

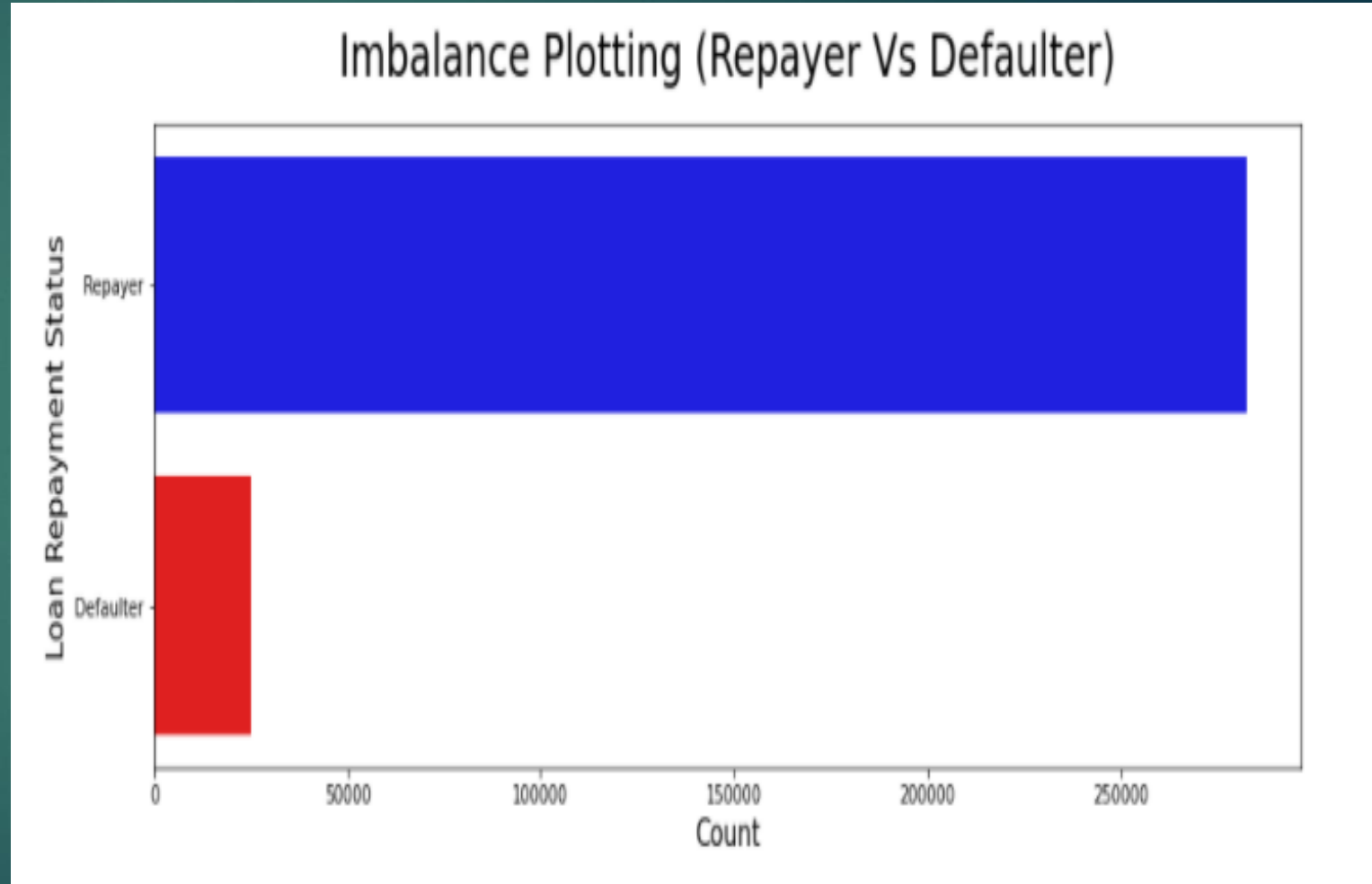
# CREDIT EDA CASE STUDY



# ANALYSIS BASED ON APPLICATION DATA

# IMBALANCE IN DATA

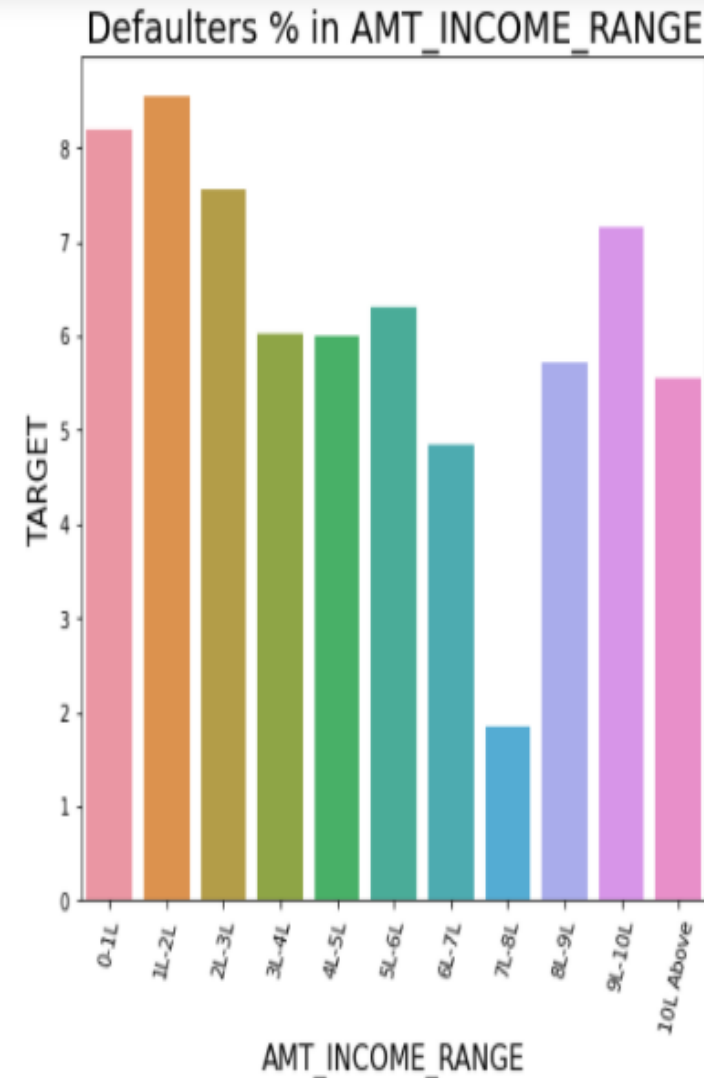
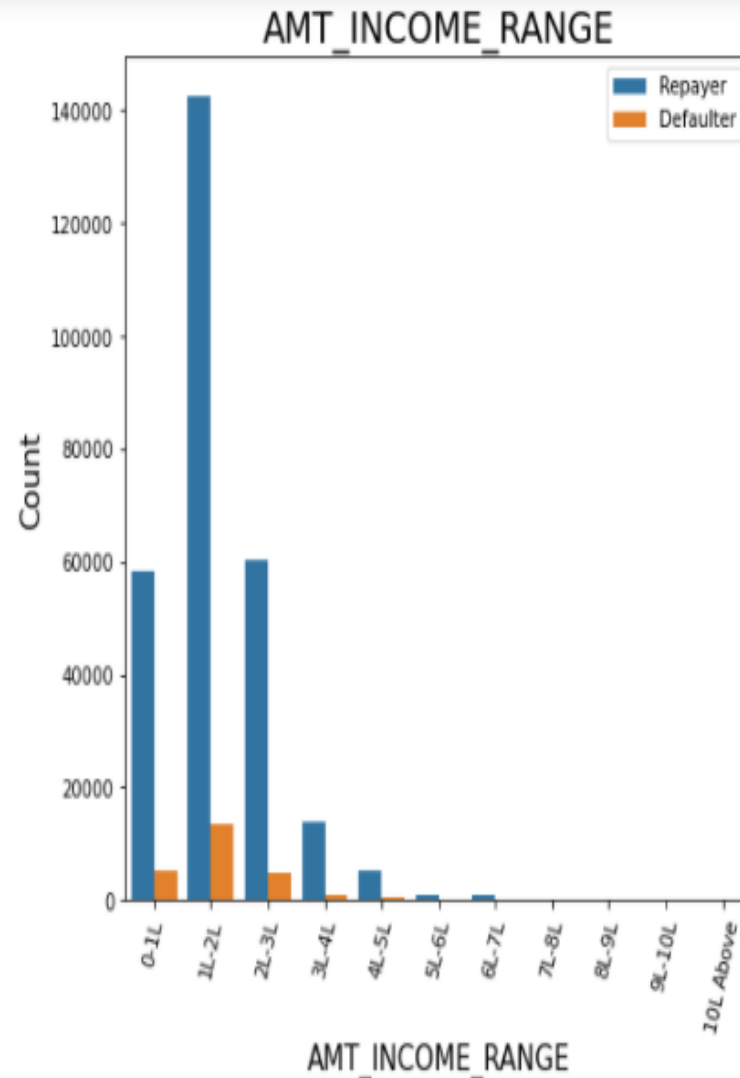
- Repayer Percentage is 91.93%
- Defaulters percentage is 8.07%
- Imbalance ratio with respect to repayer and defaulter is given: 11.39/1 (approx)



# UNIVARIATE ANALYSIS

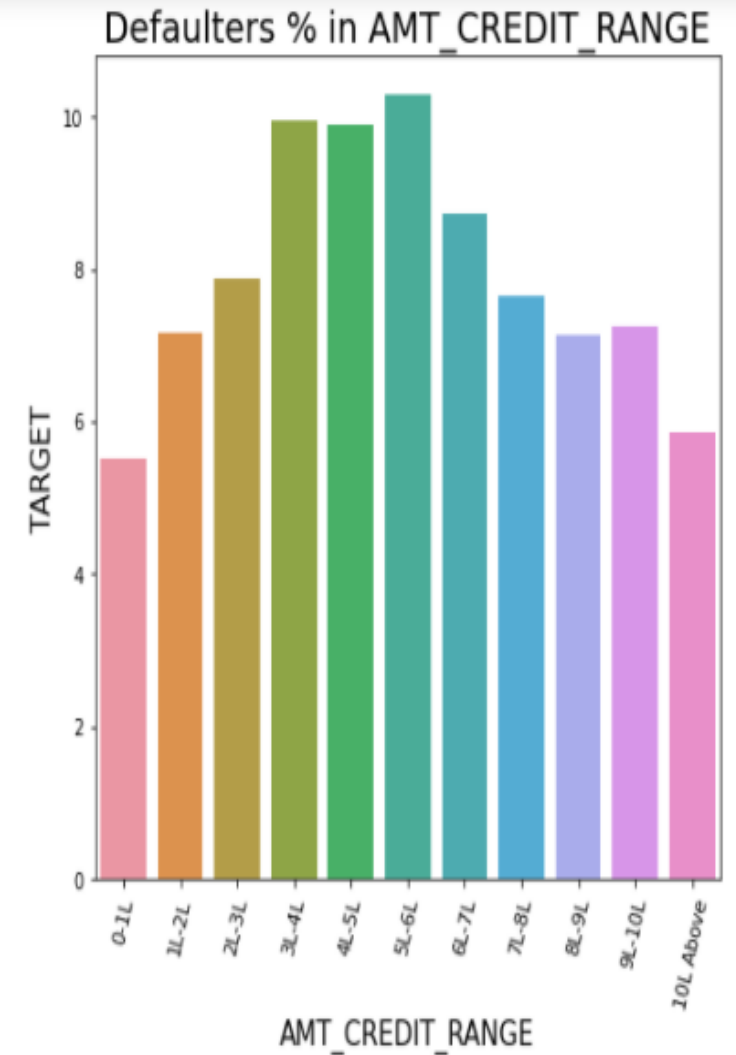
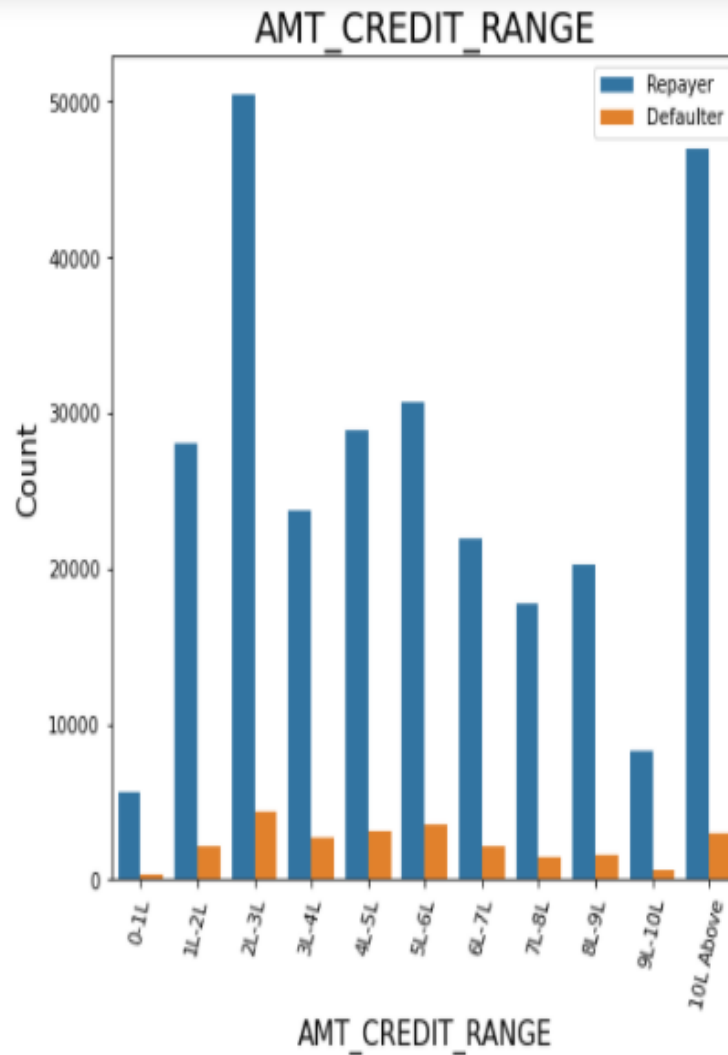
# DISTRIBUTION OF INCOME RANGE

- Majority of the applications have Income total less than 3 Lakhs.
- Application with Income less than 3 Lakhs has high probability of defaulting
- Applicant with Income 7-8 Lakhs are less likely to default.



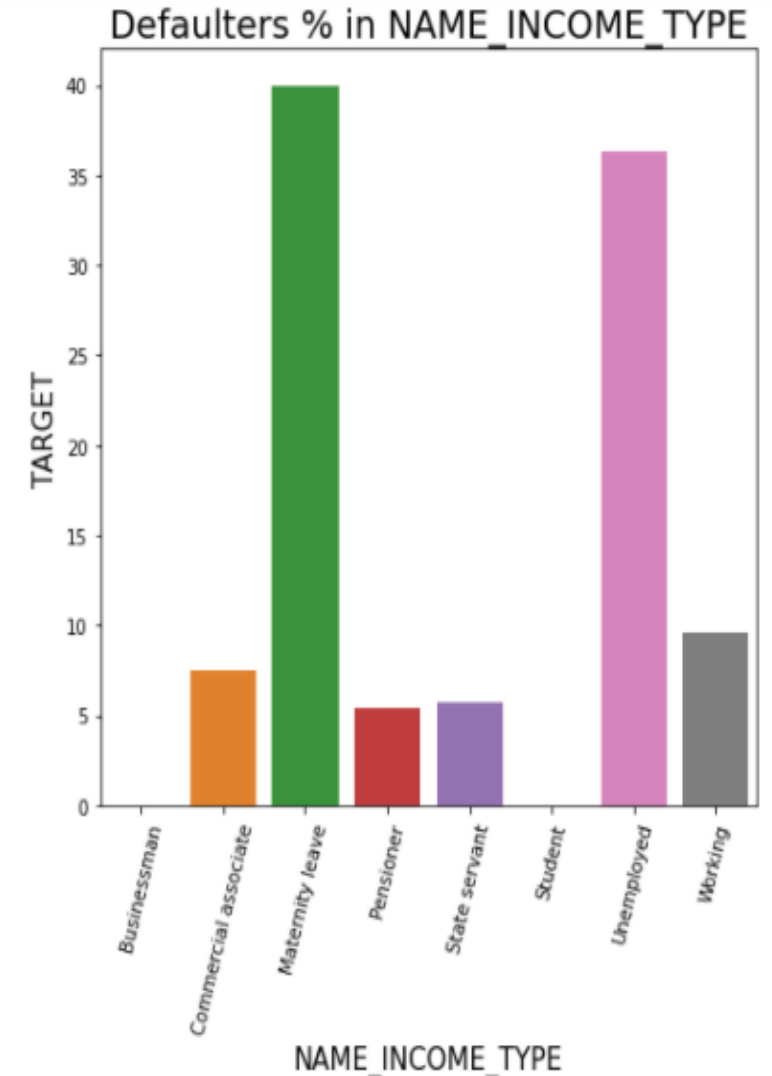
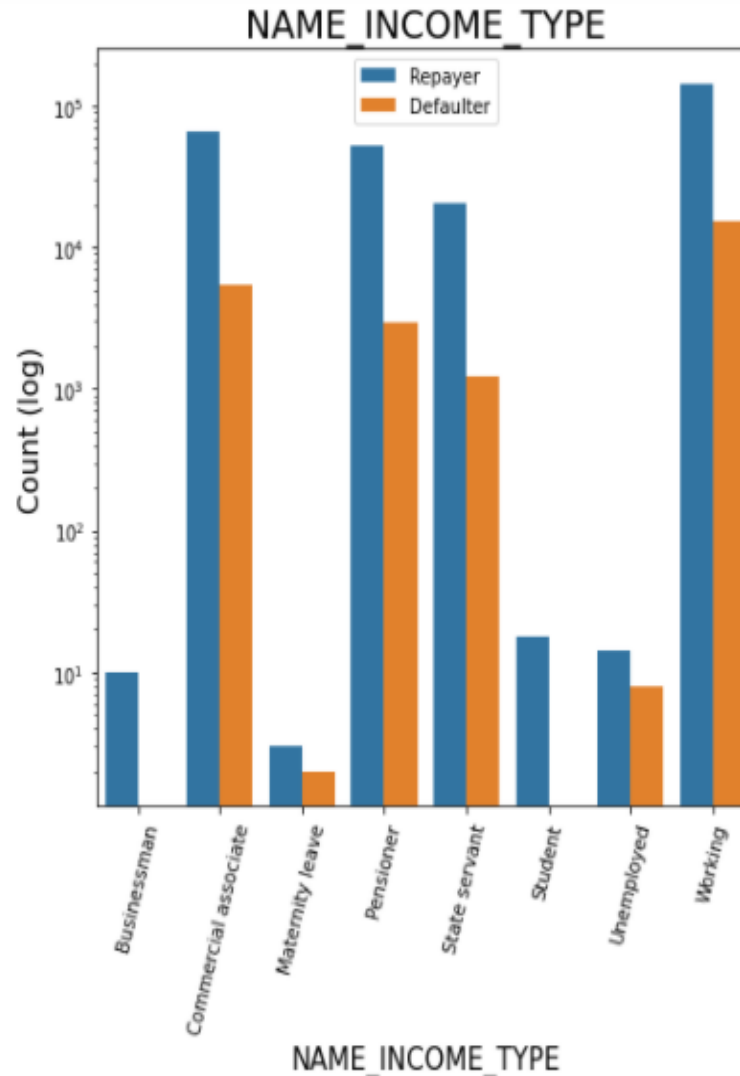
# DISTRICUTION OF LOAN AMOUNT

- There are high number of applicants have loan in range of 2-3 Lakhs followed by 10 Lakh above range
- People who get loan for 3-6 Lakhs have most number of defaulters than other loan range.



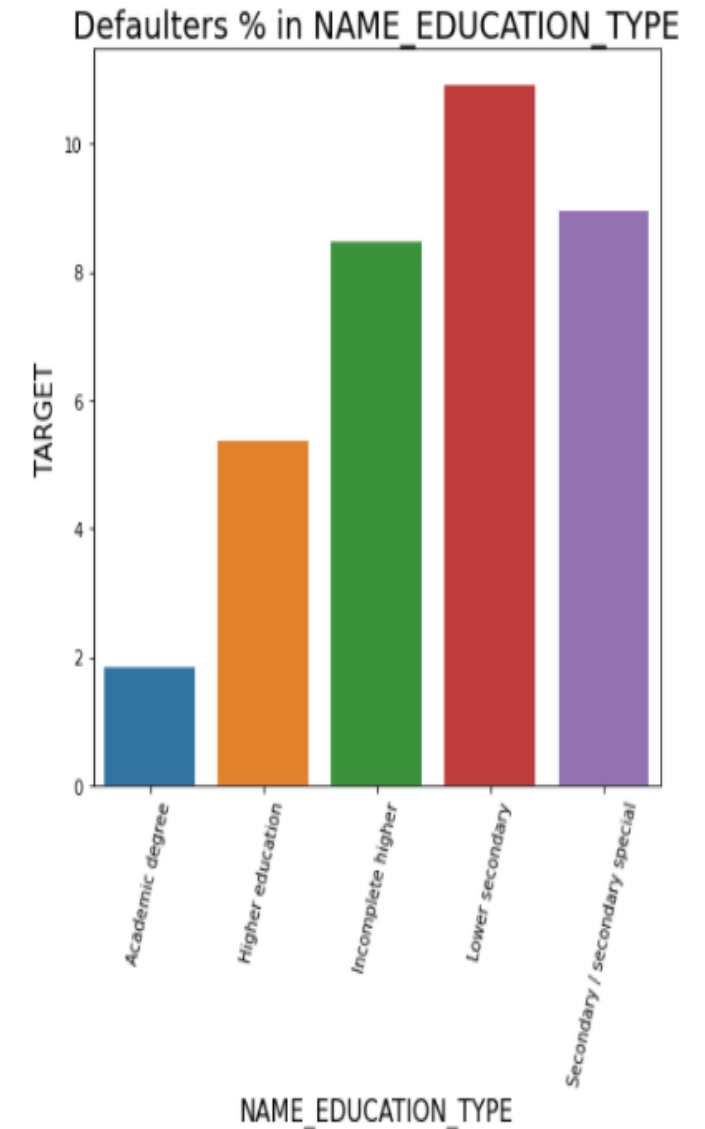
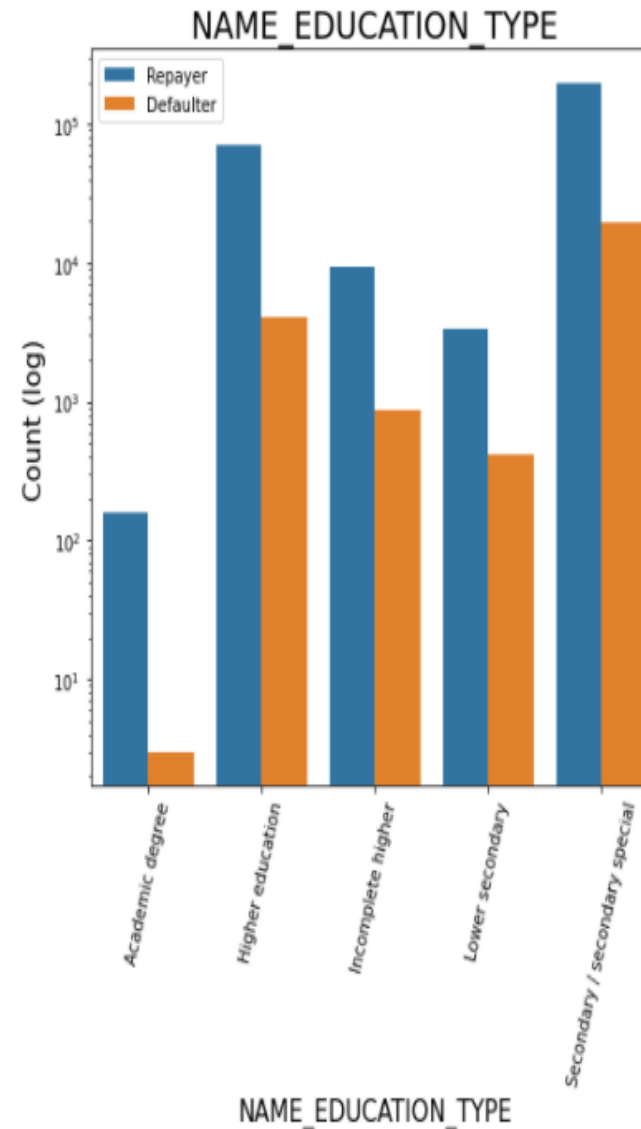
# DISTRIBUTION OF INCOME TYPE

- Most of applicants for loans income type is Working, followed by Commercial associate, Pensioner and State servant.
- The applicants who are on Maternity leave have defaulting percentage of 40% which is the highest, followed by Unemployed (37%). The rest under average around 10% defaulters.
- Student and Businessmen though less in numbers, do not have default record. Safest two categories for providing loan.



# DISTRIBUTION OF EDUCATION TYPE

- Majority of clients have Secondary/secondary special education, followed by clients with Higher education.
- Very few clients have an academic degree
- Lower secondary category have highest rate of defaulting around 11%.
- People with Academic degree are least likely to default.

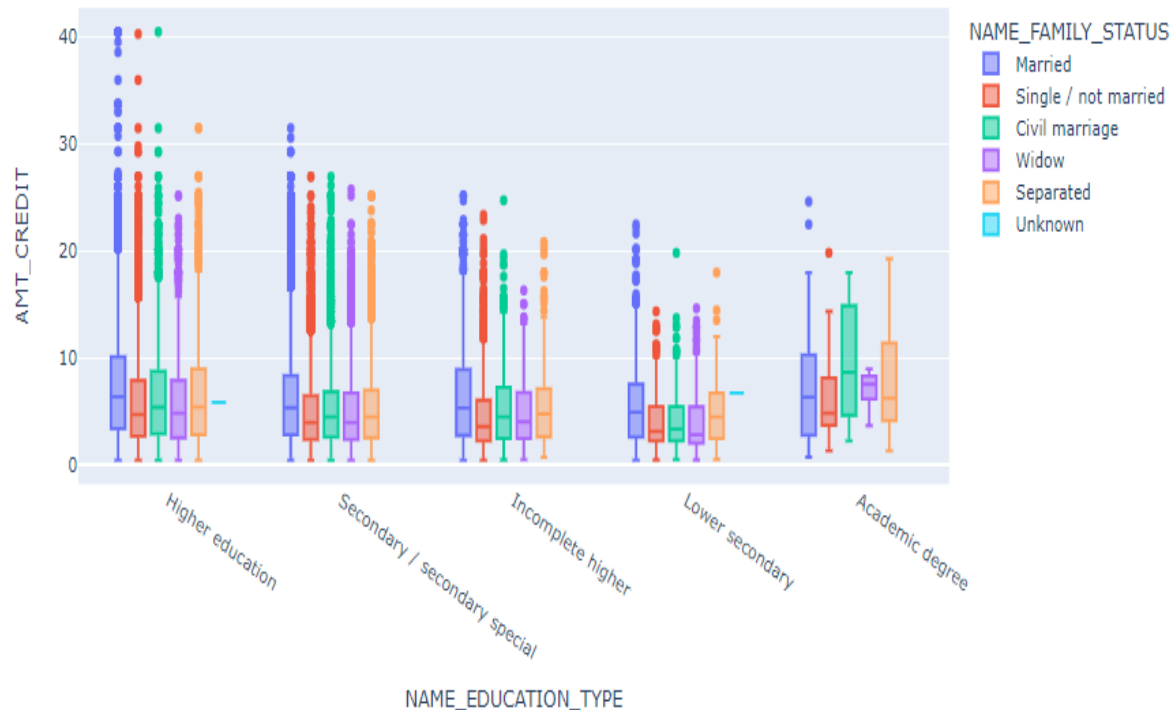




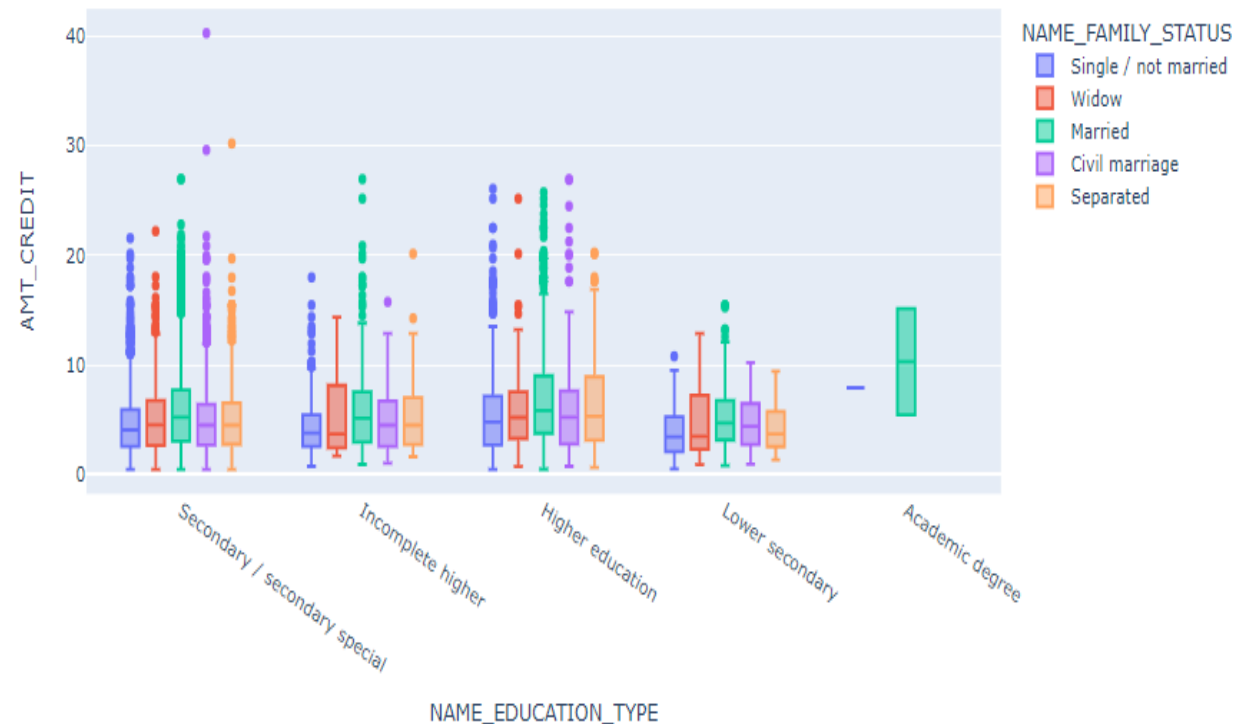
# BIVARIATE ANALYSIS

# COMPARISON OF CREDIT AMOUNT VS EDUCATION OF LOAN FOR PAYMENT DIFFICULTIES AND NON PAYMENT DIFFICULTIES

Credit amount vs Education of Loan- Non Payment Difficulties



Credit amount vs Education of Loan Payment Difficulties



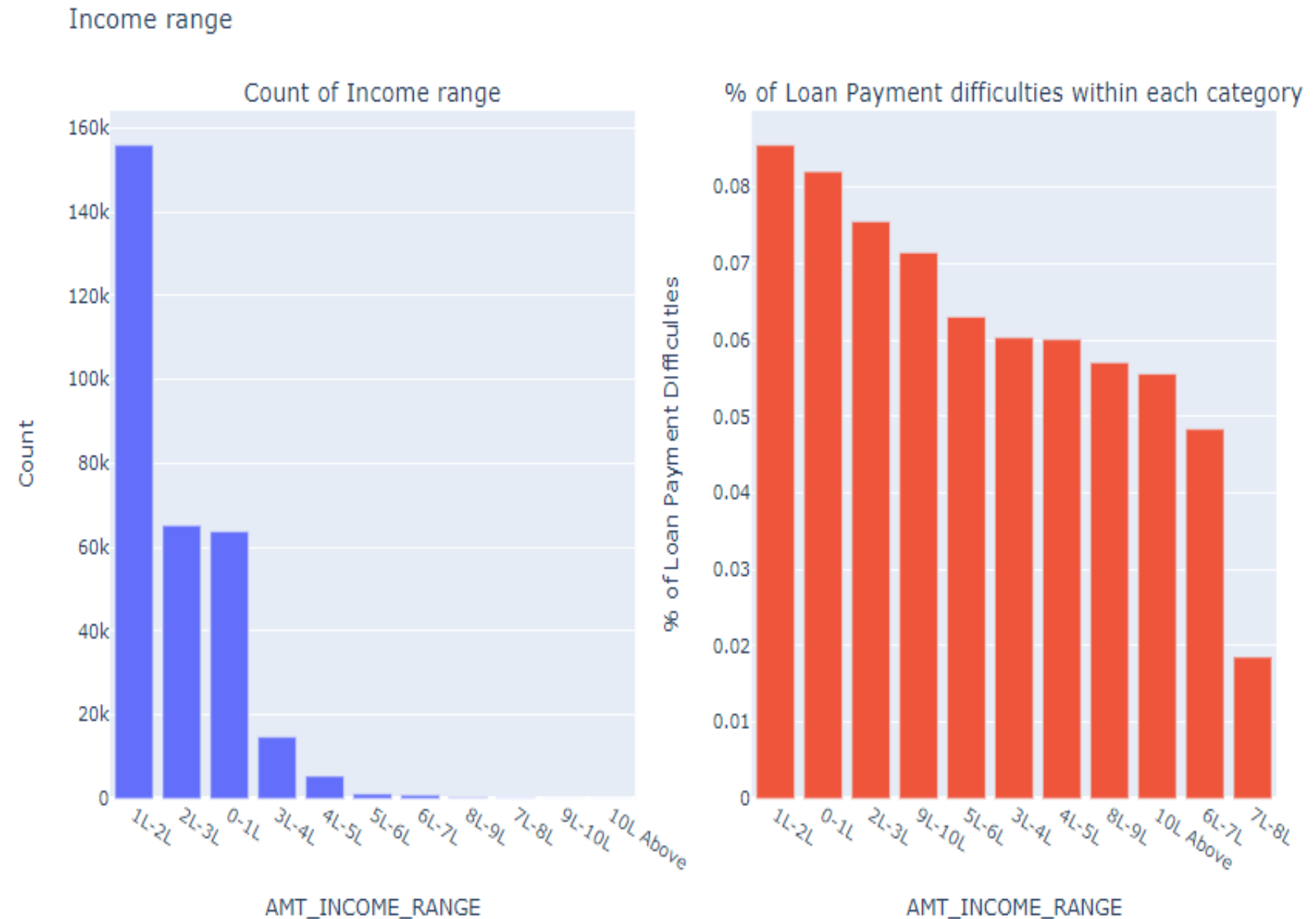
- The graphs for Loan Payment Difficulties and Loan Non-Payment Difficulties appears to be similar. We observe that Family status of 'civil marriage', 'marriage' and 'separated' of Academic degree education are having higher number of credits than others. Most of the outliers are from Education type 'Higher education' and 'Secondary'. Civil marriage for Academic degree is having most of the credits in the third quartile.



# CATAGORICAL BIVARIATE ANALYSIS

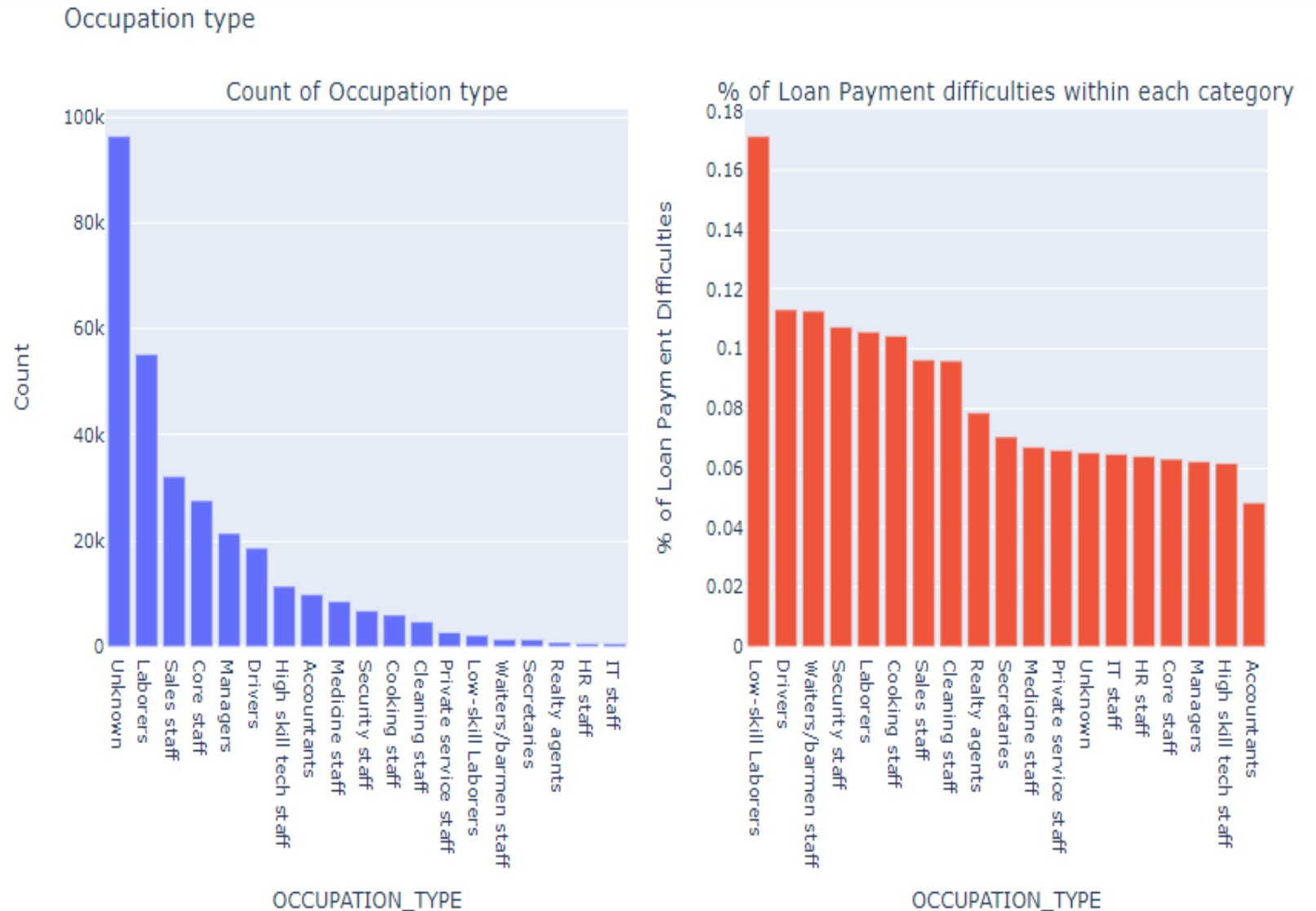
# INCOME RANGE

- From the plot above we can say that clients with 'LOW' Income range have maximum % of Loan-Payment Difficulties.



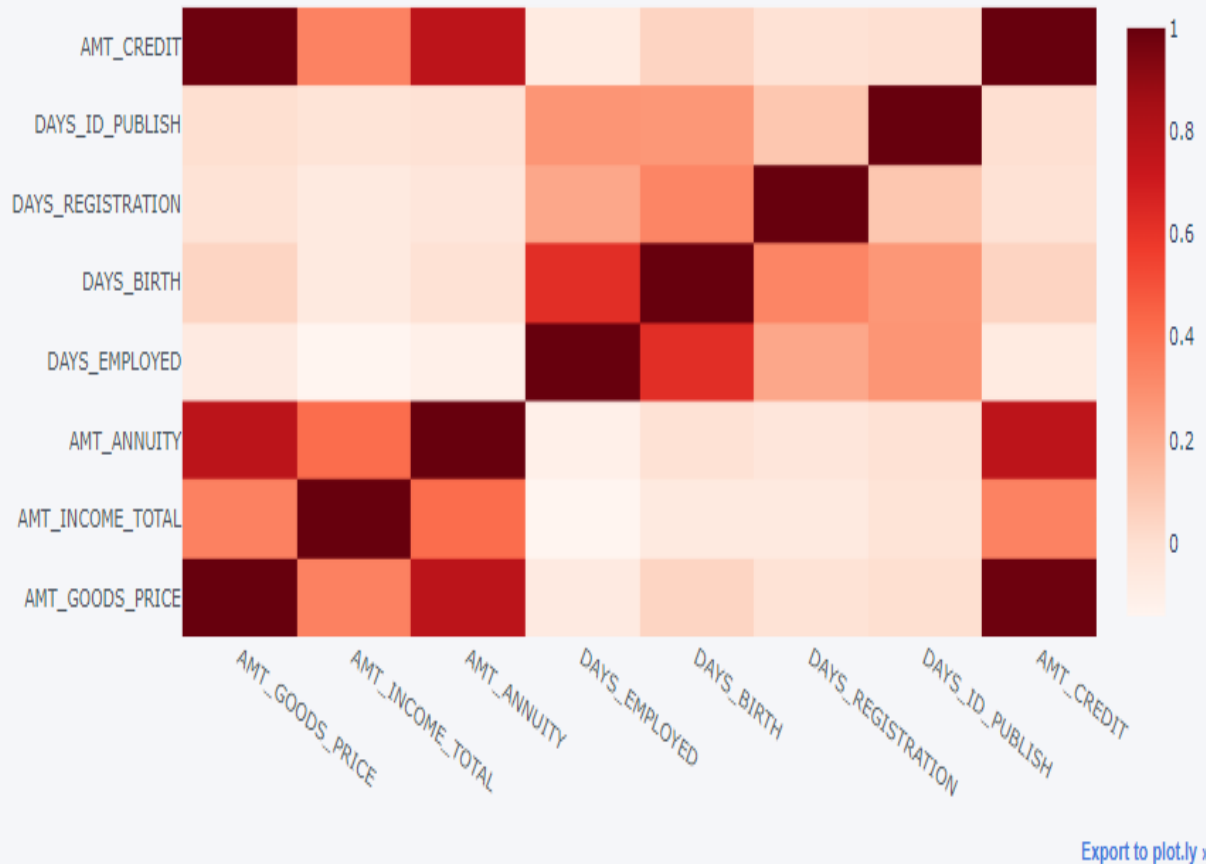
# ANALYSIS OF PAYMENT DIFFICULTY WITH OCCUPATION TYPE

- From the plot above we can say that clients with 'Lower skill Labourers' occupation type have maximum % of Loan-Payment Difficulties.

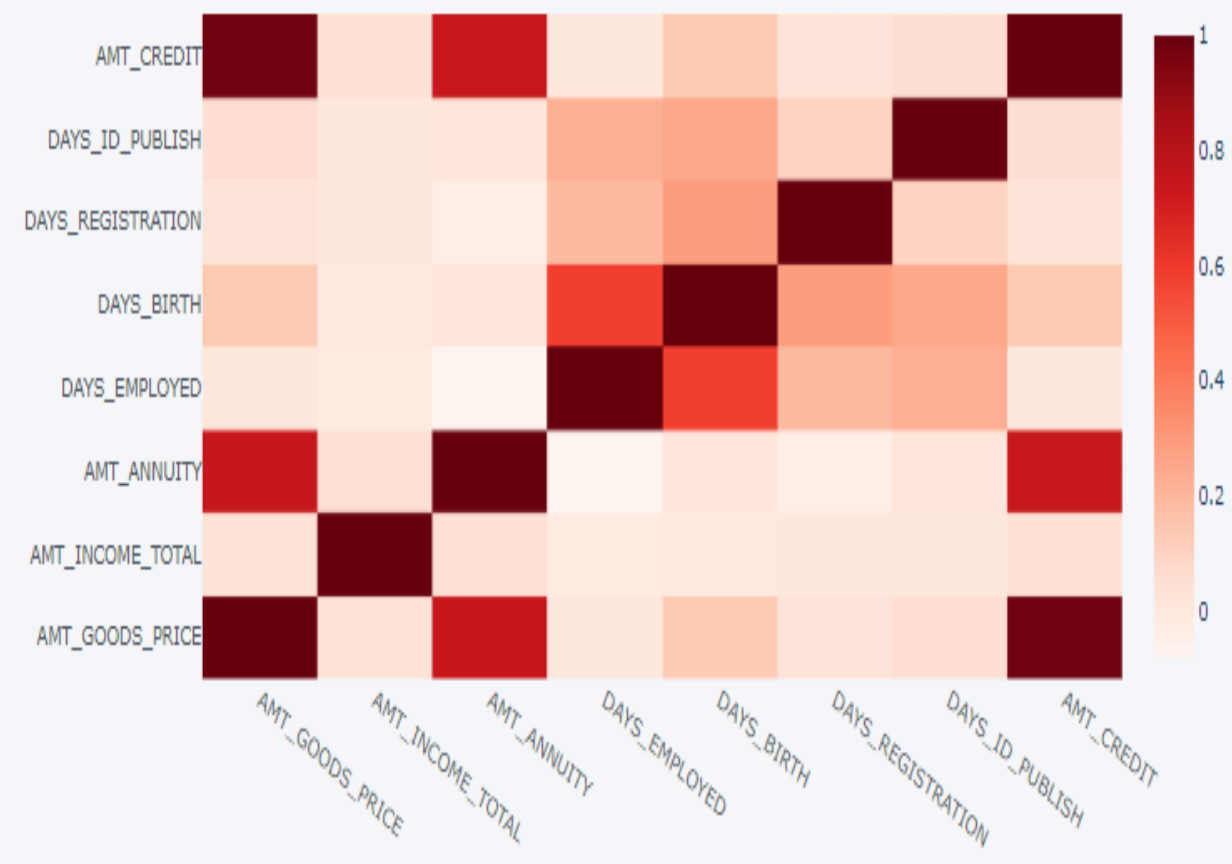


# CORRELATION

Correlation Heatmap of Loan- Non Payment Difficulties



Correlation Heatmap of Loan-Payment Difficulties



- We observe that there is a high correlation between credit amount and goods price. There appears to be some deviancies in the correlation of Loan-Payment Difficulties and Loan- Non Payment Difficulties such as credit amount v/s income

# TOP 10 CORRELATION WITH PAYMENT DIFFICULTY

	VAR1	VAR2	CORRELATION	CORR_ABS
56	AMT_CREDIT	AMT_GOODS_PRICE	0.983103	0.983103
16	AMT_ANNUITY	AMT_GOODS_PRICE	0.752699	0.752699
58	AMT_CREDIT	AMT_ANNUITY	0.752195	0.752195
35	DAYS_BIRTH	DAYS_EMPLOYED	0.582185	0.582185
44	DAYS_REGISTRATION	DAYS_BIRTH	0.289114	0.289114
52	DAYS_ID_PUBLISH	DAYS_BIRTH	0.252863	0.252863
51	DAYS_ID_PUBLISH	DAYS_EMPLOYED	0.229090	0.229090
43	DAYS_REGISTRATION	DAYS_EMPLOYED	0.192455	0.192455
32	DAYS_BIRTH	AMT_GOODS_PRICE	0.135810	0.135810
60	AMT_CREDIT	DAYS_BIRTH	0.135316	0.135316





# **ANALYSIS BASED ON PREVIOUS APPLICATION'S DATA**

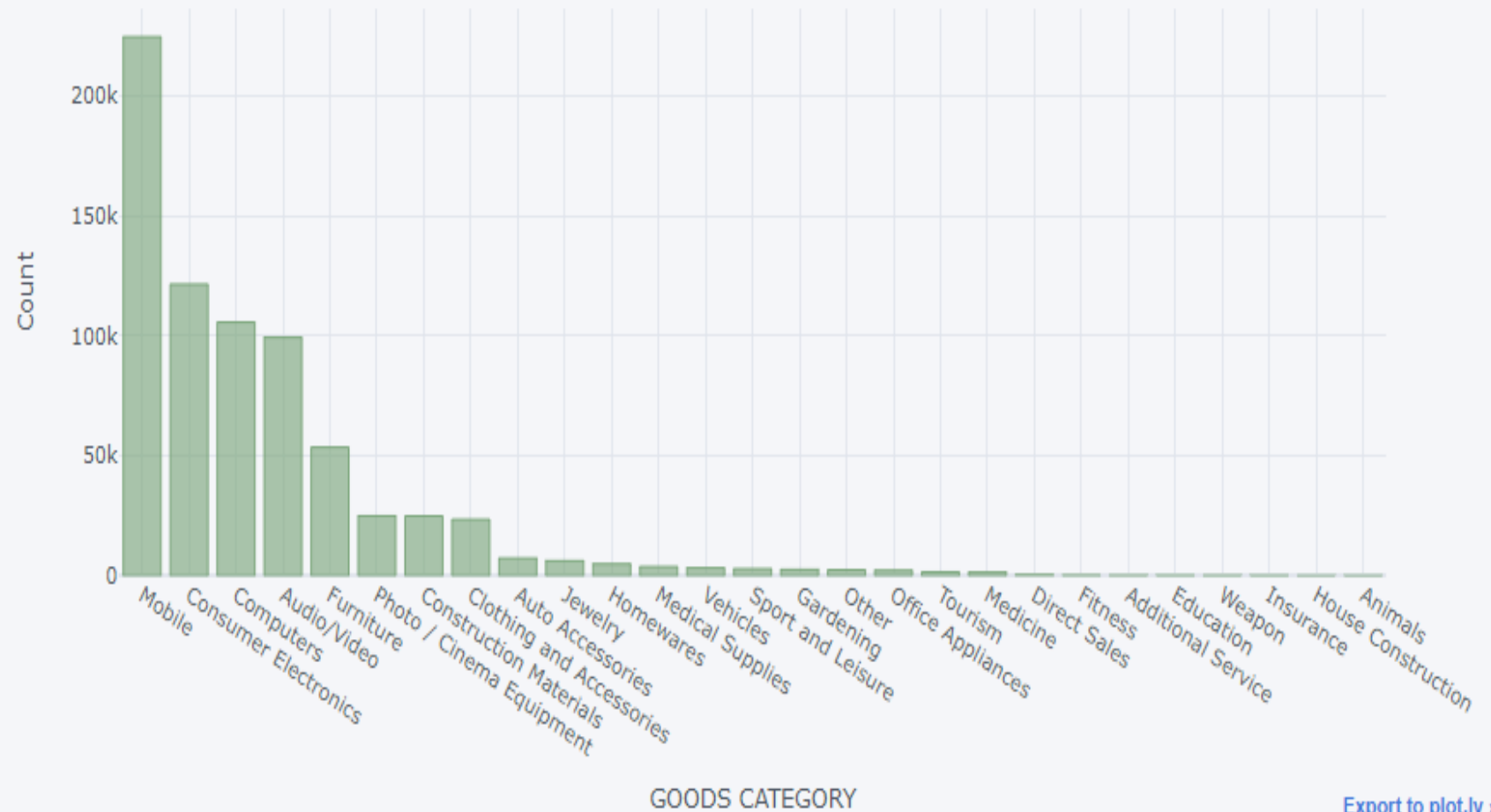


# UNIVARIATE ANALYSIS

# TYPE OF GOODS

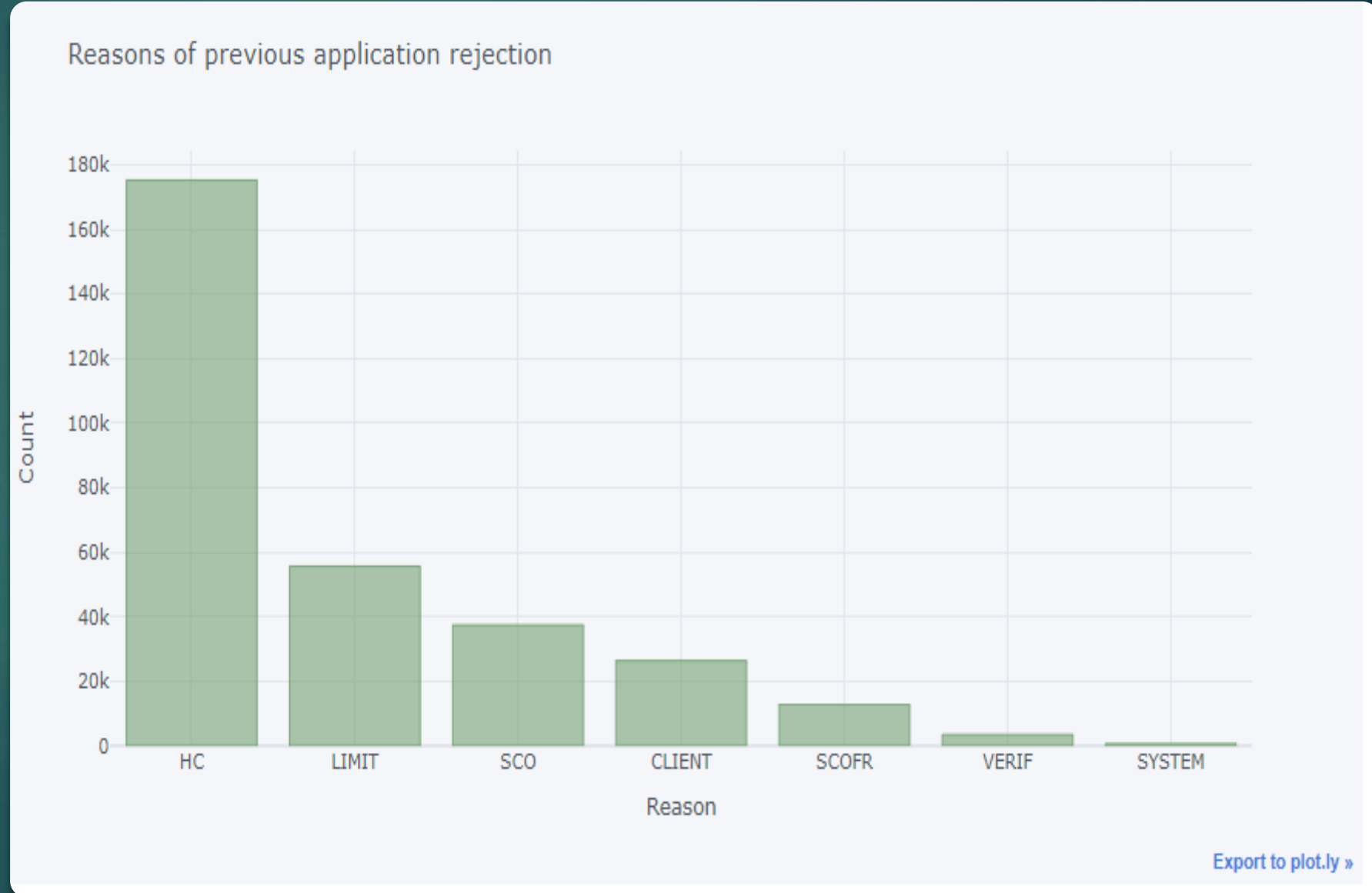
- We observe that the majority of loans are for mobiles, consumer electronics, computers and furnitures.

What kind of goods did the client apply for in the previous application



# REASONS OF PREVIOUS APPLICATION REJECTION

- We observe that HC is the reason majority of applications got rejected.



# BIVARIATE ANALYSIS

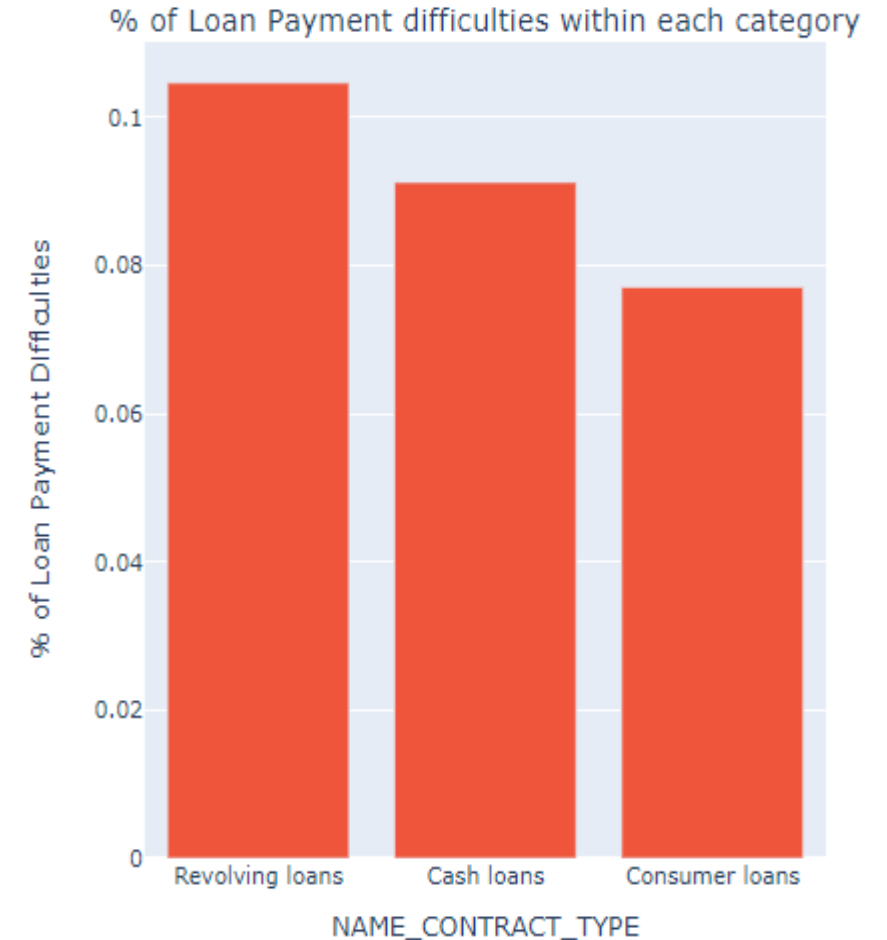
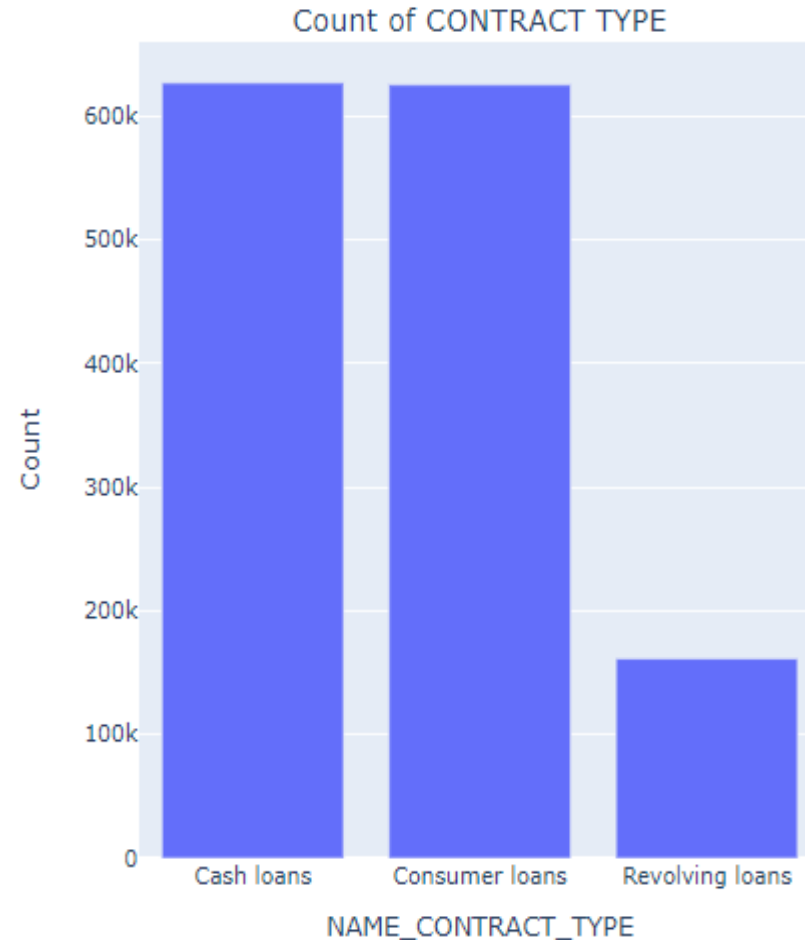
# CONTRACT TYPE

From the first graph it can be seen that most of the contract type from previous application was 'Cash loans'

It can be clearly seen from the second graph that the

- 'Revolving Loans' contracts from previous application are the ones who have maximum % of Loan-Payment Difficulties from current application.
- 'Consumer loans' contracts from previous application are the ones who have minimum % of Loan-Payment Difficulties from current application.

CONTRACT TYPE



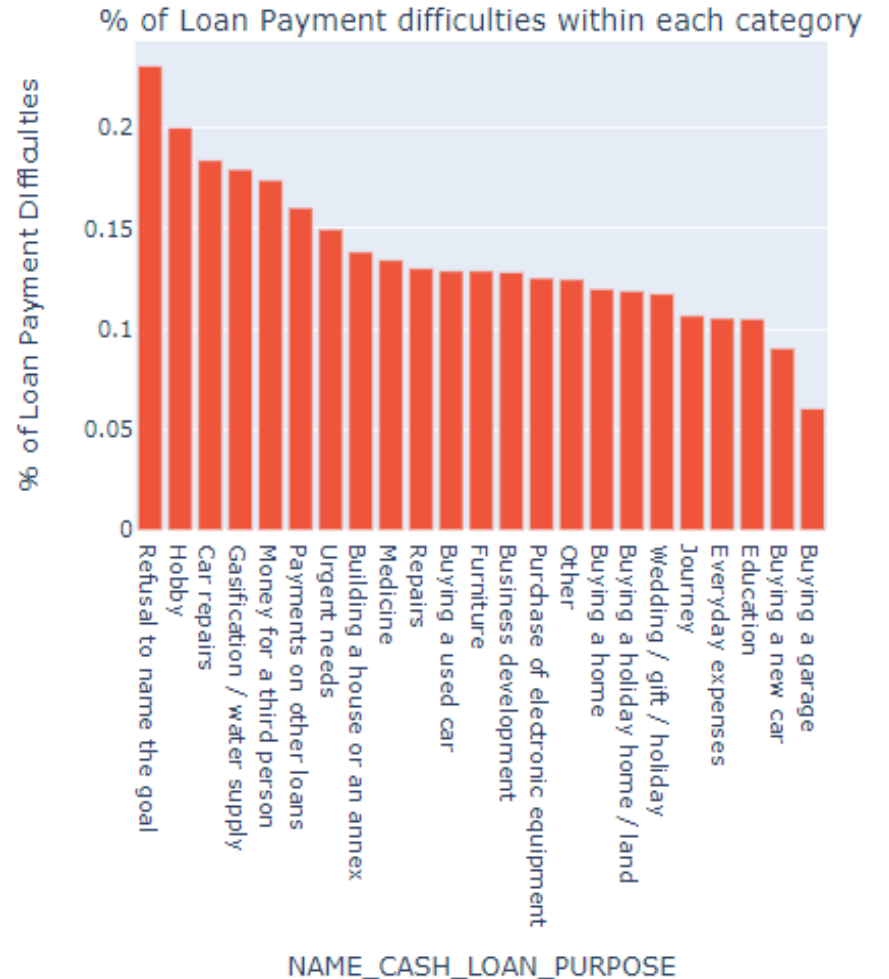
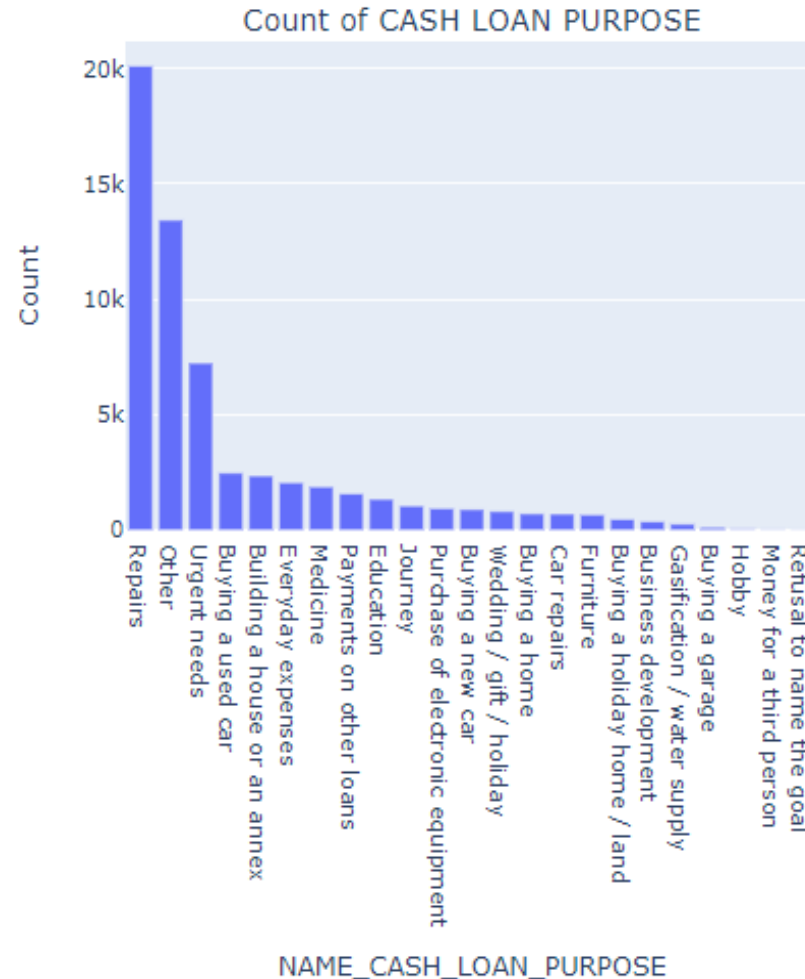
# CASH LOAN PURPOSE

From the first graph it can be seen that purpose of cash loan from previous data was maximum for 'Repairs'

It can be clearly seen from the second graph that the

- 'Refusal to name the goal' for cash loan from previous application are the ones who have maximum % of Loan-Payment Difficulties from current application.

CASH LOAN PURPOSE



# CONCLUSION



# APPLICATION DATA



- The count of Maternity Leave in NAME INCOME TYPE is very less and it also has maximum % of payment difficulties around 40%. Hence, client with income type as Maternity leave are the driving factors for Loan Defaulters
- The count of Low skilled Labourers in OCCUPATION TYPE is comparatively very less and it also has maximum % of payment difficulties around 17% Hence, client with occupation type as Low skilled Labours are the driving factors for Loan Defaulters
- The count of Lower Secondary in NAME EDUCATION TYPE is comparatively very less and it also has maximum % of payment difficulties-around11%. Hence, client with education type as Lower Secondary are the driving factors for Loan Defaulters.

# PREVIOUS APPLICATION DATA

- The count of Refusal to name the goal in NAME CASH LOAN PURPOSE is comparatively very less and it also has maximum % of payment difficulties-around 23%. Hence, clients who have Refused to name the goal for cash loan in previous application are the driving factors for Loan Defaulters.
- The count of 'Refused in NAME CONTRACT\_STATUS' is comparatively less and it also has maximum % of payment difficulties- around 12%. Hence, client with contract status as 'Refused in previous application are the driving factors for Loan Defaulters.
- The count of Revolving Loans in 'NAME CONTRACT TYPE is comparatively very less and it also has maximum % of payment difficulties around 10%. Hence, client with contract type as 'Revolving loans' in previous application are the driving factors for Loan Defaulters
- It can be observed from the graph that Clients with 'Revolving loans' and with 'Refused' previous application tend to have more % of payment difficulties in current application. Since the count of both 'Revolving loans' and 'Refused' is comparatively less(from the graphs in previous slide), clients with Revolving Loans and Refused previous application are driving factors for Loan Defaulters

# FACTORS THAT LOAN CAN BE GIVEN ON CONDITION OF HIGH INTEREST RATE TO MITIGATE ANY DEFAULT RISK LEADING TO BUSINESS LOSS:

- **NAME\_HOUSING TYPE:** High number of loan applications are from the category of people who live in Rented apartments & living with parents and hence offering the loan would mitigate the loss if any of those default.
- **AMT CREDIT.** People who get loan for 3-6 Lakhs tend to default more than others and hence having higher interest specifically for this credit range would be ideal.
- **AMT INCOME:** Since 90% of the applications have Income total less than 3Lakhs and they have high probability of defaulting, they could be offered loan with higher interest compared to other income category.
- **CNT\_CHILDREN & CNT\_FAM\_MEMBERS:** Clients who have 4 to 8 children has a very high default rate and hence higher interest should be imposed on their loans.
- **NAME\_CASH\_LOAN\_PURPOSE:** Loan taken for the purpose of Repairs seems to have highest default rate. A very high number applications have been rejected by bank or refused by client in previous applications as well which has purpose as repair or other. This shows that purpose repair is taken as high risk by bank and either they are rejected, or bank offers very high loan interest rate which is not feasible by the clients, thus they refuse the loan. The same approach could be followed in future as well.

# SUGGESTIONS

- 90% of the previously cancelled client have actually repaid the loan. Record the reason for cancellation which might help the bank to determine and negotiate terms with these repaying customers in future for increase business opportunity.
- 88% of the clients who were refused by bank for loan earlier have now turned into a repaying client. Hence documenting the reason for rejection could mitigate the business loss and these clients could be contacted for further loans.

**THANK YOU**