Presentation on Credit EDA Case Study

Prepared BY:

PUJARANI SAMAL

Abstract

- Credit risk analysis will assist the business in deciding whether to approve a loan based on the applicant's profile. This limits commercial loss to the corporation and prevents financial loss for the business.

- I have followed several steps such as Understanding Business problems by creating sub-problems, preparation of clean data, and Exploratory Data Analysis. I have added the relevant results along with Problem Solving framework in the subsequent slides.

Steps

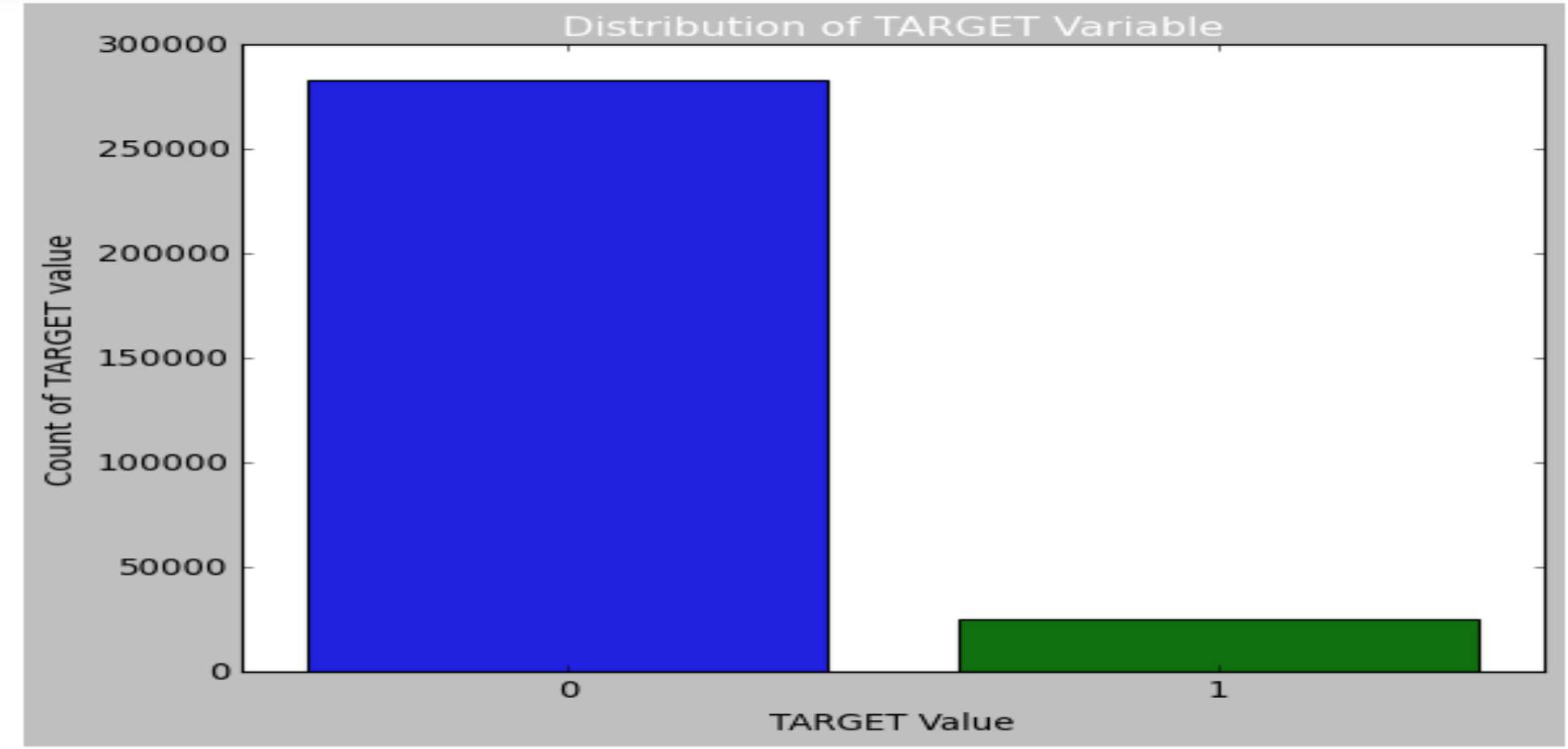
- Understanding of Raw Data
- Data Cleaning and Manipulation
- Look for data imbalance, univariate, and bivariate analysis, as well as correlation from Application Data
- Data merging with application data with previous application data.
- data analysis by univariate, bivariate analysis, and correlation in previous application data
- Recommendations and Risks

Preparation of Raw Data

- Preparation of Raw Data using two different files *application_data i.e 'df' and previous_application i.e 'df2'*.
- Selected the Required Clean data by removing Null values Missing Value and filter by categorical and numeric type.
- Based on categorical and numeric we have to do univariate and bivariate analysis. .
- Below Graph mentioned payment difficulties on Application_data and after merging both data we can find out approval, canceled, refused, and unused status.

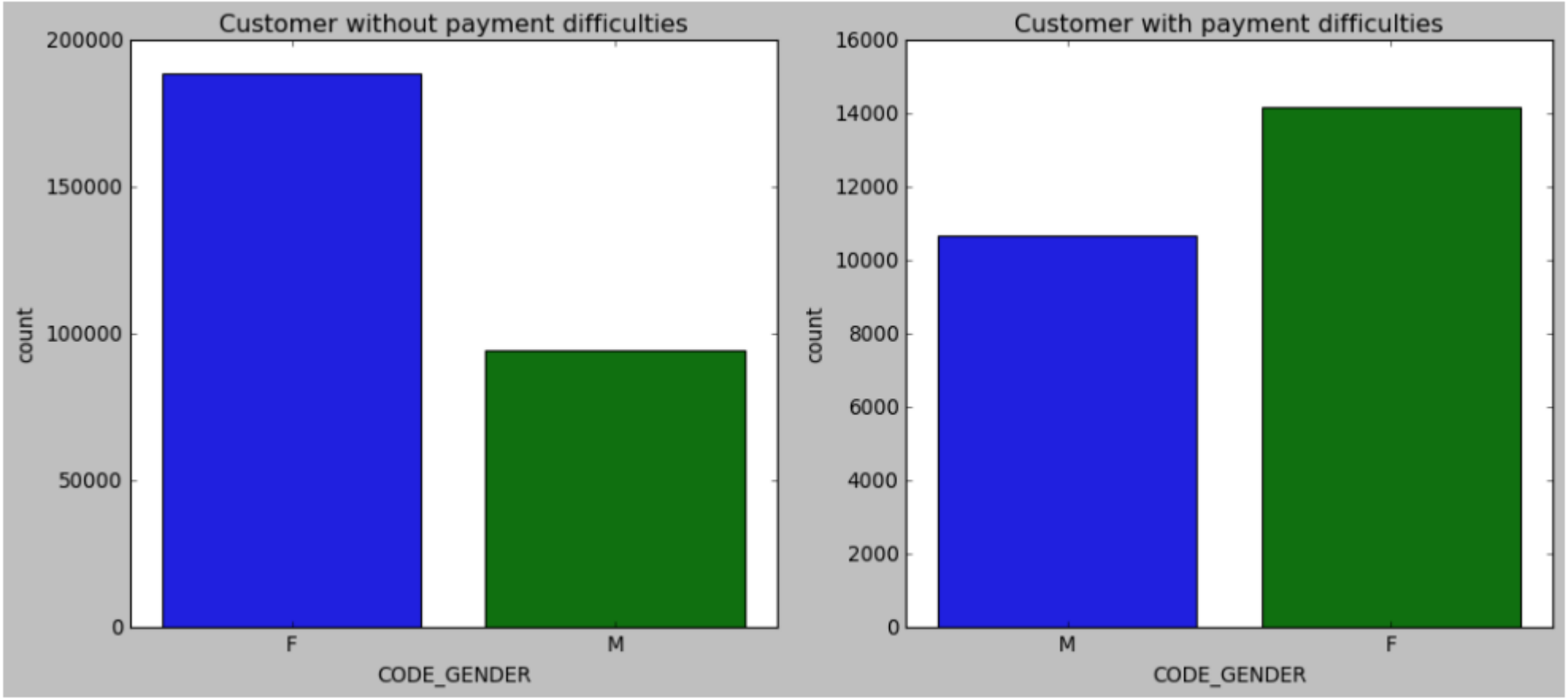
Distribution of TARGET Variable

- Below Graph Represents default and non-default value



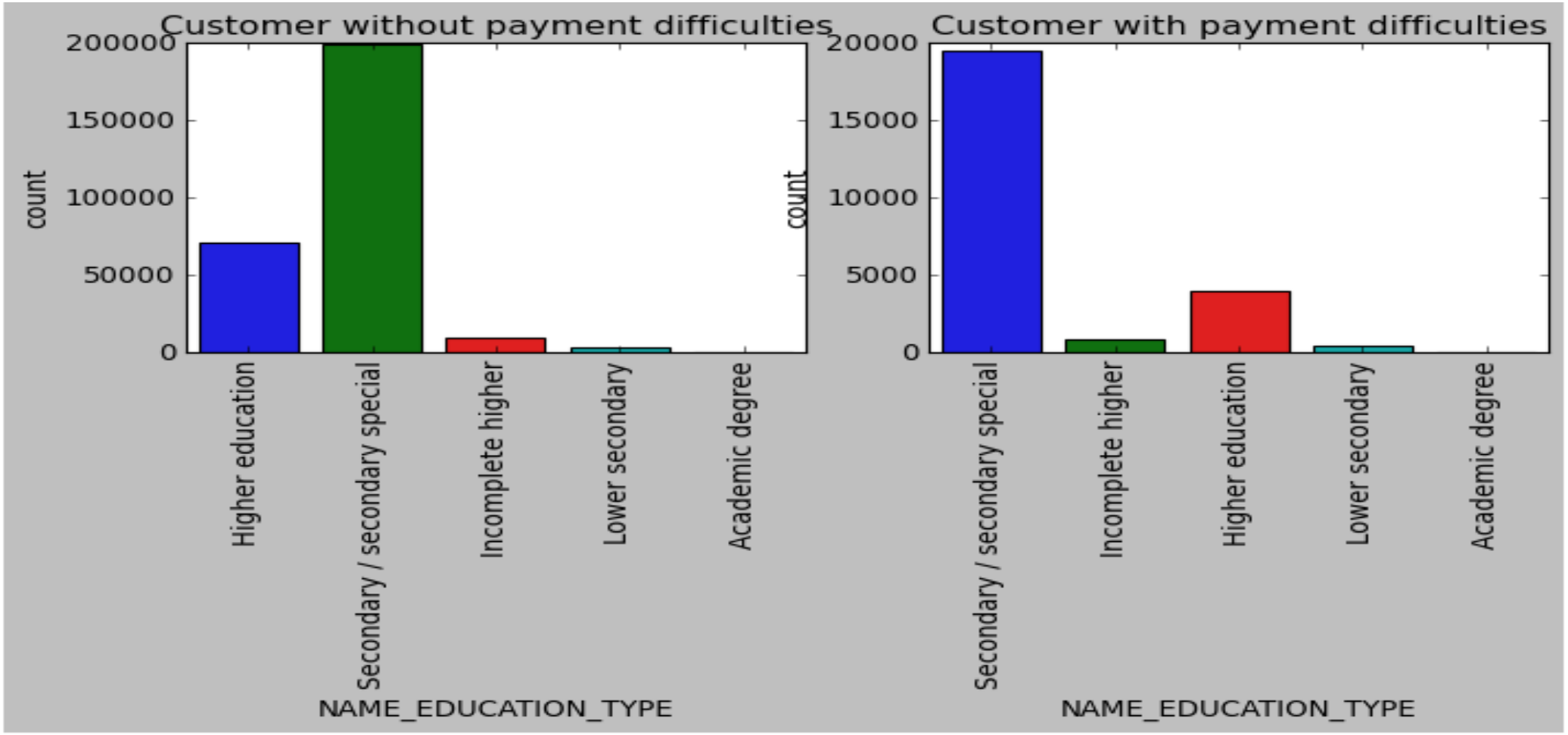
Payment difficulties: Code_Gender

- Below is the default and nondefault comparison with CODE_GENDER payment status



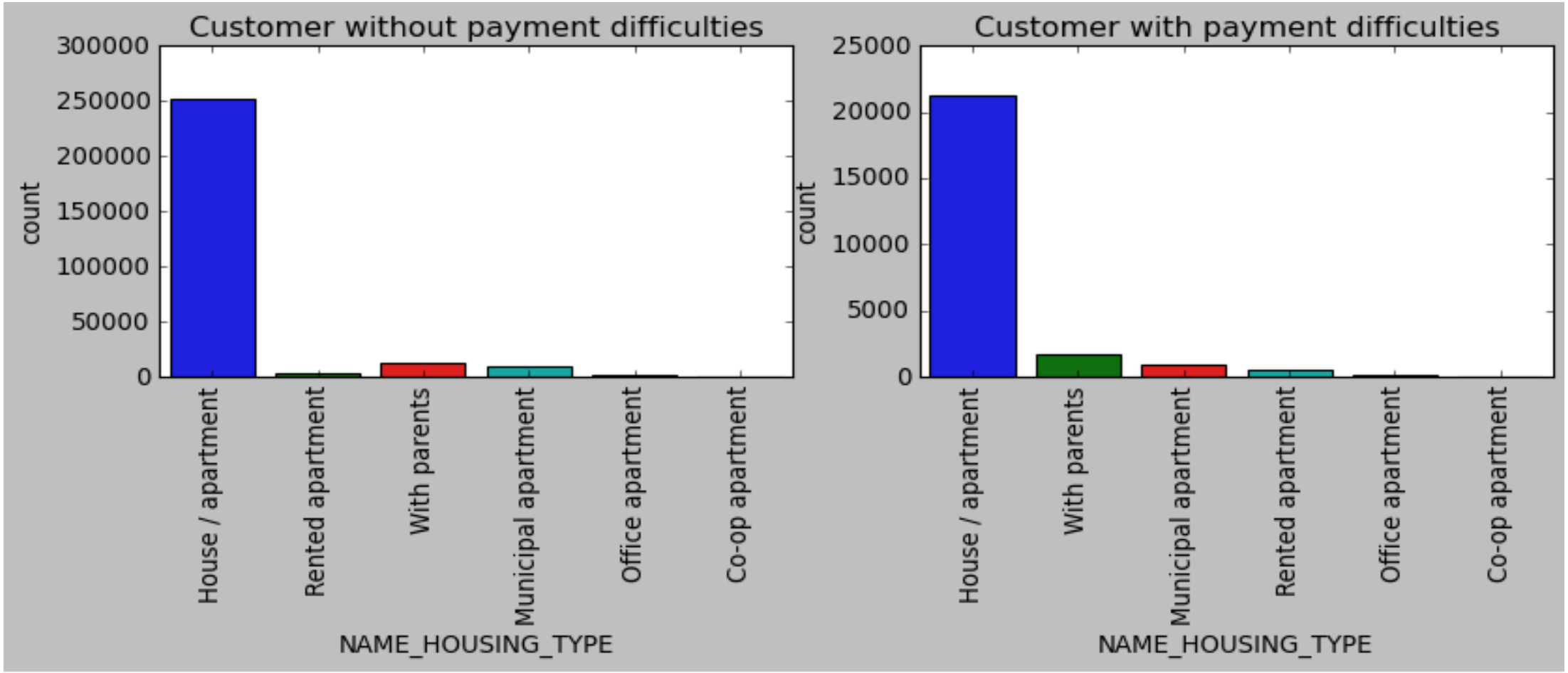
Payment difficulties: NAME_EDUCATION_TYPE

- Below is the default and non-default comparison with education type payment status



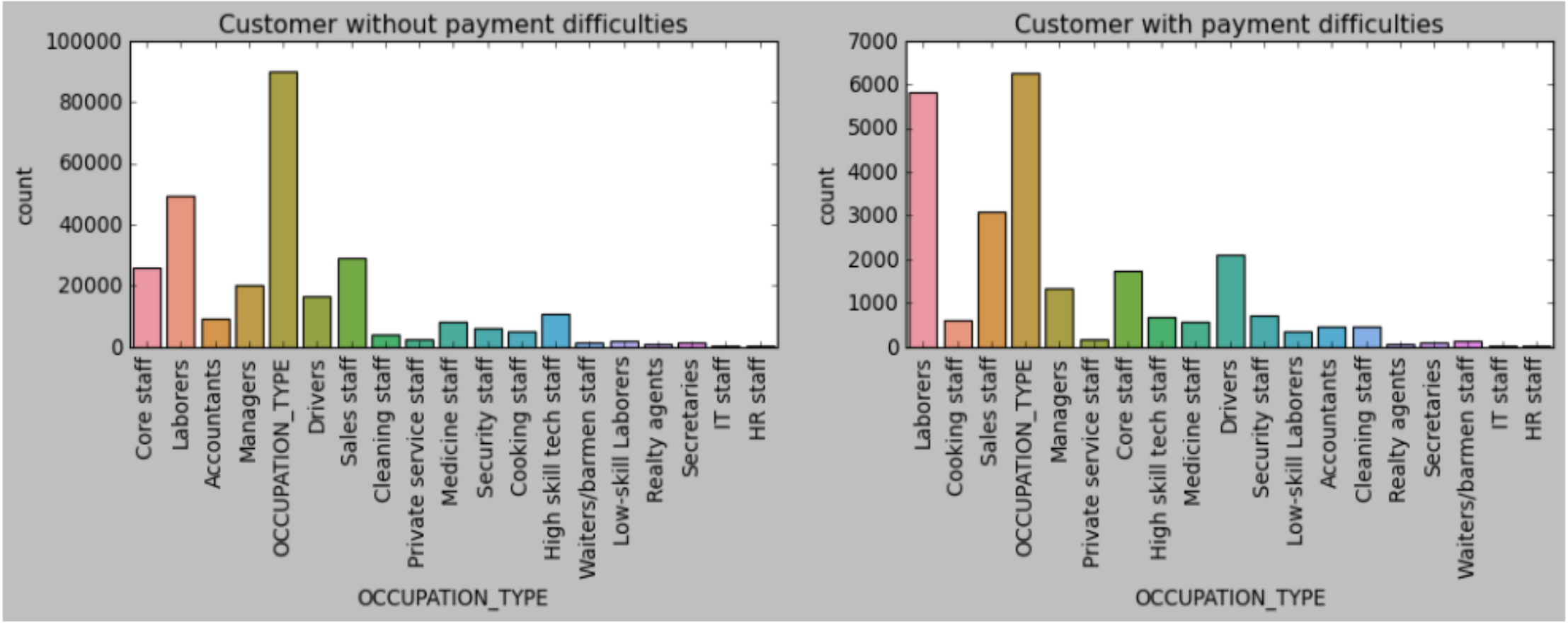
Payment difficulties: NAME_HOUSING_TYPE

- Below is default and non-default comparison with Housing Loan status



payment difficulties: OCCUPATION_TYPE

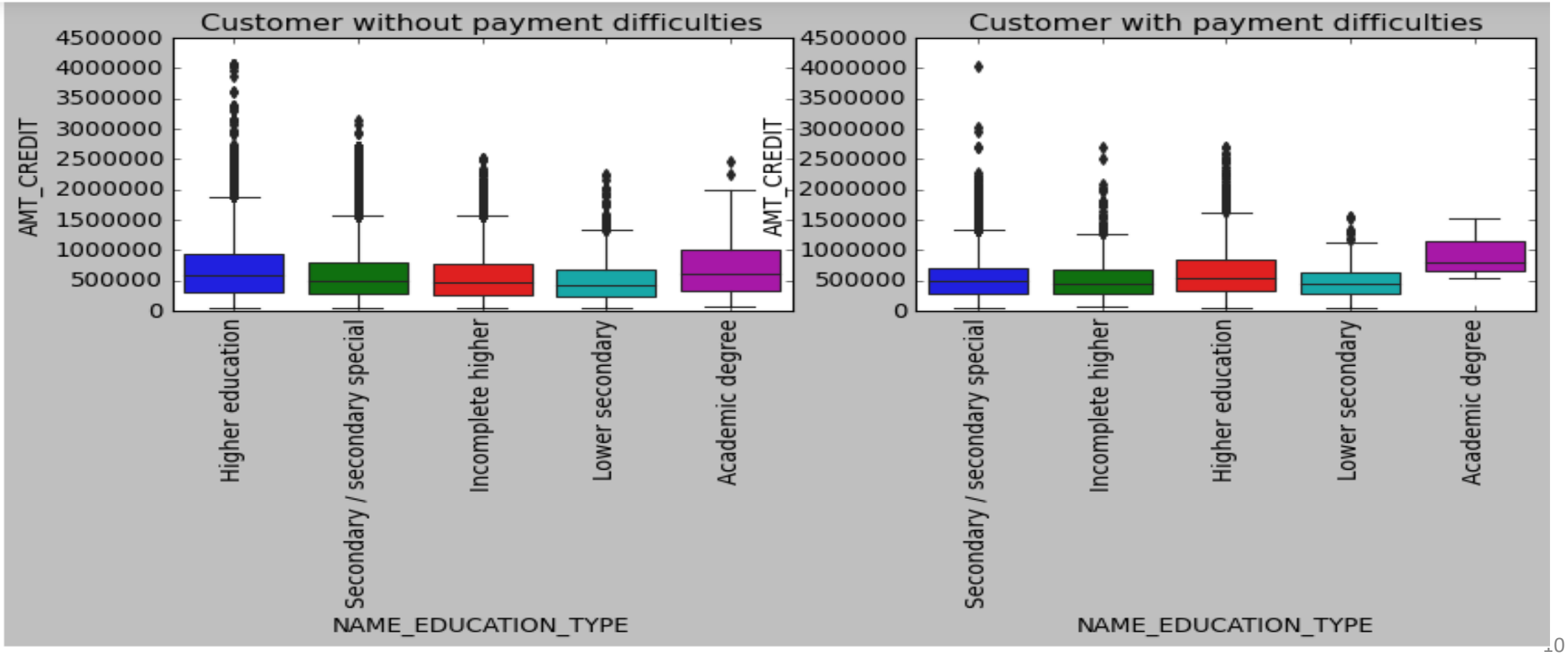
- Below is default and nondefault comparison with Occupation_type Payment status



Bivariate Analysis

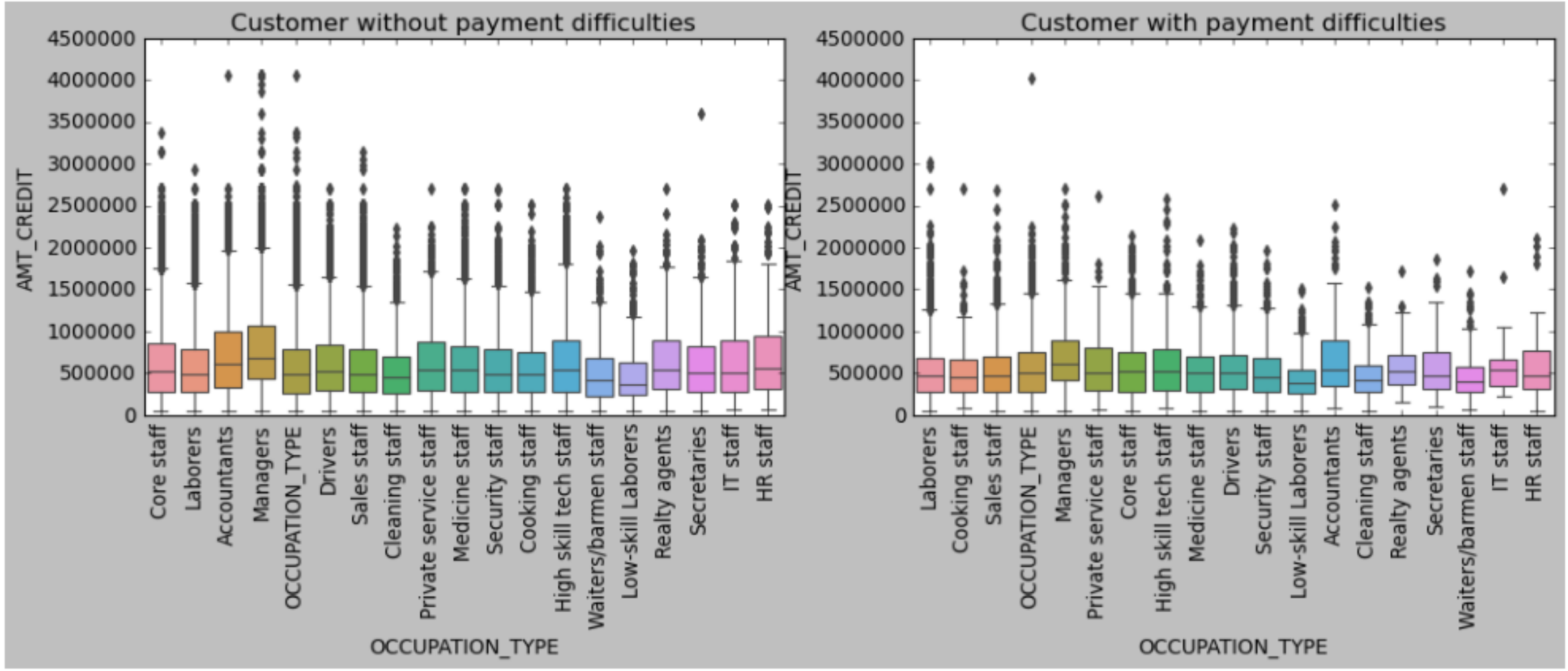
payment difficulties:
NAME_CONTRACT_TYPE

- Below is default and nondefault comparison with Education_name Payment status



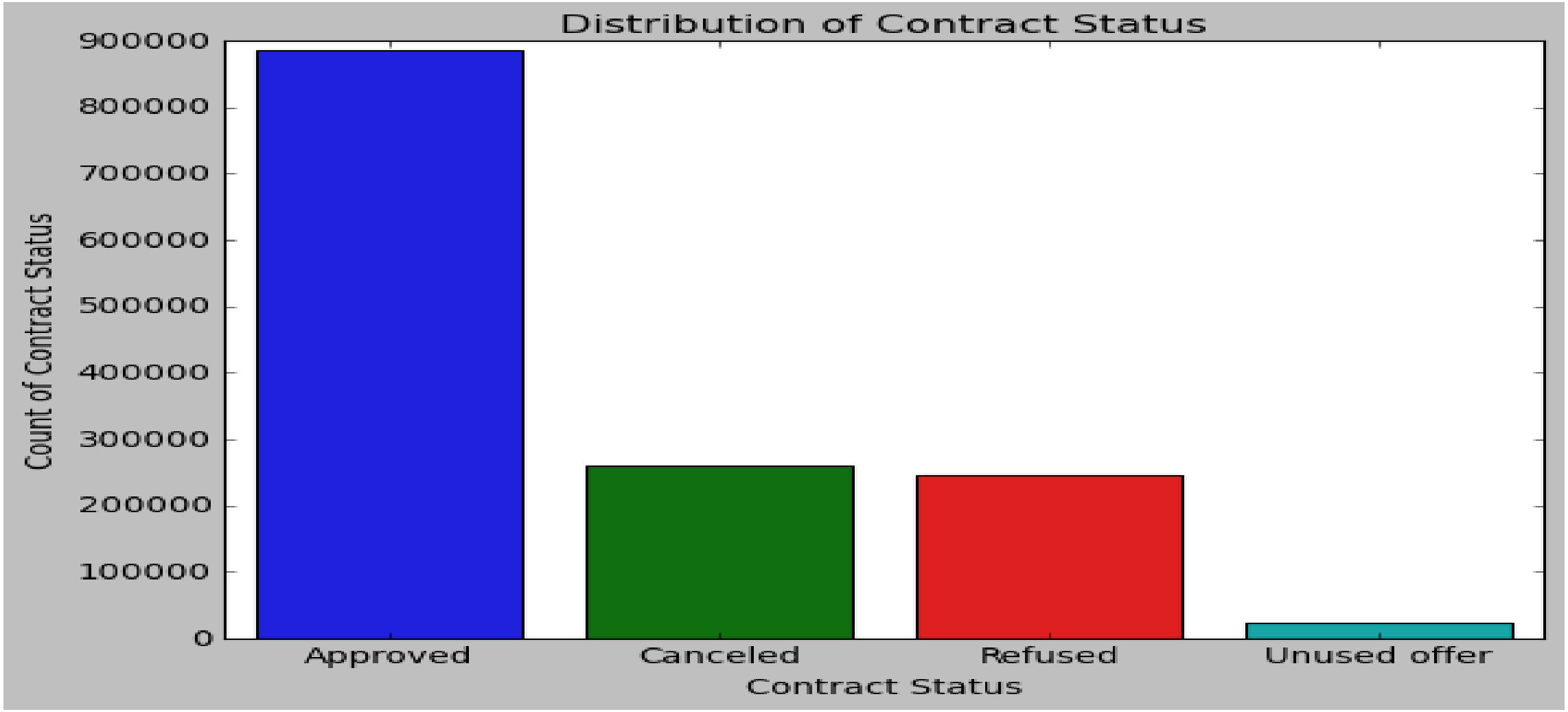
payment difficulties: OCCUPATION_TYPE

- Below is default and nondefault comparison with Occupation_type status

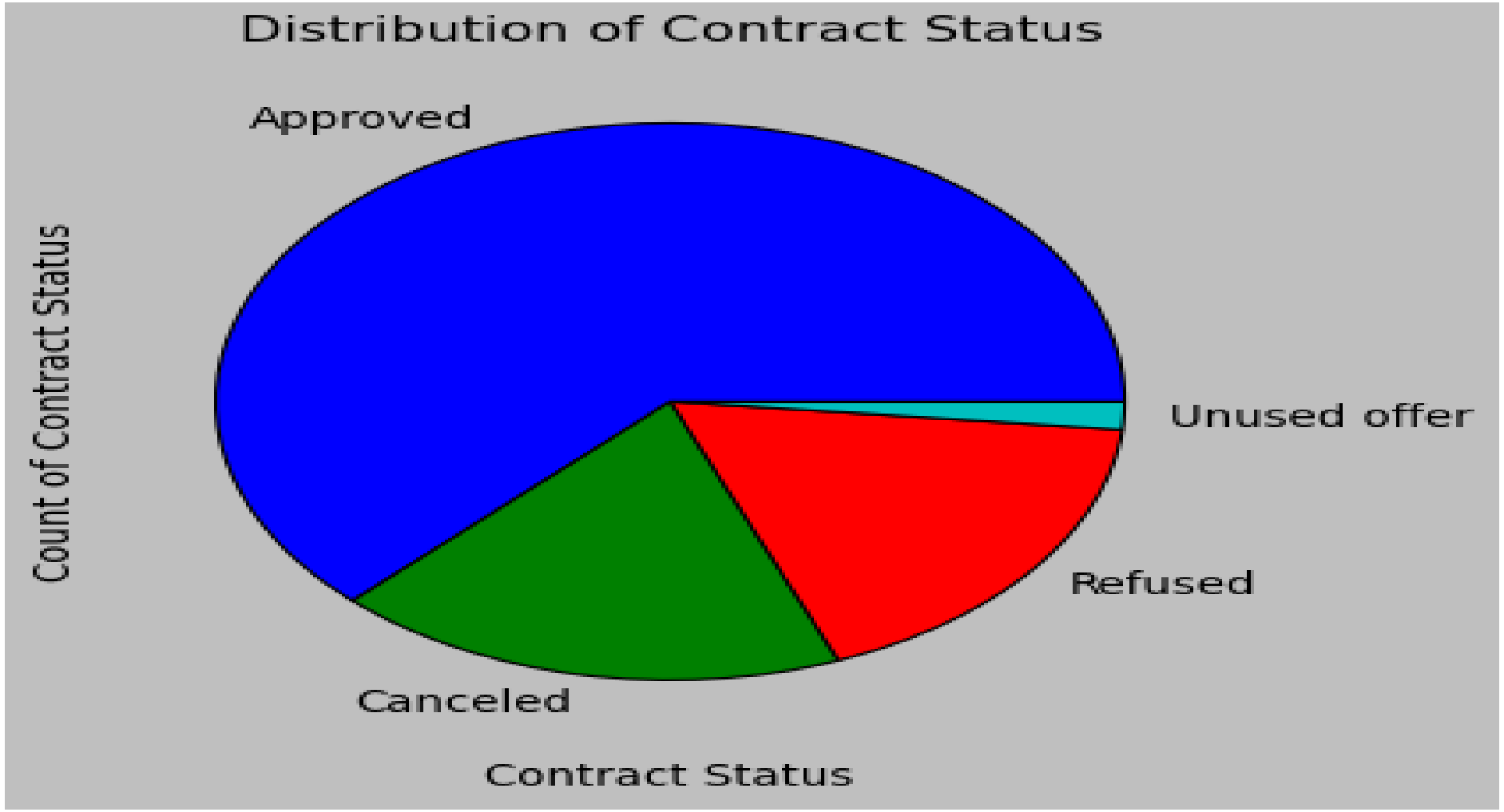


Previous_application data: Distribution of Contract Status

- Below graph shows the previous_application Loan Approval/refused/cancelled/unused status

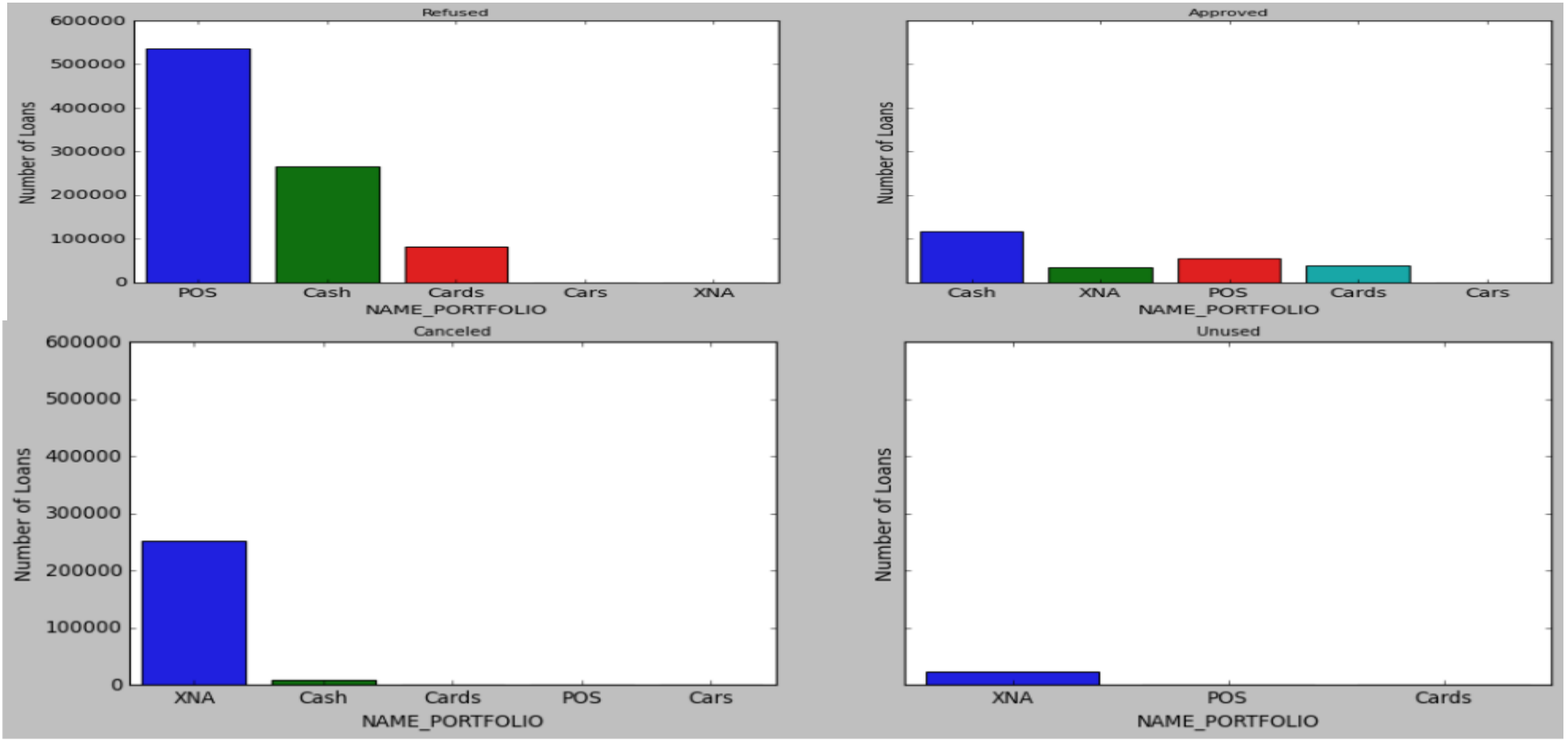


Previous_application_data: Distribution of Contract Status



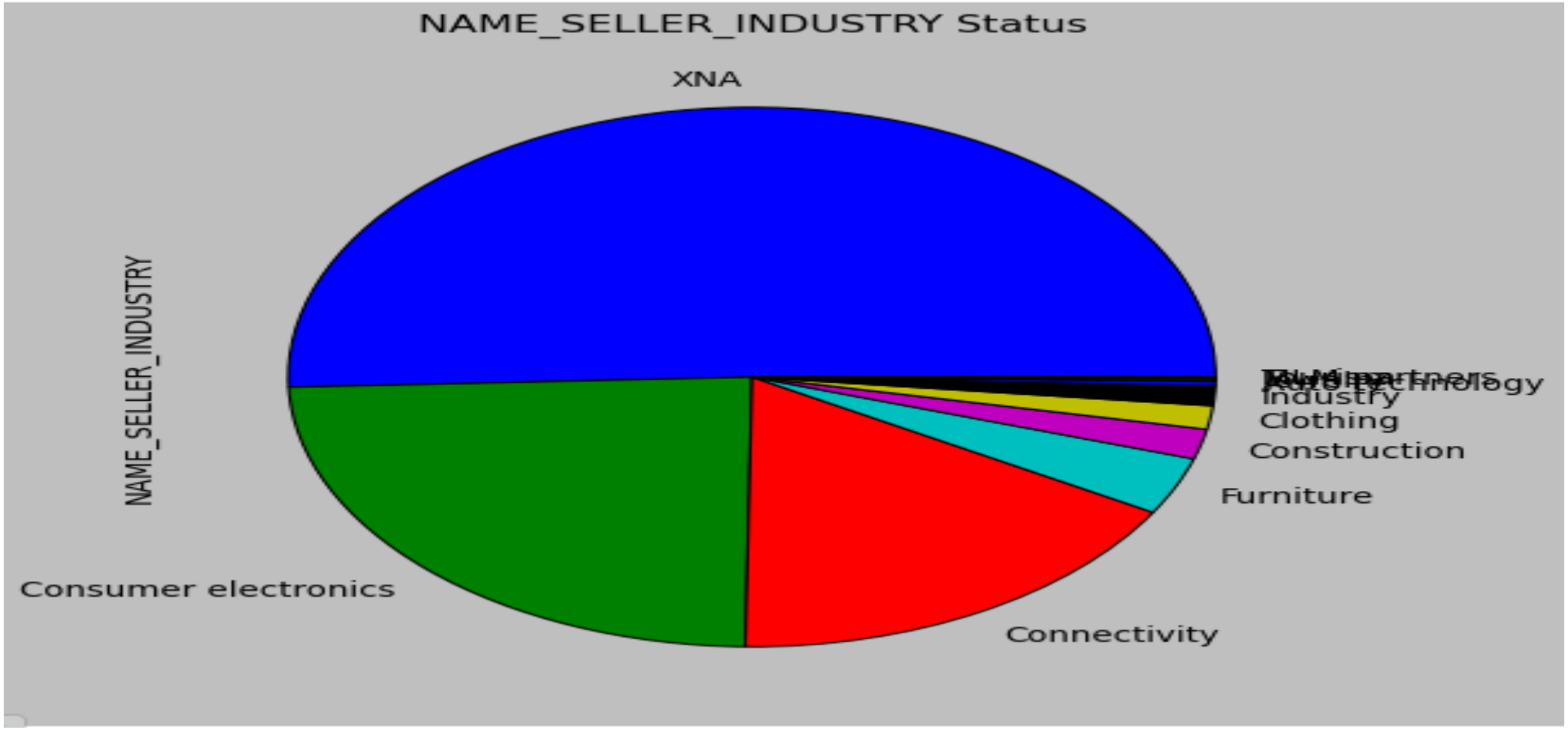
Merge of two dataset NAME_PORTFOLIO Status

- Below graph shows the Merge of two datasets Loan statuses with Approval/refused/canceled/unused



NAME_SELLER_INDUSTRY Status

- The below graph shows the Merge of two dataset NAME_SELLER_INDUSTRY Status



Recommendations and Risks

- Distribution of Contract Status Loan Approval is more.
- Most Loan Approved by Cash.
- Female is getting more Refused more approved more canceled more unused but in the case of male it is having average in every category
- Working-type people are applying for more loans as compared to others and also Commercial associates people are taking more loans.
- Married people are applying for and taking loans more than other people are taking loans in the format of cash through the bank
- laborers are getting the most refused and most approved loans, also Sales staff is getting the second most refused and approved loans.


THANK
YOU !