CREDIT EDA ASSIGNMENT

- Parag Ranjan Behera (DS C39)

CONTENTS

INTRODUCTION- PROBLEM STATEMENT AND GOALS

- When a client applies for loan, the company has to decide for loan approval based on the applicant’s profile.
- Two types of risks are associated with the company’s decision:
- If the applicant is likely to repay the loan (Non- Defaulter), then not approving the loan results in a loss of business to the company.
- If the applicant is not likely to repay the loan(Defaulter), then approving the loan to such a client will lead to a financial loss for the company.
- Our goal is to analyze the new and previous application data to study the pattern of defaulters (TARGET) and give recommendations based on the observations which should help the company to make informed decision based on past data, and reduce the number of defaulters in future.

EDA METHODOLOGY

STEP-1 :- Loading of data and understanding data.

STEP-2 :- Data Cleaning and Handling.

STEP-3 :- Analysis of the data and obtaining inferences.

STEP-4 :- Analysis of all the inferences and suggesting business recommendations.

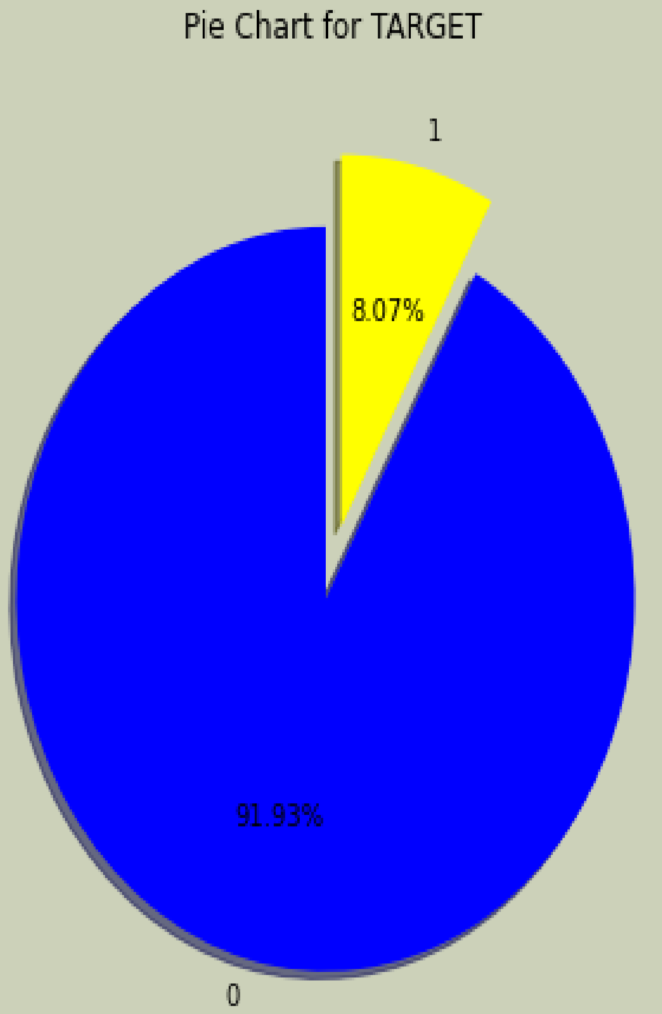
DATA CLEANING

- Before any data is used, it needs to be cleaned and treated to get the best possible views for analysis.
- Following steps were performed on each data set.
- The structure of the data is observed and unwanted (Data with nulls) or data not important for analysis are removed.
- The data types of columns that are misrepresented are corrected.
- Data is converted to usable formats or converted to other formats as needed(days to years) when needed.
- Outliers and Imbalance in the data are detected and reported for imputation.
- Imputation is done on the data as per column by choosing values to impute based on the outliers.
- Plotting and analysis is done after cleaning the data.

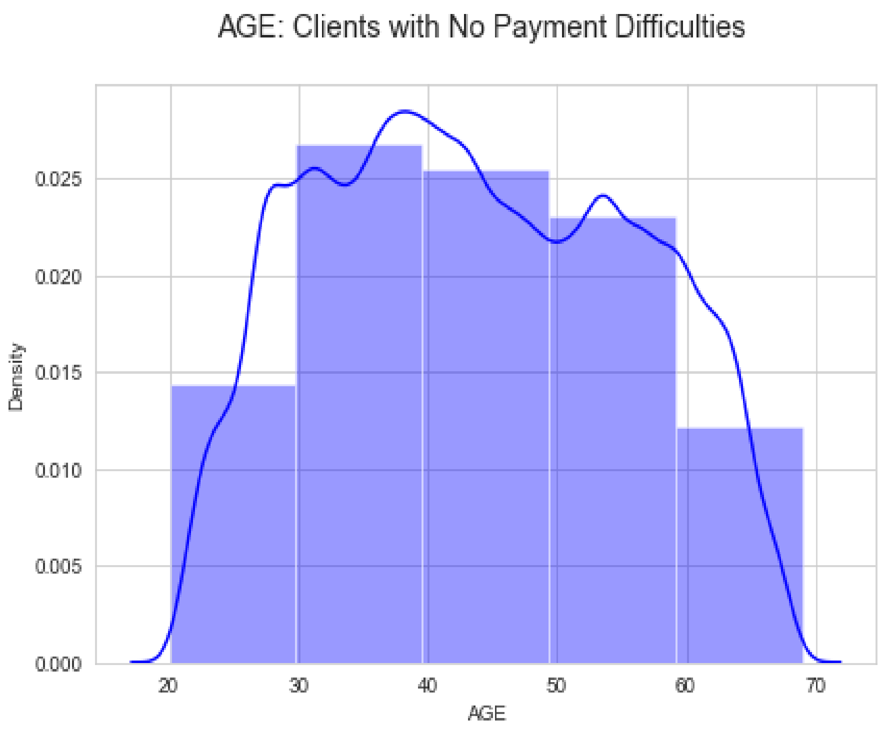
ANALYSIS OF NEW APPLICATION DATA AND PREVIOUS APPLICATION DATA

DATA IMBALANCE

- **As we can see that, there is high data imbalance between clients with payment difficulties and clients with no payment difficulties.**
- **Approx. 92% of clients have paid the loan on time, while approx. 8% of clients faced difficulties in paying the loan on time.**
- **The ratio of data imbalance for TARGET variable is 91.93 : 8.07.**
- **The data is divided into two parts, Clients with Payment Difficulties (TARGET= 1) and Clients with No Payment Difficulties (TARGET = 0) for further analysis.**

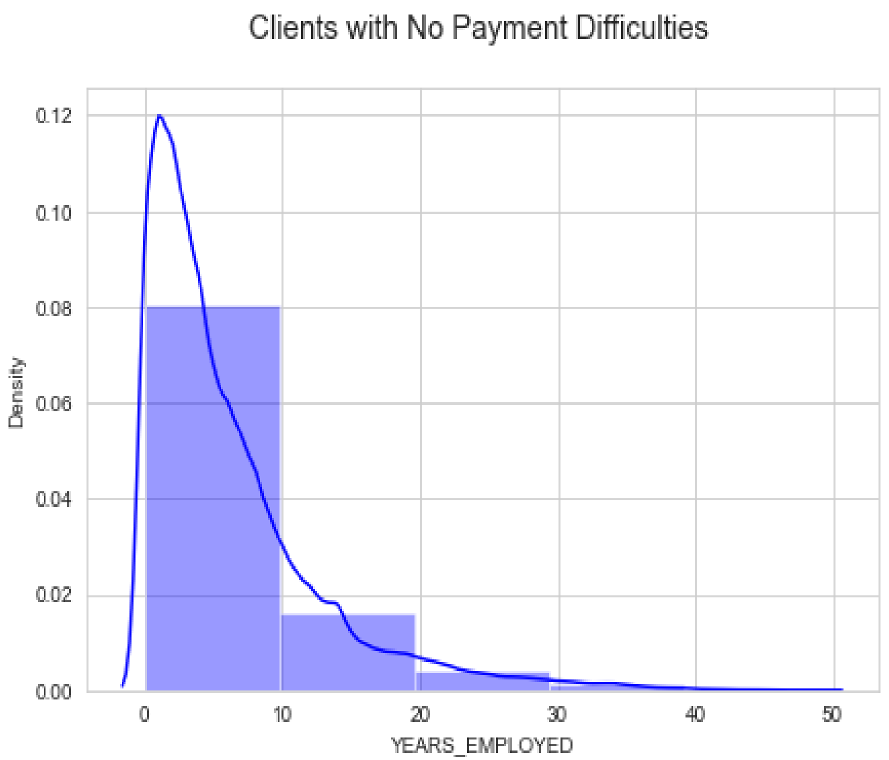


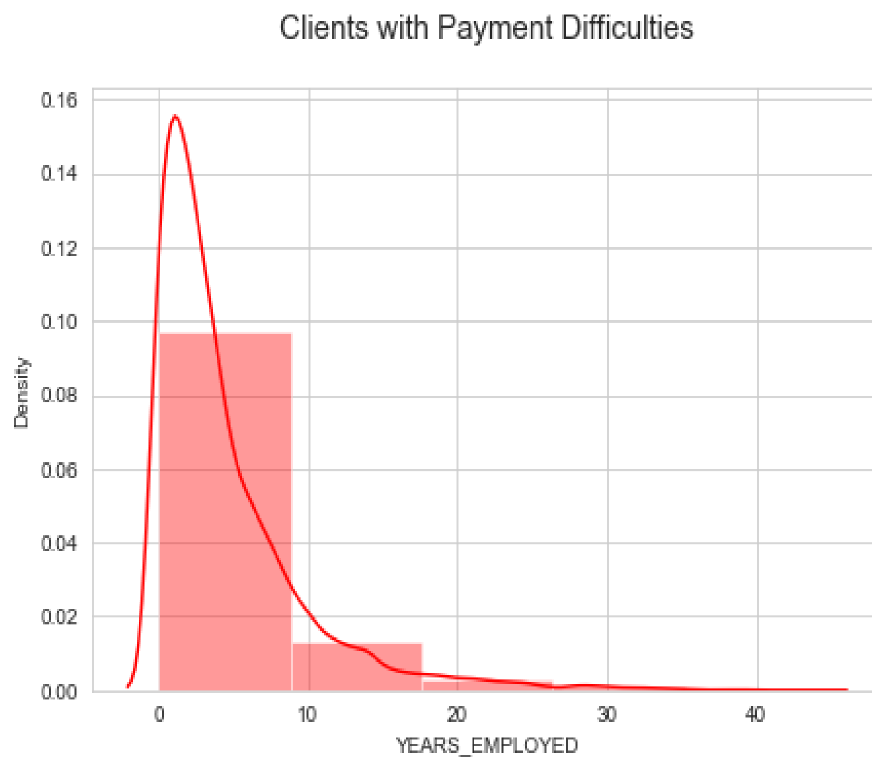
AGE DISTRIBUTION OF CLIENTS



- **Majority of the clients with no payment difficulties are between the age of 25 -50 and then decreases.**
- **The age distribution shows a much higher number of defaulters between the lower age group i.e., 25- 45.**

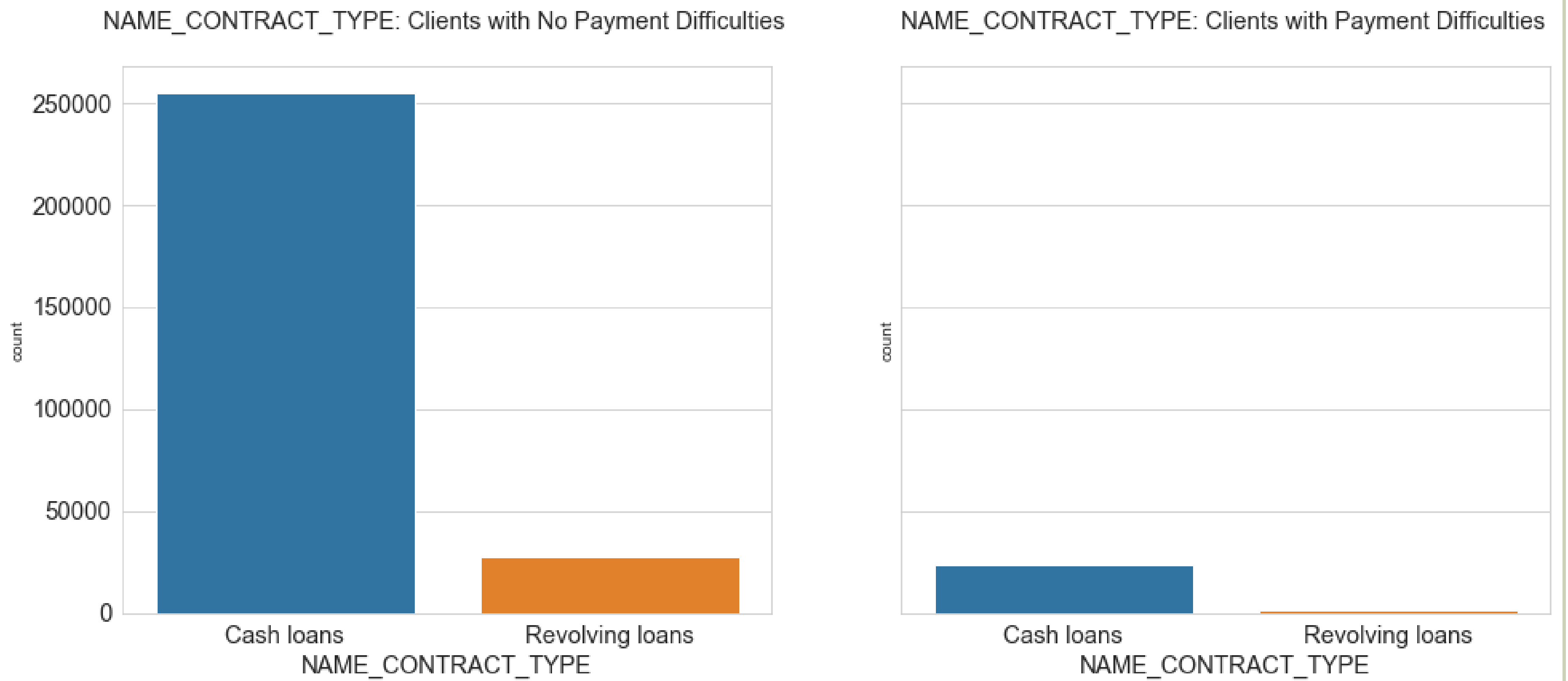
YEARS OF EMPLOYMENT OF CLIENTS





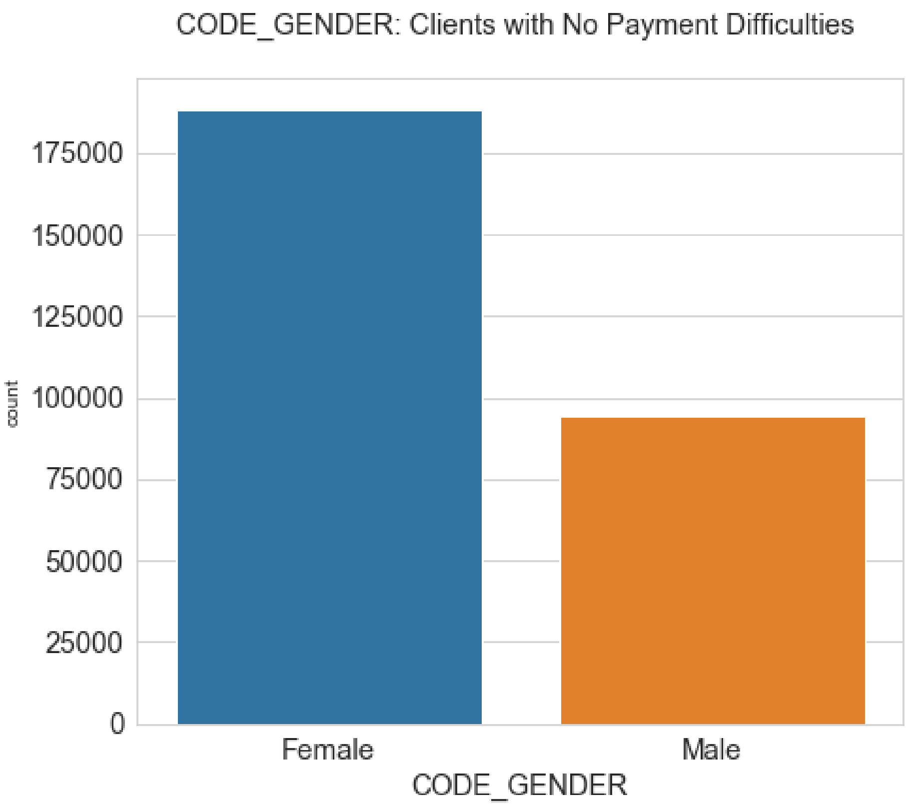
- **For both clients with or without Payment difficulties, the peak number of applicants are working with the current organization for less than 5 years of employment and drops considerably after that.**

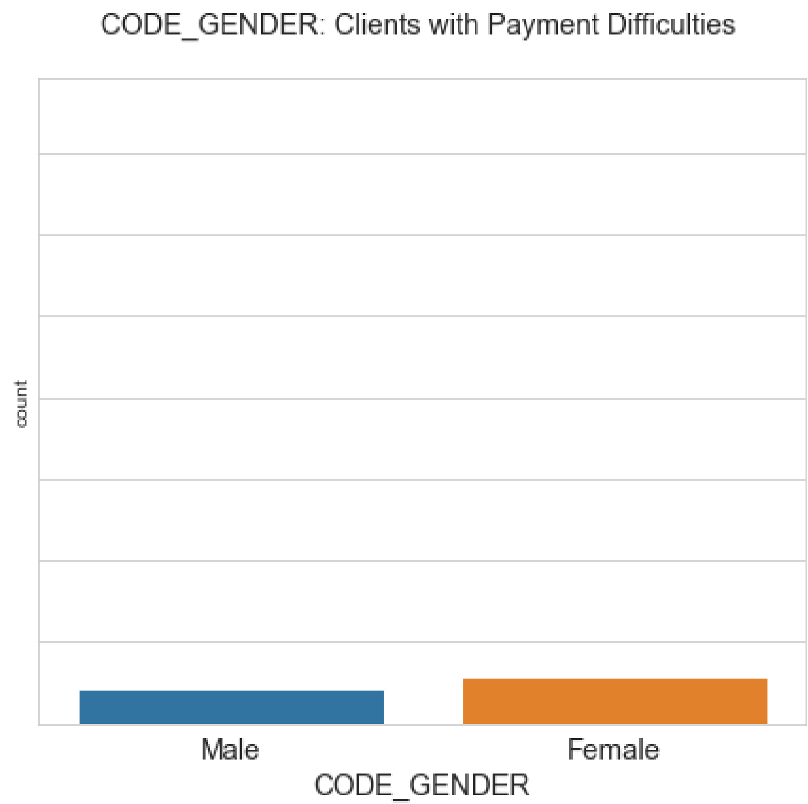
TYPE OF LOAN CLIENTS APPLIED FOR



- **Cash loans overall are more popular than the revolving loans.**
- **While resolving loans have a relatively higher number of defaulters.**

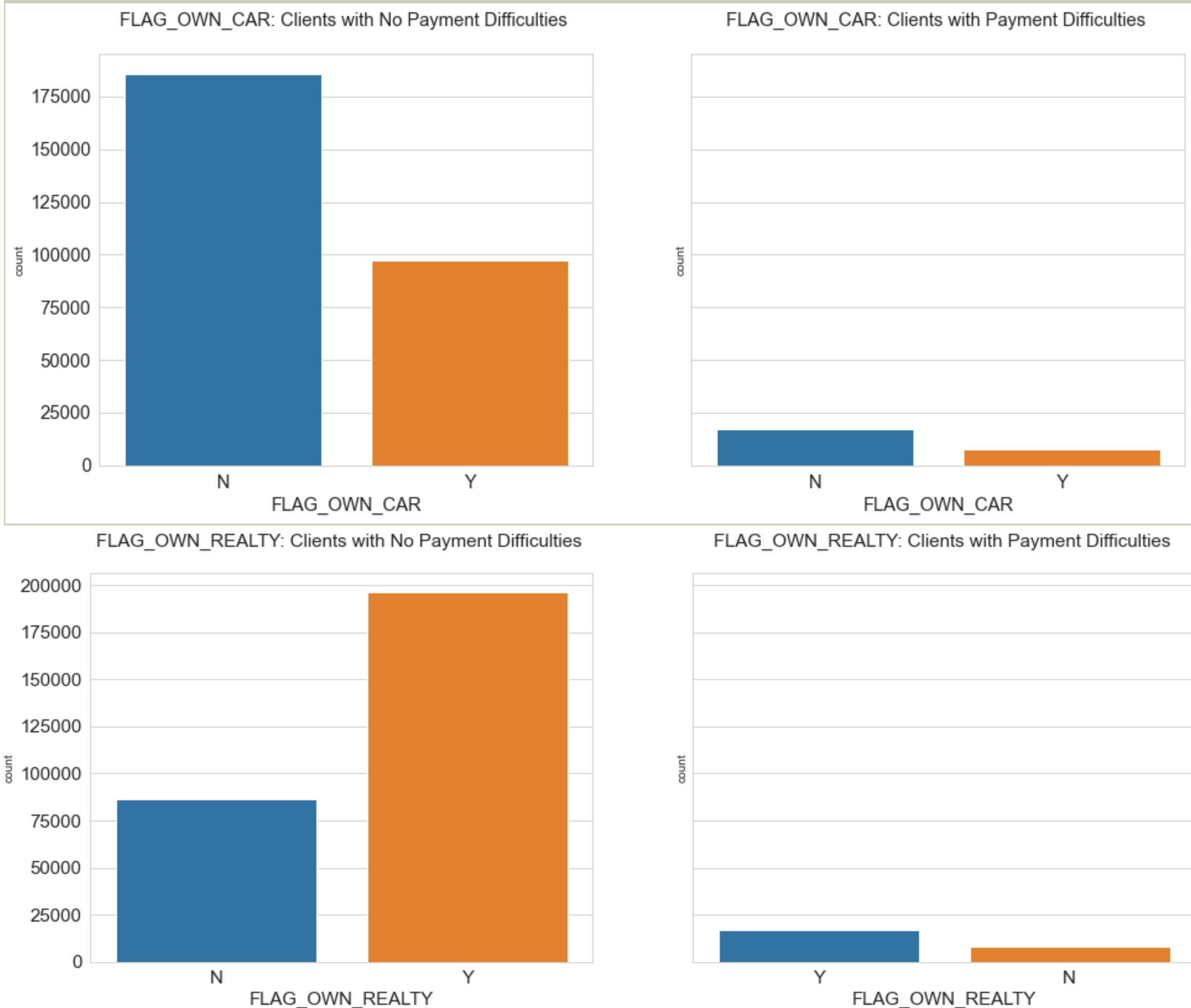
GENDER OF CLIENTS





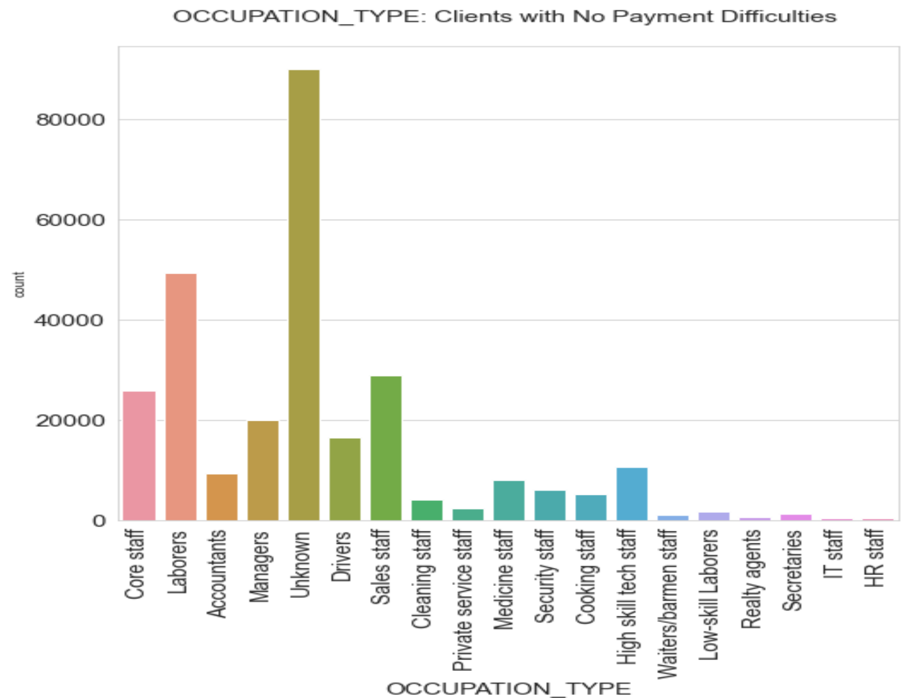
- Females are paying credit on time and at the same time they are defaulting on loans.
- But number of males defaulting are higher than females.

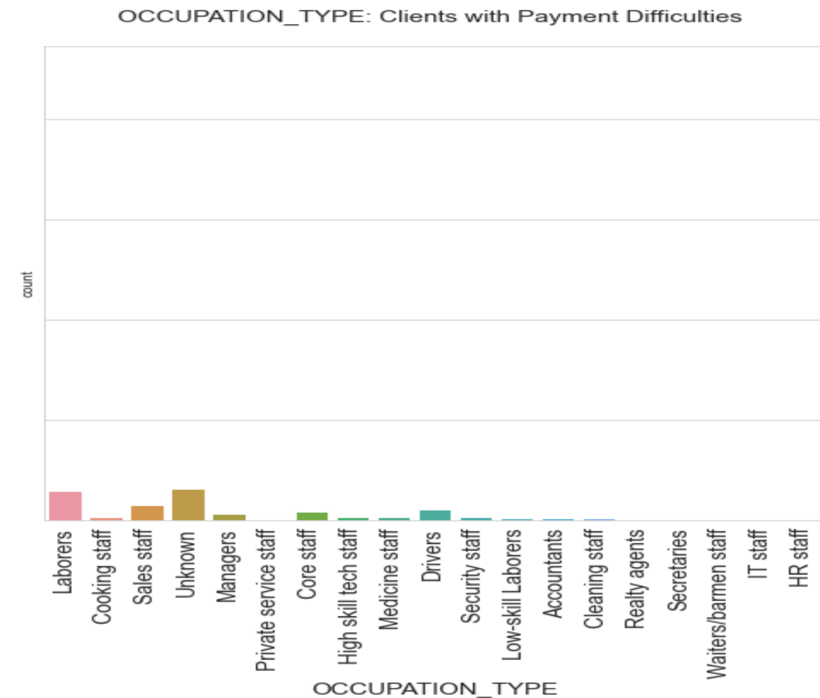
ASSETS OF CLIENTS



- The client without owning a car have defaulted more than those who have cars.
- Number of the clients with payment difficulties already own some kind of realty and this could mean that they might have some existing loans on that own realty.

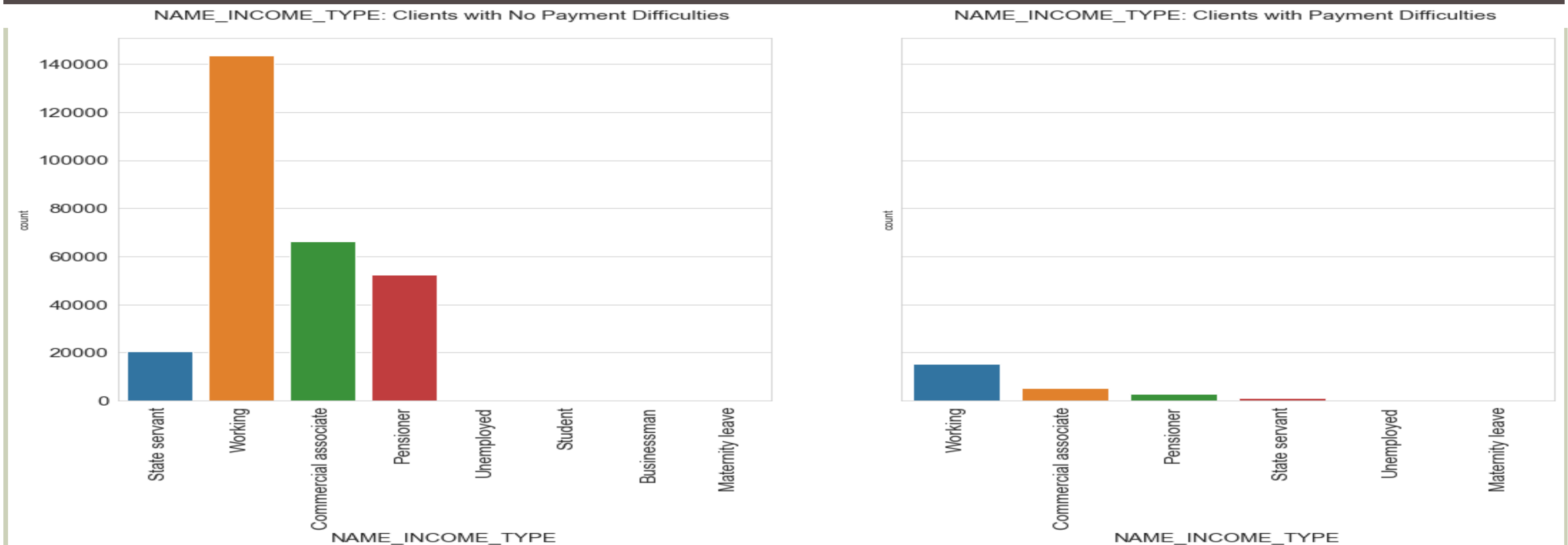
OCCUPATION OF CLIENTS





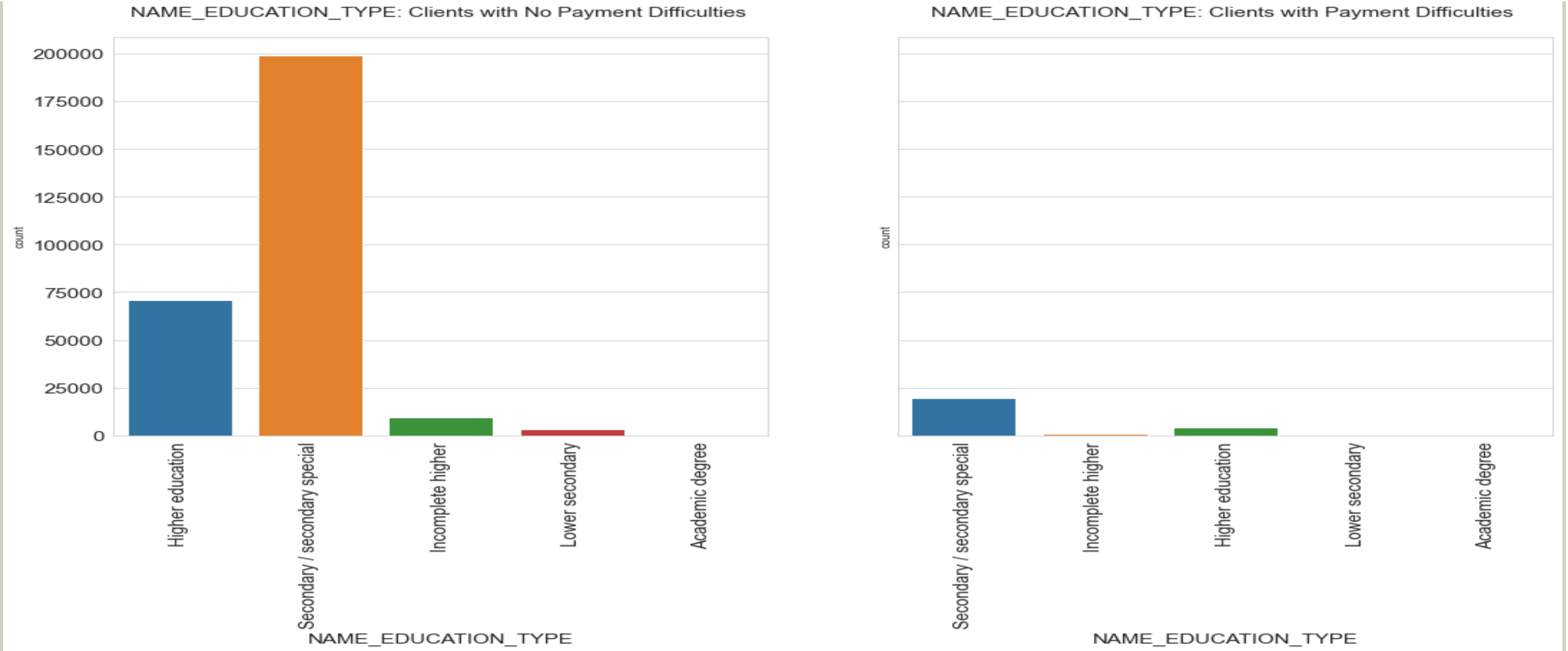
- A big portion of the client's occupation is form of Laborers, Unknown, Sales Staff and core staff.
- Most of the defaulters are from Laborers and Unknown occupation type.
- But relatively low-skilled laborers, Drivers and Security staff are major defaulters.

INCOME TYPE OF CLIENTS



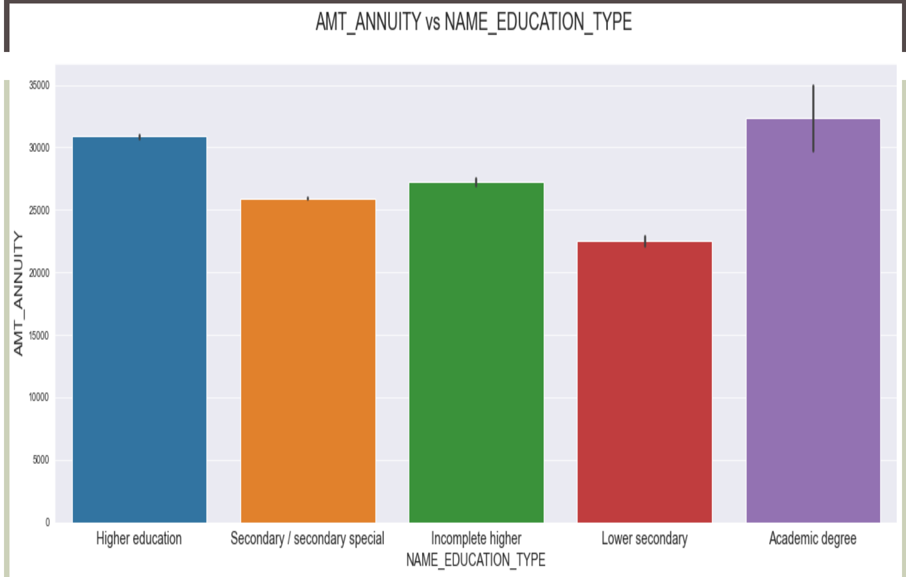
- Majority of the defaulter applicants are from ‘Working’ income type, followed by Commercial associate and pensioners.
- But percentage of defaulters is very high among state servant.
- Also, for non-defaulter's the applicants are same like defaulter's. They are from ‘Working’ income type, followed by Commercial associate and pensioners.

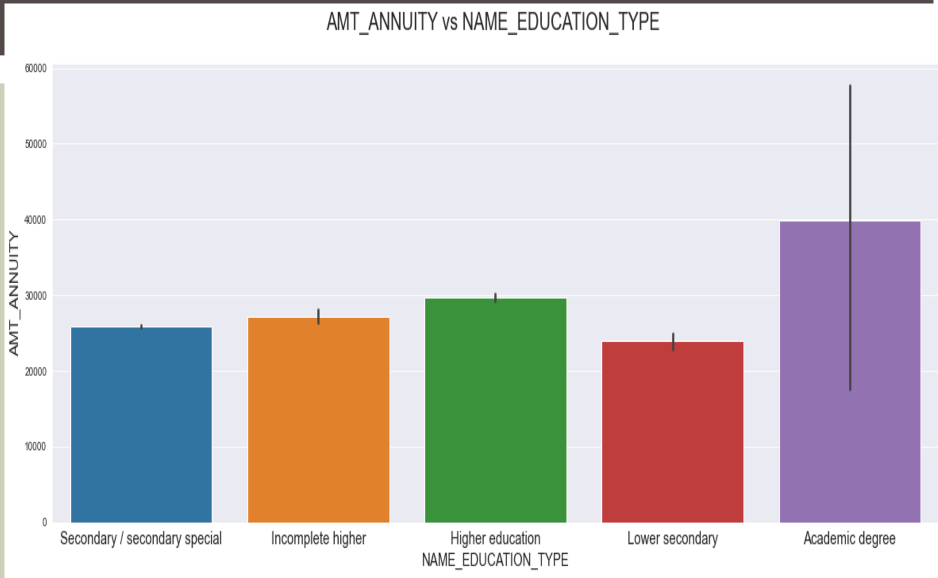
EDUCATION TYPE OF CLIENTS



- Secondary/ Secondary special education have the highest count in non-defaulters as well as in defaulters.
- But percentage of defaulters in incomplete higher education is high among others.

EDUCATION TYPE VS MONTHLY INTETREST

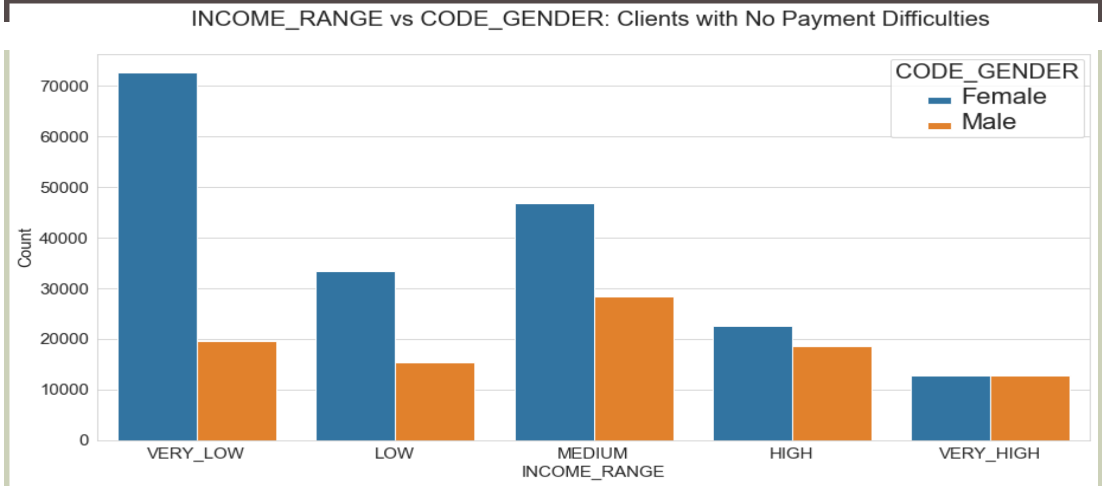


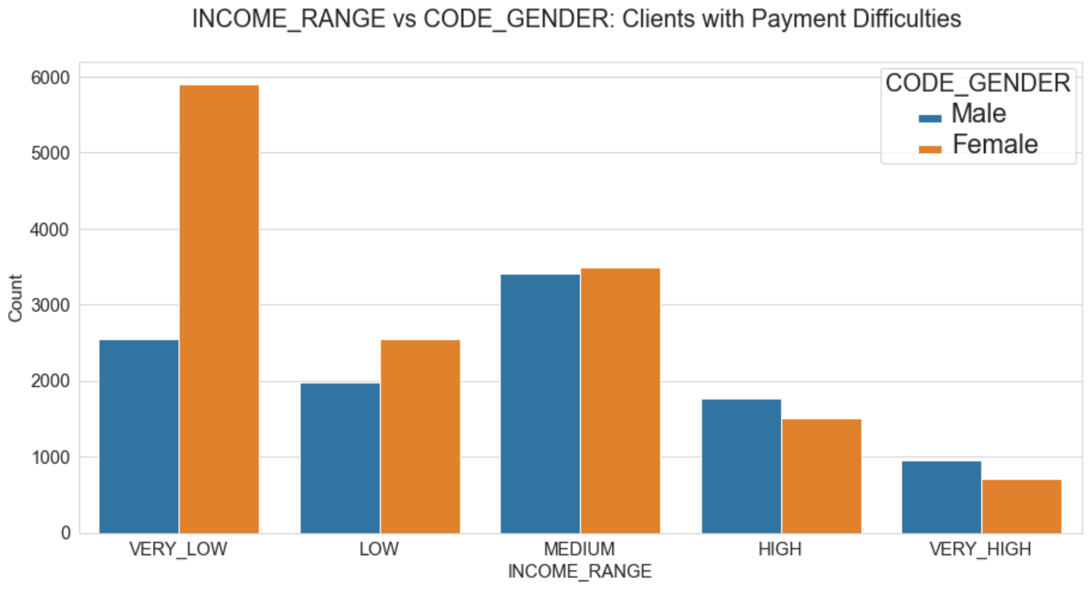


Non-Defaulters **Defaulters**

- In case of Clients with No Payment Difficulties, clients with academic degree can pay highest monthly annuity compared to others.
- In case of Clients with Payment Difficulties, clients with academic degree have to pay high monthly annuity compared to others.
- Company can provide higher loans to clients with academic degree.

INCOME RANGE VS GENDER



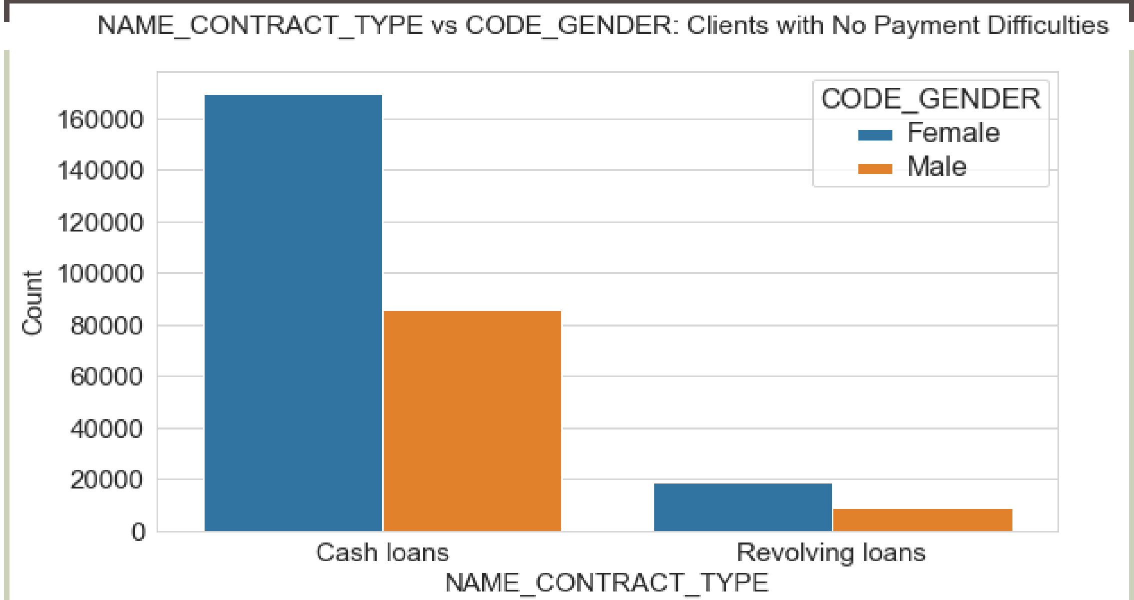


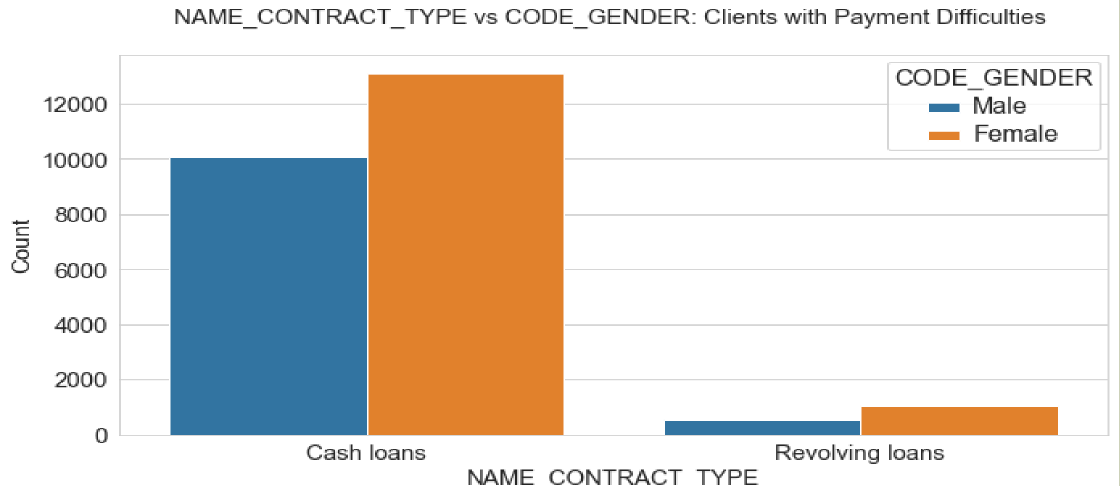
- Clients with No Payment Difficulties, the count of females in the very low income category is very high than males.

- While in the very high income range, females and males are equal. Whereas in all other cases, the count of females are more than males.

- In case of defaulters, males counts are higher than females in medium, high and very high income range.

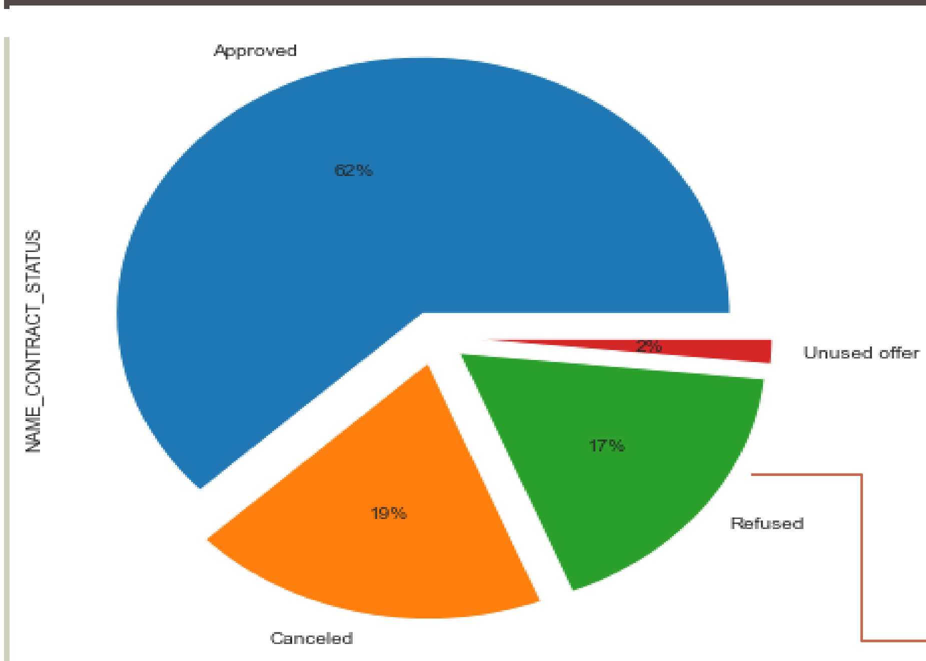
LOAN TYPE VS GENDER



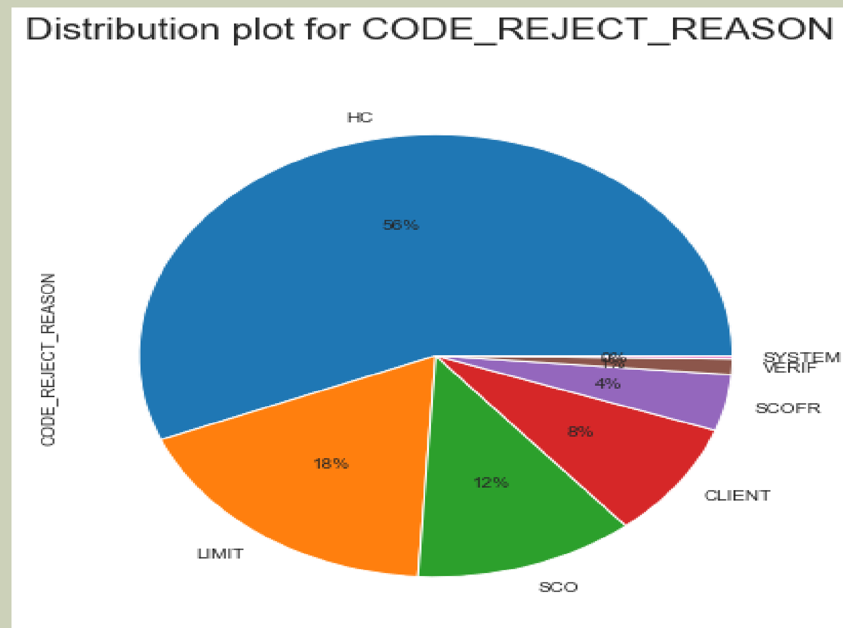


- Clients with No Payment Difficulties, in the cash loans the females have the highest count than in males as well as in revolving loans.
- For Clients with Payment Difficulties, females have the highest counts in the cash loans as well as in the resolving loans as compared to males.
- But the percentage of defaulters are males in both cash and revolving loans.

PREVIOUS LOANS

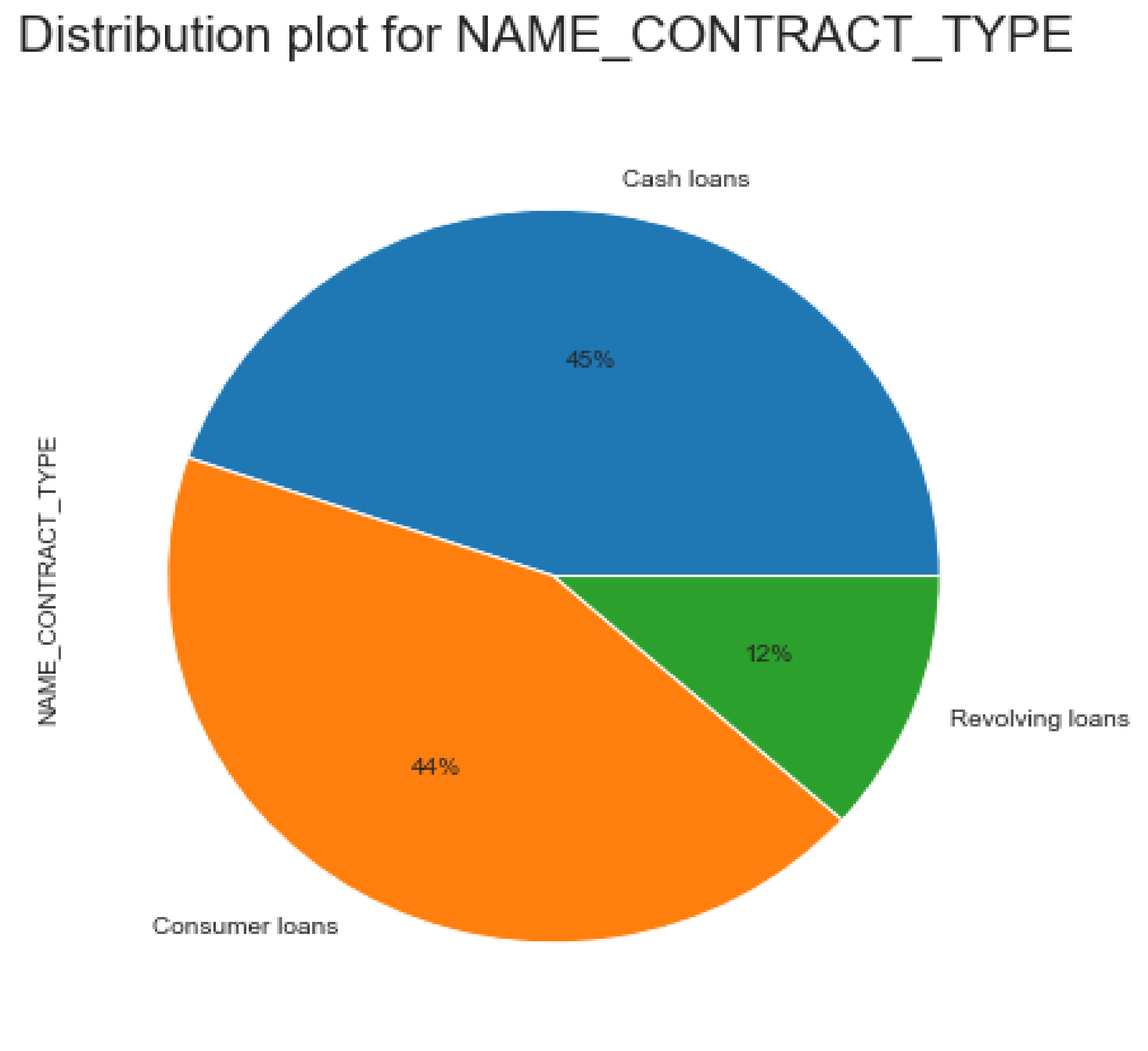


- 62% of the previous applications were approved.
- Only 17% of the applications were refused.
- 19% were cancelled by the applicant.



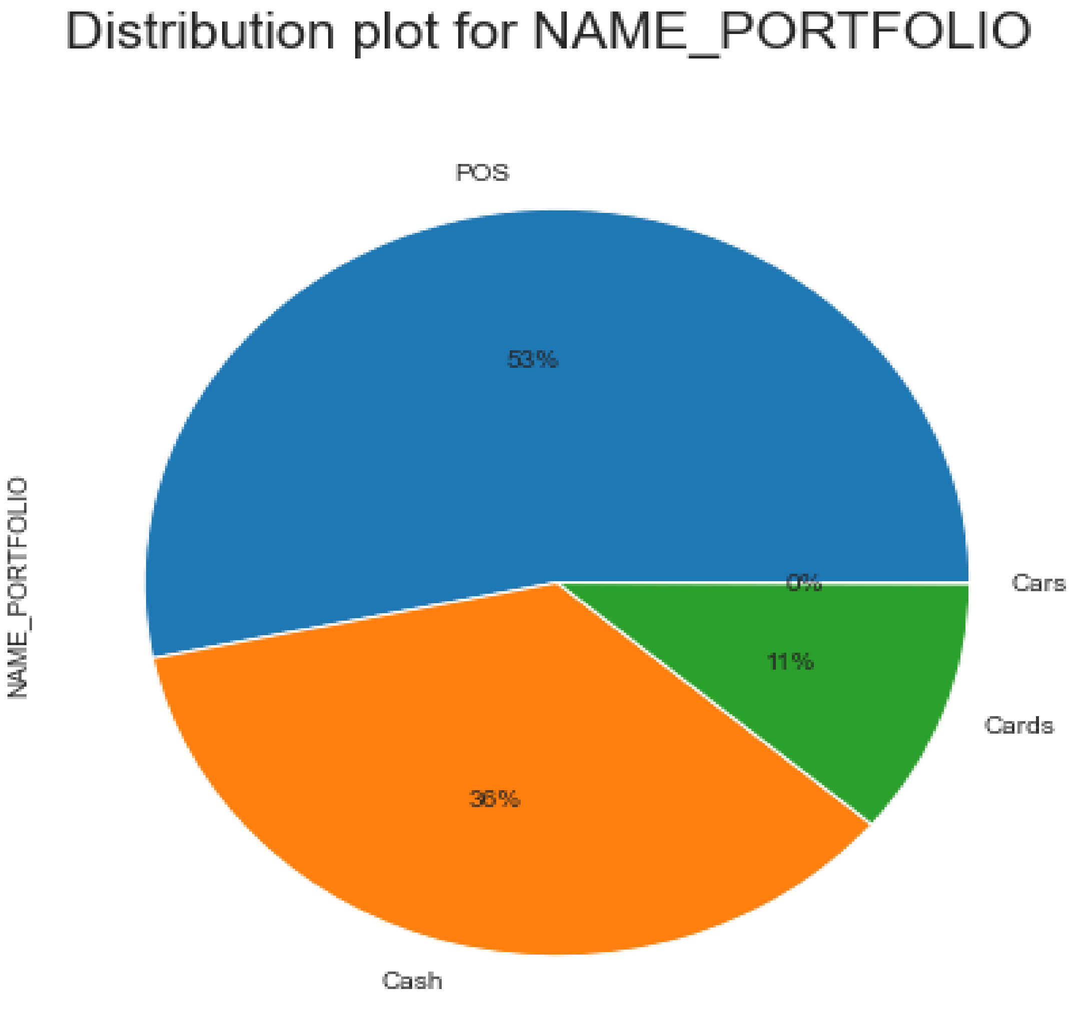
- 56% of the rejections were due to HC.
- LIMIT and SCO together accounted for 30% of rejections.

LOAN TYPES IN PREVIOUS APPLICATIONS



- Cash and Consumer loans were the most popular contract types in all the previous applications, accounting for over 89% of all previous applications.
- Only 12% of clients in the past applied for revolving loans.

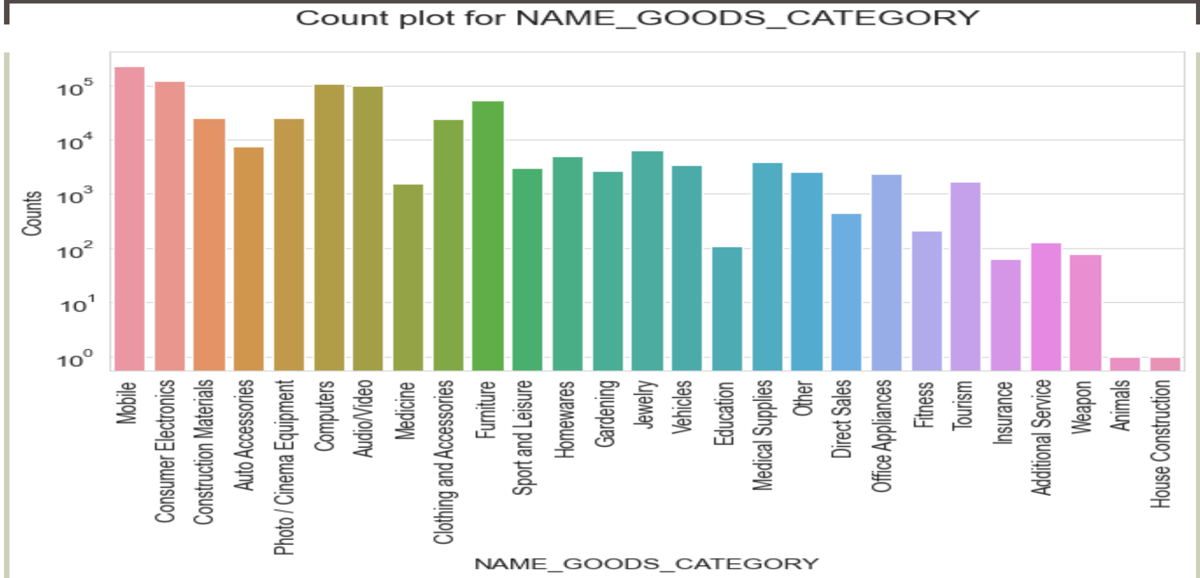
PORTFOLIO OF LOANS



- 53% of the past applications were for Point-of-Sale (POS) credit, followed by Cash at 36%.

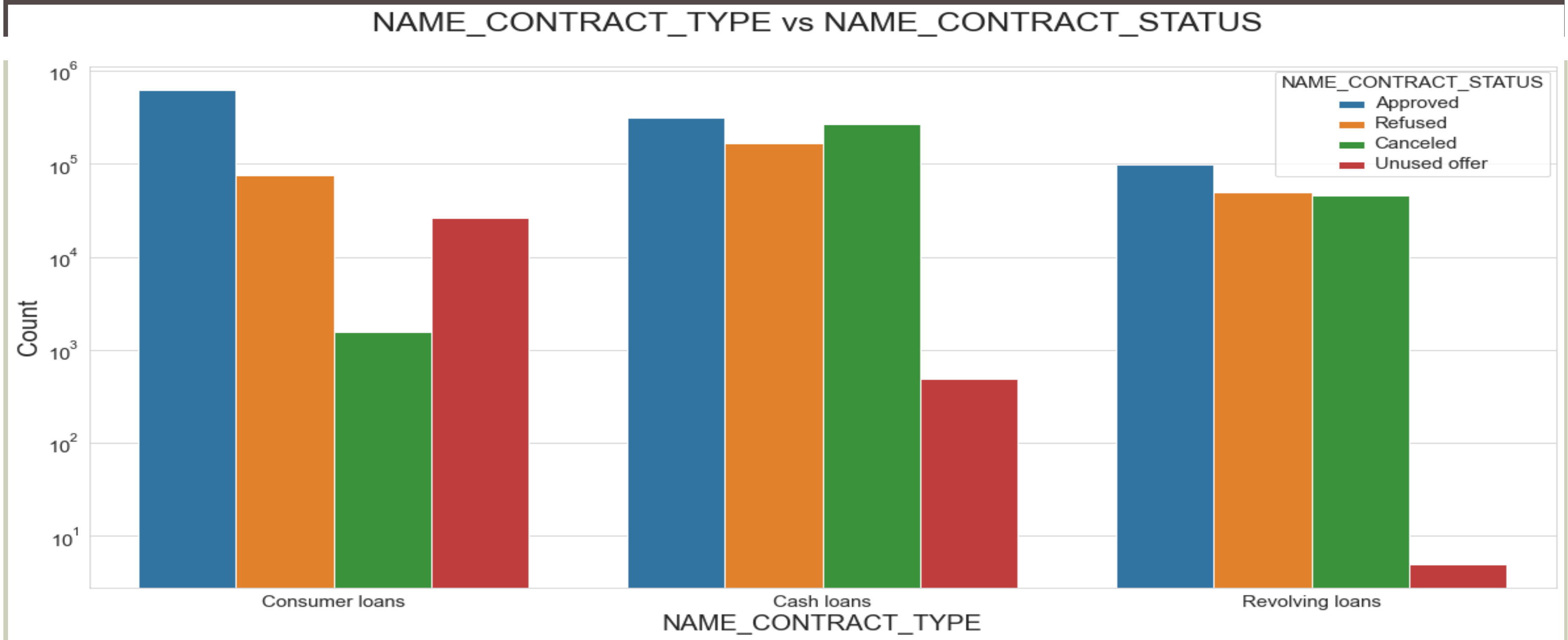
- Very negligible applications were for cars.

CATEGORY OF GOODS.



- **Clients those have applied for credit to purchase mobiles are highest among others.**
- **Consumer electronics, computers and audio/video have the second highest number of application.(approx.)**
- **The highest in goods category is House construction with goods price above 10 lakhs.**
- **The second highest is Tourism category with goods price above 1 lakhs.**

LOAN TYPE VS LOAN STATUS



- **The consumer loans have the highest approval count amongst the other cash loans and revolving loans.**

- **While cash loans have highest refusing count and cancelled count as compared with other types of loans.**

RECOMMENDATIONS

- Payment difficulties are more prevalent mainly in the younger demographic of 25-35 age - Checks on past records must be done.

- Employment status checks for work history with the current employer must be done, along with the client's current work status.

- Cash loans are popular and are safer for the company as revolving loans have considerably high % of defaulters and can become worrisome in the long run.

- Clients who already own a realty may have problems in repaying, which might be due to existing credits. Background check on other loans must be done on such clients to prevent issues in repayment.

- Academic Status of client must be assessed and considered before approval.
- Clients who were previously refused and reapplied, checks should be made not only for their current info/status but also the reason of their previous rejection.
- While gender shouldn’t be a determinant factor on who gets a loan, men have shown tendency to have higher defaults and a stricter approval might reduce those cases.