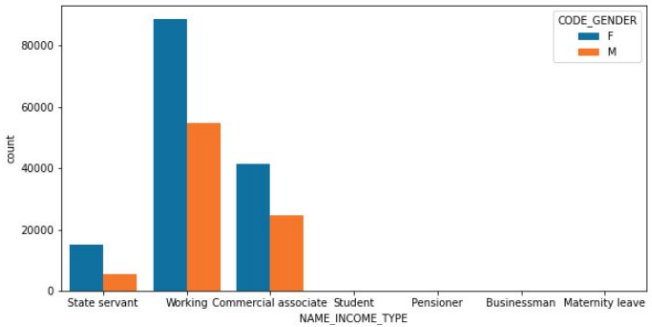
Distribution of Income Range for target 0

1. Very less count for income range high and very high.
2. Female counts are higher than male.
3. This graph show that females are more than male in having credits for that range



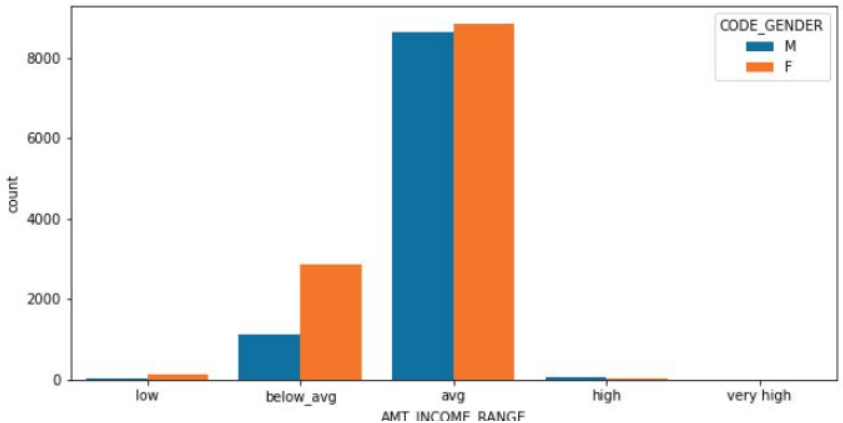
Distribution on Income type for target 0

1. Number of credits is in order Working, Commercial Associate and State Servant.
2. Almost negligible credits for Student, Pensioner, Businessman,Maternity leave



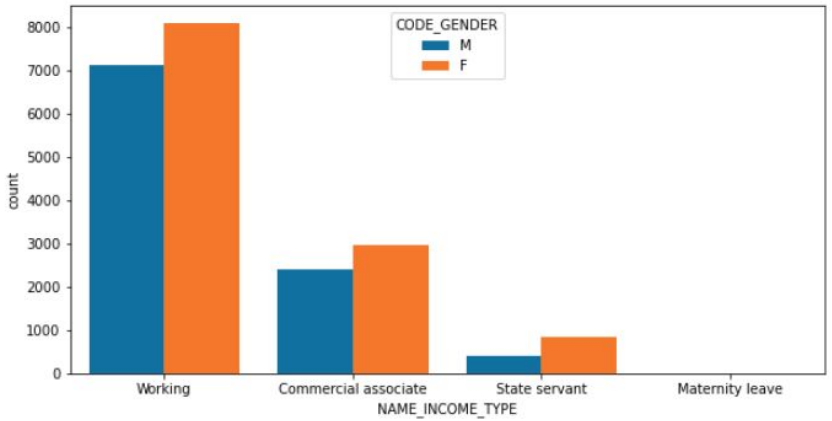
Income range for target 1

1. Female count is more than male
2. Income range between 100000-500000 has more number of credits than others
3. Females have more credit than males.

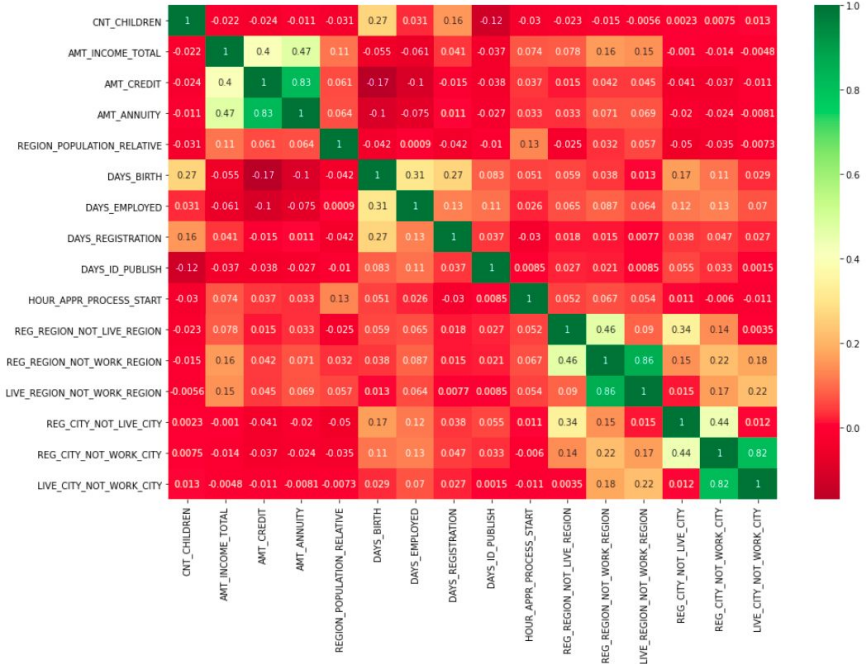


Income type for target 1

1. Females have more credit than males
2. Working has highest number of credits than others.
3. Almost negligible credits for Maternity leave compared to others



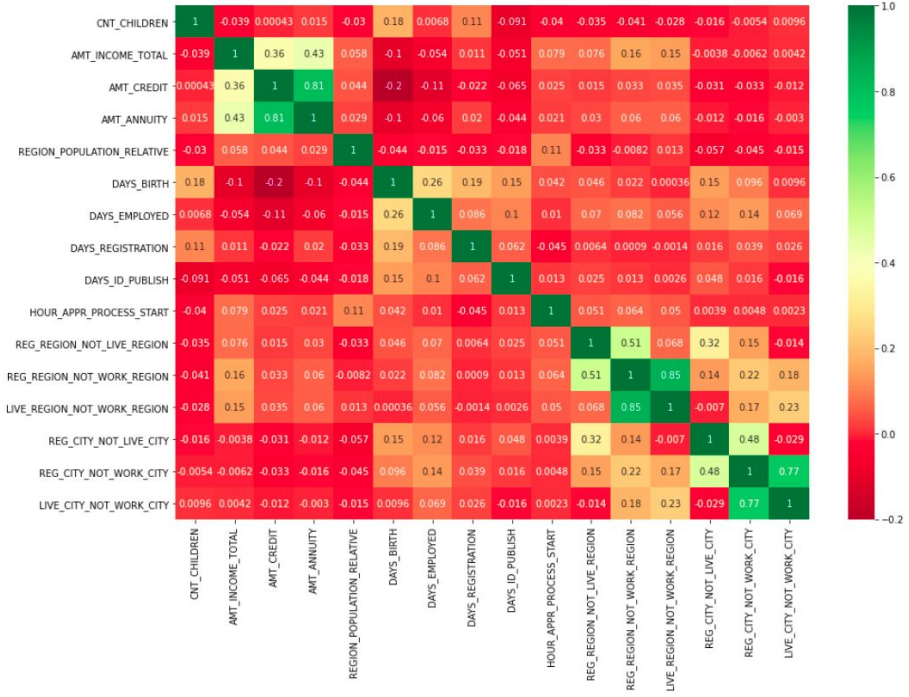
Correlation for target 0



Inferences for correlation target 0

1. Credit amount is higher in densely populated area
2. Income amount is also higher in densely populated area
3. In densely populated area, we've less child clients

Correlation for target 1

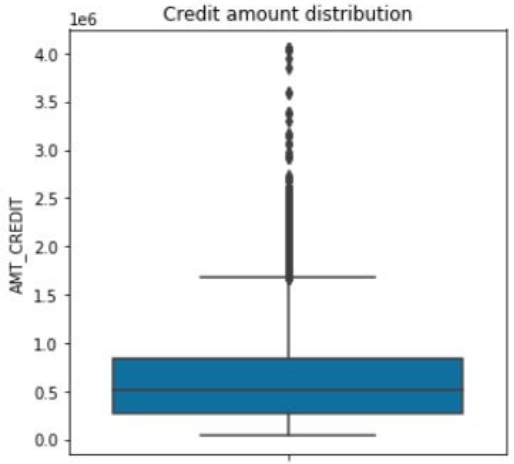


Inferences for correlation for target 1

1. Client's are having less children and permanent address does not match with contact address

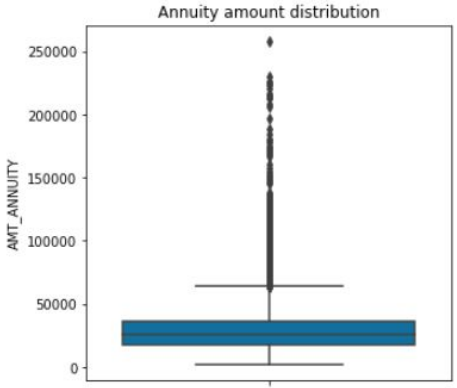
Box plot for credit amount

1. In credit amount, there are some outliers

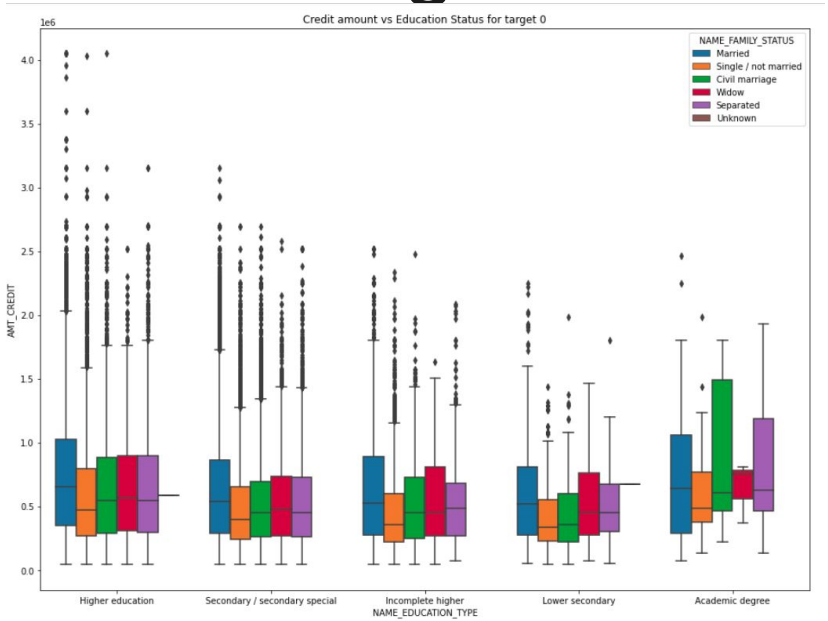


Box plot for annuity amount.

1. Outliers are detected in annuity amount.
2. 1st quartile is smaller than 3 rd quartile



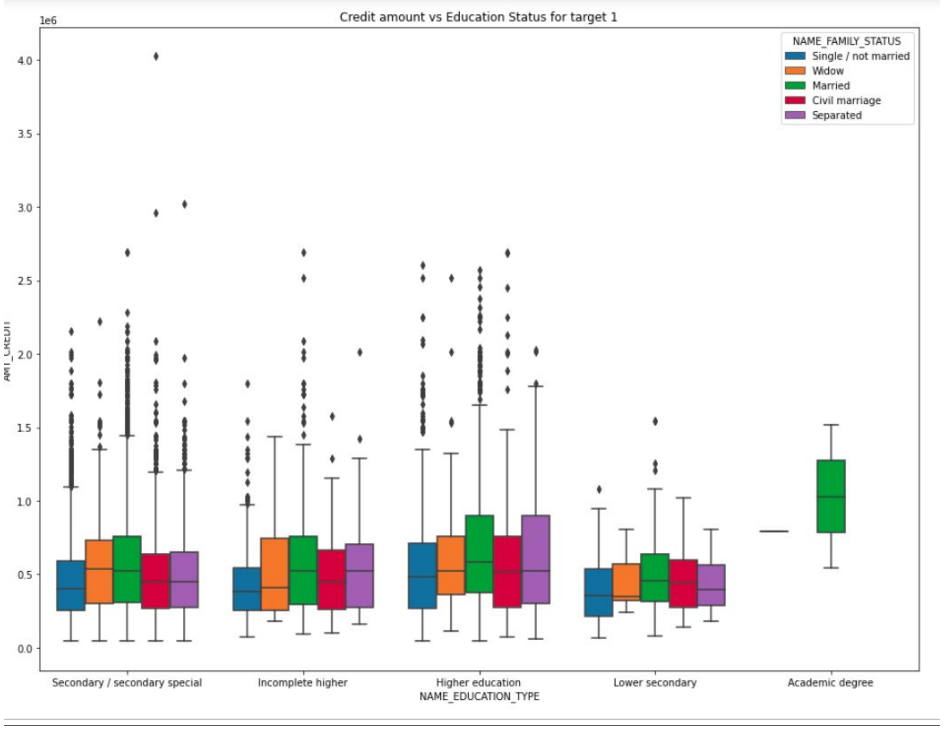
Bivariate analysis for Credit amount vs Education Status for target 0



Inferences

1. Higher education of family status of 'marriage', 'single' and 'civil marriage' are having more outliers.
2. Higher number of credits for the family staus of 'civil marriage', 'marriage' and 'separated'

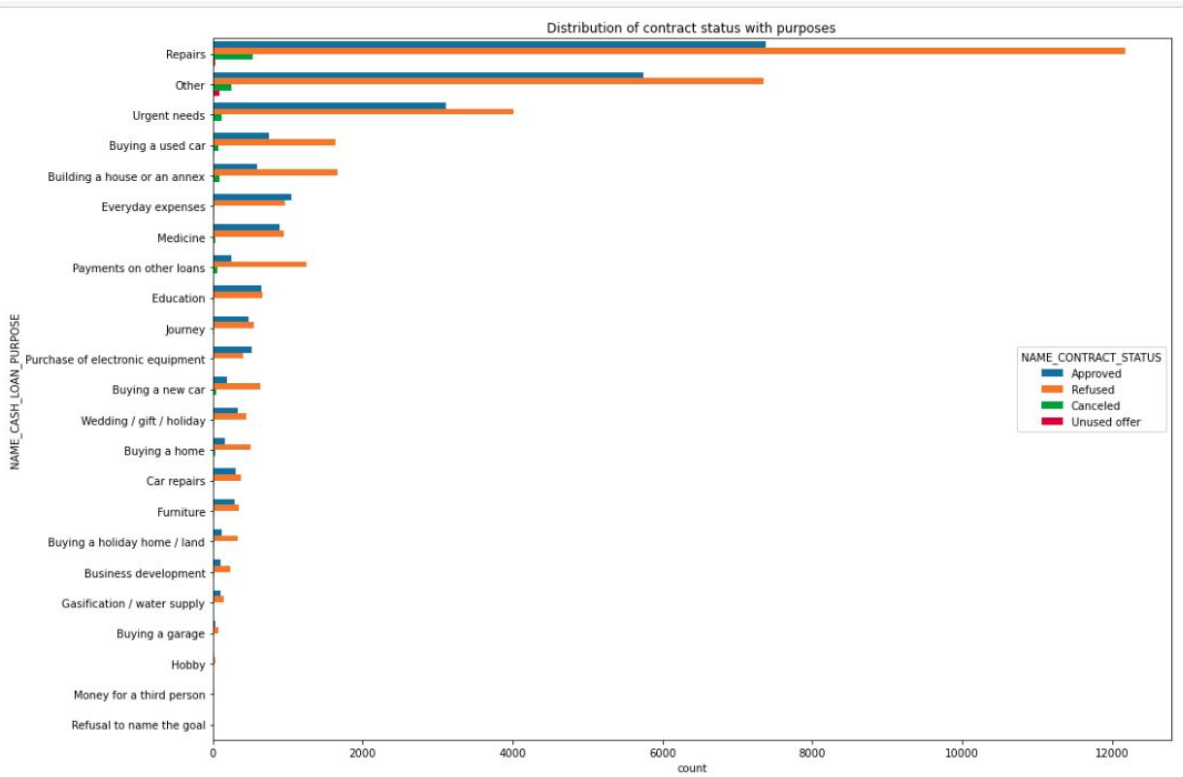
For Credit amount vs Education Status for target 1



Inferences

1. Most of the outliers are from Education type 'Higher education' and 'Secondary'.
2. Quite similar with Target 0 From the above box plot we can say that Family status of 'civil marriage', 'marriage' and 'separated' of Academic degree education are having higher number of credits than others.

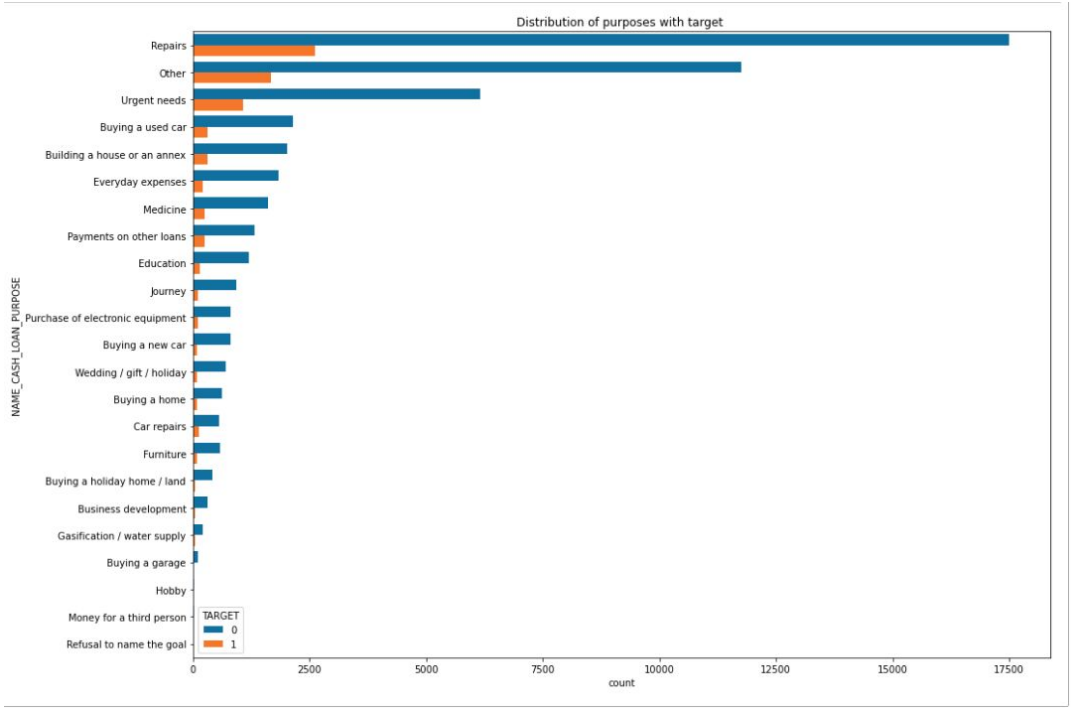
Distribution of contract status with purposes



Inferences

1. Most rejection of loans came from purpose 'repairs'.
2. we have equal number of approves and rejection for education purposes

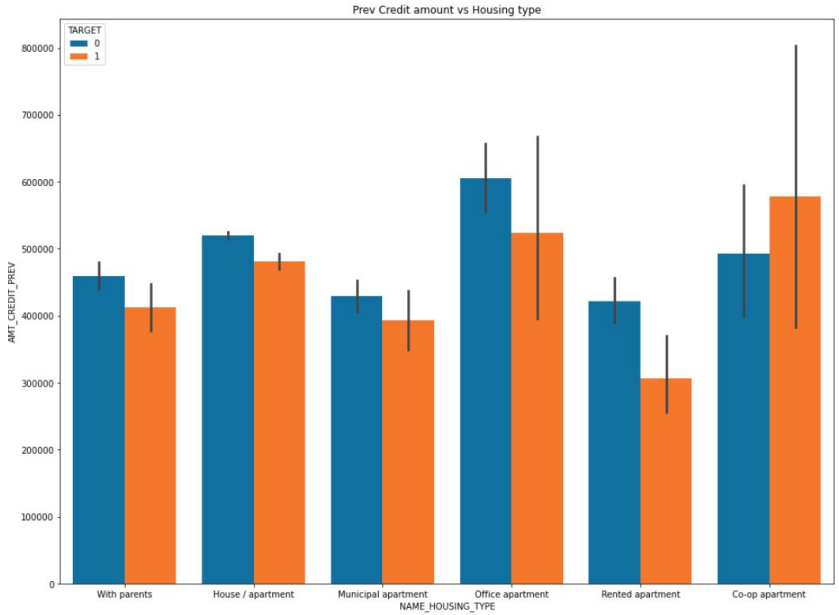
Distribution of purposes with target



Inferences

1. Loan purposes with 'Repairs' are facing more difficulites in payment on time.
2. There are few places where loan payment is significant higher than facing difficulties. They are 'Buying a garage', 'Business developemt', 'Buying land','Buying a new car' and 'Education' Hence we can focus on these purposes for which the client is having for minimal payment difficulties.

Prev Credit amount vs Housing type



Inferences

1. So, we can conclude that bank should avoid giving loans to the housing type of co-op apartment as they are having difficulties in payment.
2. Bank can focus mostly on housing type with parents or House\appartment or miuncipal appartment for successful payments.

Conclusion

1. As there are least unsucessful payments for 'With Parents' get as much as clients from housing type
2. Banks should focus more on contract type ‘Student’ ,’pensioner’ and ‘Businessman’ with housing ‘type other than ‘Co-op apartment’ for successful payments.
3. Also with loan purpose ‘Repair’ is having higher number of unsuccessful payments on time.