

Credit EDA - Case Study

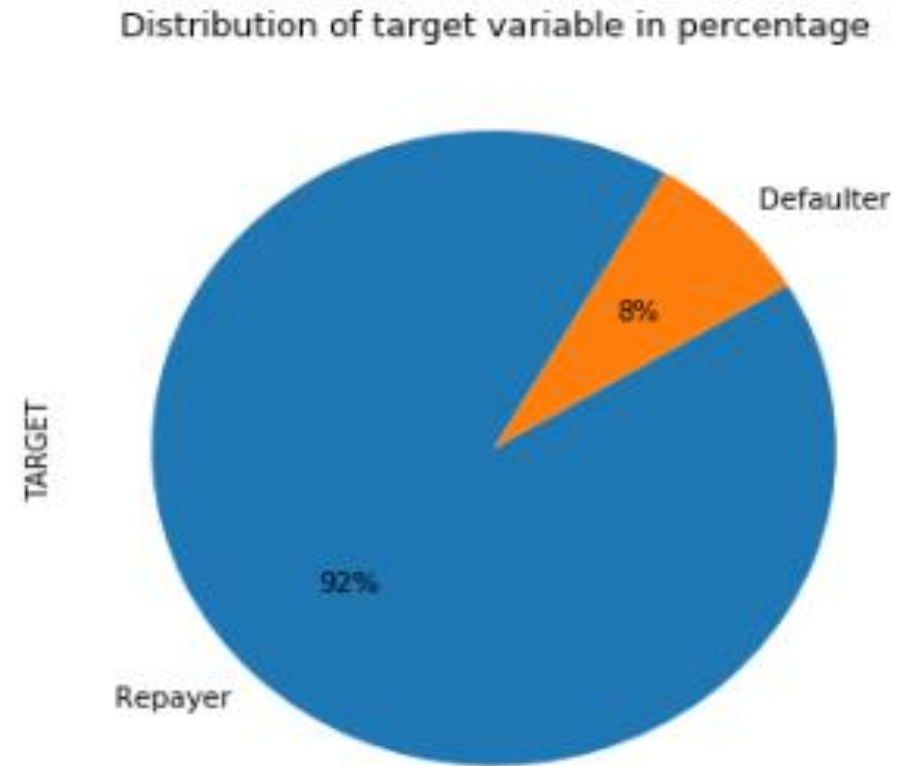
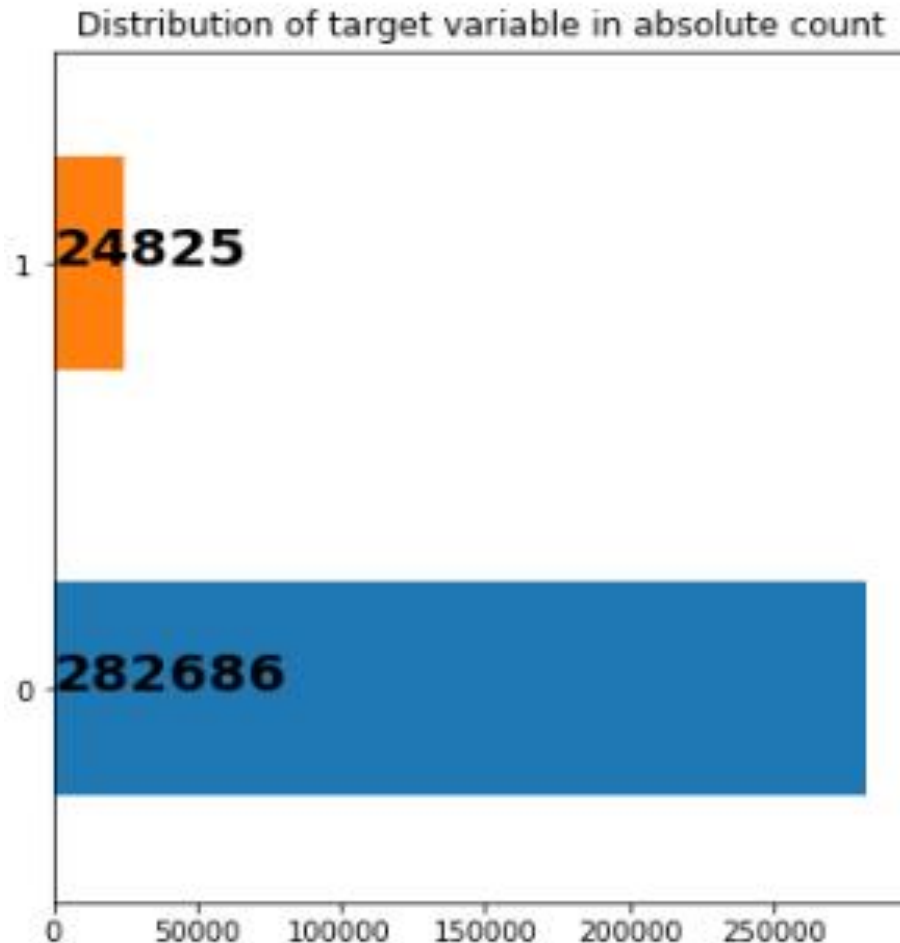
- **Problem Statement**

There 2 types of risks associated with the bank's decision

1. Applicants likely to repay the loan, but not approving the loans results in a business loss
2. Applicants not likely to repay the loan, but approving the loans results in a financial loss

Data Imbalance Check :-

From below plot we conclude that data is imbalance as 92% are with Target 0 i.e. client with no payment difficulty while rest 9% are clients with payment difficulty.



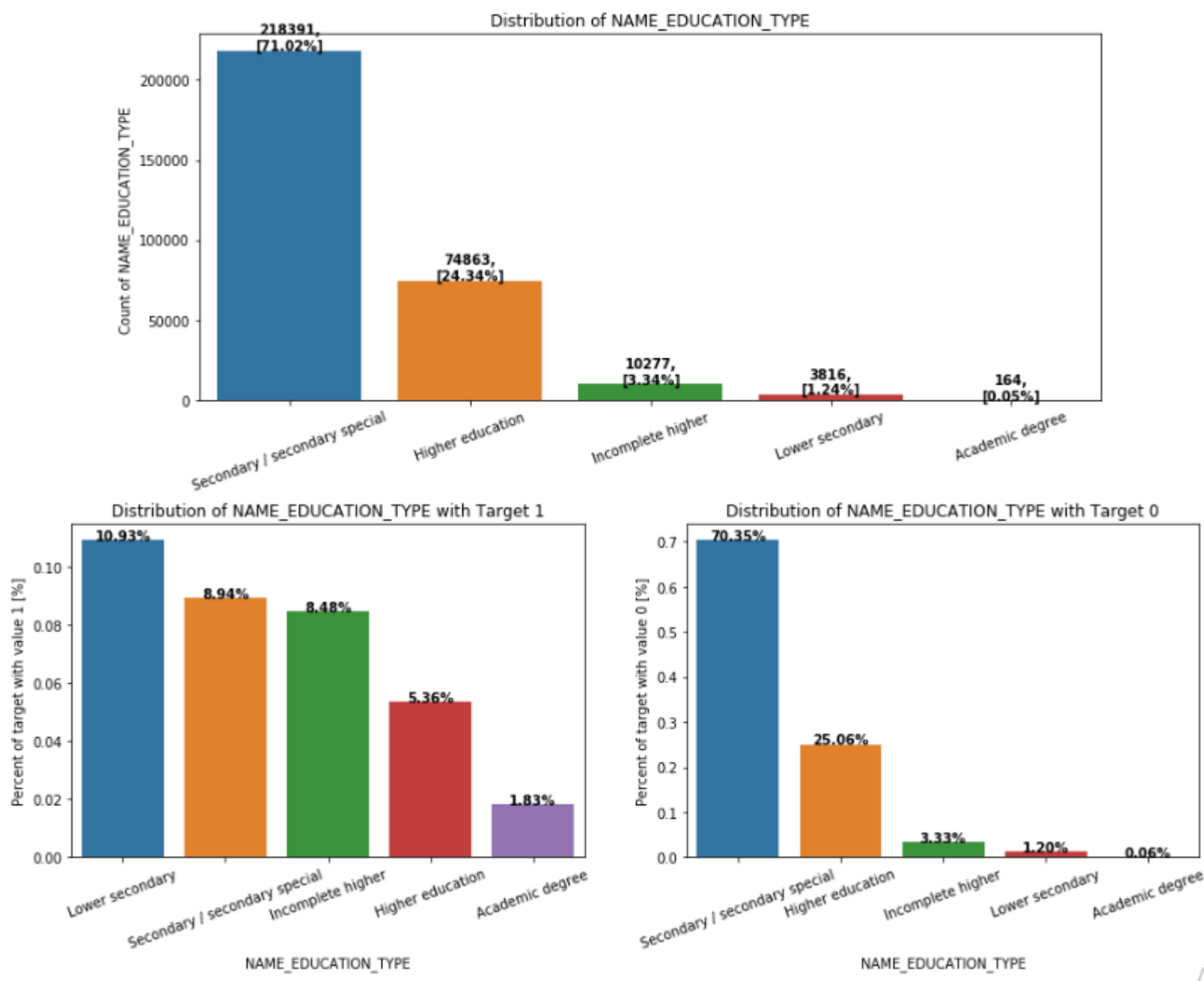
1. Following are important variables for deciding rate of approval by bank whose plot and reason explained in subsequent slides

variables & its categories where rate of acceptance should be higher

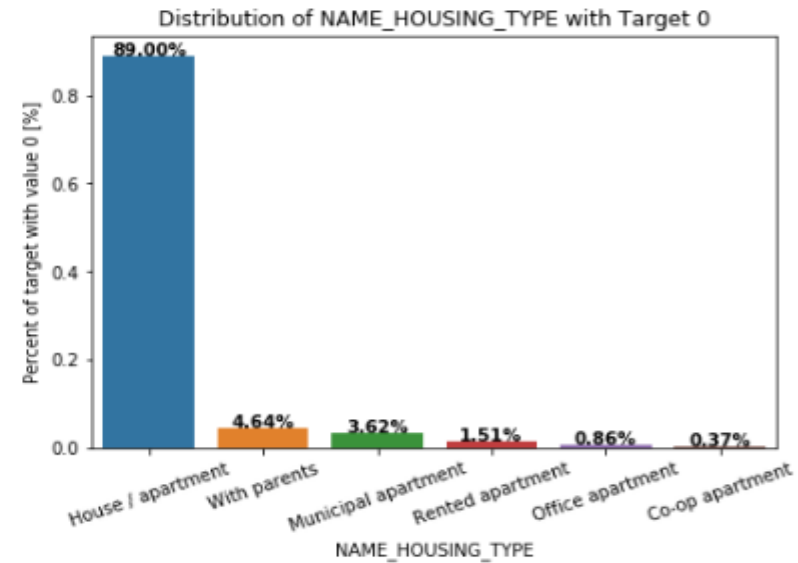
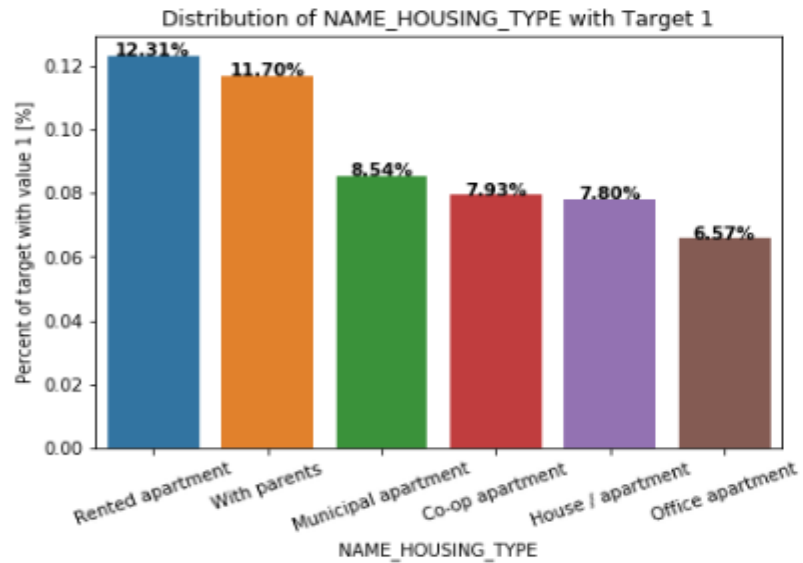
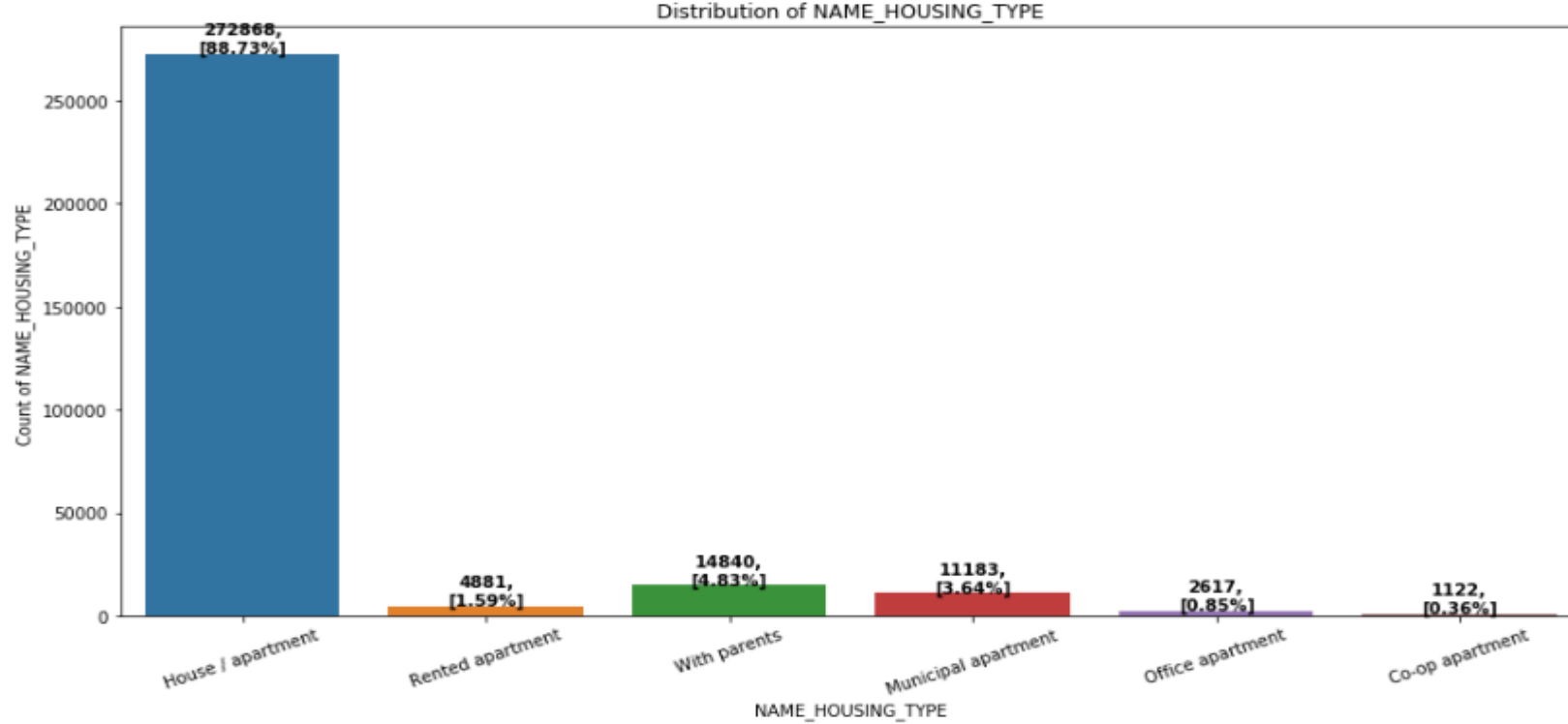
1. House/Apartment type of Housing Type
2. Secondary/Secondary Special Category in Education Type
3. In Type_Suite Unaccompanied type
4. In contract type Cash Loans

variables & its categories where rate of rejection should be lower

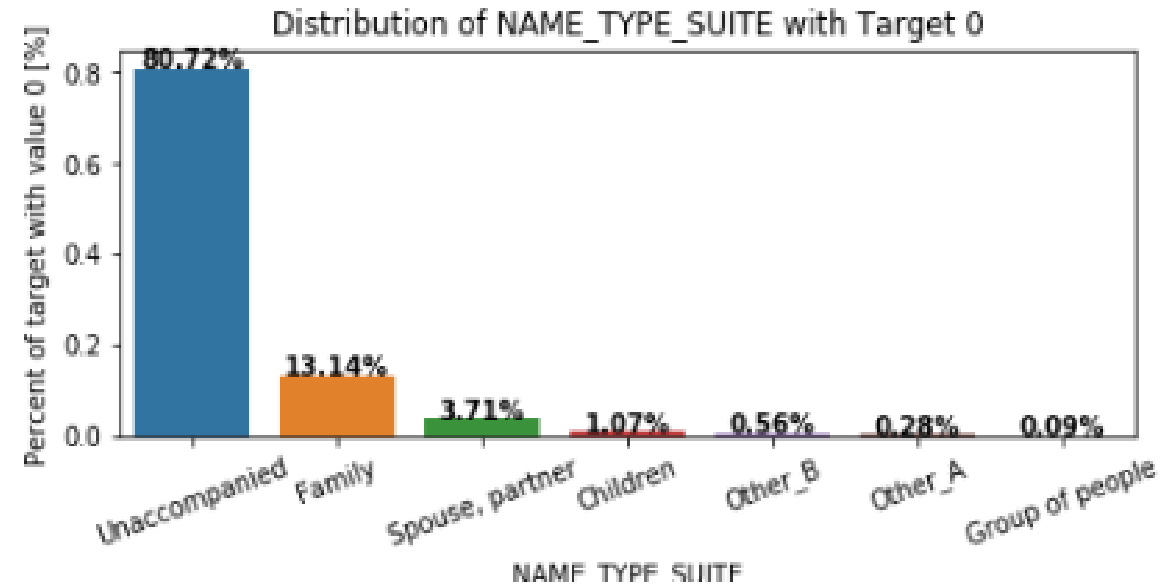
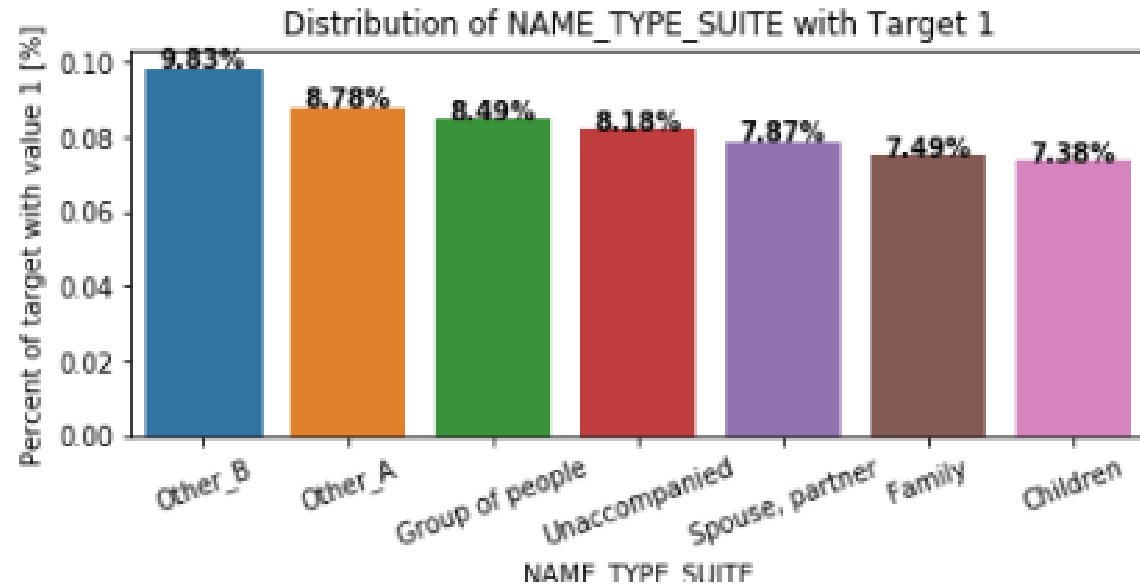
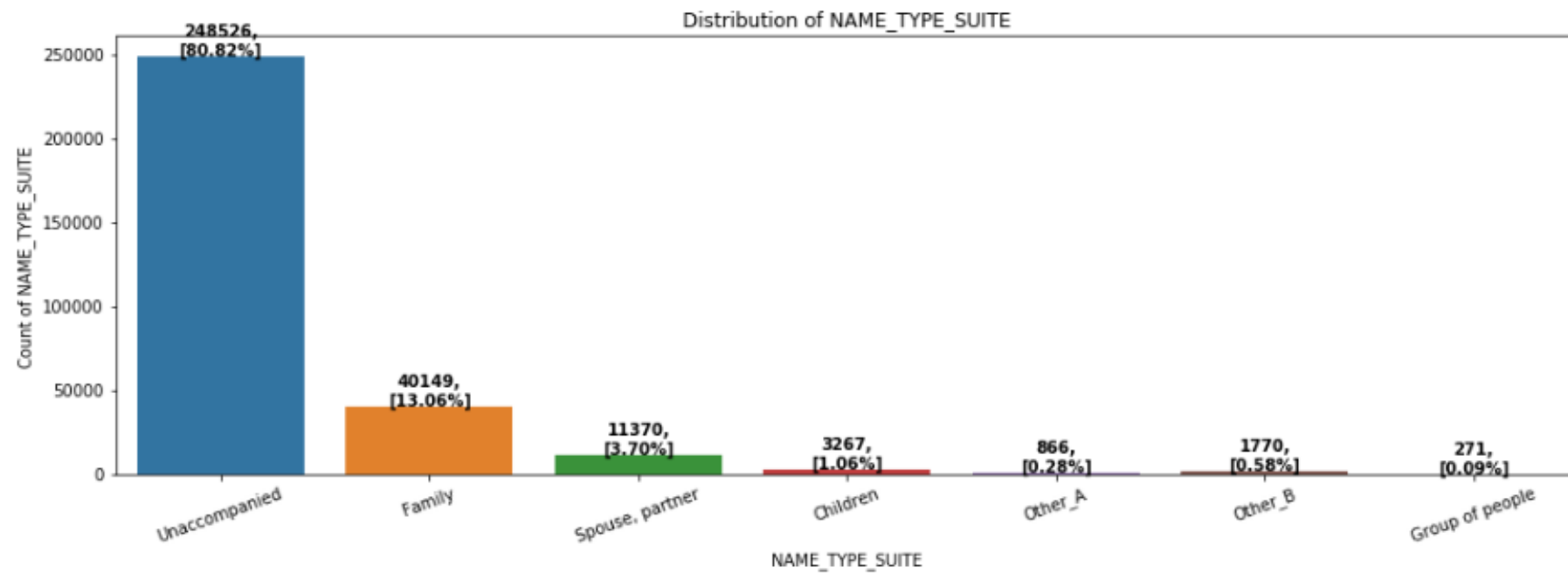
1. In Income_Type Unemployed & Maternity Leave
2. Low-Skilled Laborer in Occupation Type
3. In housing type Rented category
4. In education type lower secondary



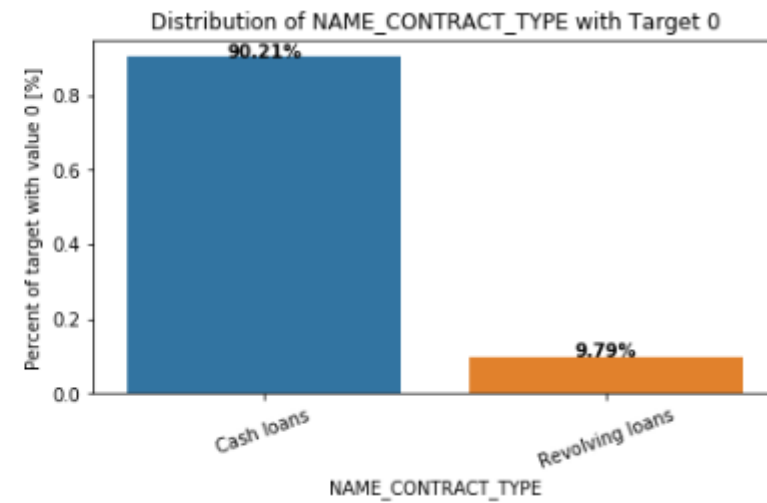
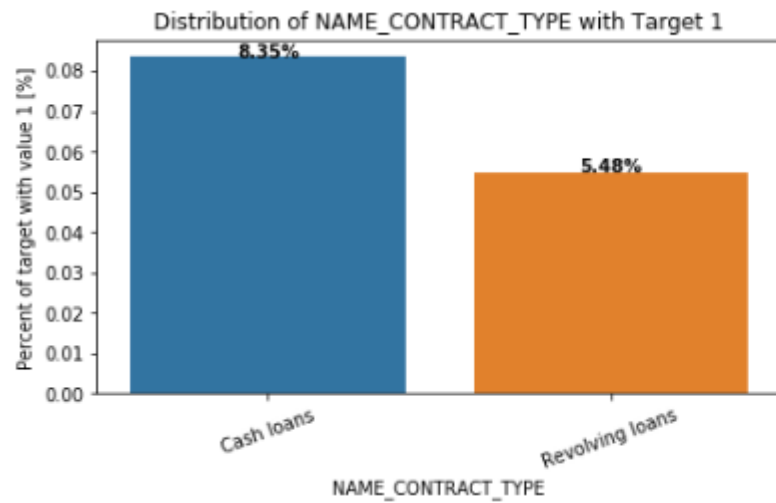
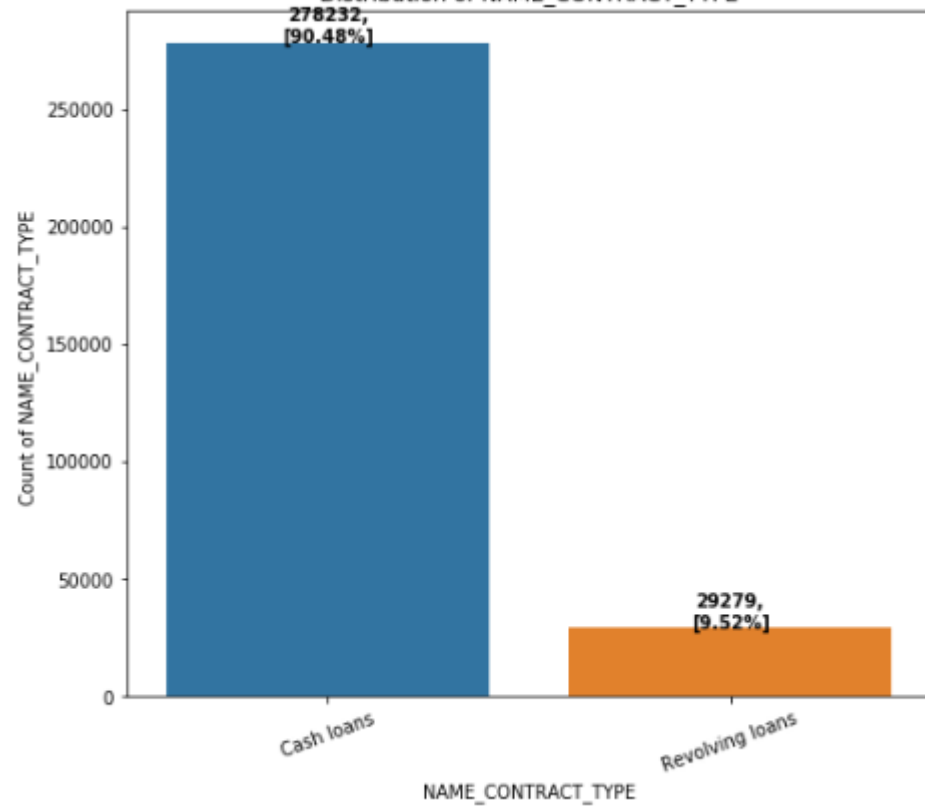
71% of data in education category is from Secondary/Secondary Special and the rate of default is very low is only about 8.9%



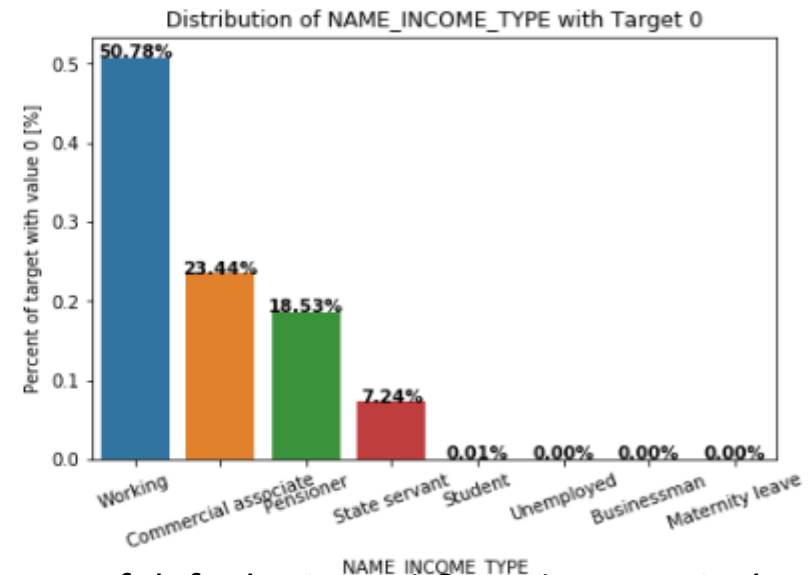
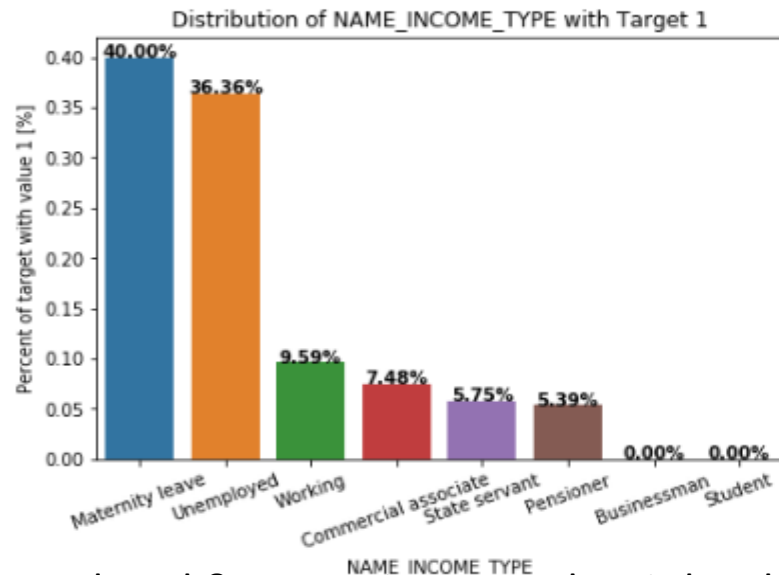
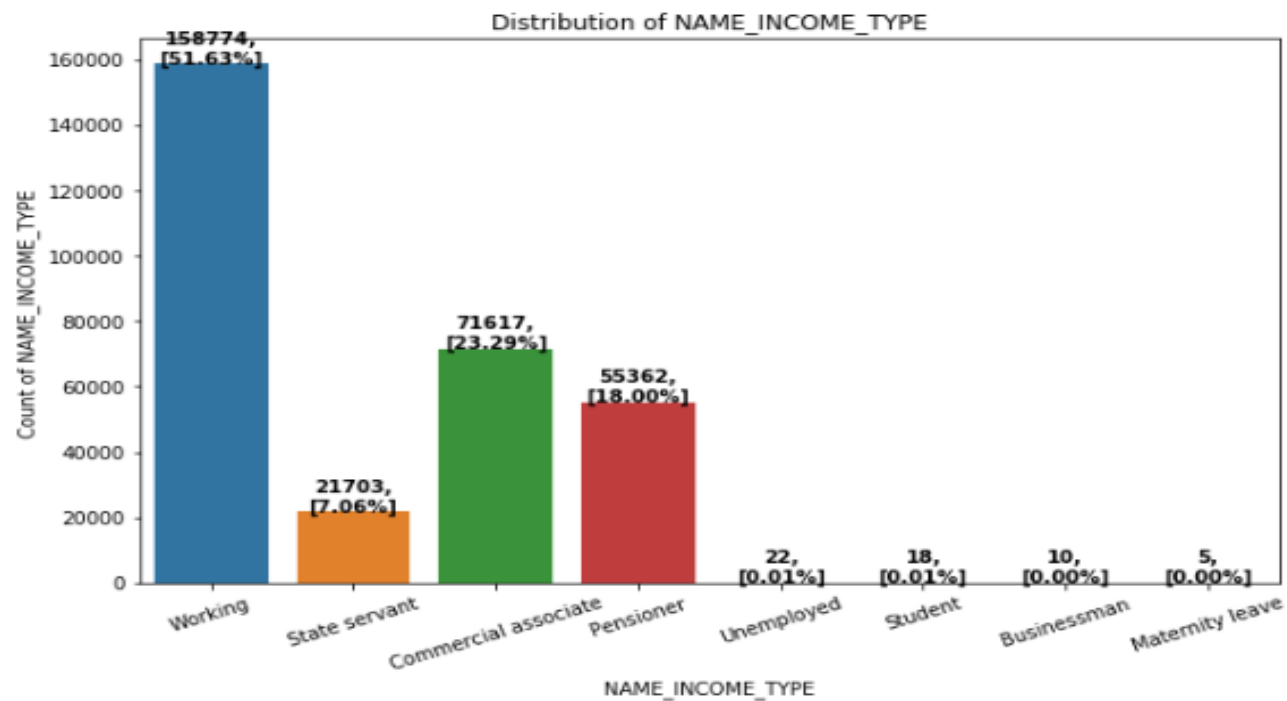
In Housing Type House/Apartment type contributes to 88% of data but the rate of default is very low is only 7.8% which is at 5th position in category.



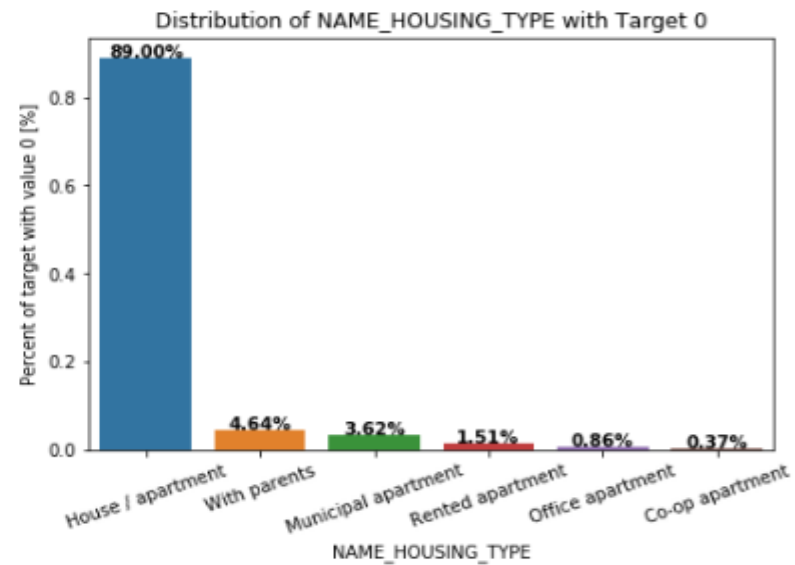
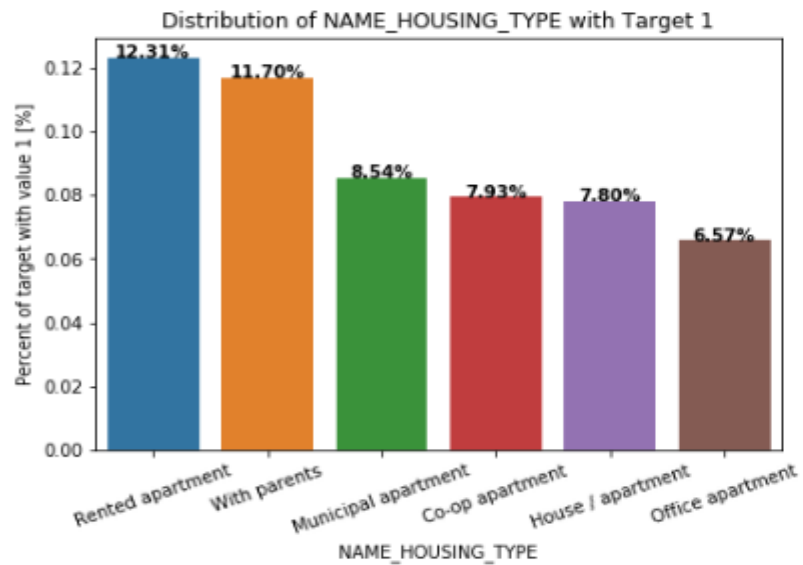
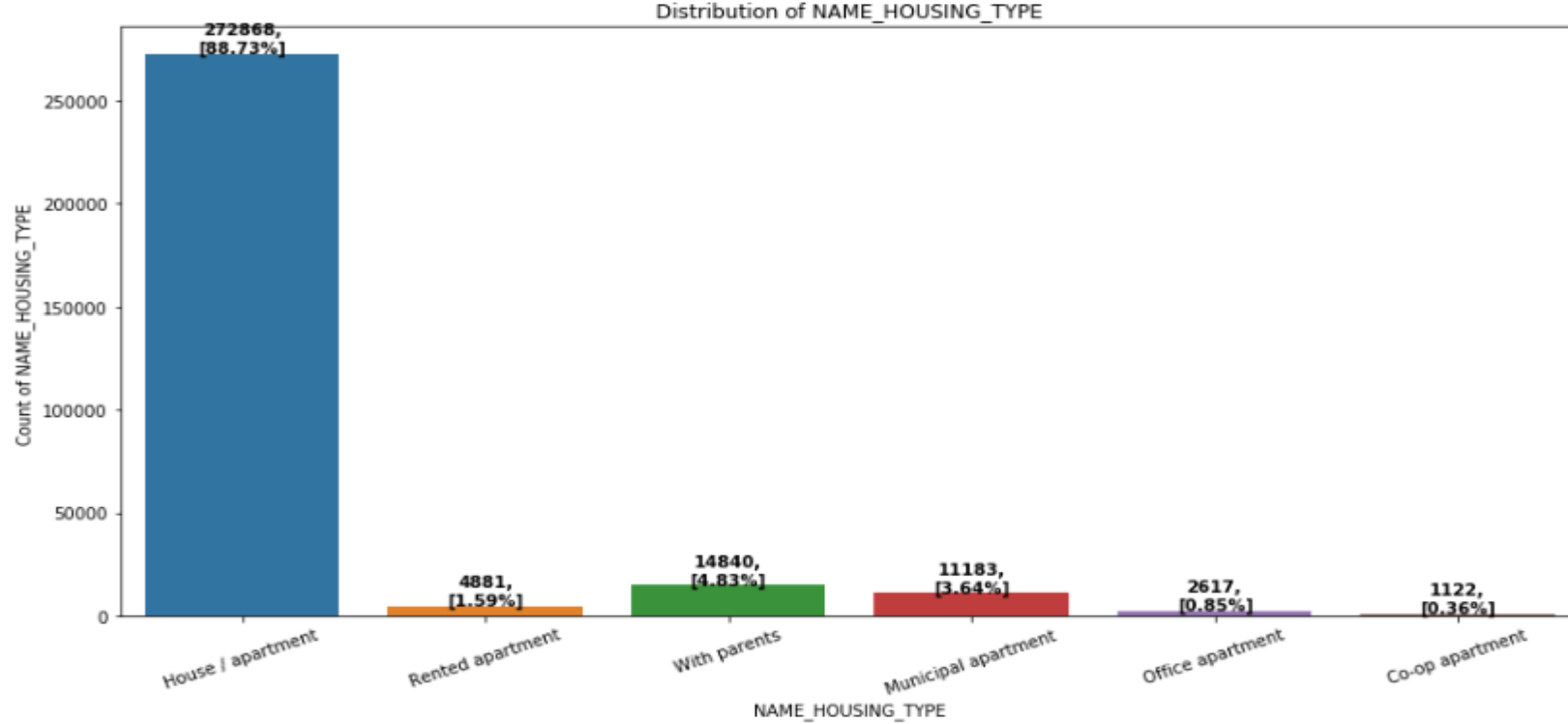
In Type_Suite Unaccompanied type is having 80% of data but 8.14% defaulter which is 4th position in this category.



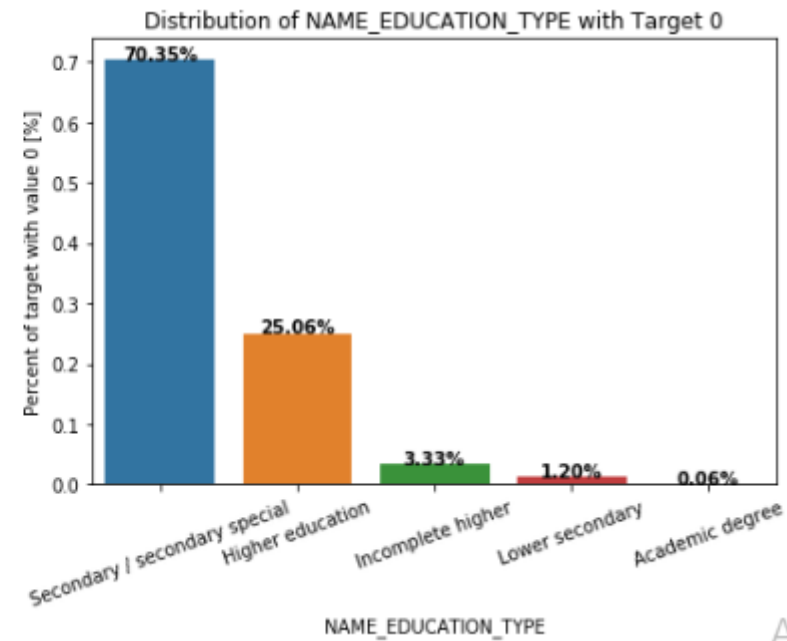
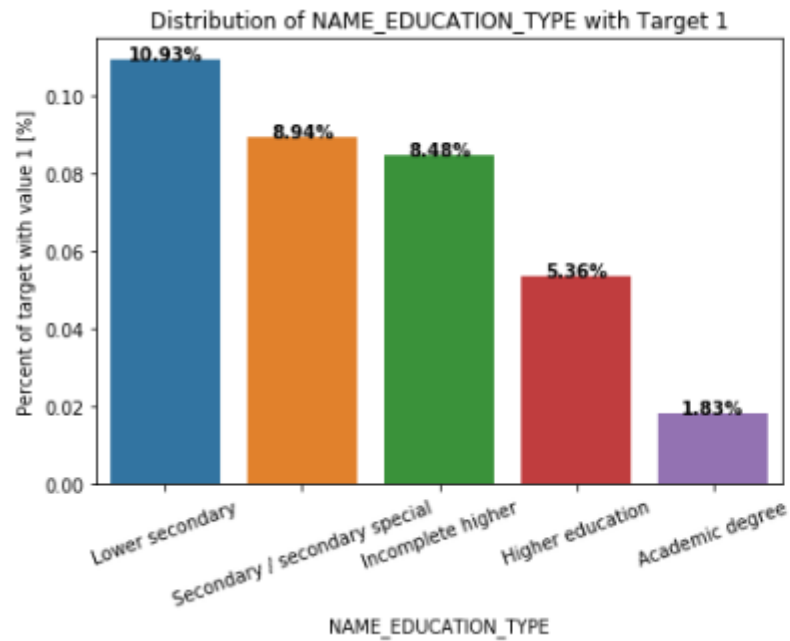
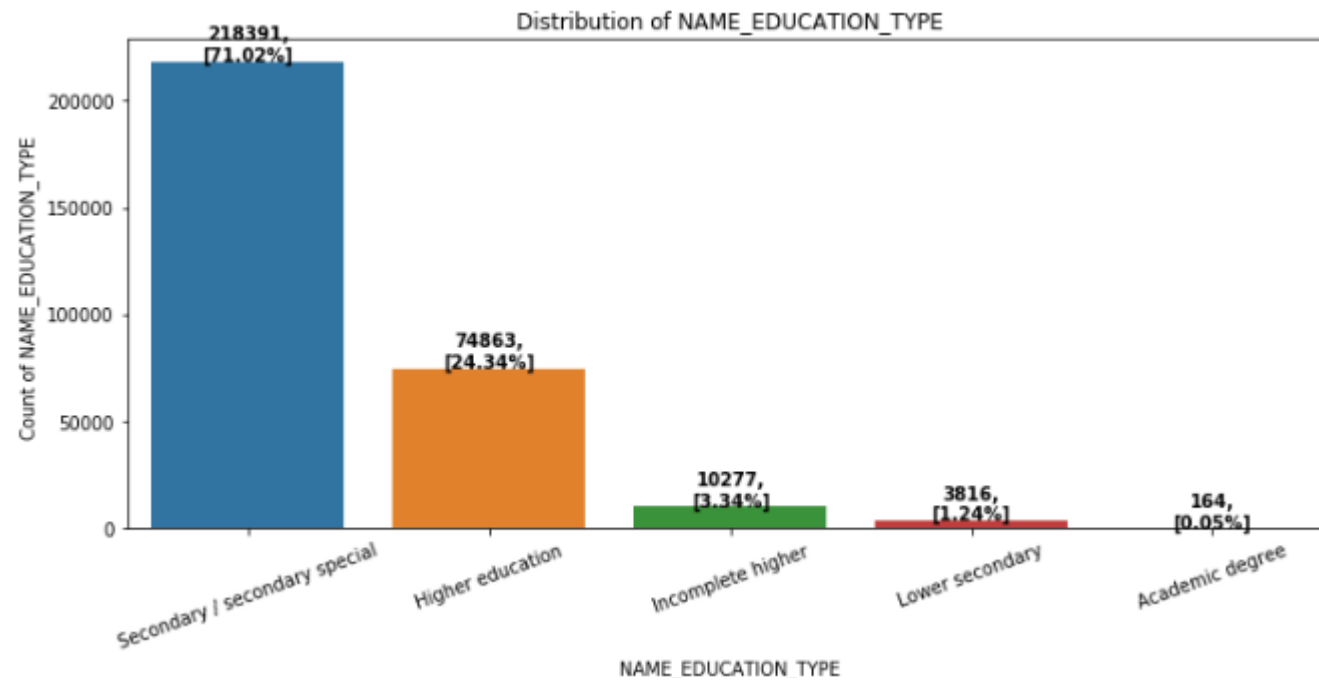
In contract type 90% of contribution is from Cash Loans & rate of defaulter is only 8.35%.



In Income_Type Unemployed & Maternity Leave data is less but rate of default is 36% & 40% respectively which is very high.



In housing type Rented category is having data of 1.59% but rate of default is 12.31%



In education type lower secondary is having 1.24% data contribution but rate of default is 10.93% which is very highest in category.