



Credit Risk Analysis - EDA Case Study

Sravani Thogata

Business Objectives

- Identifying patterns that indicate if a client has difficulty in paying their instalments to help the Banks in finalizing the actions to take like Denying the loan, reducing the loan amount, or lending at higher interest rates for risky customers.
- Identifying the customers who are capable of repaying the loan.
- Identifying the risky customers who have difficulty to repay the loan.

Problem Statement

Any loan application has two types of risk associated:

- If the applicant repays the loan, then not approving would be a loss to business.
- If the applicant can't repay the loan, then approving such loans would lead to financial loss.

Overall Analysis Approach:

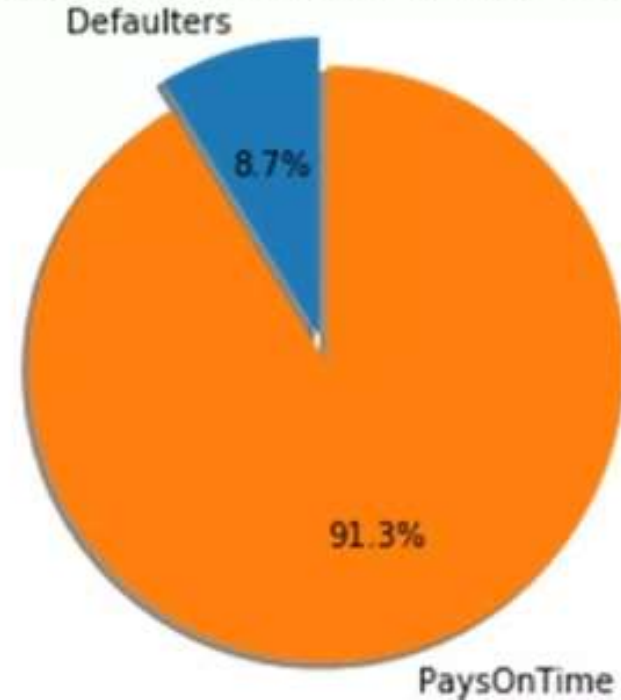
- Understanding the given data for analysis
- Ignoring the irrelevant features for the moment for effective analysis
- Analyzing the feature patterns using different graphs
- Analyzing the factors which have high impact in determining the defaulters
- Determining the attributes of defaulters ie, not likely to pay

Exoploratory Data Analysis

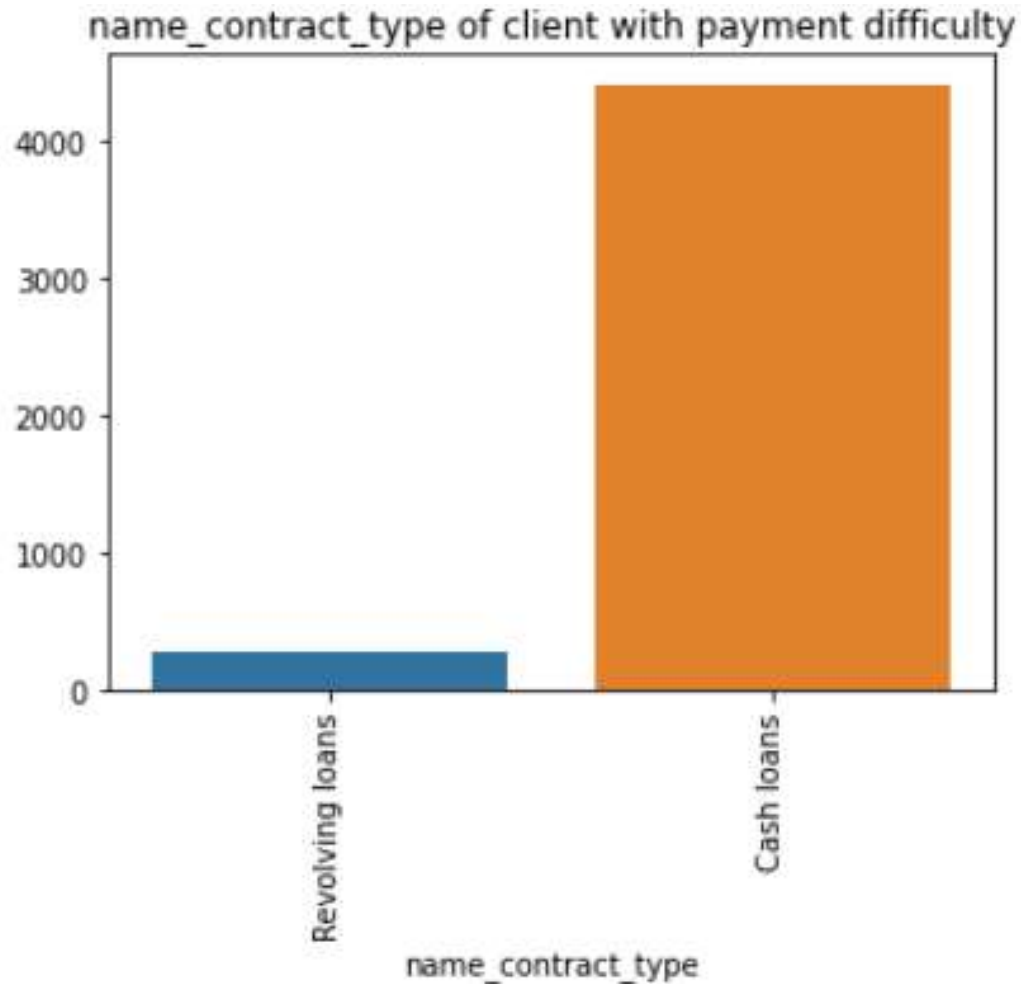
Defaulters proportion in the data set

- Defaulters are nearly 8.7% of the total customers.

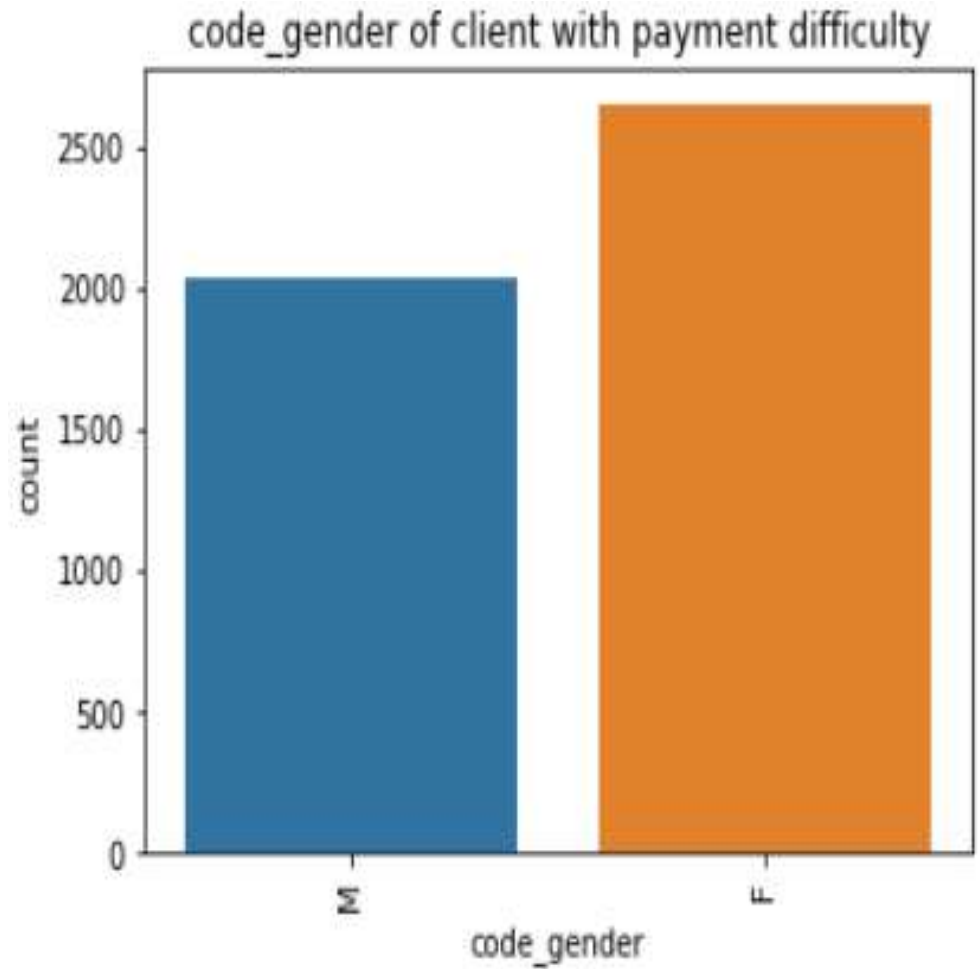
Proportion of Defaulters in Total Loans



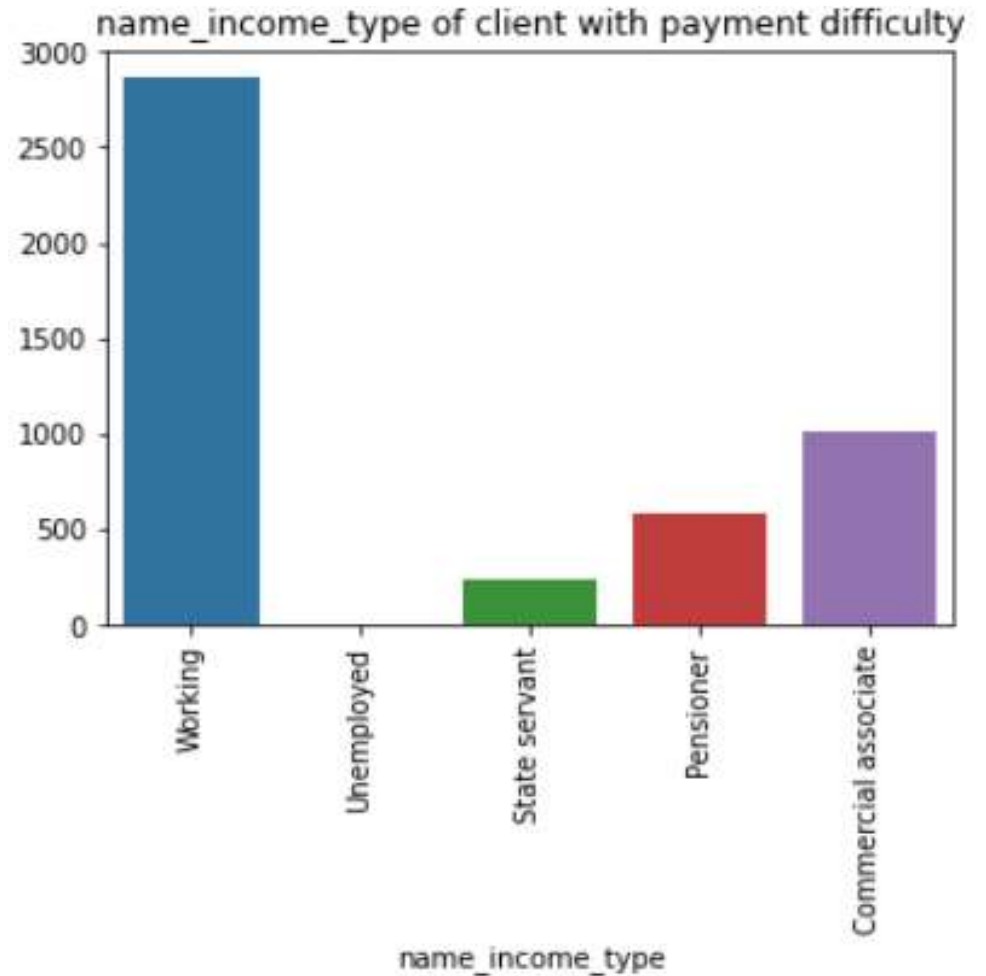
- Cash loans are have difficulty in repaying. They have more defaulters



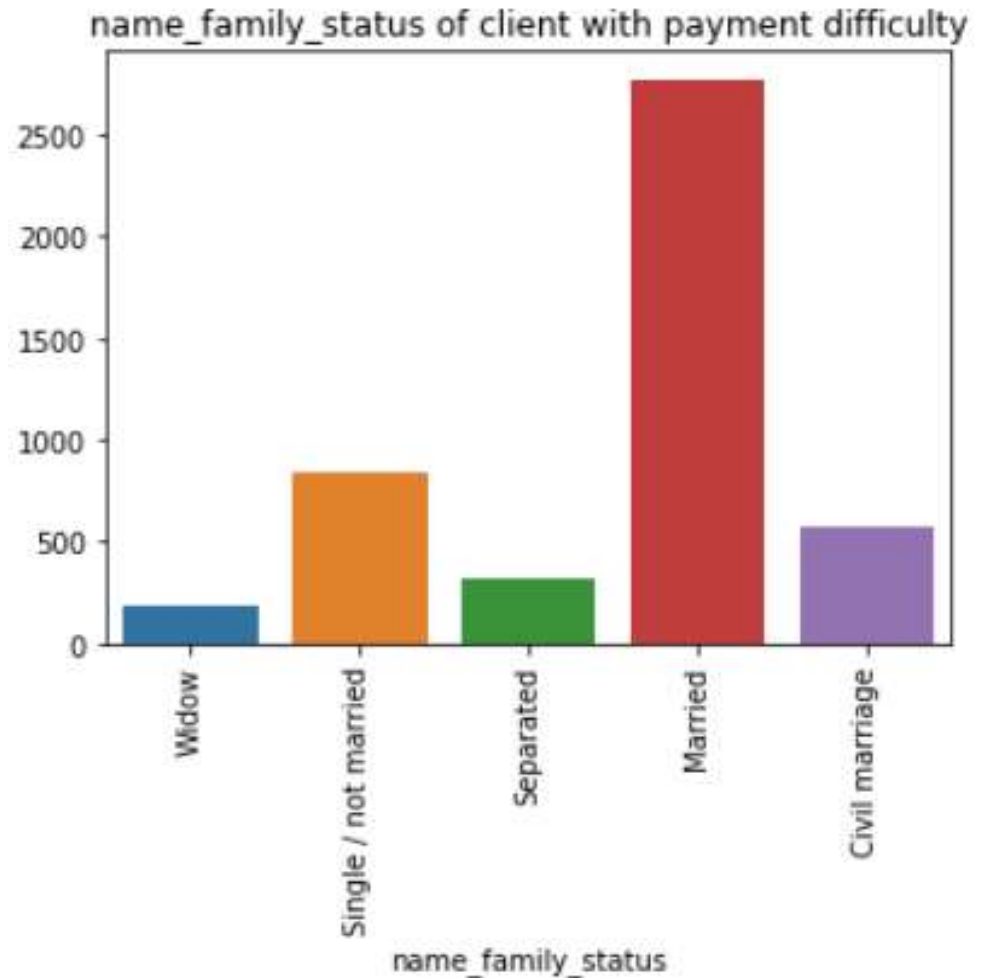
- Female clients often have difficulty in repaying the loans. They constitute most of the defaulters



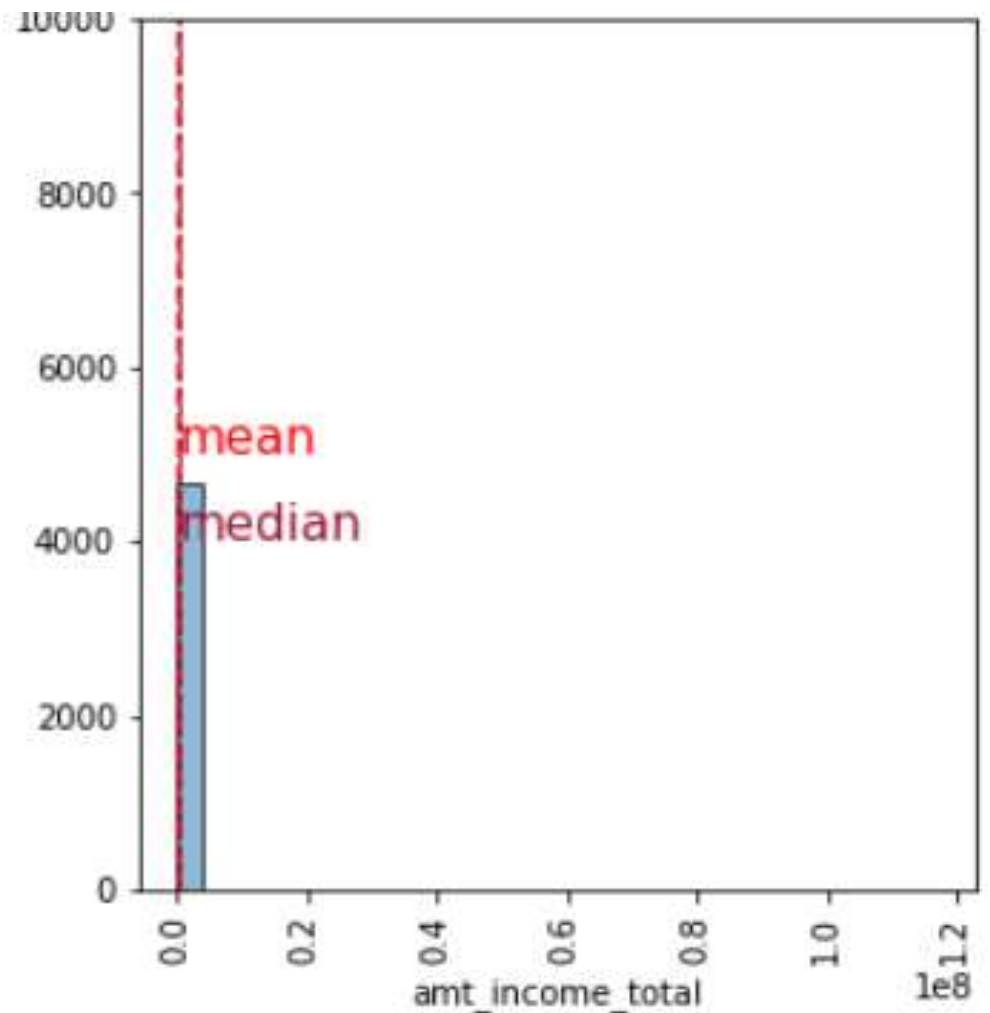
- Working class clients constitute majority of the defaulters



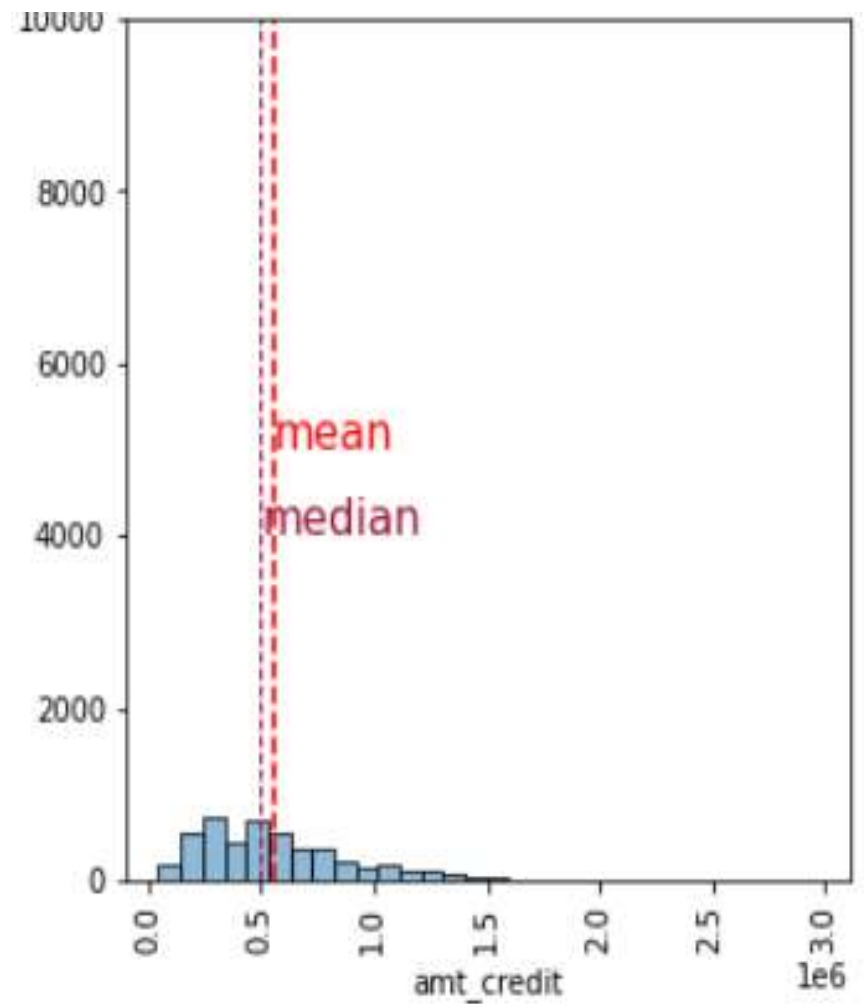
- Married clients are having difficulty in repaying and are majority of defaulters



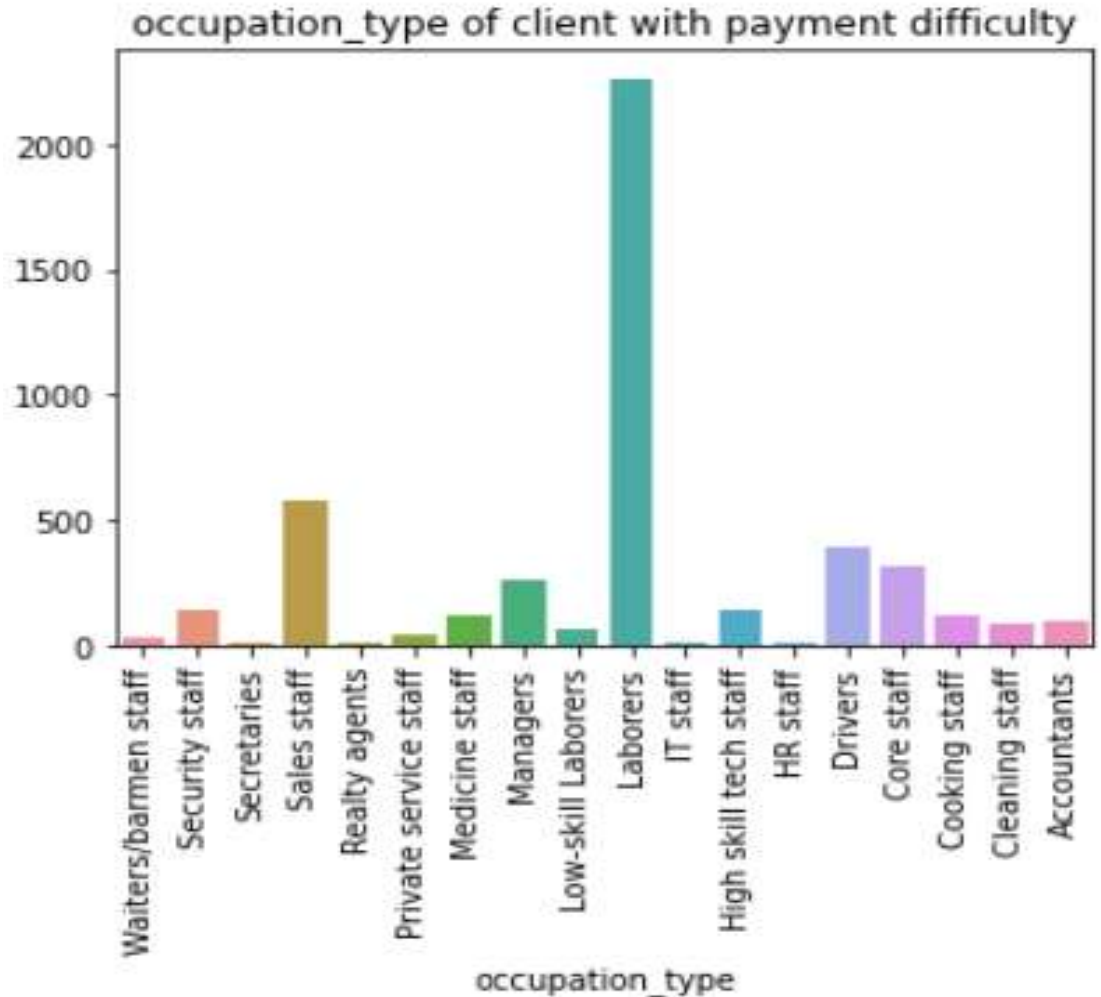
- Clients who earn low income often have high probability of being defaulters



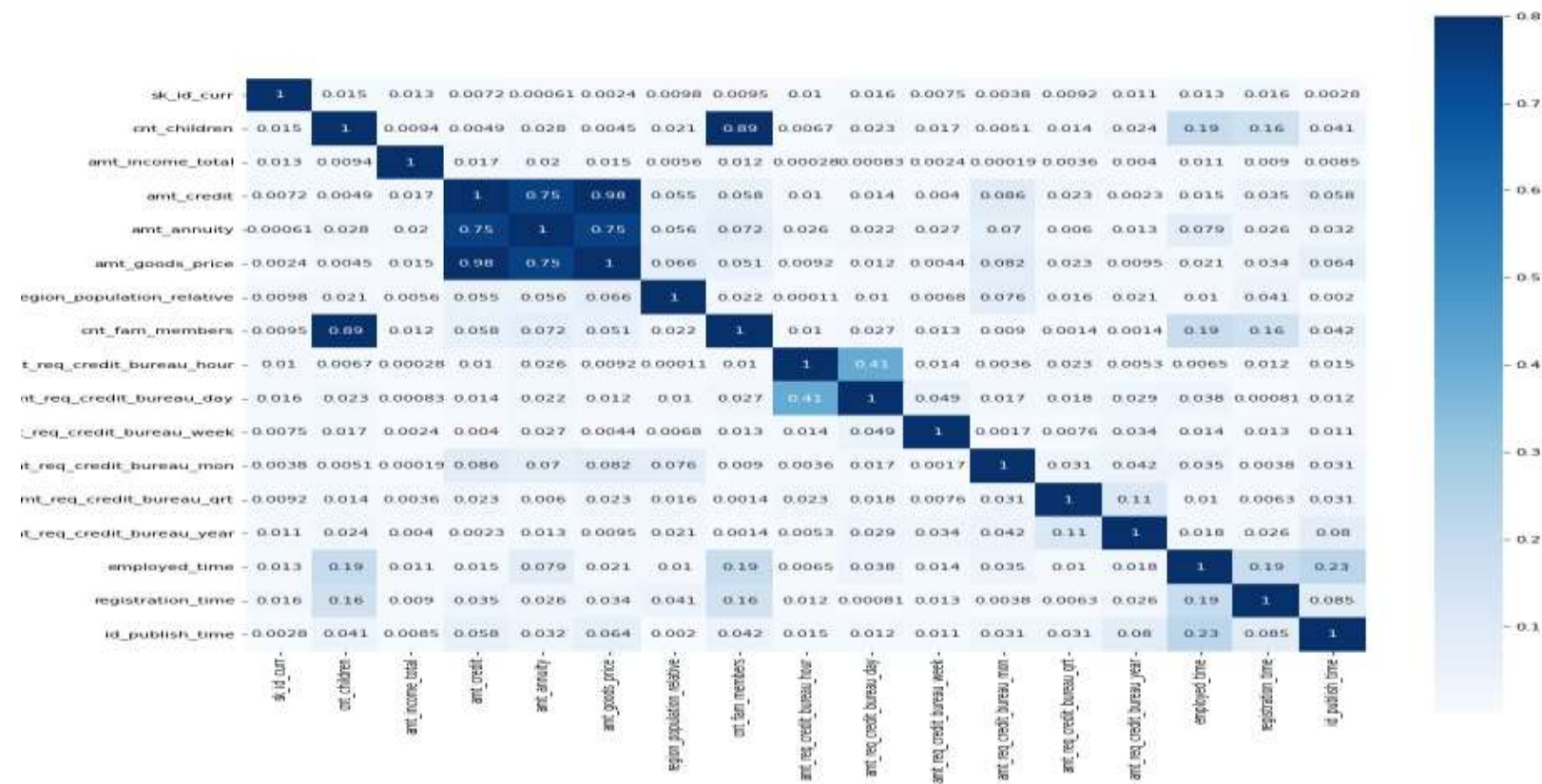
- Applications for low credit amount have high probability of being defaulters compared to high credit amount applications.



- Labourers constitute majority of the defaulters



Correlation of features with respect to payment difficulties



Inference from correlation matrix :

- Goods price and credit amount are highly correlated in case of defaulters, so applications for high credit amount with high good price or vice versa can be a defaulter.
- Family members and children count are highly correlated in defaulters which can be concluded as applicants with more family members and children have high probability of being defaulters.

Conclusions From Analysis:

- 8.7 % of customers have difficulty in paying the loans in time
- Customers who applied for Cash loans have high probability of being defaulters
- Females constitute the majority of the defaulters
- Working class are the majority of the defaulters
- Married clients often have difficulty repaying the loan and constitute most of defaulters

cont....

- Customers who earn less have high chance of being defaulters
- Applications for low credit amount have high probability of being defaulters compared to applications for high credit amount.
- Daily labour class constitute majority of defaulters
- Clients with big family and more children are often defaulters than people with small families
- Applications with high goods price and high credit amount and applications with low goods price and low credit amount combinations often can be defaulters.

Thank you