

Data Toolkit Case Study

Bank Loan Analysis

Objective

This case study aims to identify patterns which indicate if a client has difficulty paying their installments which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc. This will ensure that the consumers capable of repaying the loan are not rejected. Identification of such applicants using EDA and Hypothesis is the aim of this case study

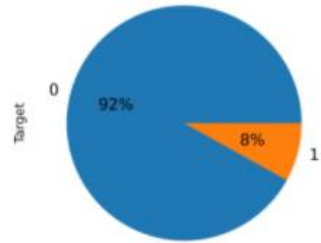
Hypotheses

- H_0 = Applicant is likely to repay the loan
- H_1 = Applicant is NOT likely to repay the loan

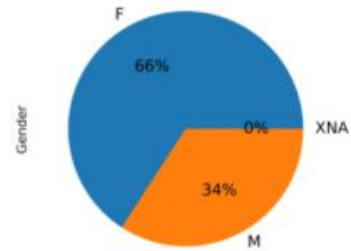
	H_0 is True	H_0 is False	Description of errors
We decide to reject H_0	Alpha (New Business Loss)	Correct Decision	Type 1(Alpha): If the applicant is likely to repay the loan, then not approving the loan results in a loss of business to the company
We fail to reject H_0	Correct Decision	Beta (Financial Loss)	Type 2(Beta): 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.

Understanding Variables

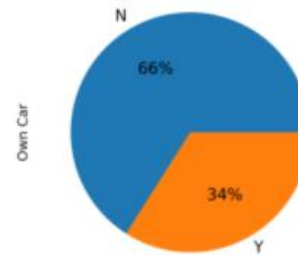
Pays on Time vs Defaulters



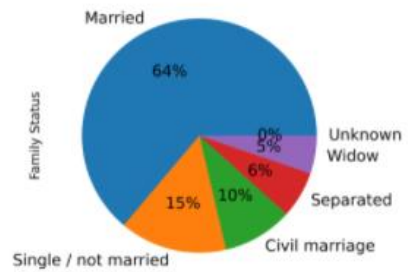
Gender Composition



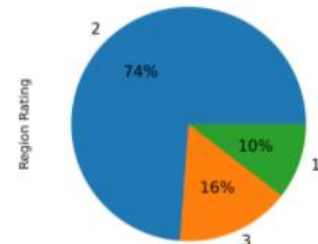
Car Ownership of Customers



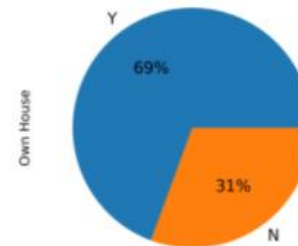
Family Status of Customers



Region Rating of Customers



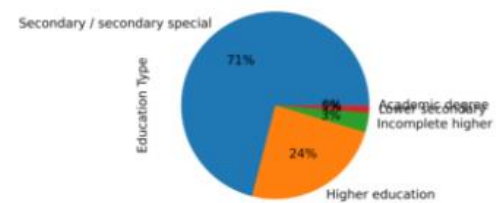
House Ownership of Customers



Income Type of Customers

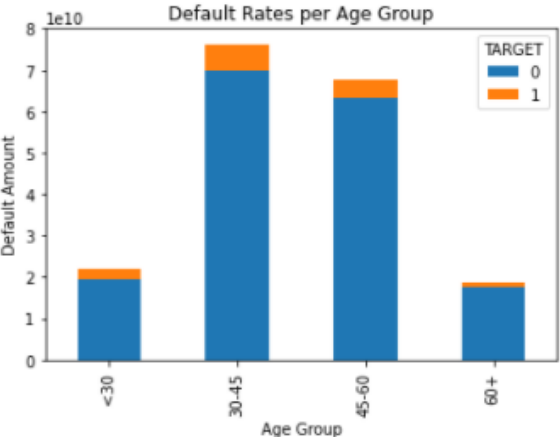
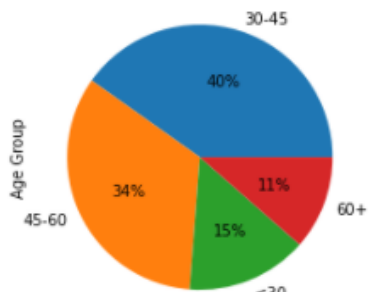


Education Type of Customers

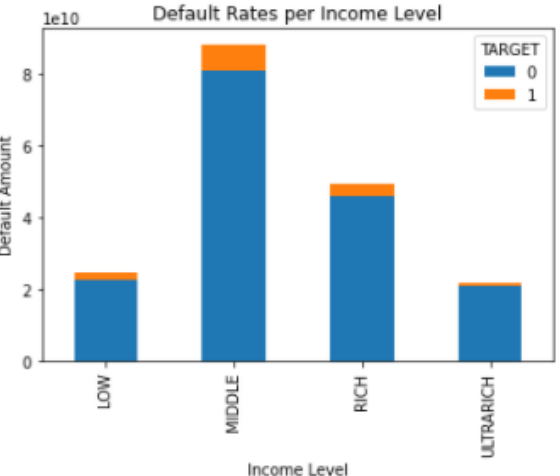
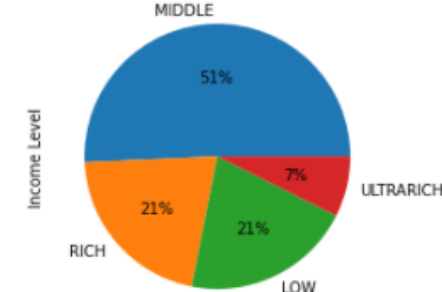


Understanding Variables

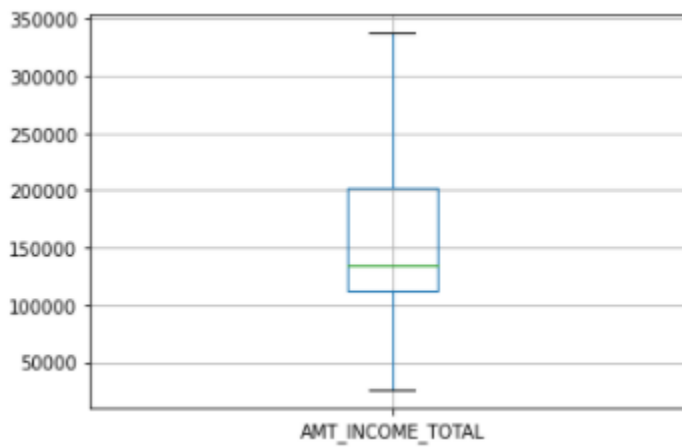
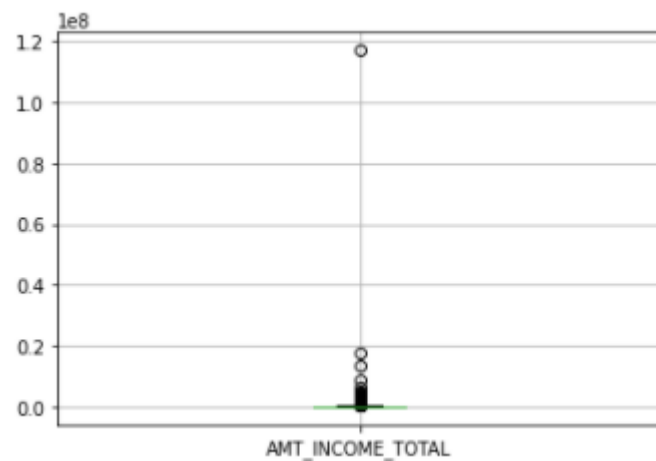
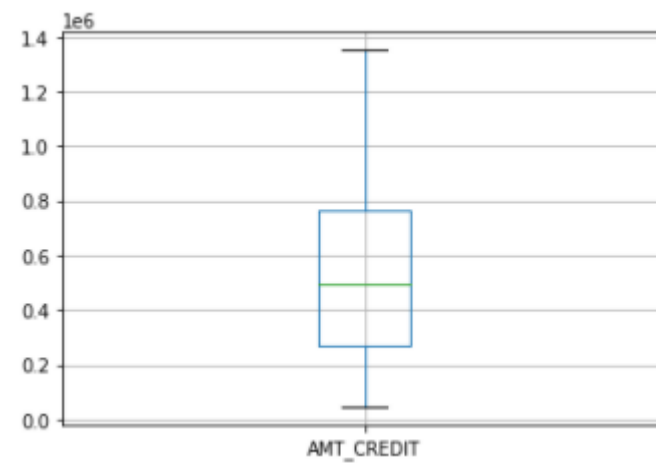
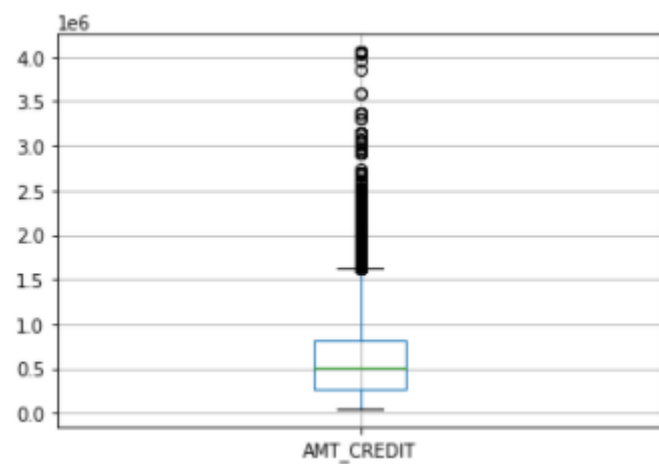
Customer Composition based on Age Group



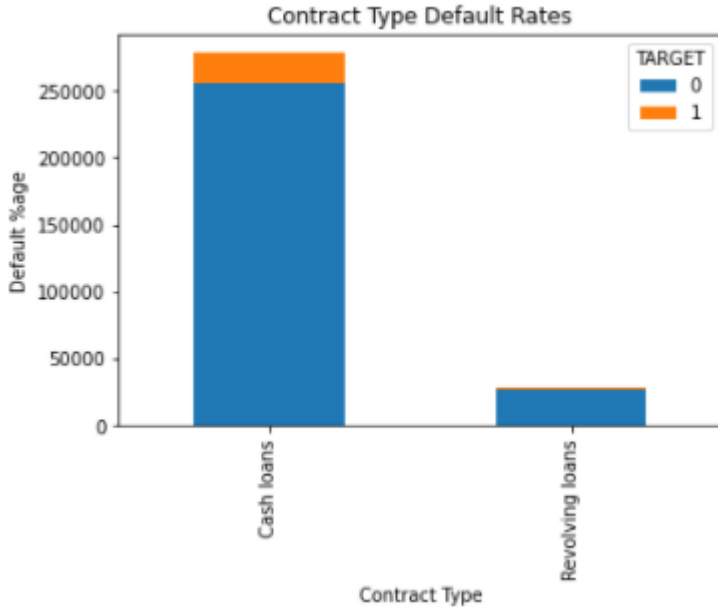
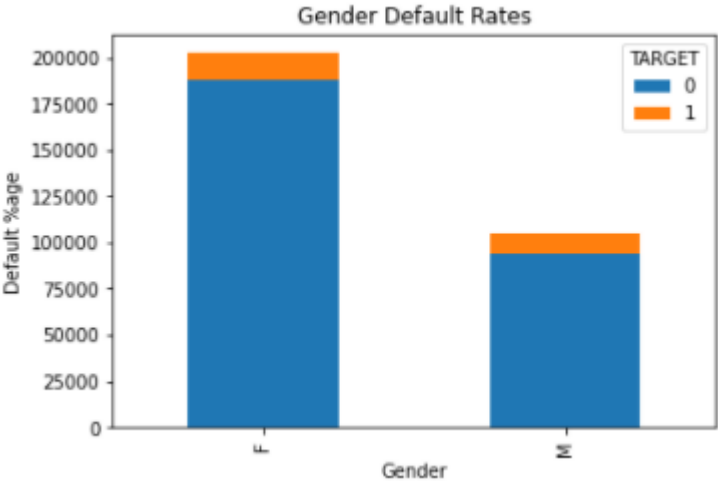
Customer Composition based on Income Level



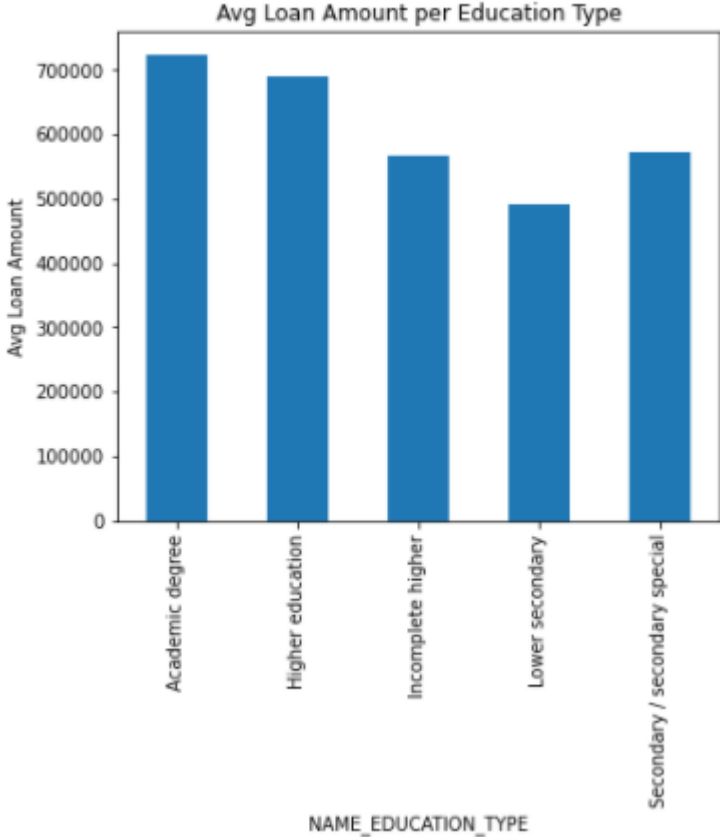
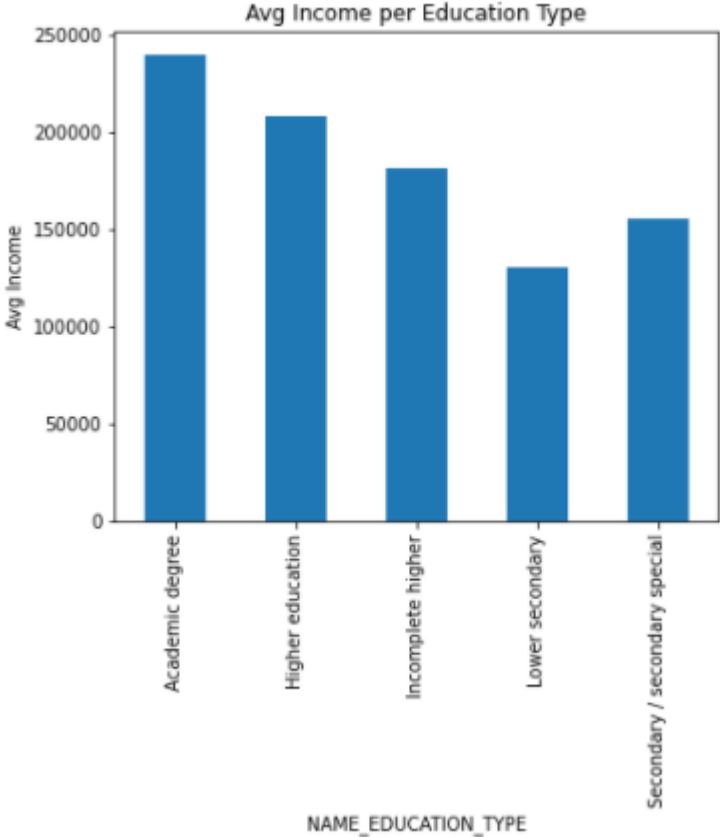
Removing Outliers



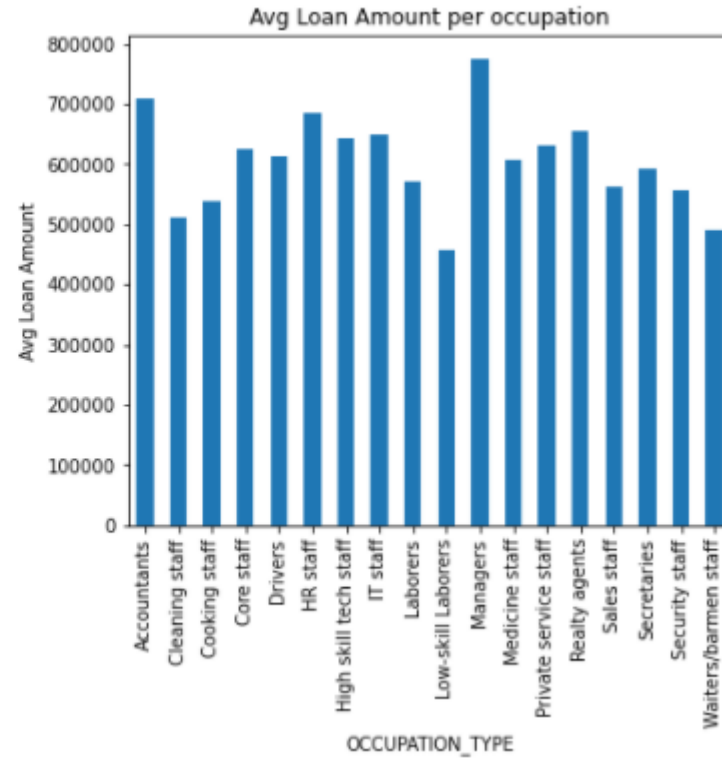
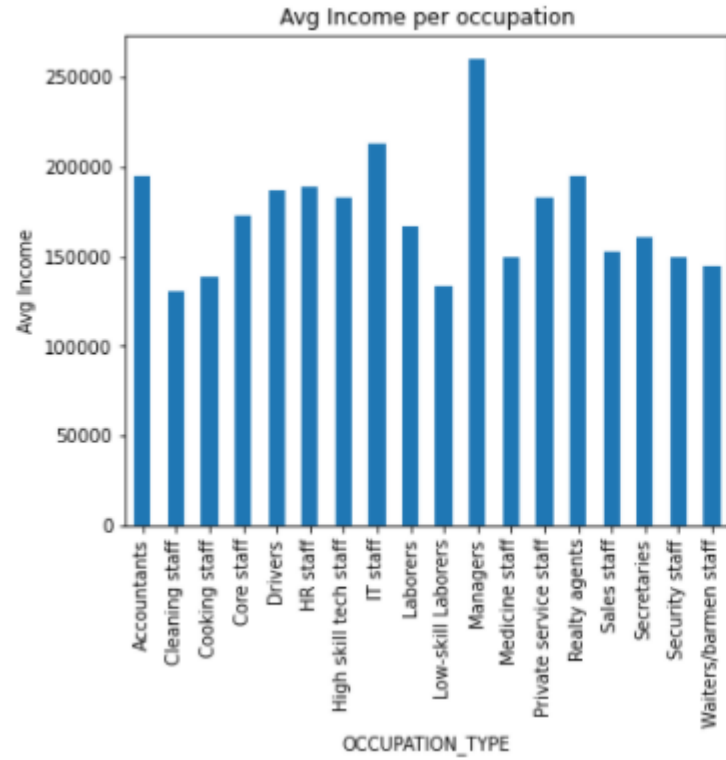
Default Rates



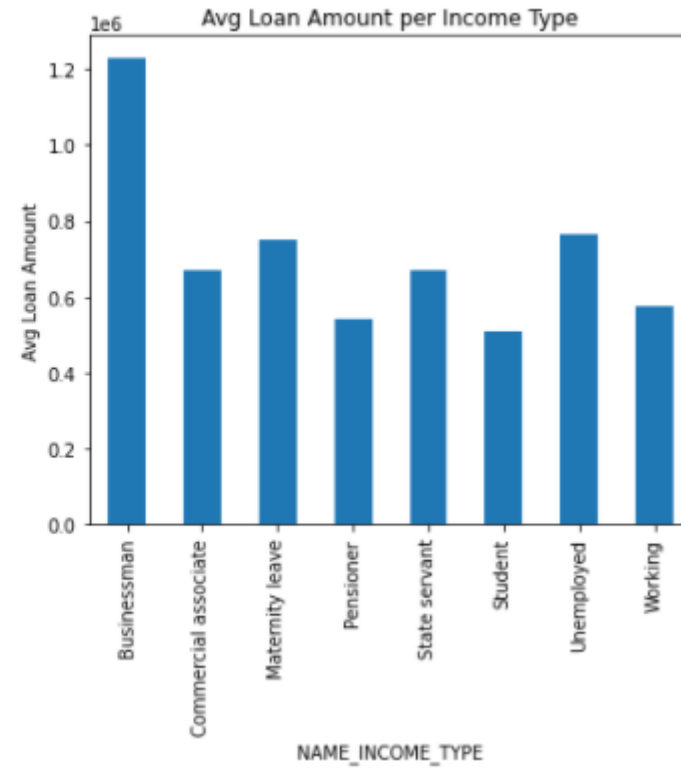
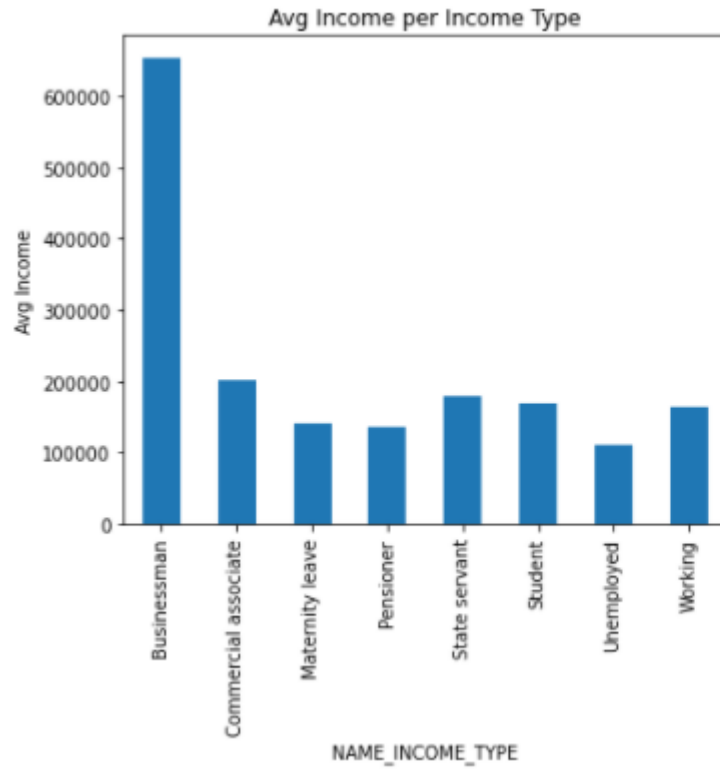
Income and Loan Amount basis Education type



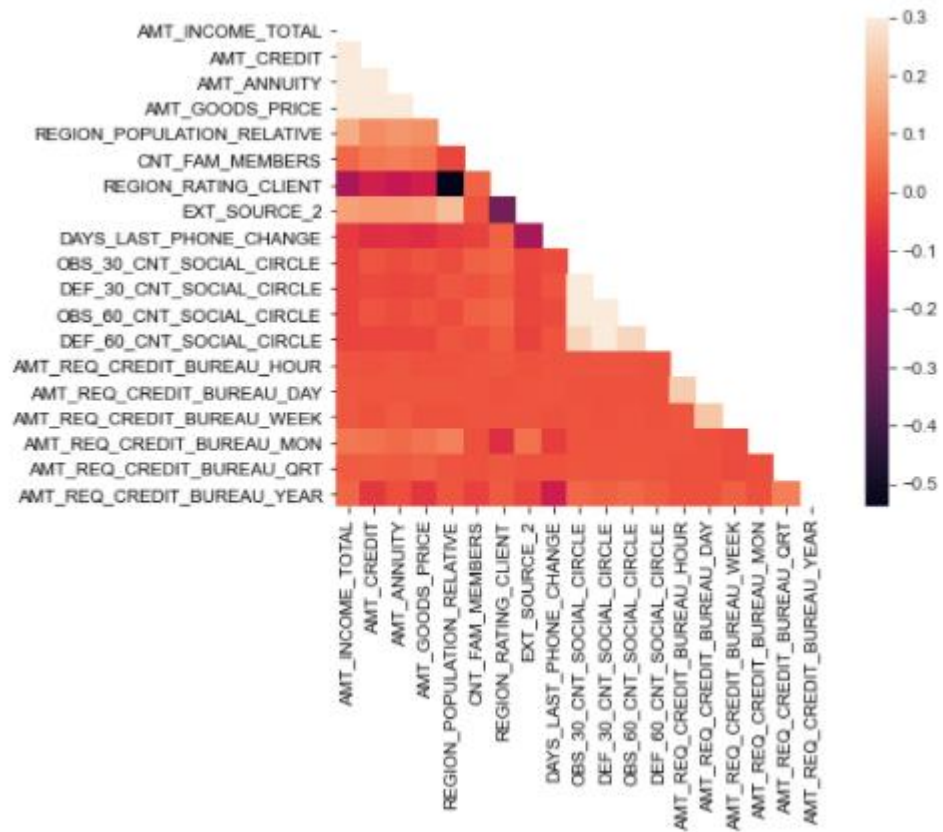
Income and Loan Amount basis occupation type



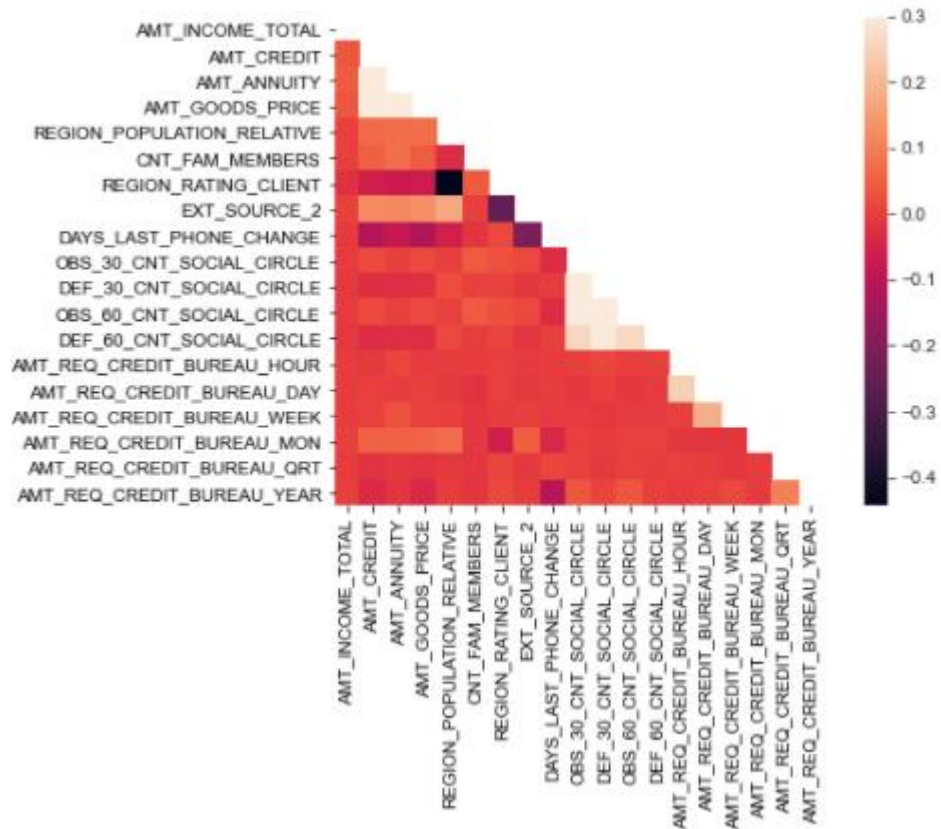
Income and Loan Amount basis occupation type



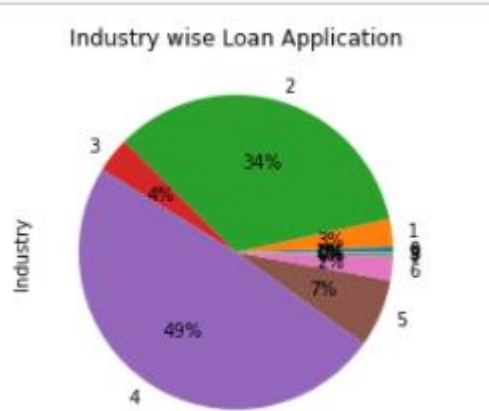
Correlation for customers who pay on time



Correlation for customers who default



Industry wise Loan application



	NAME_SELLER_INDUSTRY	SK_ID_PREV
4	Consumer electronics	341696
2	Connectivity	238341
5	Furniture	48963
3	Construction	25417
1	Clothing	19896
6	Industry	16567
0	Auto technology	4080
7	Jewelry	2107
8	MLM partners	921
9	Tourism	374

Client wise Loan application and status

