CREDIT EDA ASSIGNMENT.

BALAJI V

INTRODUCTION

- This case study aims to give you an idea of applying EDA in a real business scenario. In this case study, apart from applying the techniques that you have learnt in the EDA module, you will also develop a basic understanding of risk analytics in banking and financial services and understand how data is used to minimize the risk of losing money while lending to customers.

- The given data contains the information about the loan application at the time of applying for the loan. It contains two types of scenarios:

1. The client with payment difficulties: he/she had late payment more than X days on at least one of the first Y instalments of the loan in our sample,

2. All other cases: All other cases when the payment is paid on time.

Problem Statement.

- When a client applies for a loan, there are four types of decisions that could be taken by the client/company):

1. Approved: The Company has approved loan Application

2. Cancelled: The client cancelled the application sometime during approval. Either the client changed her/his mind about the loan or in some cases due to a higher risk of the client he received worse pricing which he did not want.

3. Refused: The company had rejected the loan (because the client does not meet their requirements etc.).

4. Unused offer: Loan has been cancelled by the client but on different stages of the process. In this case study, we will use EDA to understand how consumer attributes and loan attributes influence the tendency of default.

Business Understanding

Steps Followed for CREDIT EDA

1 Data understanding

2 Data Cleaning and Manipulation

3 Checking Outliers

4 Data analysis

5 Presentation and Recommendations

Libraries Used in python For EDA

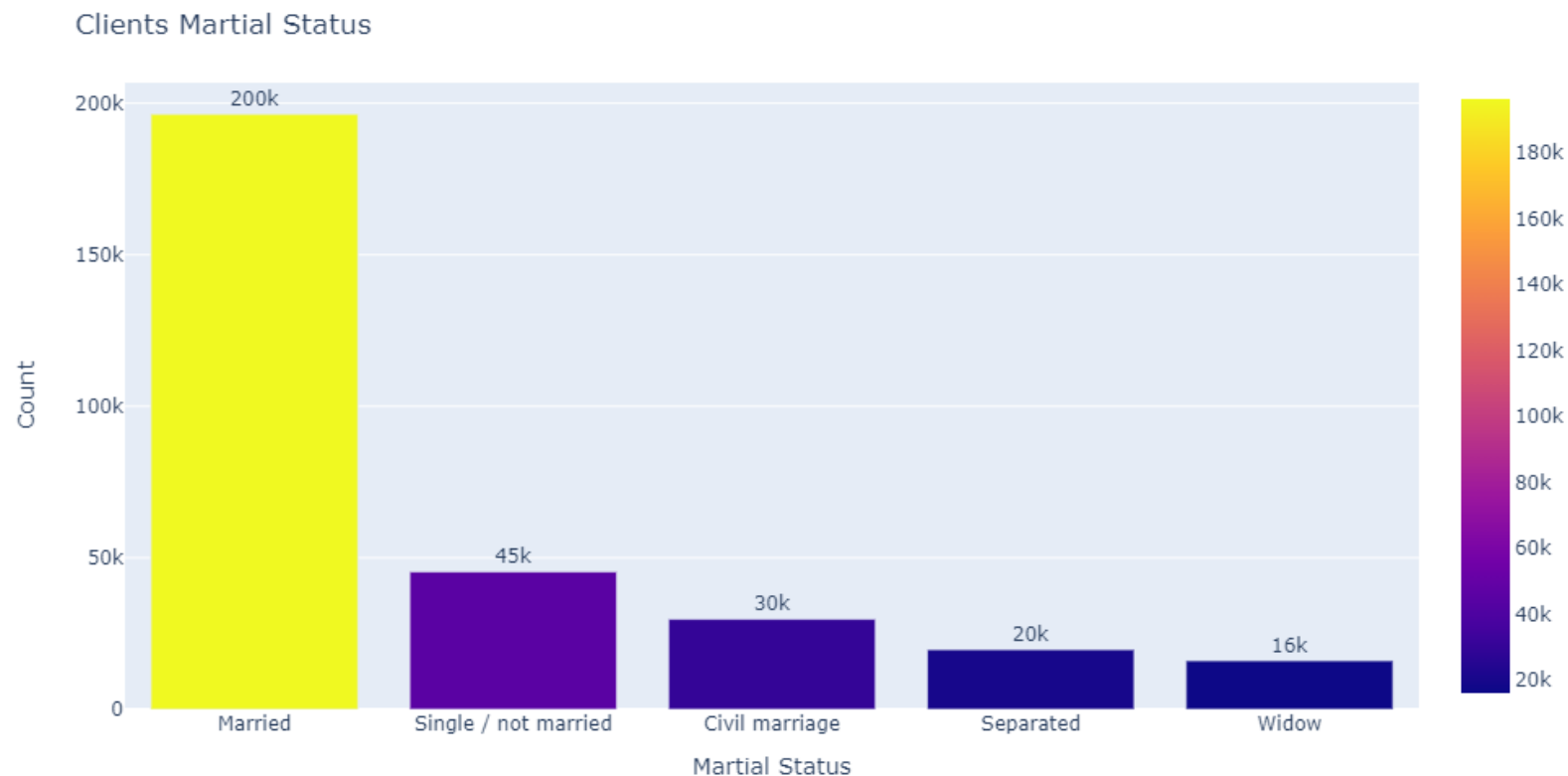|		
Numpy	Pandas	Matplotlib(Plotlib)
Seaborn	Plotly (express ,graph objects, offline,subplots)	Cufflins

UNIVARIATE ANALYSIS

BASED ON CLIENTS WITH PAYMENT DIFFICULTIES AND CLIENTS ON TIME PAYERS.
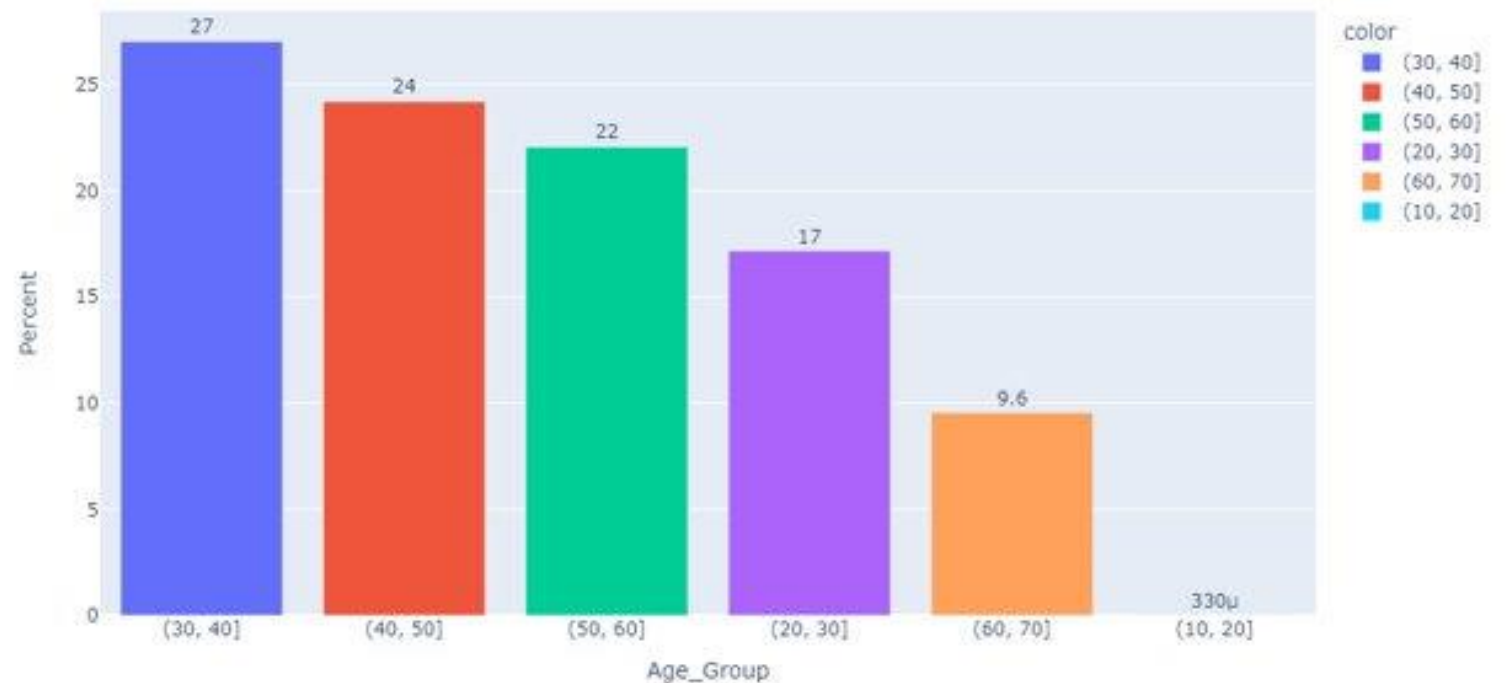
Univariate Analysis

MARTIAL STATUS

- Most of the clients who applied for loan are Married.

Clients Martial Status

AGE GROUPS

UNIVARIATE
Analysis

- 27 % of applicants are in the age group of 30-40

- Followed by
  40-50(24%) ,
  50-60(22%).

INCOME GROUPS

UNIVARIATE
Analysis

- 27 % of applicants are in the Income group of Middle class.

Followed by

- Upper Middle class(25%) ,
- Lower Middle class (22%).
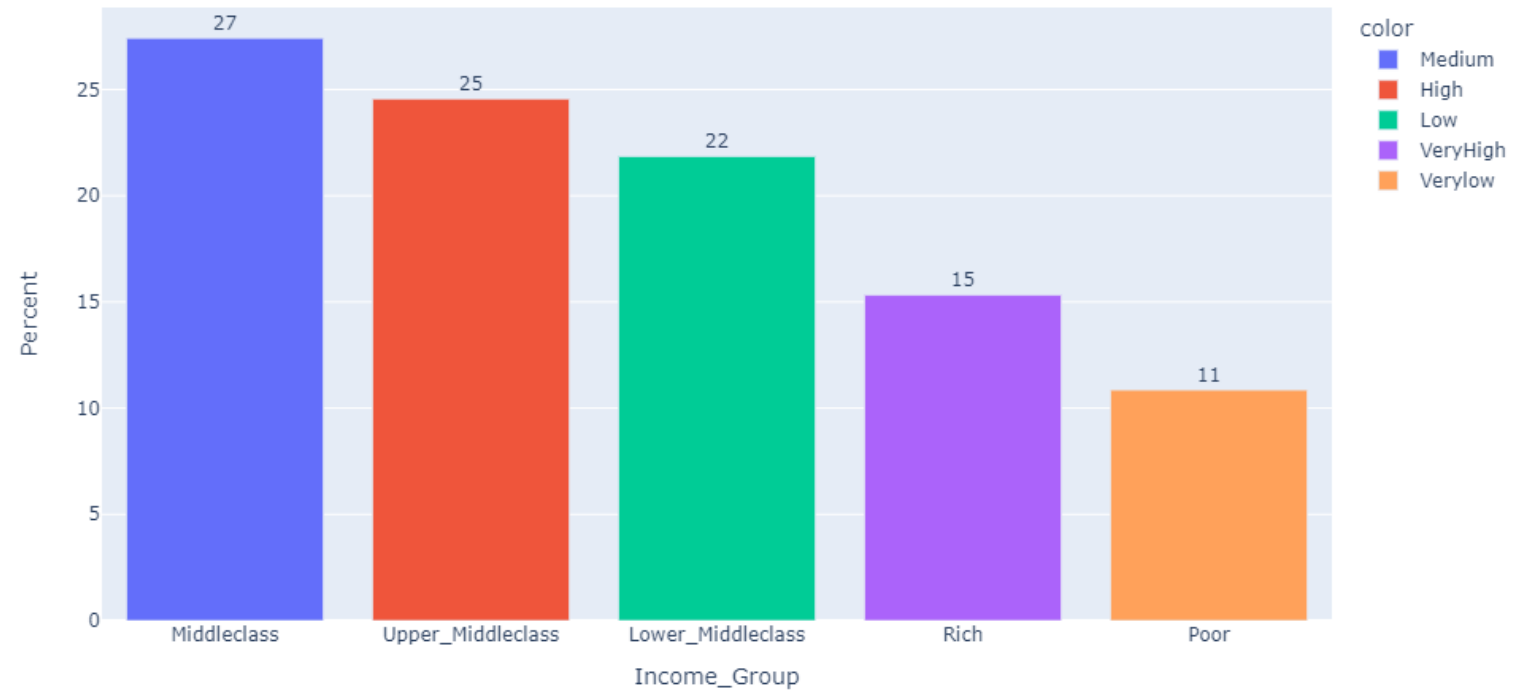
UNIVARIATE
Analysis

OCCUPATION TYPE

- Most applicants has not mentioned there occupation Type so I have considered as Unknown.

- If we drop those Unknown. Laborers are the most applicants for loan.

CLIENTS OCCUPATION_TYPE
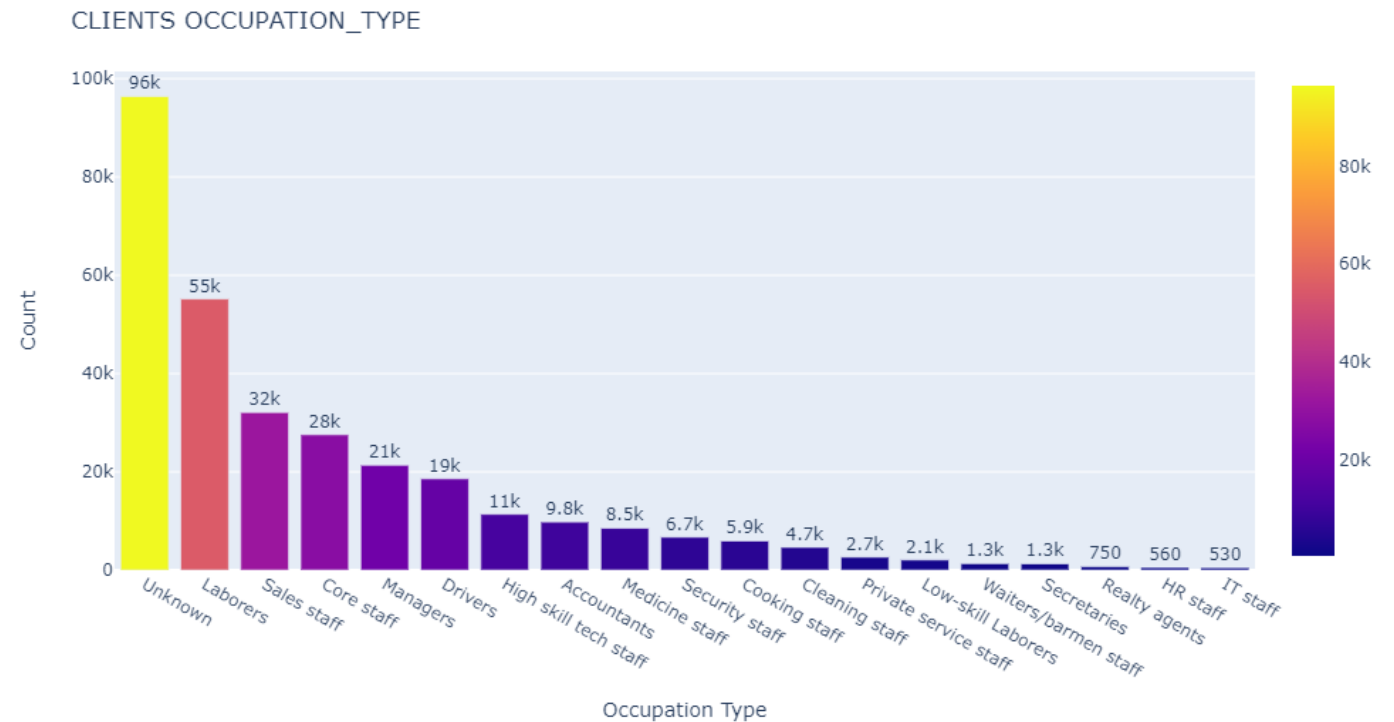
UNIVARIATE Analysis

Target

- 91.9% of applicants has made on time payment
- Whereas only 8.07% Has payment difficulties.
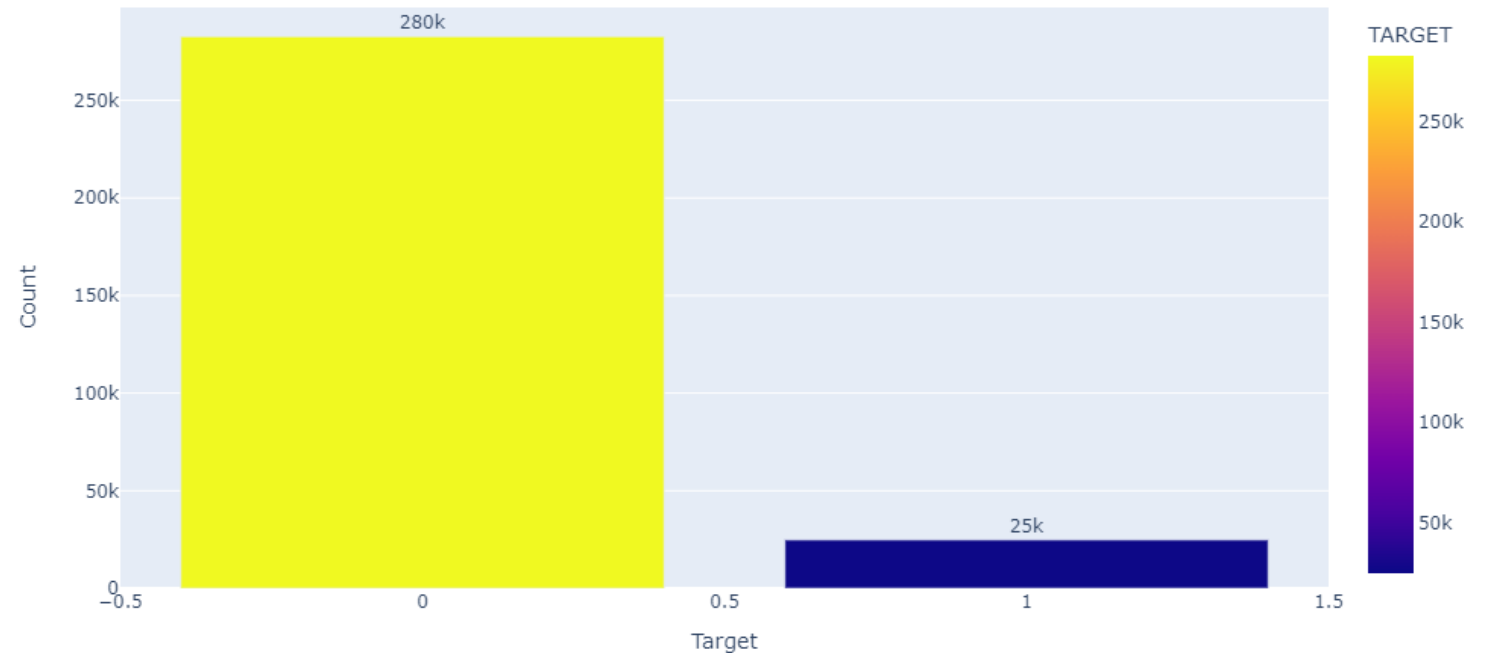
The Imbalance ratio 1:11,Whereas 1 in every 11 applicants has payment difficulties.

BIVARIATE ANALYSIS

&

MULTIVARIATE ANALYSIS

BASED ON CLIENTS WITH PAYMENT DIFFICULTIES AND CLIENTS ON TIME PAYERS.

CONTRACT TYPE.

Mostly Cash loans are Approved and Mostly Cash loans are processed on the both the conditions at the rate of:

- Clients With Payment Difficulties. Cash loans – 93.6%,Revolving loans – 6.4%
- Clients With On Payments. Cash loans – 90.3%,Revolving loans – 9.7%
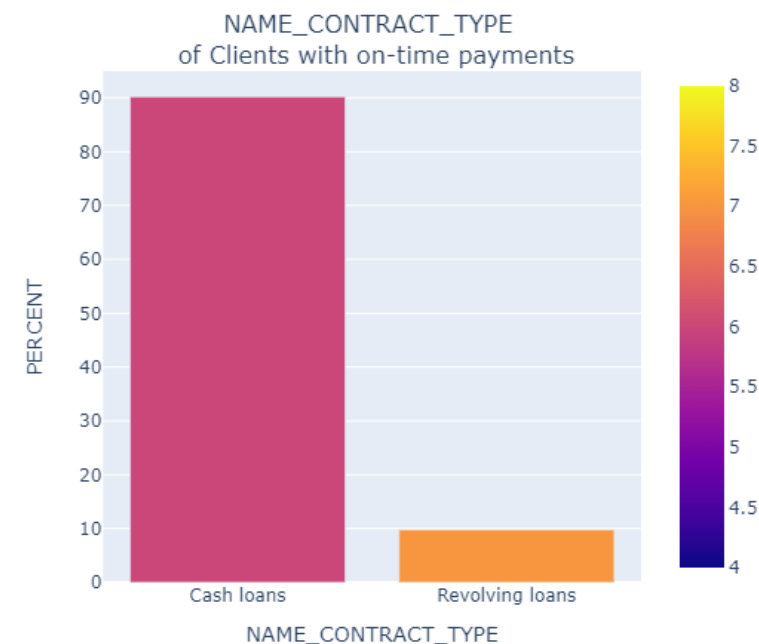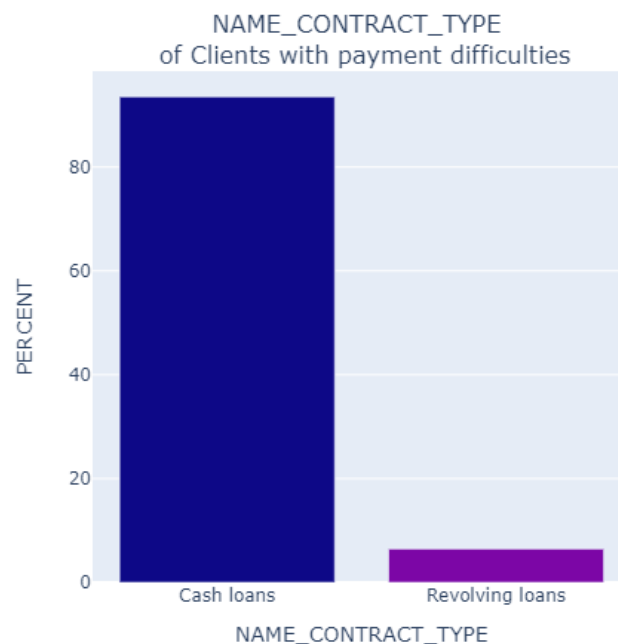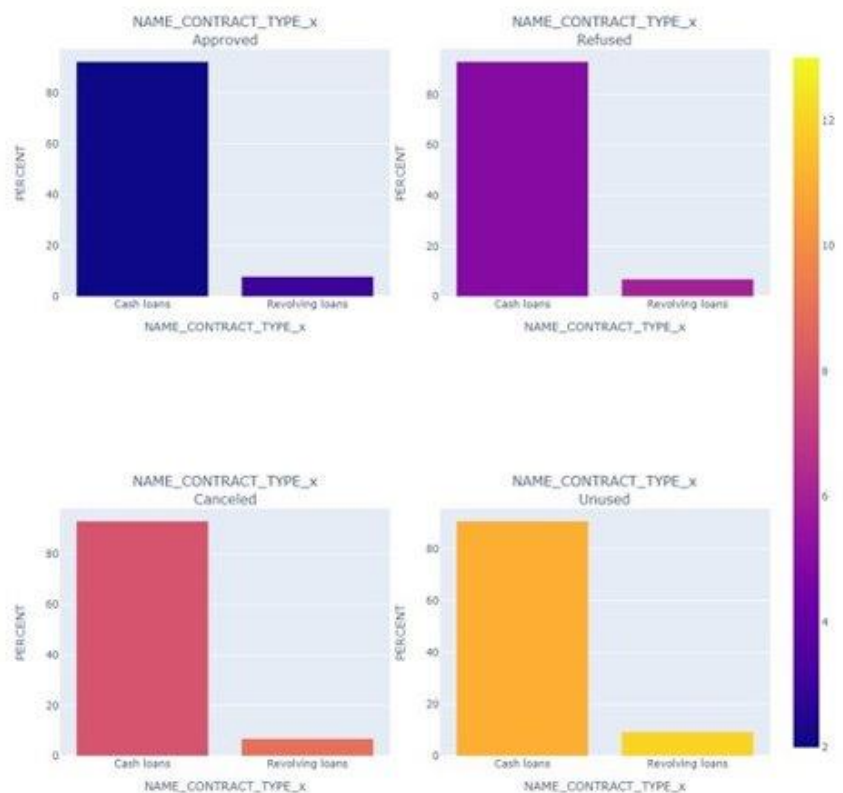
Code Gender.

Mostly Females clients are processing more loans on the both the conditions at the rate of:

- Clients With Payment Difficulties.

    1. Females – 57.1%
    2. Males – 42.9%

- Clients With On Payments.

    1. Females – 66.6%
    2. Males – 33.4%

- Female Clients Who has income as Working has better On Payments then Males and other Income type.
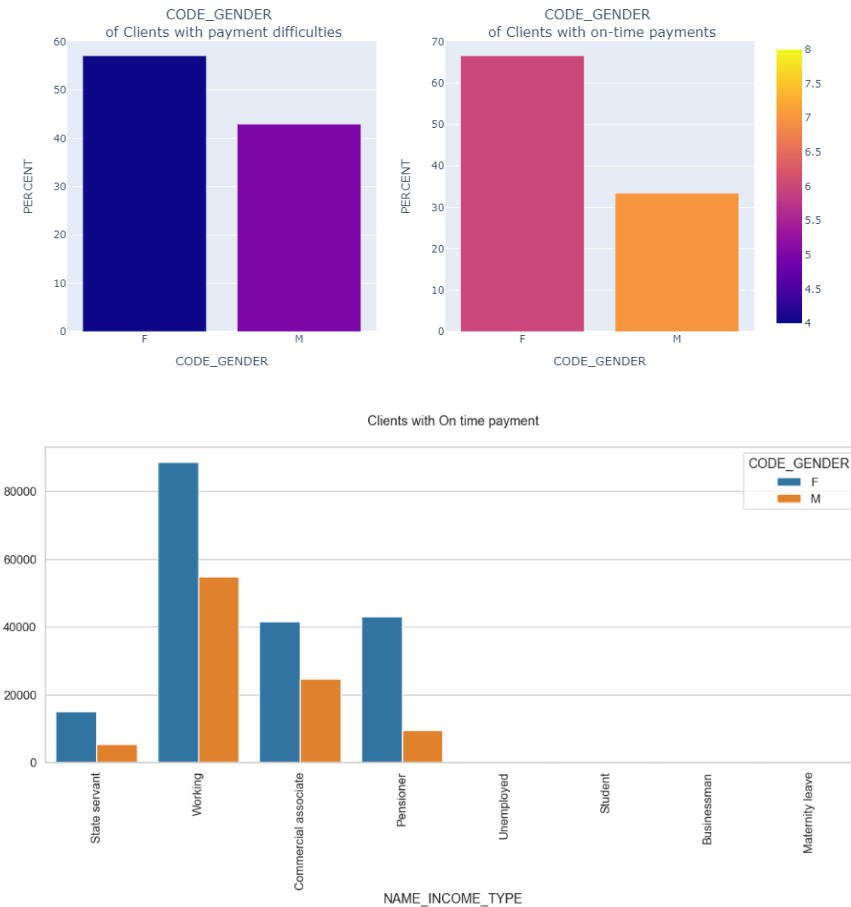
Income Factor.

I have grouped Clients based upon the income.

- Most of the applications for loan is from the clients who fall under Middleclass and Upper Middleclass.
- But the Clients who are under Rich income group has better on time payments
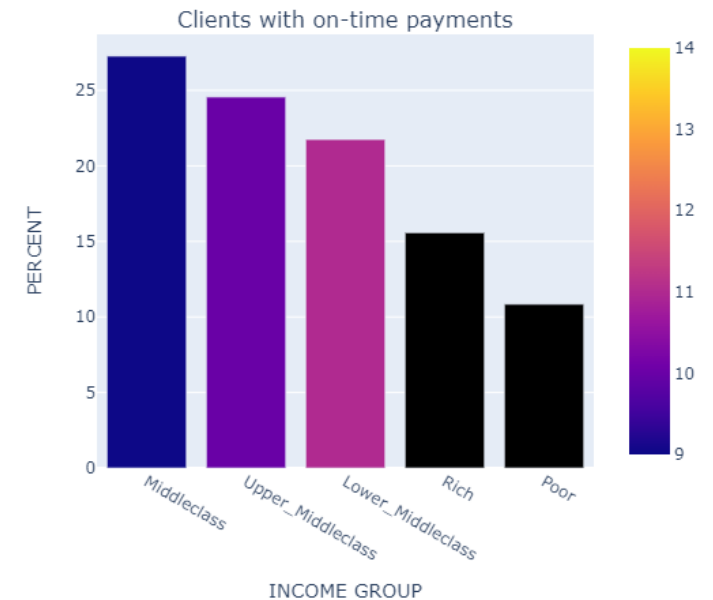
Income Type.

I have grouped Clients based upon the income.

- Students and Businessmen has no payment difficulties.
- Pensioners has better on payment.
- Working class has processed more loans in comparison with others.
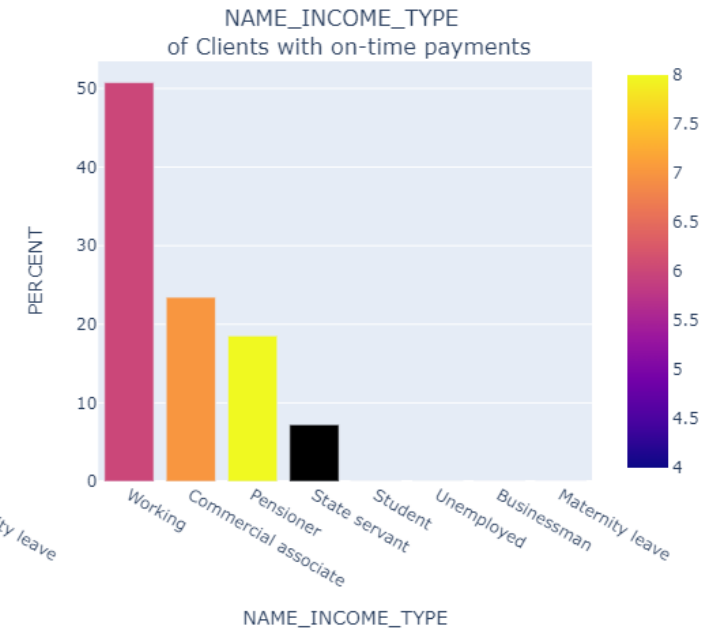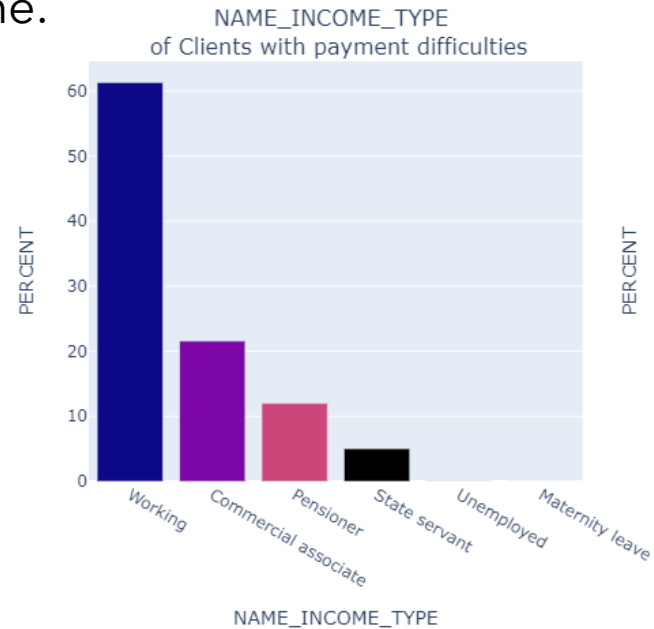
FAMILY STATUS.

- Clients who are Married has better on time payment then others.

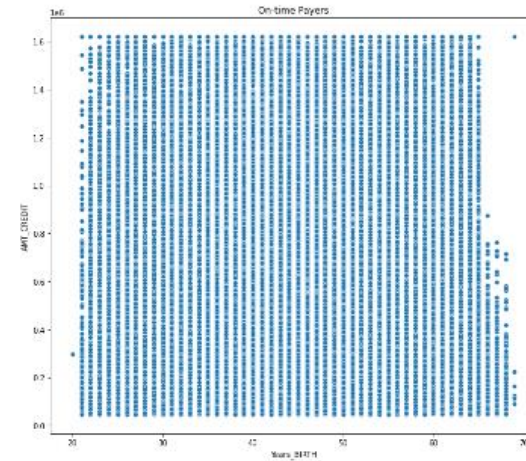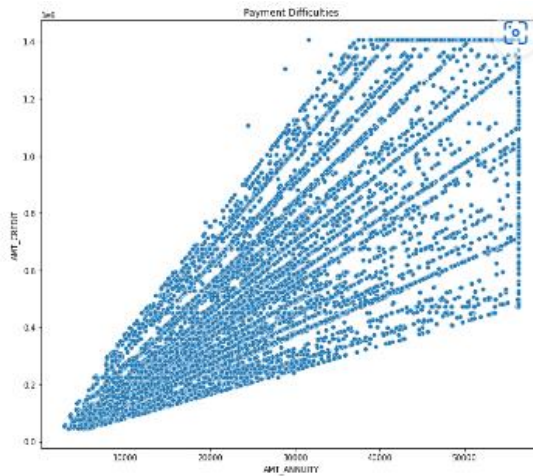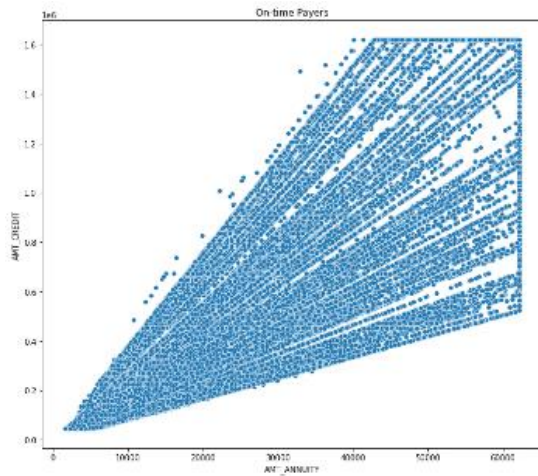- As most of the clients are married they have most Approved and Refused loans.
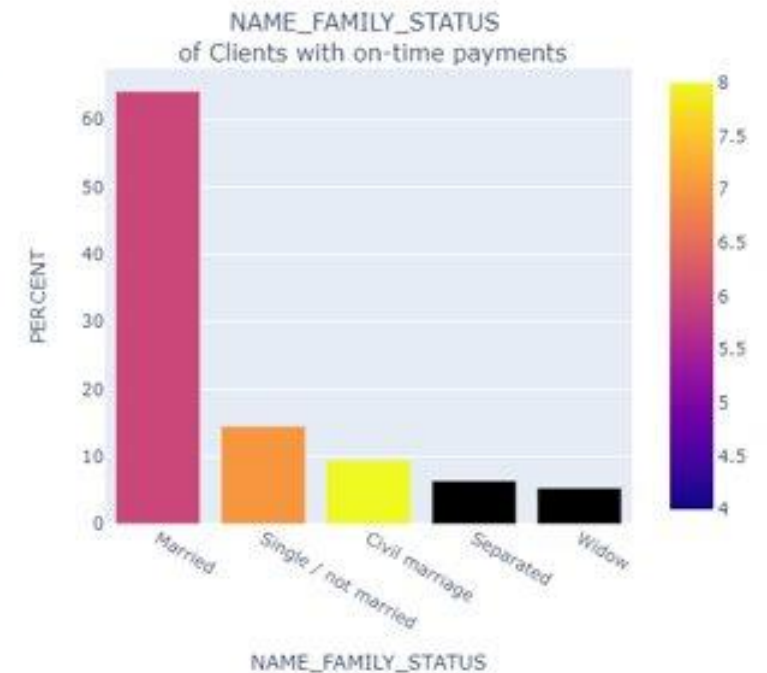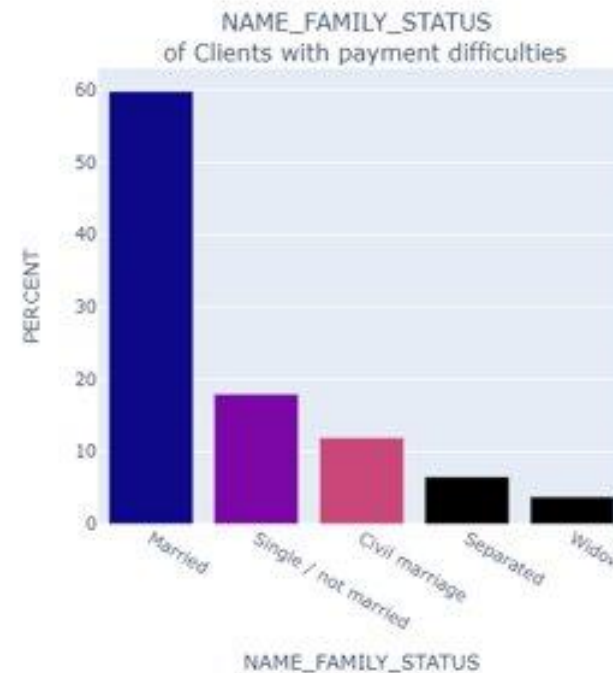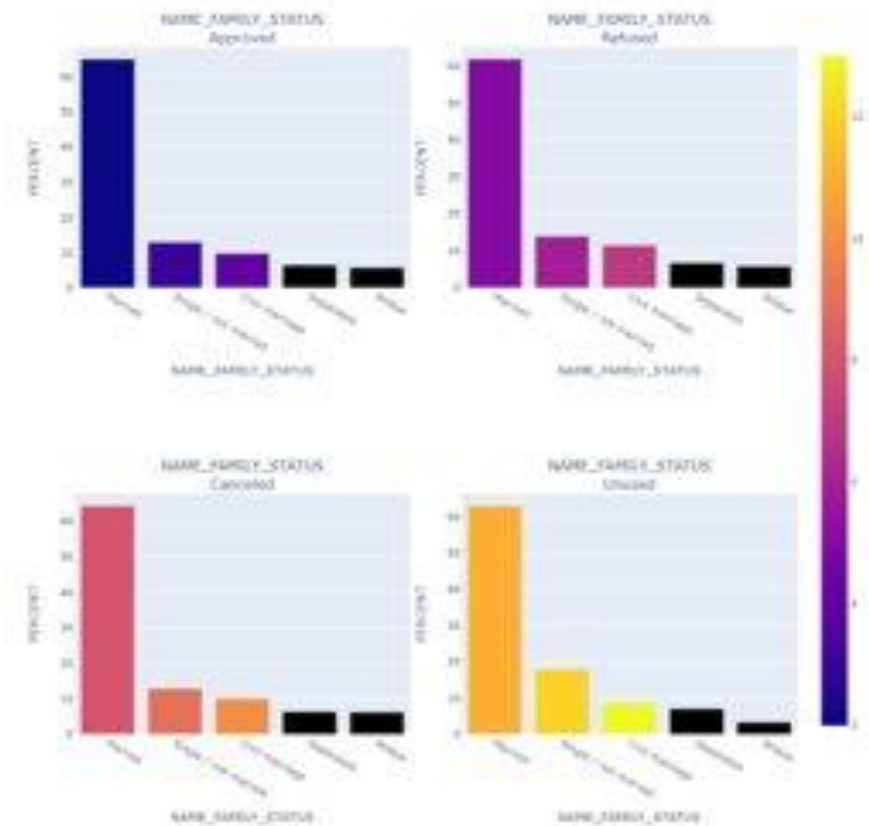
EDUCATION LEVEL.

- Clients with Higher Education has better on time payment then others.

- Clients with Higher education and Academic Degree should be given preference.
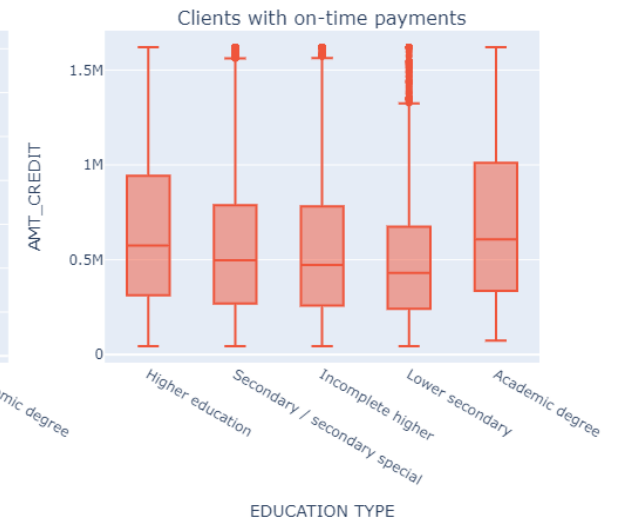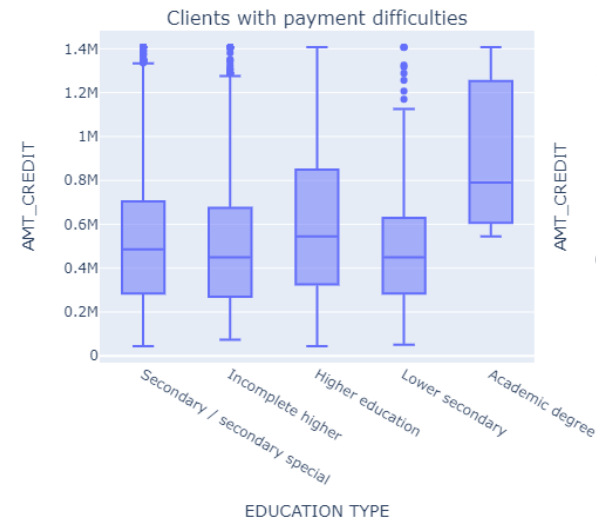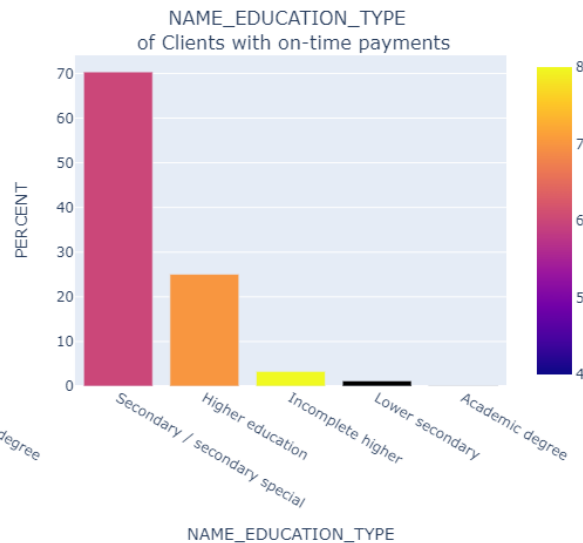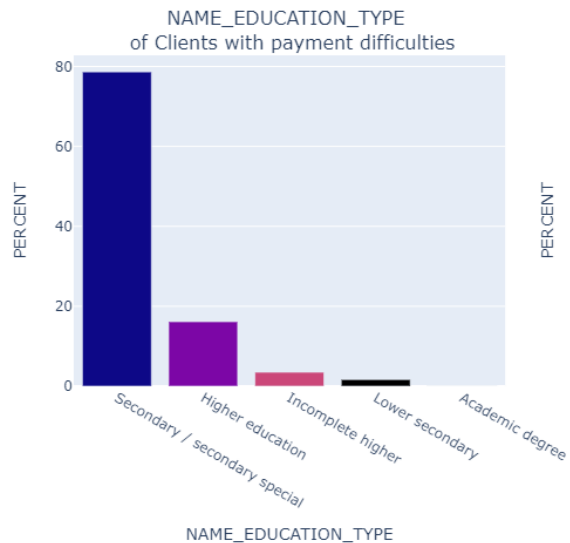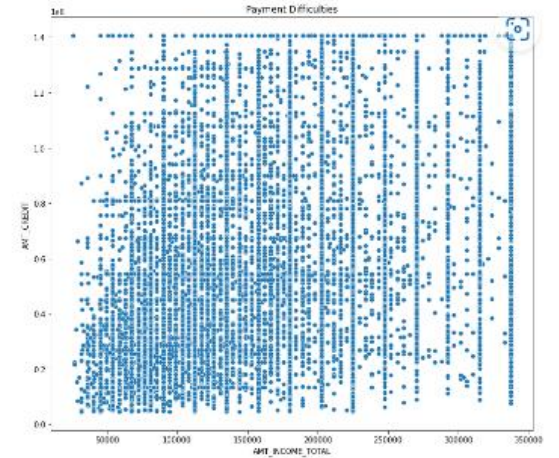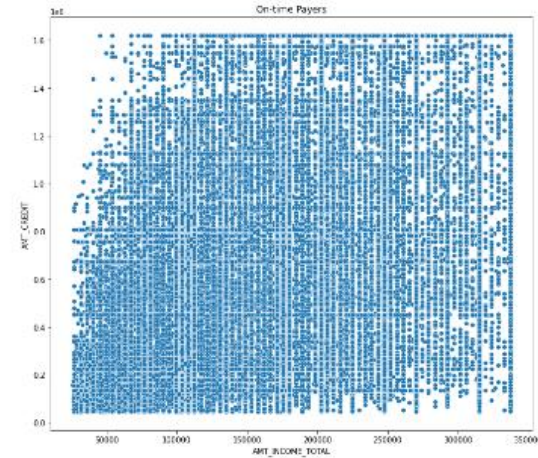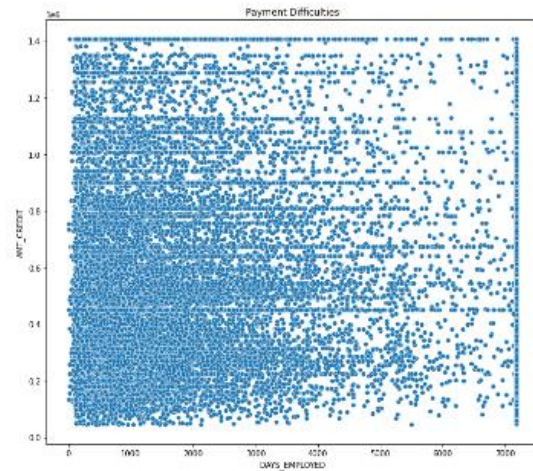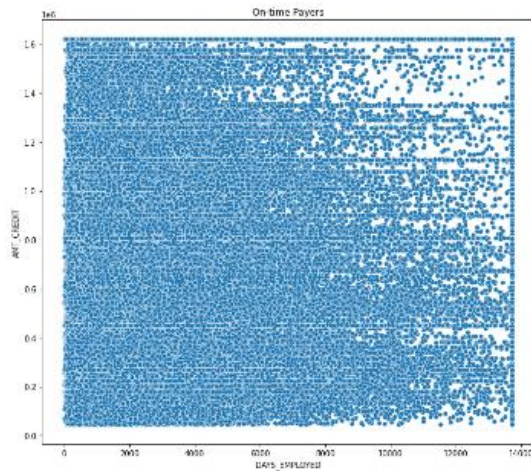
Top Correlation based on AMT_CREDIT

- AMT_CREDIT AND YEARS_BIRTH
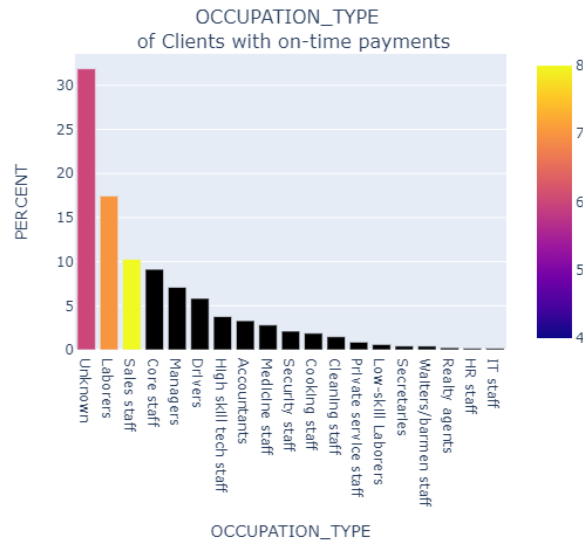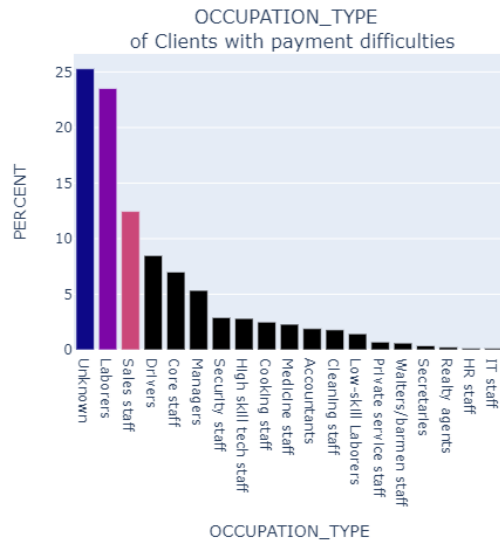- AMT_CREDIT AND AMT_ANNUITY
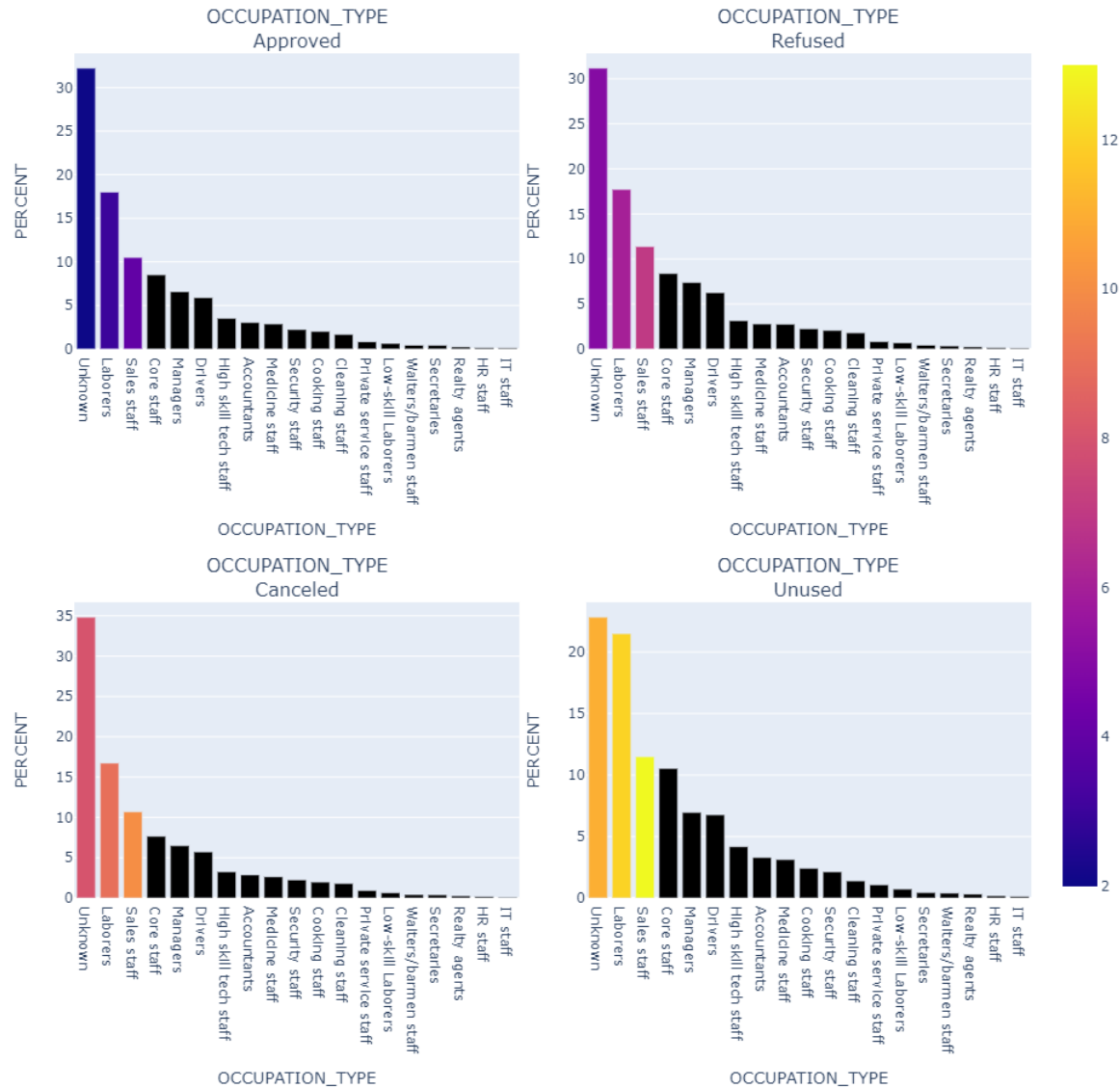
TOP CORRELATION BASED ON AMT_CREDIT

- Clients employed for long period of time has better rate of On time payment

- AMT_CREDIT and AMT_INCOME

OCCUPATION TYPE.

- Clients who's occupation type Core staff has Better AMT_CREDIT rate and better On time payments.

- Where Clients who has Occupation type as Laborers has Payment Difficulties .
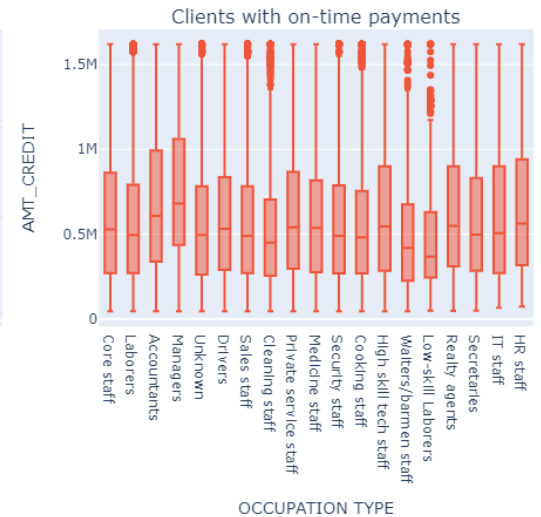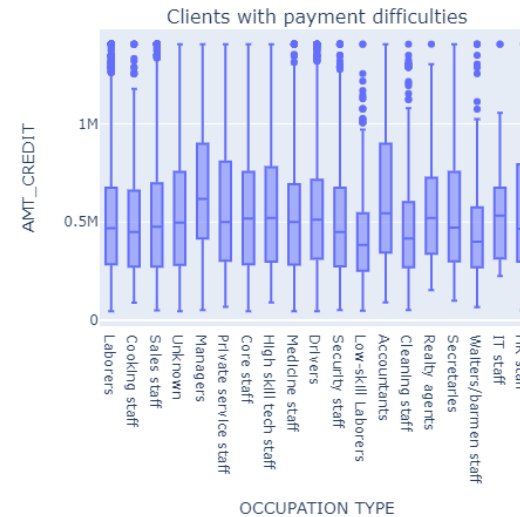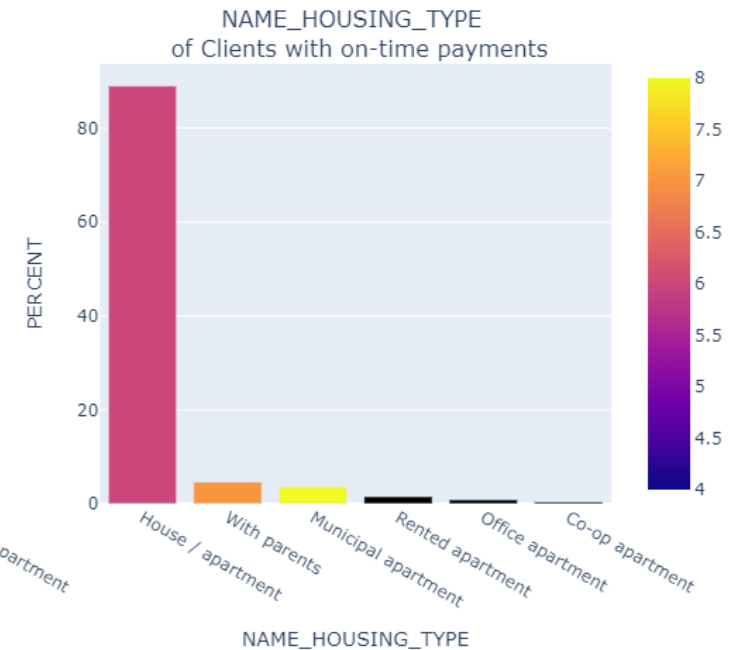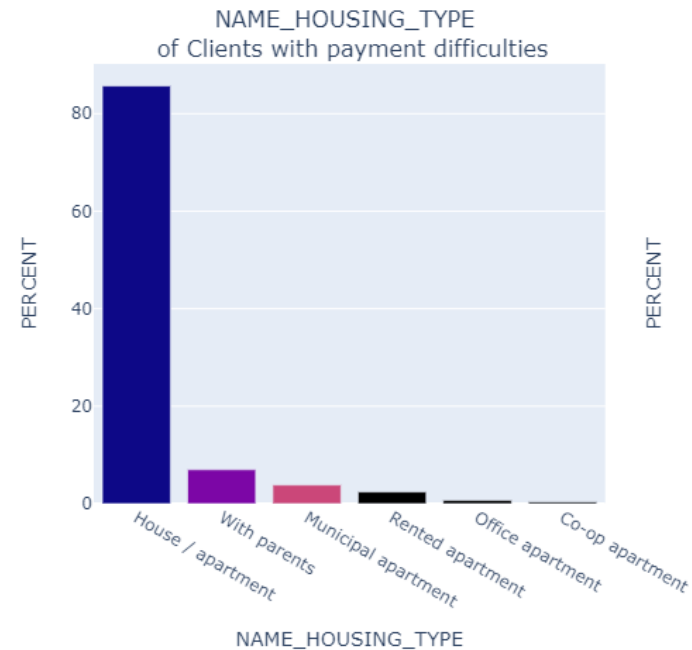
- Clients who's occupation type as Laborers has got loans

1. Approved:17.9%
2. Refused:17.7%
3. Canceled:13.4%
4. Unused:21.7%

HOUSING TYPE.

- Clients who's occupation type Core staff has Better AMT_CREDIT rate and better On time payments.
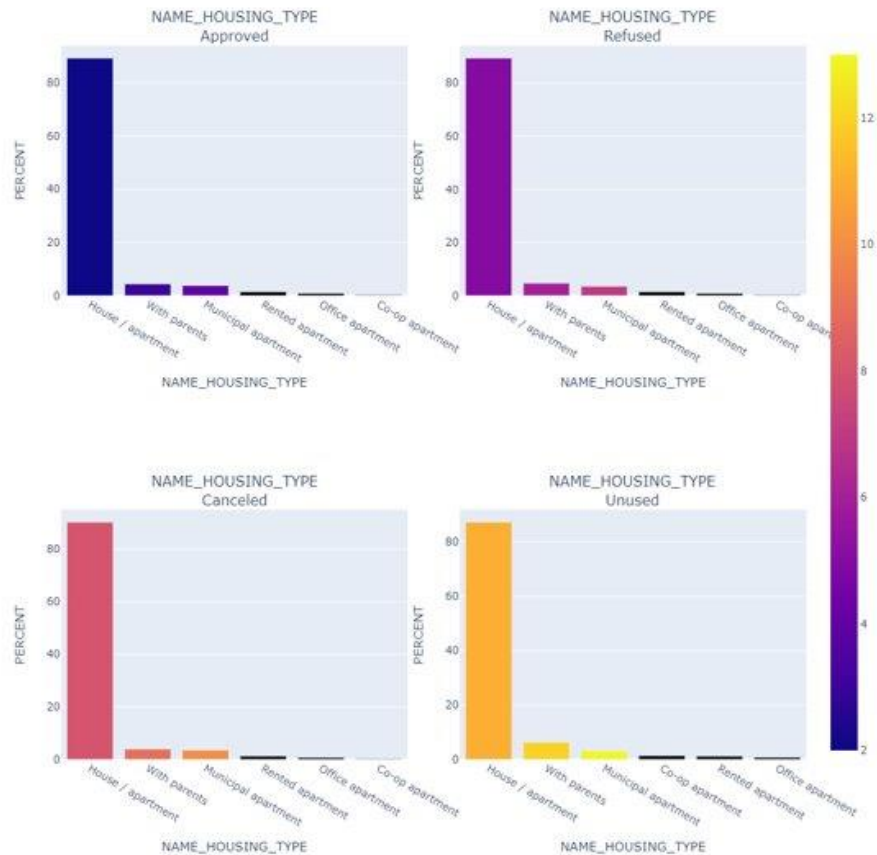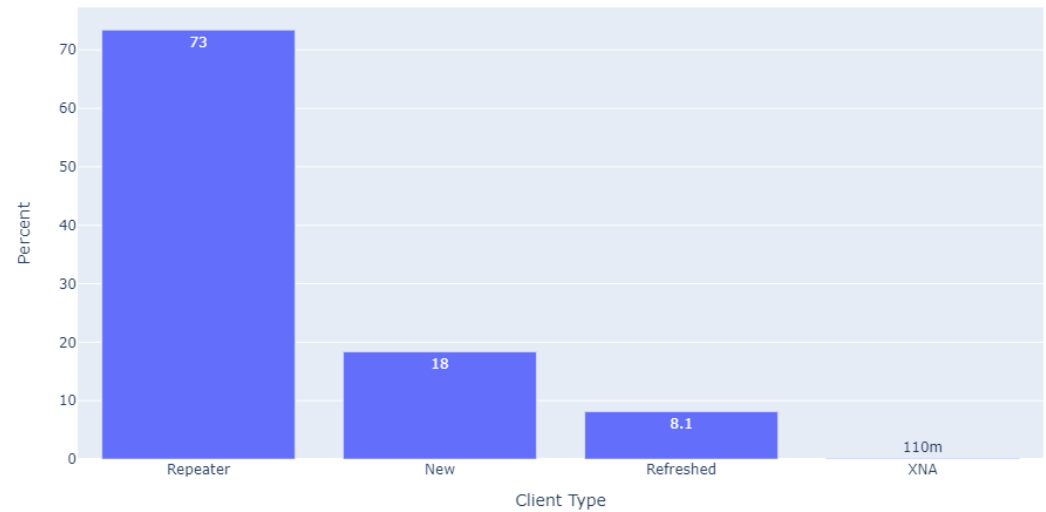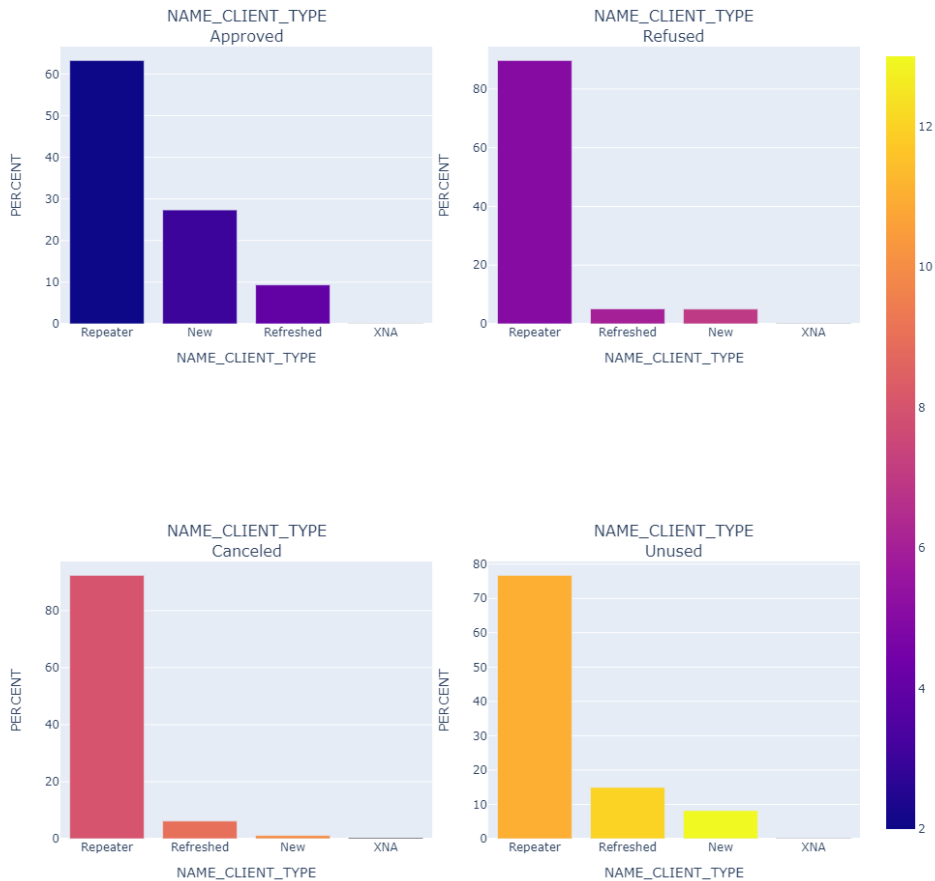- Where Clients who has Occupation type as Laborers has Payment Difficulties .
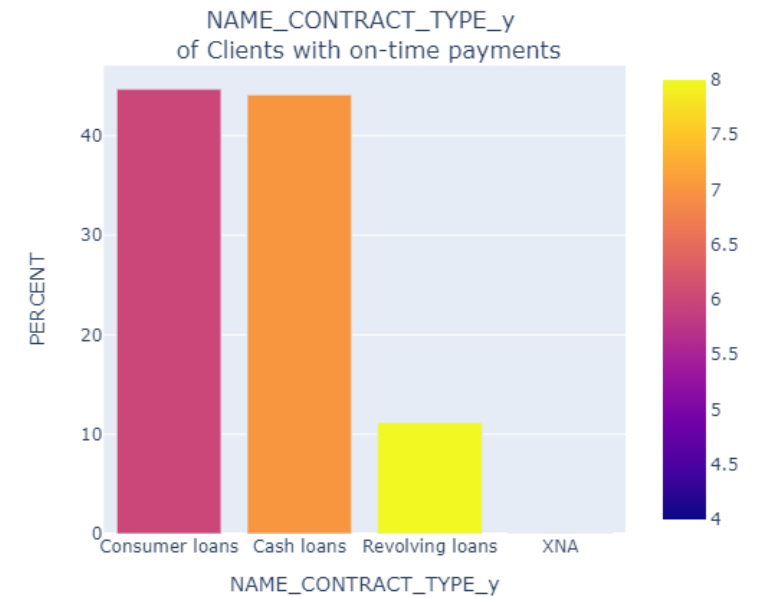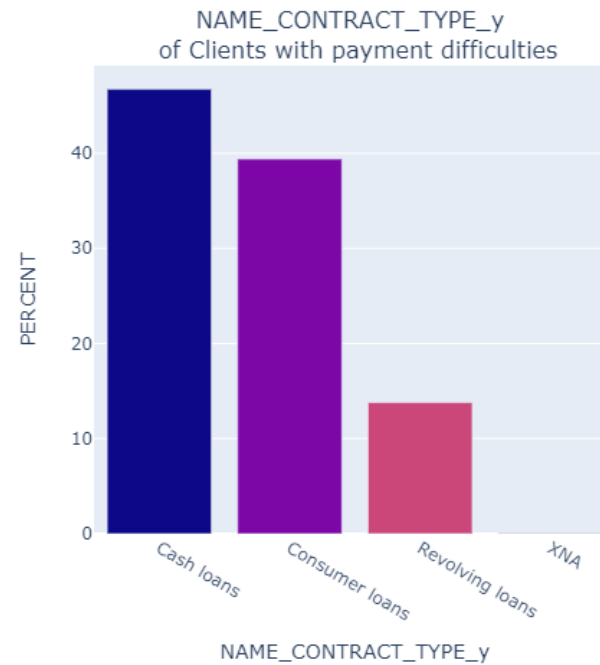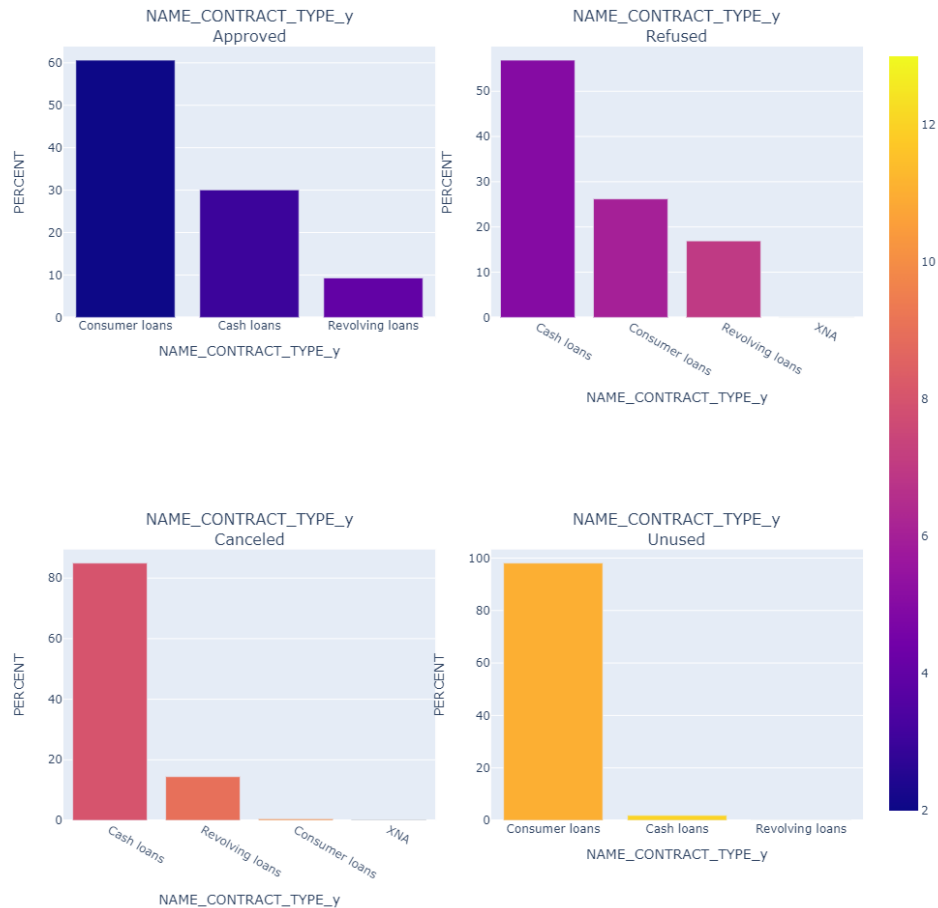
CLIENT TYPE.

- Most of the clients are Repeaters at 76.3% Whereas 18.3% are New clients.

- The percentage of loans are Approved for repeater is 63.24% and for New Clients 27.35%

CONTRACT TYPE.

- Clients Who have Consumer loans has Better On time payment
- Most approved loans are Consumer loans and Most refused loans are Cash Loans.

Conclusion 1:

Clients to be Targeted to approve the loan.

➢ Bank should focus on age (30-40),(40-50) and (50,60) as the better on time payment percentage than other group.

➢ Bank should focus on Clients who are Married and Widows.

➢ Bank should focus on clients who are Students and Businessman.

➢ The clients who has higher education has better on time Percentage than Others.

➢ The client's whose Income source as Businessman, Pensioner and Commercial associate.

➢ Female clients who are working.

➢ Clients whose housing type is with parents.al associate.

Conclusion 2:

- **Chance of Clients having payment difficulty.**

- Clients who are unemployed.

- Clients who has no own house.

- Clients whose occupation type as labourers, Salesman, Drivers.

- Age group of 60-70.

- Client in the income group of low.

- Clients who has less employment days.

THANK YOU