CREDIT EDA CASE STUDY

By

Kamil Sheikh & Sneha Soman

From DSC-34 batch

OBJECTIVE

- Identify patterns and correlation of attributes that help understand the credit-worthiness of a customer.

- Understand how consumer attributes and loan attributes influence the tendency of default.

- Identify patterns which indicate if a client has difficulty paying their installments and thereby take appropriate actions at the time of application evaluation.

About application data set:

- The application data set started with a table size of 307511 & 122 rows and columns respectively.
- After data cleaning the analysis was performed on table size 307495 & 76 rows and columns respectively.
- The analysis on this data set progressed by partitioning an variable TARGET on based on "no_difficulty_df_app" & "difficulty_df_app" referred to TARGET = 0& TARGET = 1 respectively in the tables and chart
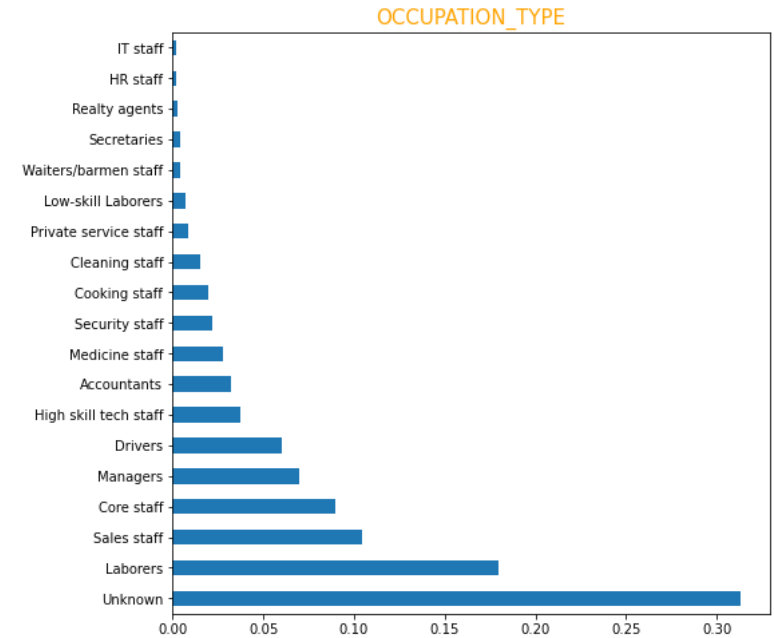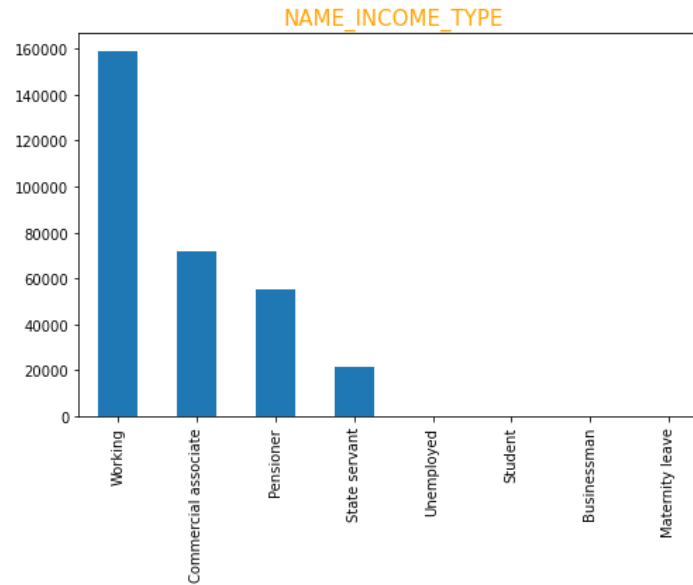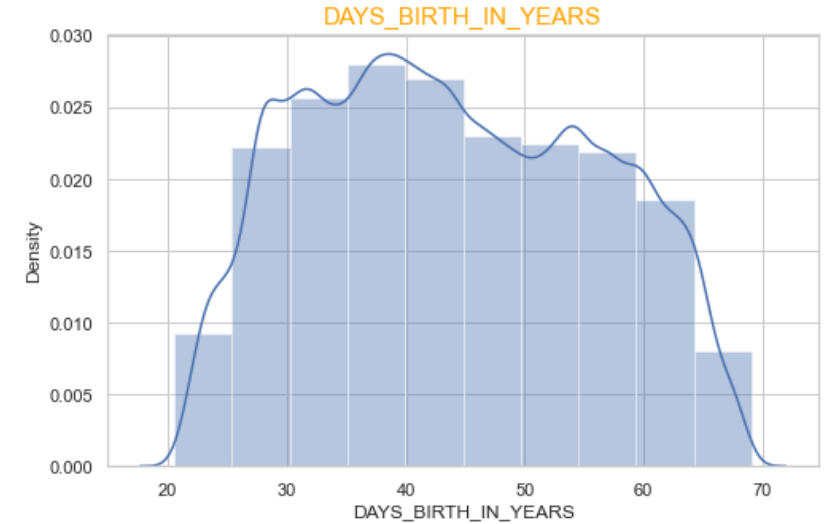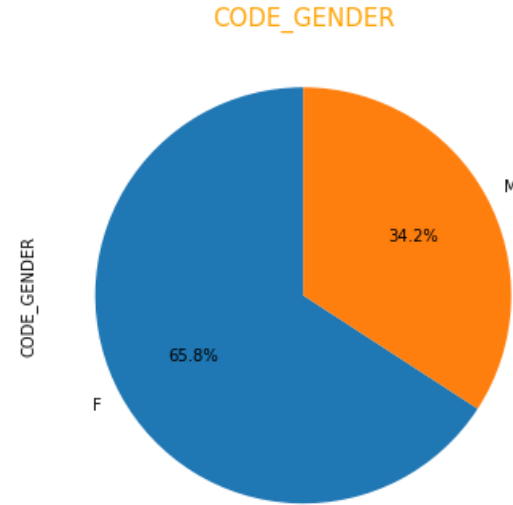
OBSERVATIONS

65.8% of loan applicants are Females & only 34.2% are Males

Age group of the loan category is higher between 35-40.

More than 50% of our loan applicants are from 'Working' category followed by 'Commercial Associate' & 'Pensioner' at the 3rd position.

Out of the 18 employment categories labors have the highest loan application

OBSERVATIONS

More than 73% of the applicant are married customers.

Most of the loan applicants are in the Secondary/secondary special segment which comprises more than 70% of the applicant.

92% of the applicants by far has been regular with their payments. Only 8% have had atleast 1 default/late payment.
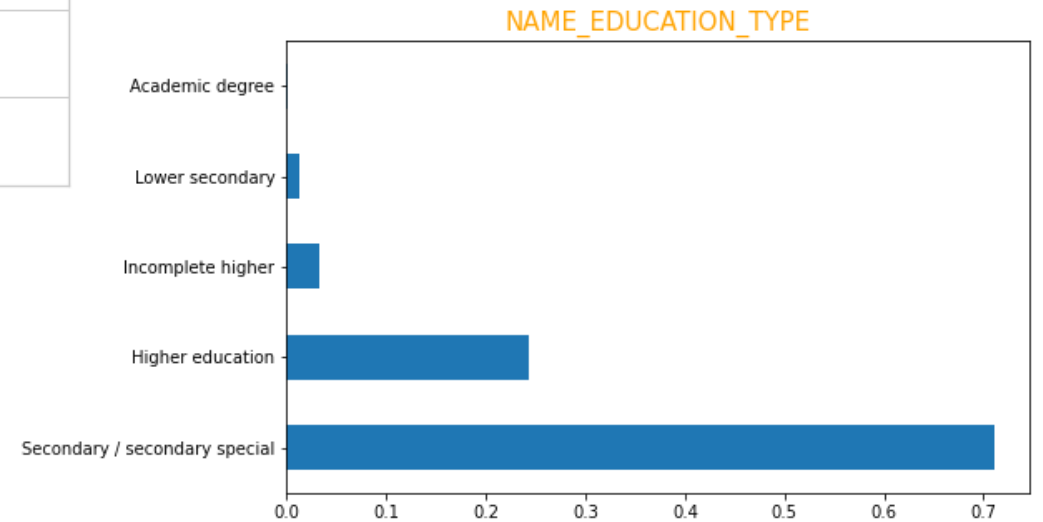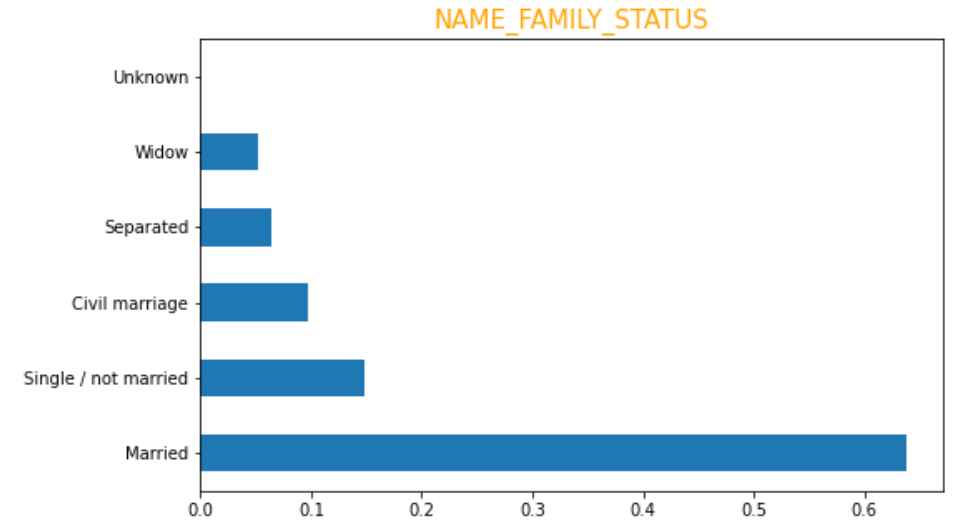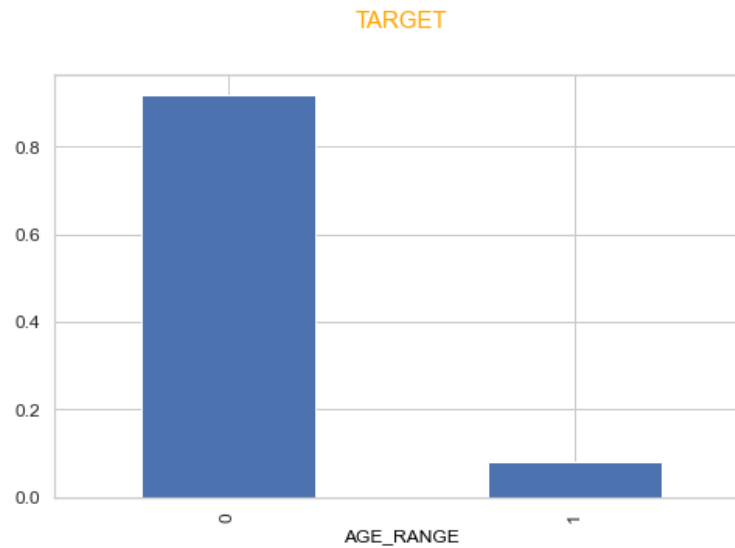
Credit Range vs Target

Clients with Credit range between 2,00,000 - 5,00,000 & 5,00,000 - 10,00,000 though have a higher number of defaulter their ratio in comparison to that of Target = 0 is much better than credit amount 0-2,00,000 & 10,00,000 - 2,00,000
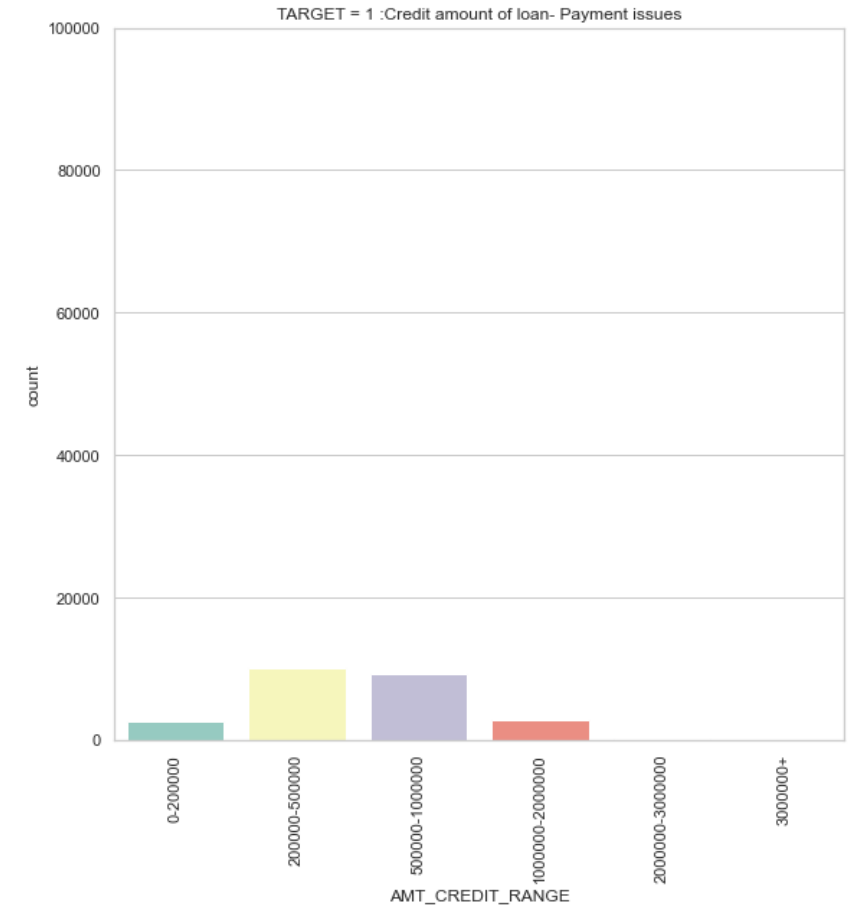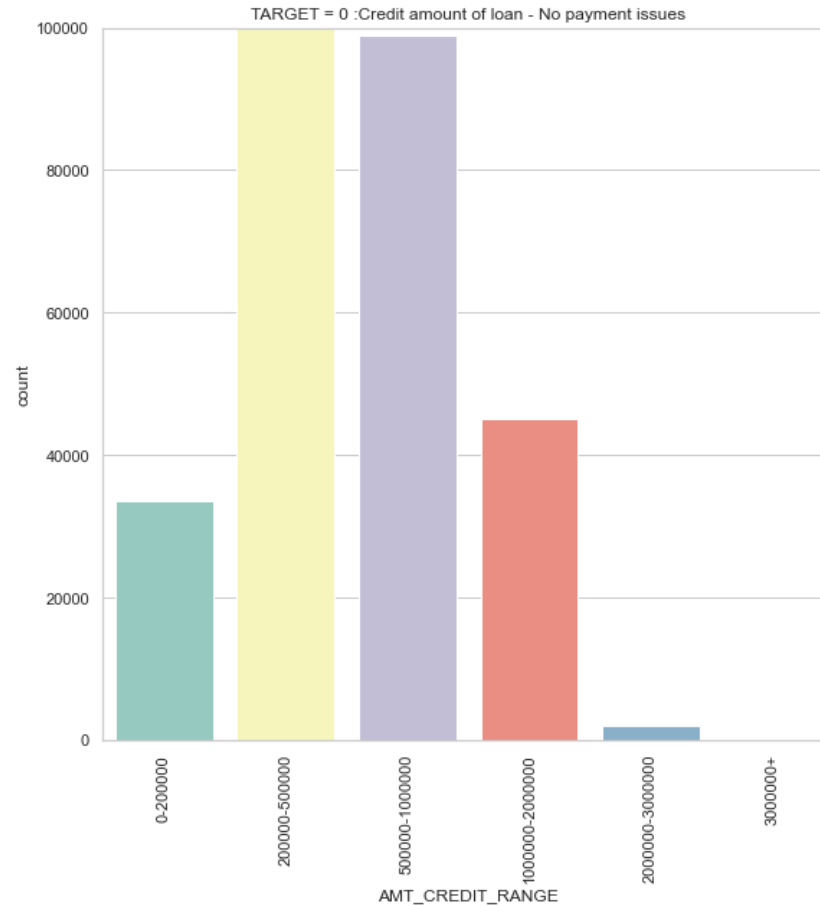
Occupation Type Vs Target

1. Categories like Managers, Sales staff, Drivers show a more unreliable tread in the payments.

2. Professions like High skill tech staff, Medicine staff, HR staff, IT staff, Secretaries, Reality agents are reliable for credits as they have low or nill payment issues.
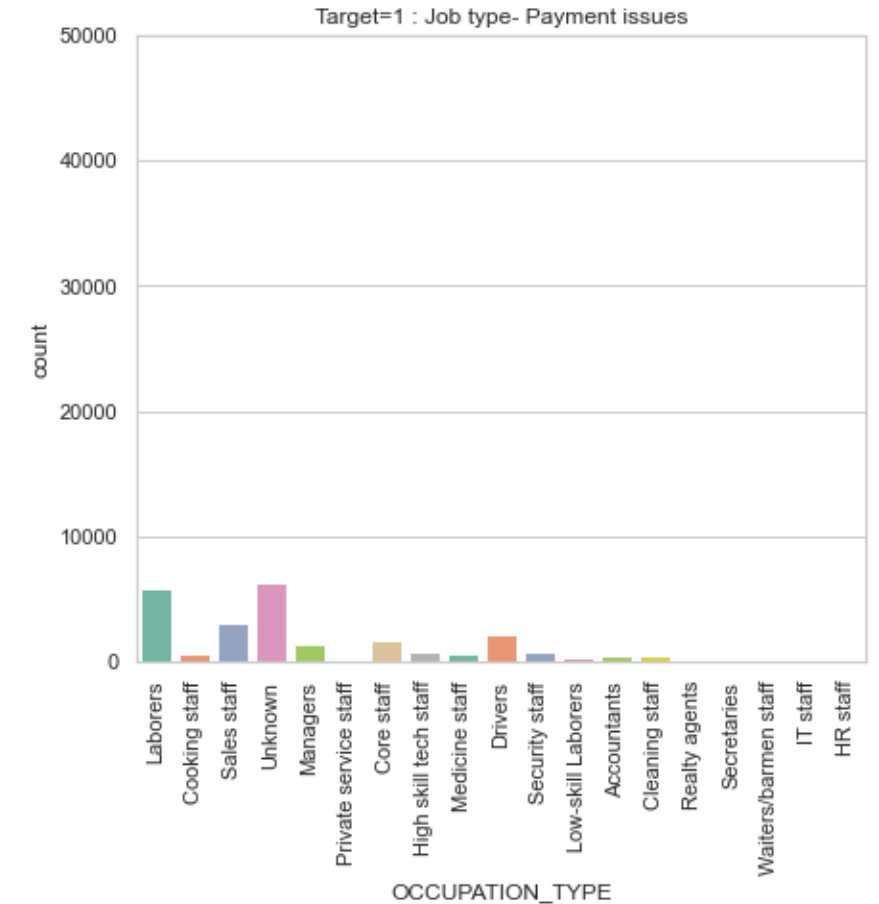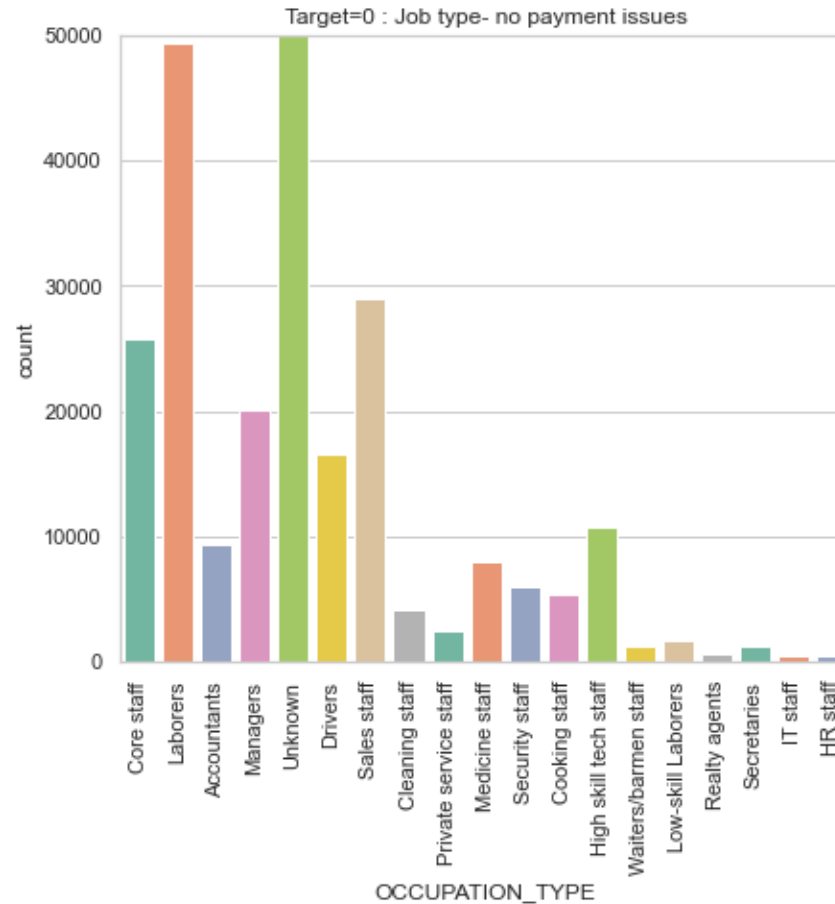
Occupation Type Vs Target

1. Categories like Managers, Sales staff, Drivers show a more unreliable tread in the payments.
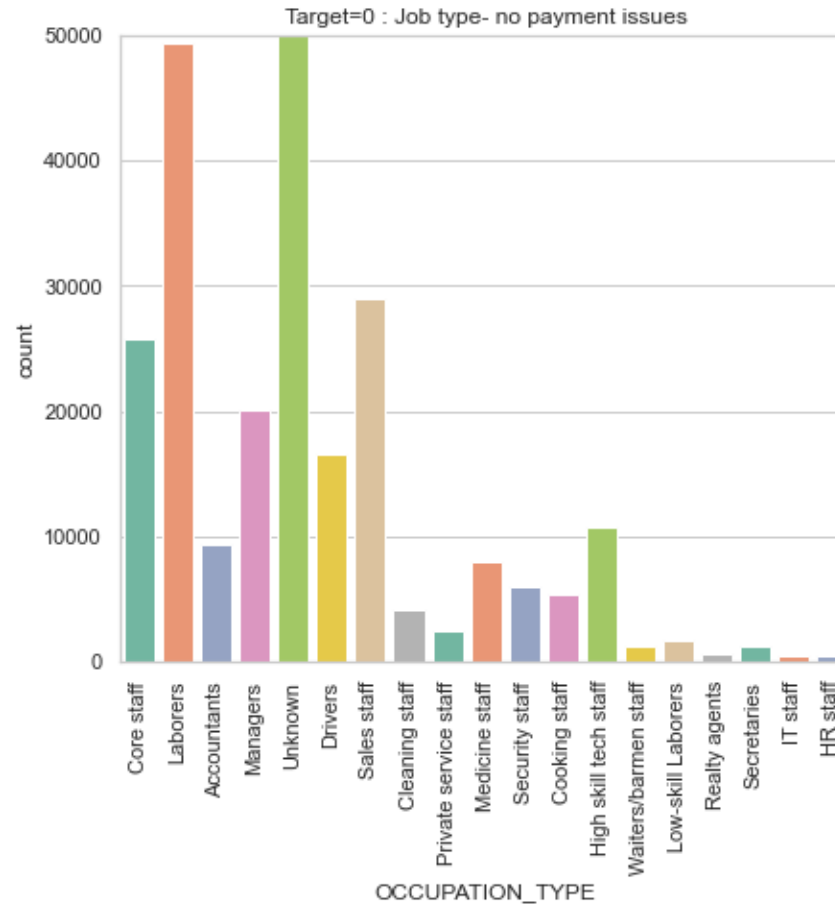
2. Professions like High skill tech staff, Medicine staff, HR staff, IT staff, Secretaries, Reality agents are reliable for credits as they have low or nill payment issues.
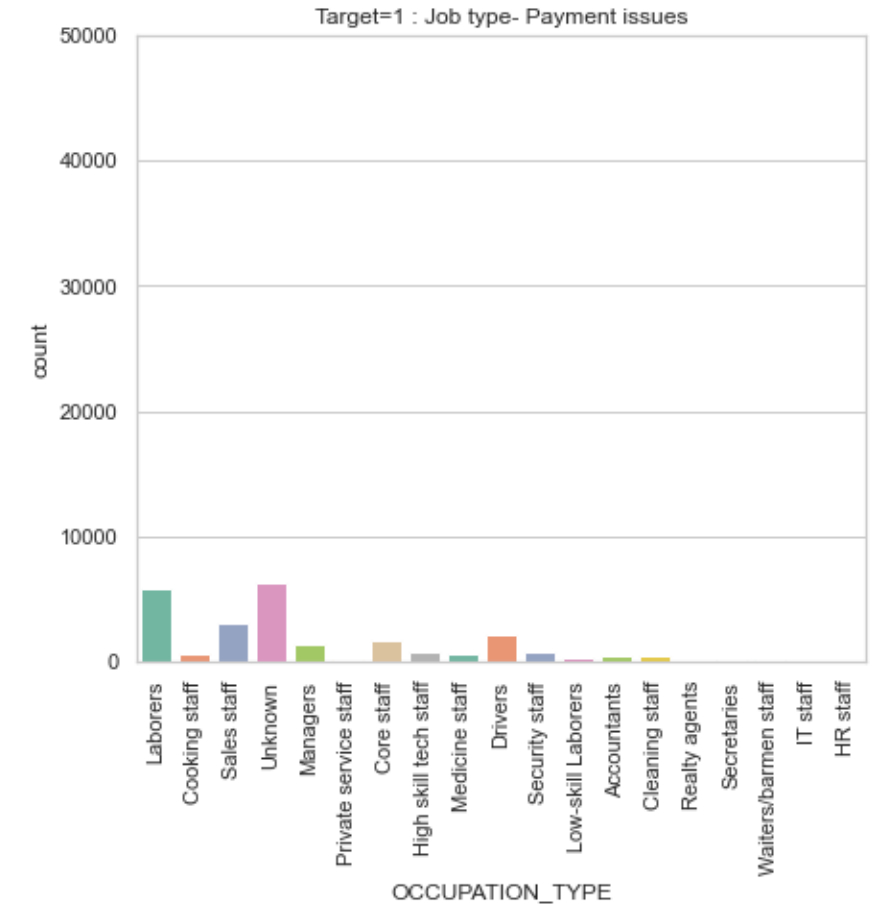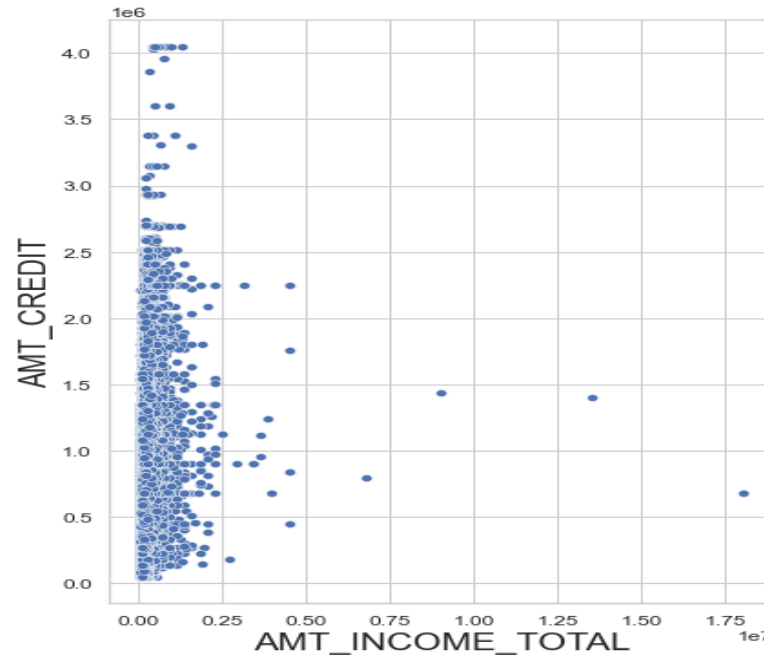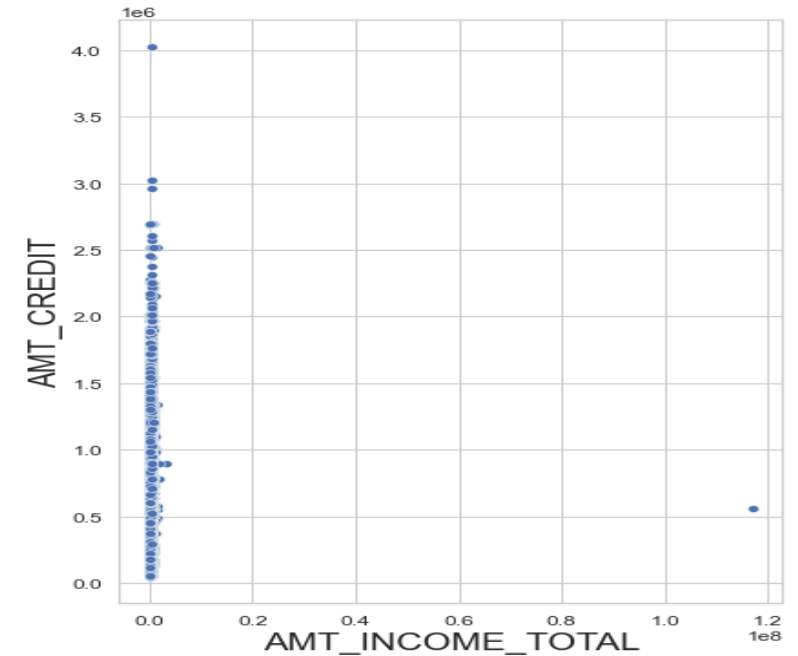
NUMERICAL _NUMERICAL ANALYSIS

Those who payback the loan on time can get higher credits. Mosty low income clients have difficulties in paying installments and therefore they have low credits. Goods_price and credit have linear relationship. Clients who have paid back on time with higher goods_price have higher credit
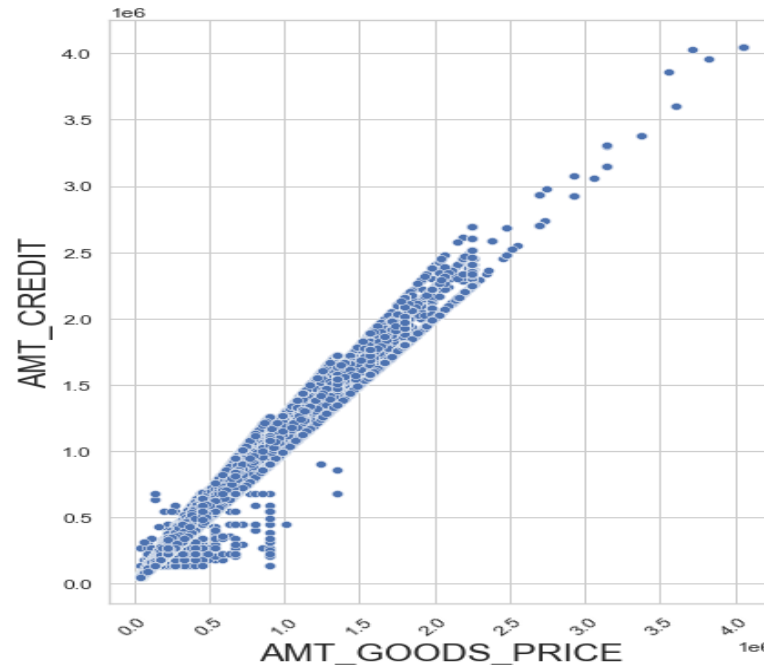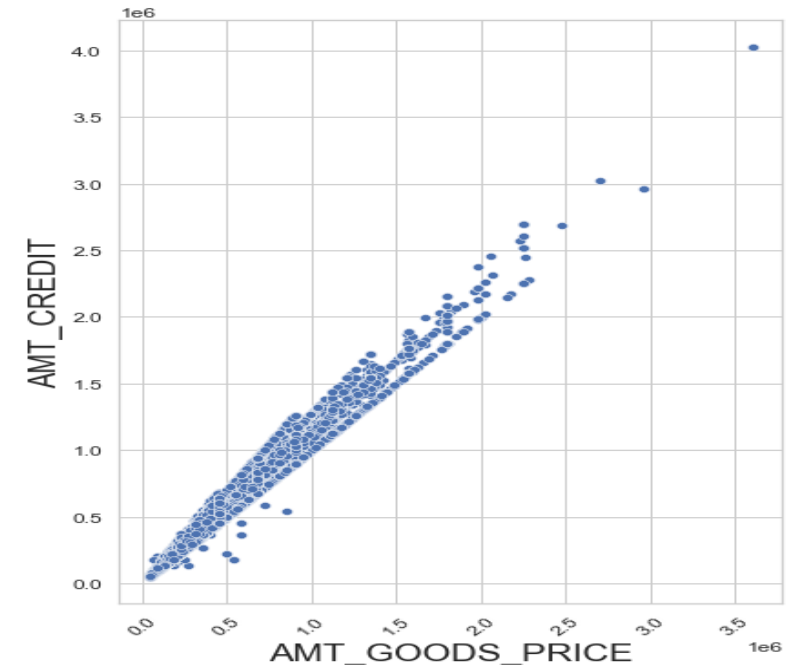
SUGGESTION

1) Customers like Businessman , Pensioner and Students with housing type House and with Parents should be aggressively pursued as they have least default chances.

2) Banks should rather focus on married couples with higher education or singles with higher education as they have least payment difficulties .

3) Customers matching above mentioned profile with the loan purpose of Repair , Urgent Need , money for third person and Hobby also have very high chance to default . Hence Bank must avoid approving their loan for assured paybacks .

4) Customers like less educated widows , workers living in co-op apartments , or males in general have high chances of defaulting . Hence Bank must avoid approving their loan for assured paybacks