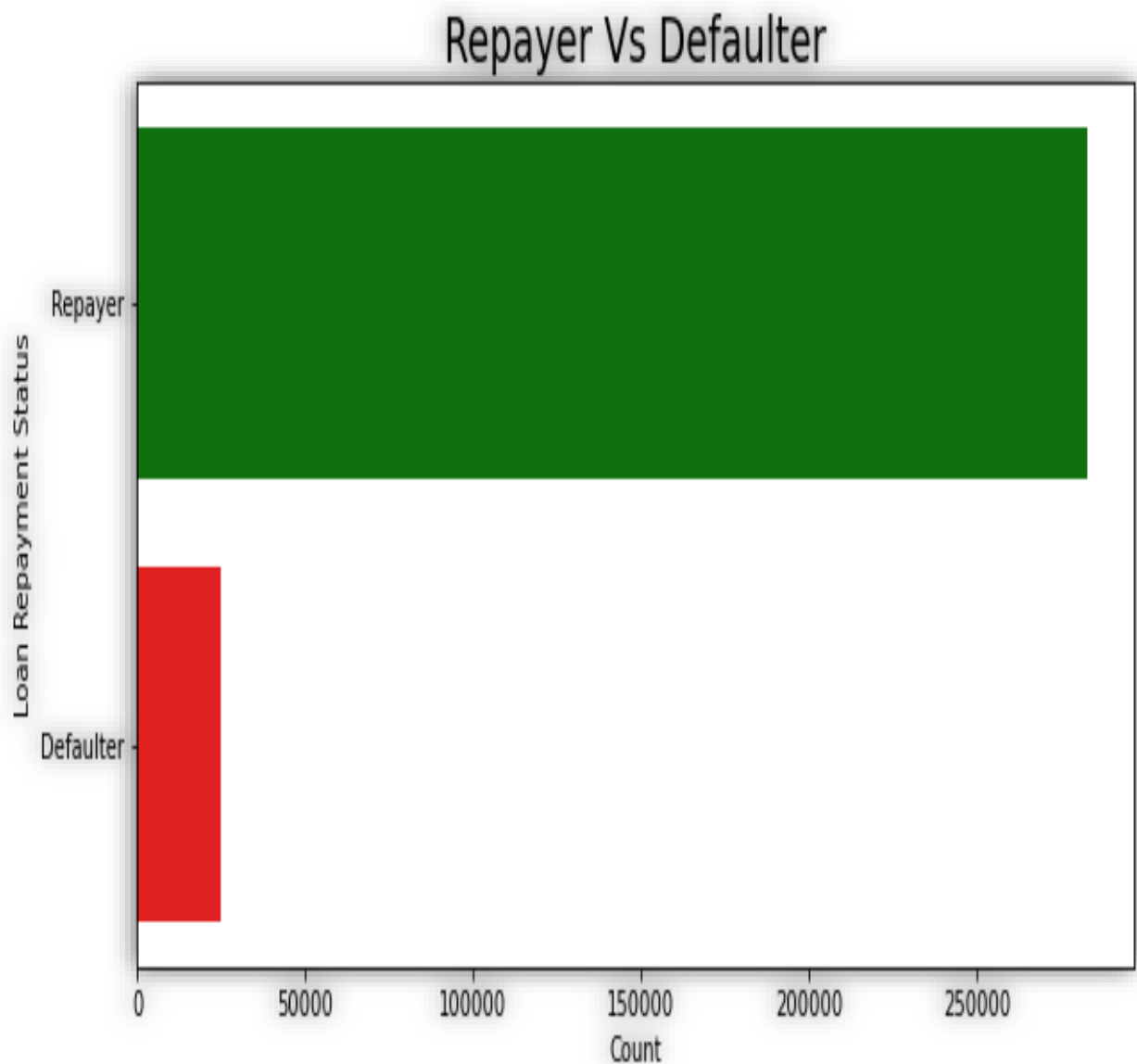# *Credit EDA Case Study*

BY-

➢ BHARATH MD

➢ RISHABH SHUKLA

Data Analysis on Dataset 1: Application Data
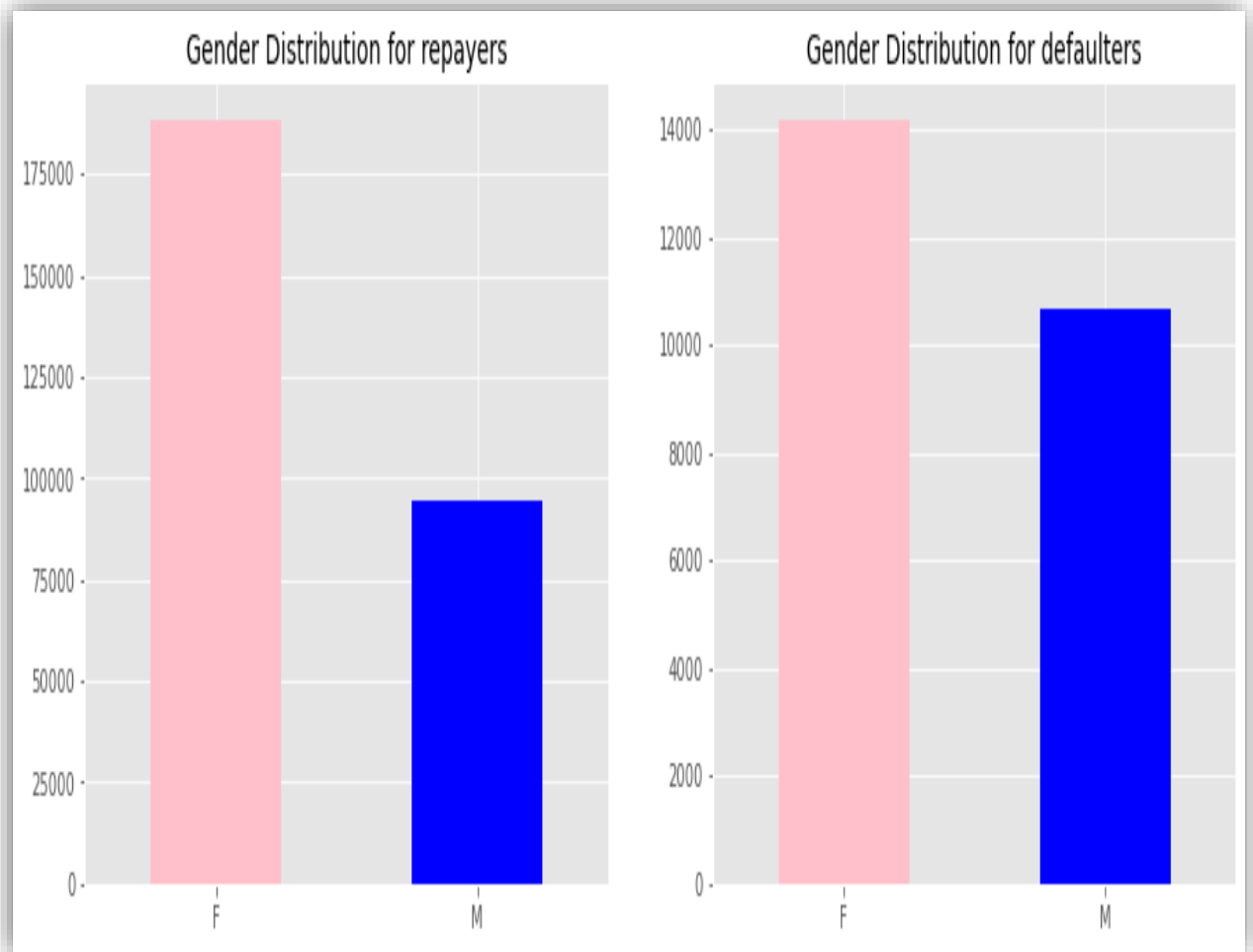
*Repayer Vs Defaulter

- Repayers > Defaulters

UNIVARIATE ANALYSIS

LET'S MAKE VISUALIZATIONS FOR THE FOLLOWING COLUMNS :

- GENDER
- CONTRACT TYPE
- INCOME TYPE
- FAMILY STATUS
- EDUCATION TYPE
- HOUSING TYPE
- TYPE SUITE
- ORGANIZATION TYPE
- OCUPATION TYPE
- INCOME RANGE
- AGE GROUP
- CREDIT RANGE TYPE
- GOODS PRICE RANGE TYPE
- CHILDREN COUNT

# Gender Distribution For Repayers & Defaulters

- FEMALES (F) are the majority in both cases of being repayers and defaulters.
- There is an increase in payment defaulters for MALE (M).

Contract Type for Repayers & Defaulters
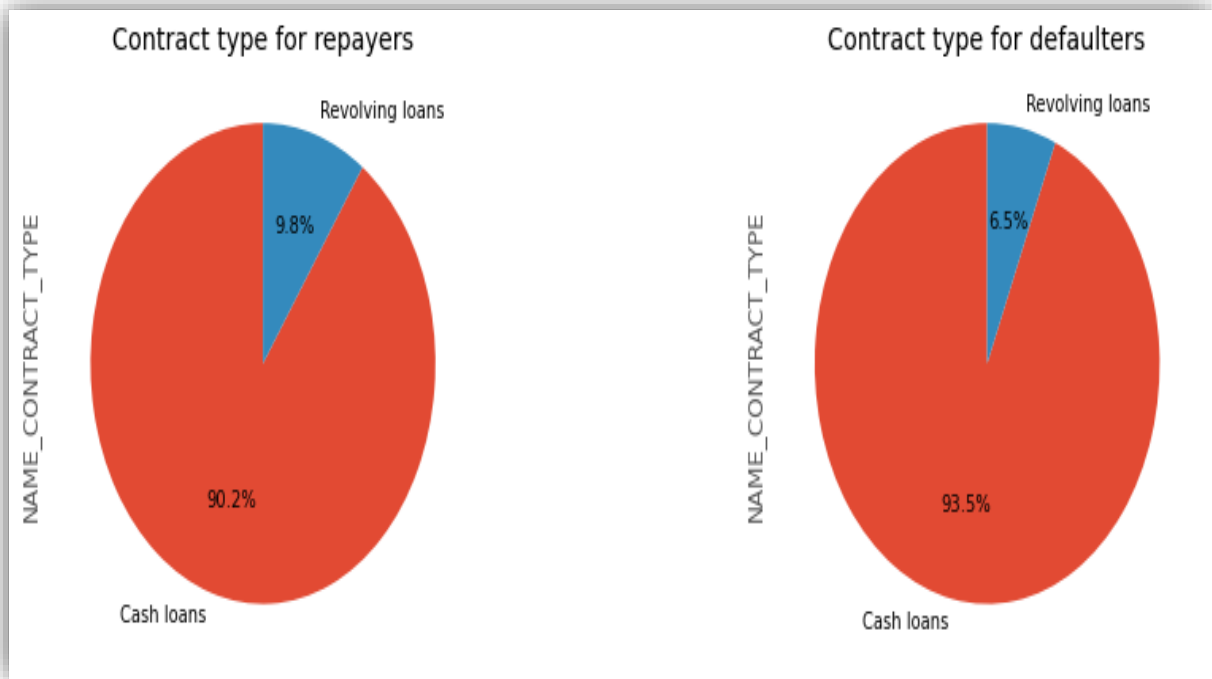
- Majority prefer getting a cash loan
- Cash loan defaulters are slightly more than the cash loan repayers
- Revolving loans defaulters are less than revolving loans repayers .

Income Type For Repayers & Defaulters
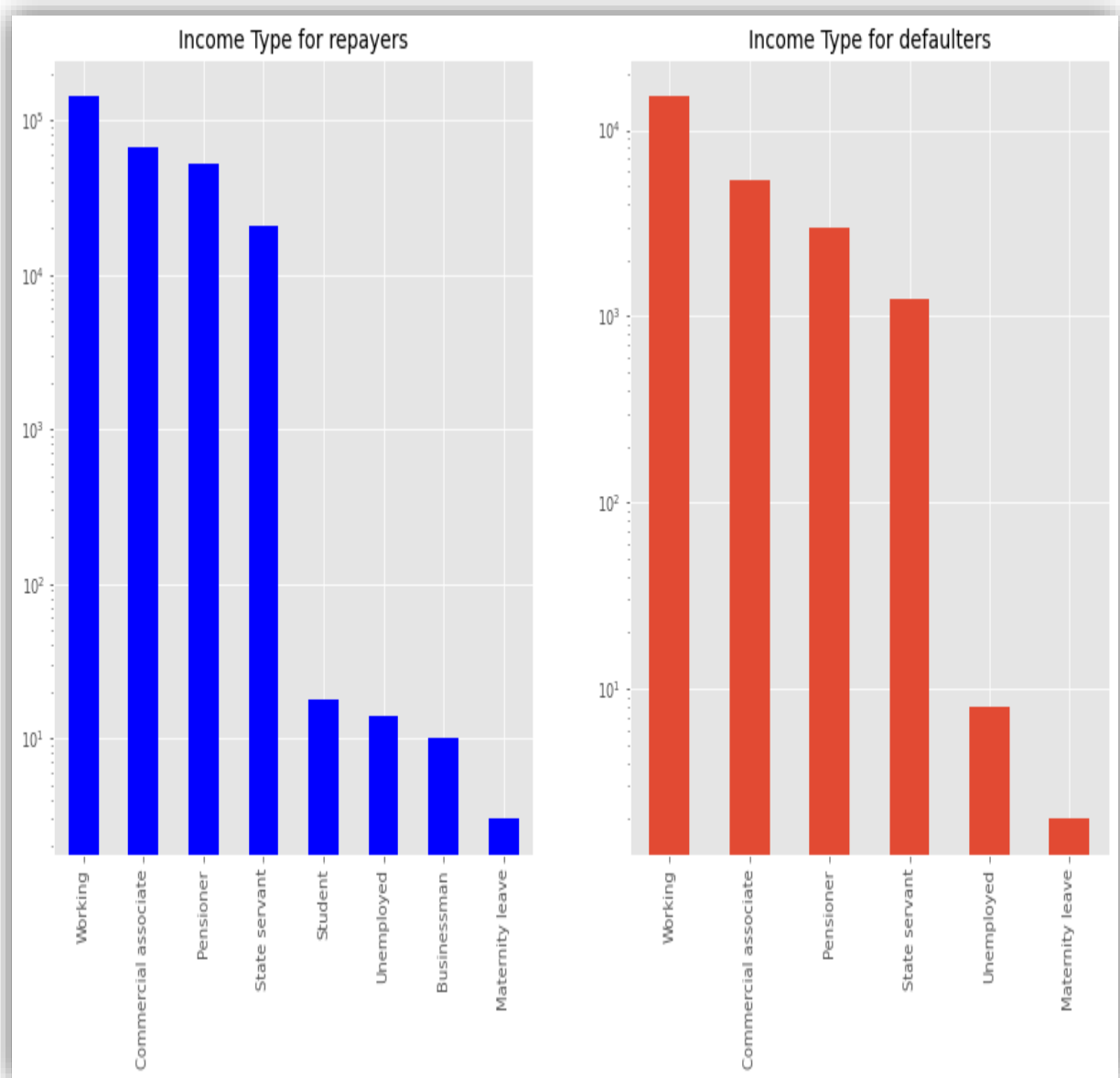
- working class people show an increase in percentage defaulting.
- pensioners show decrease in defaulting.
- Businessmen have a 100% track record of not defaulting.

Family Status For Repayers & Defaulters
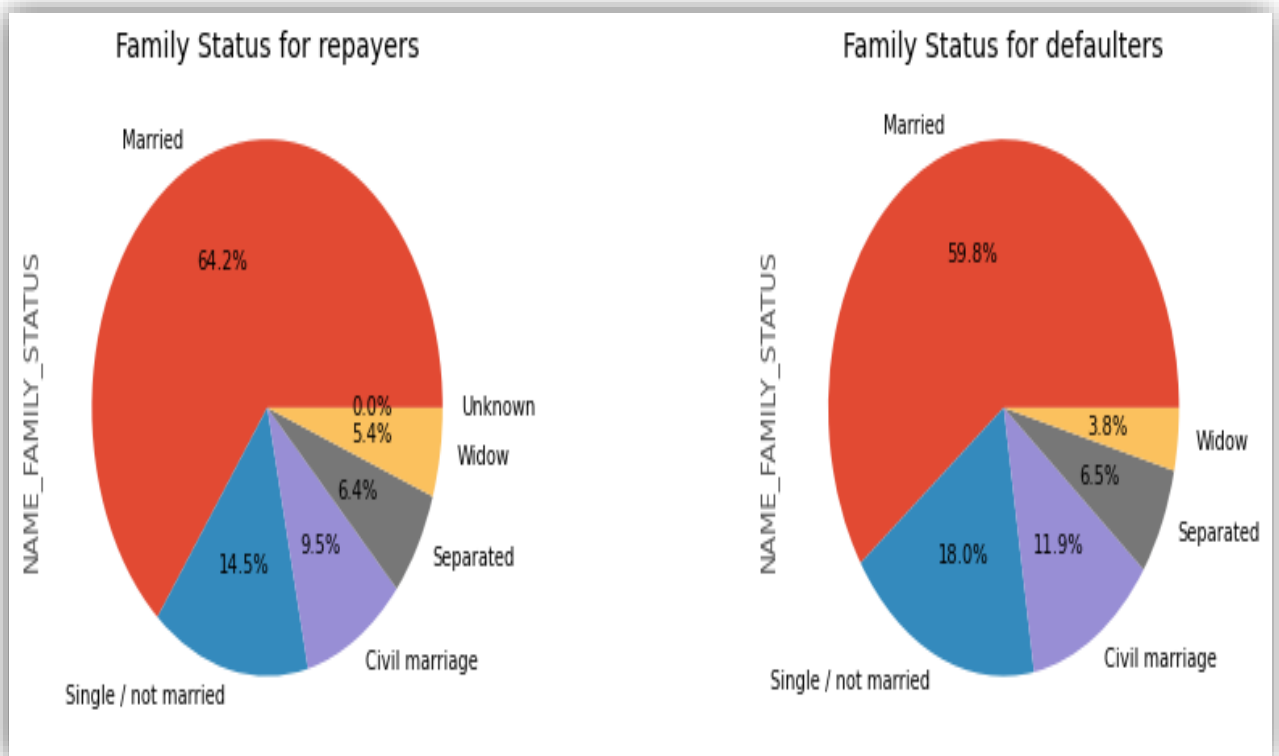
- Civil marriage category , who tend to default more increase in percentage is seen in single and
- Decrease in percentage is seen in married and widow category , fewer chances of them defaulting.

Education Type For Repayers & Defaulters

Education type for loan of the applicants
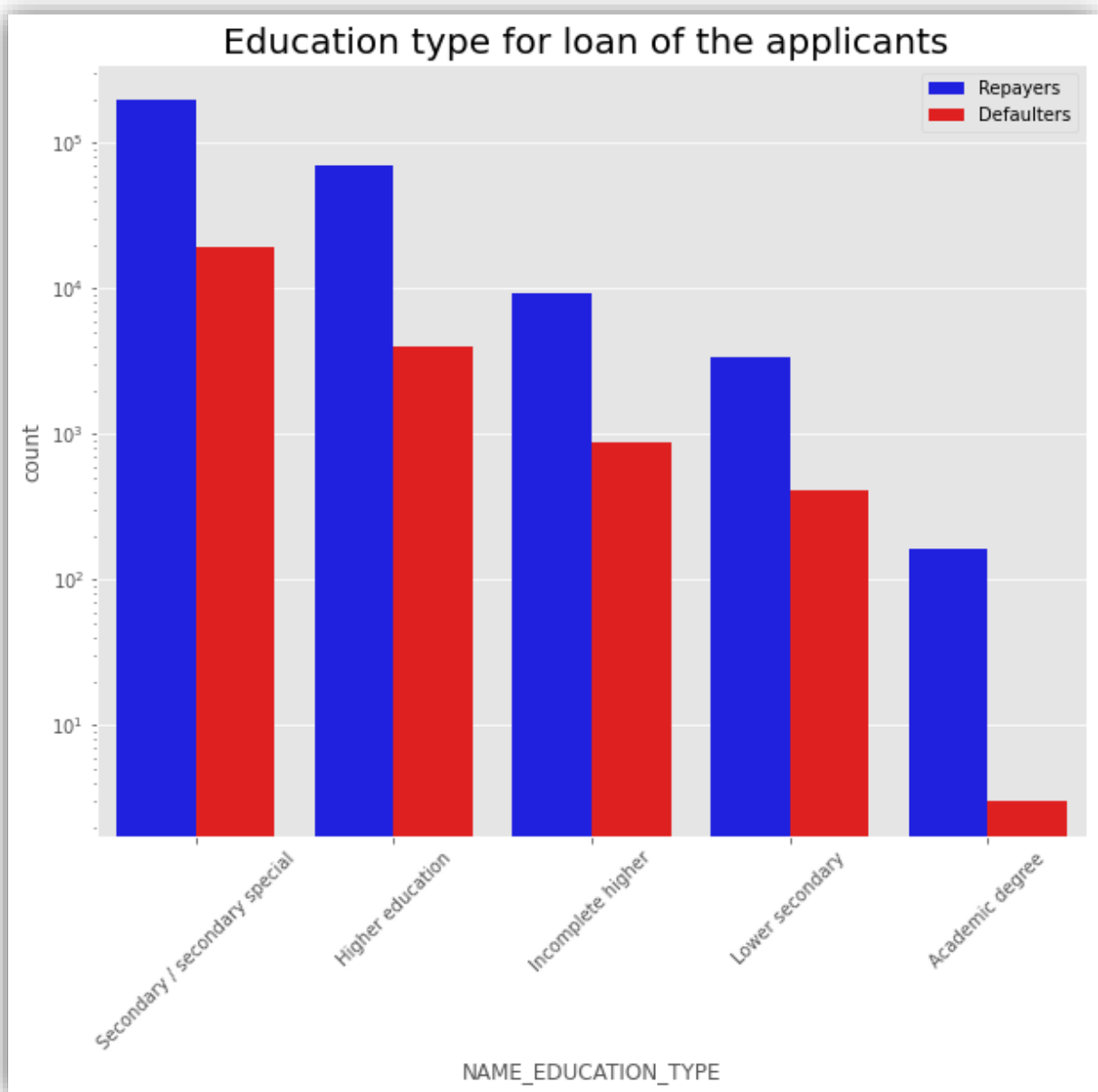
- Majority of applicants pursuing secondary education
- People with academic degree tend to default the least

Housing Type For Repayers & Defaulters

- Slight increase in percentage is seen with people defaulting who live with their parents
- Slight increase in percentage is seen with people defaulting who live in rented apartment

Suite Type For Repayers & Defaulters
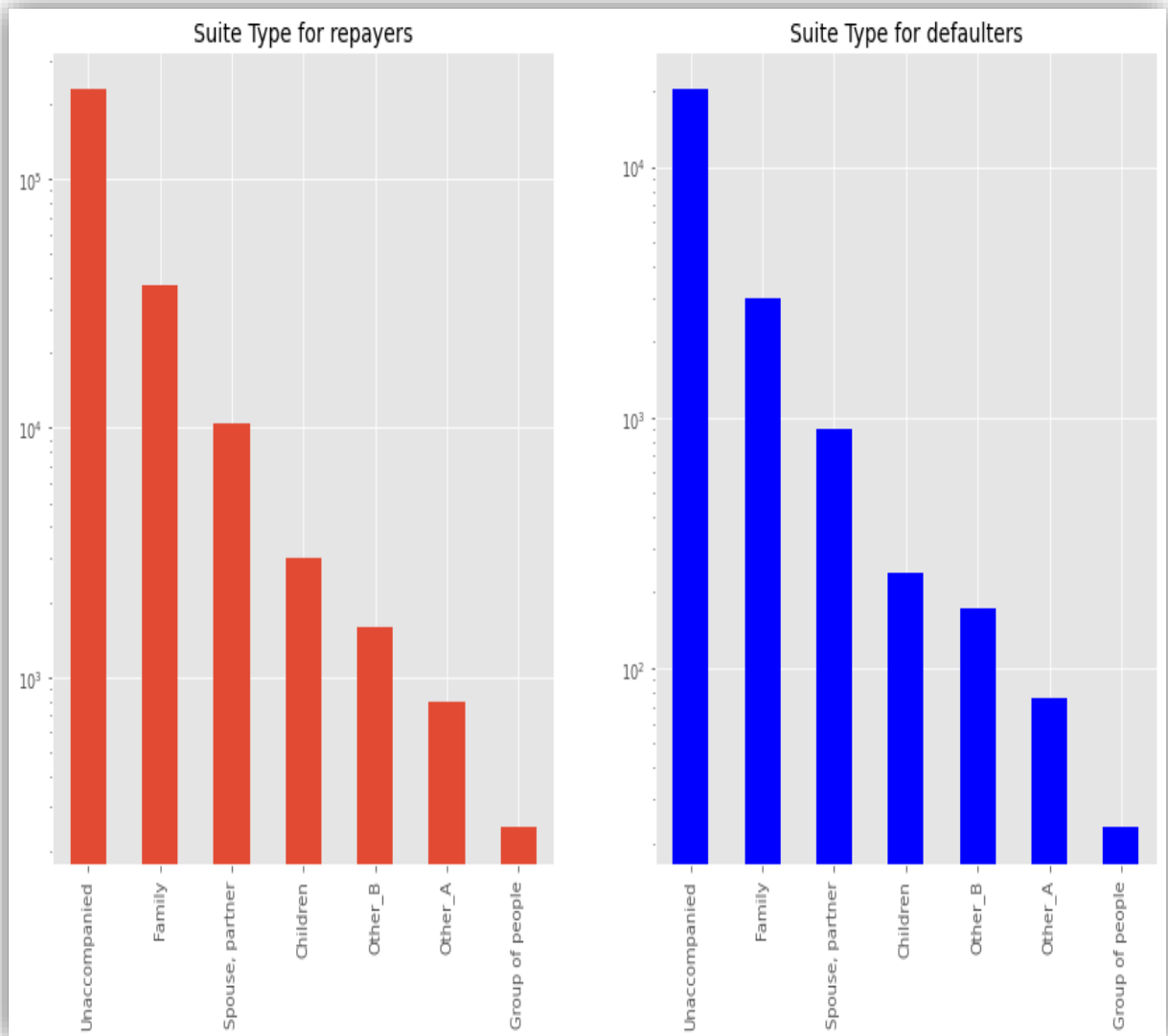
- Much difference is not seen here.

Organizational Type For Repayers & Defaulters
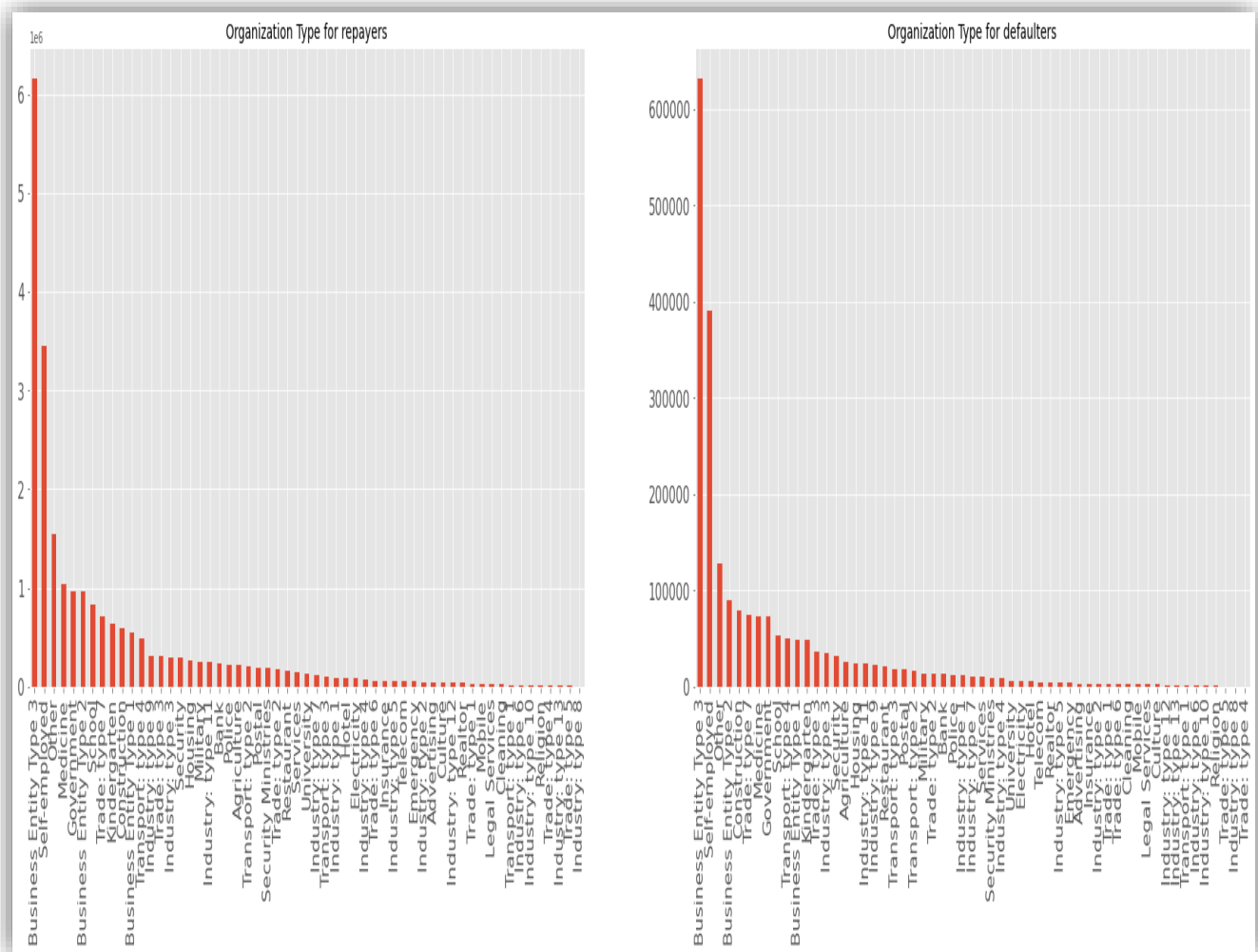
- As observed earlier business entity type 3 is the major customer and not much difference is observed between the repayer and defaulter .

Occupation Type For Repayers & Defaulters
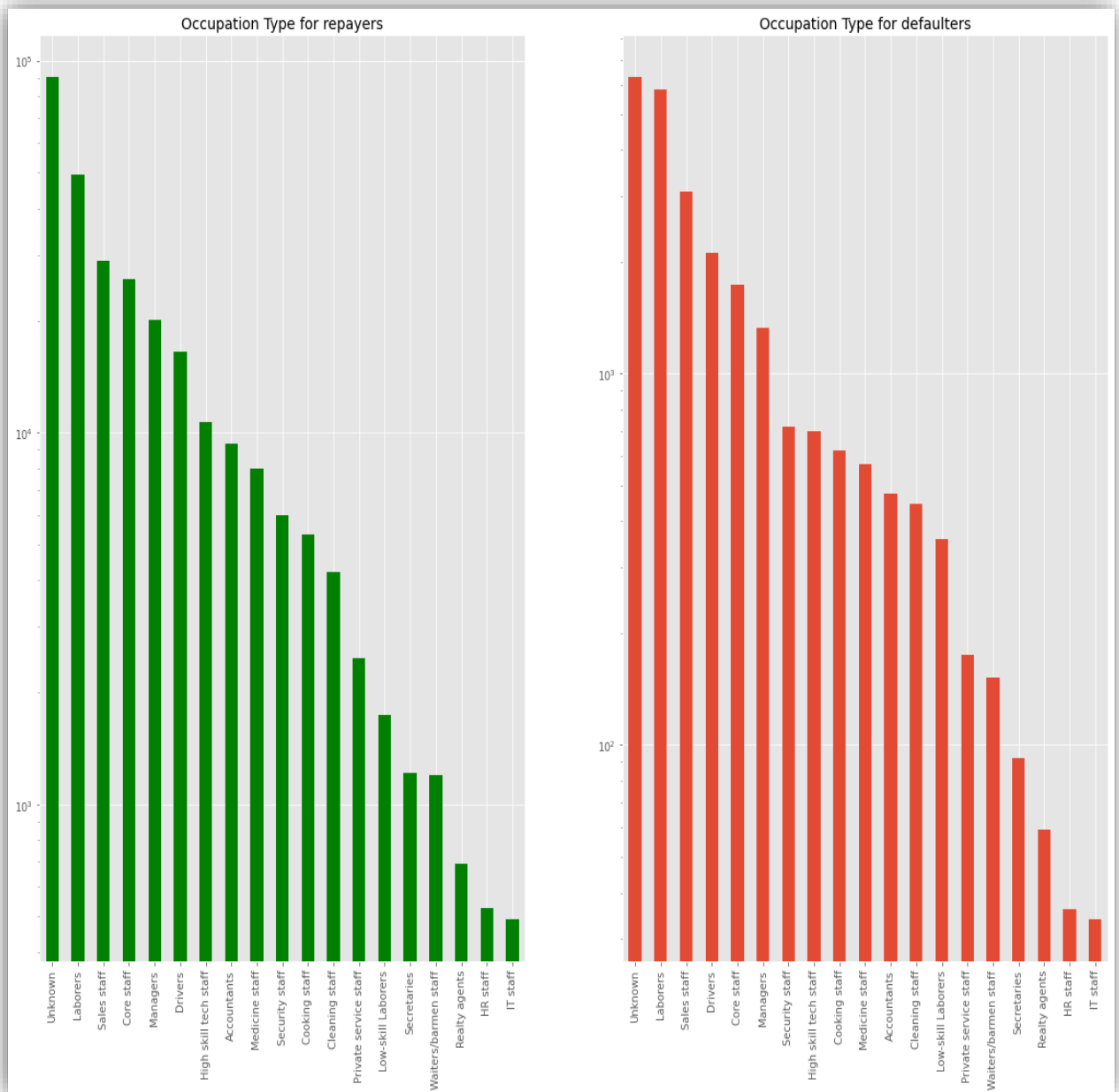
- There seems to be an increase in defaulting percentage in Laborers, Drivers

- Managers seem to repay and not default comparatively

Income Range Type For Repayers & Defaulters
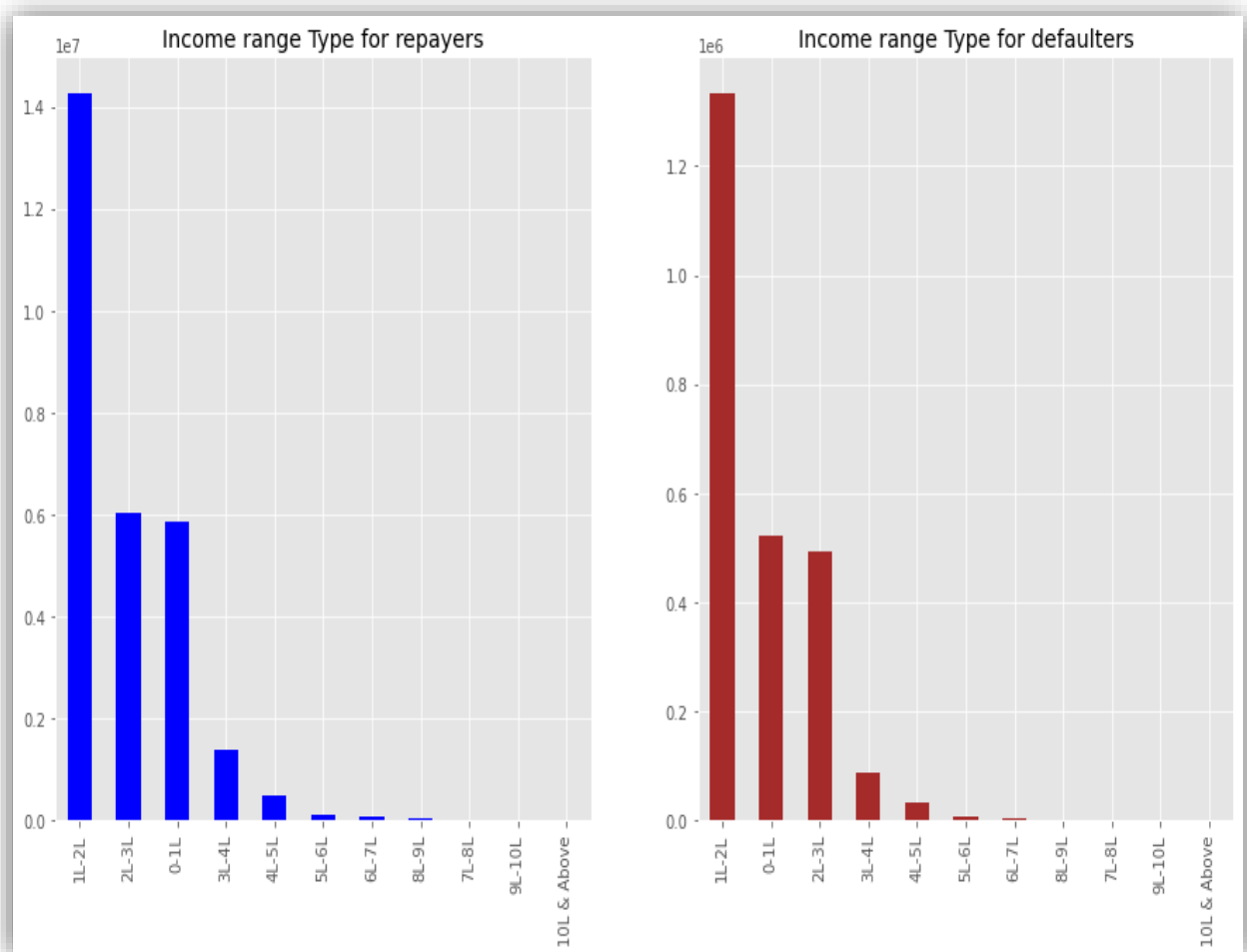
- Not much difference is observed , slight increase in percentage for defaulting with low income range which is less than 3L

Age Group For Repayers & Defaulters
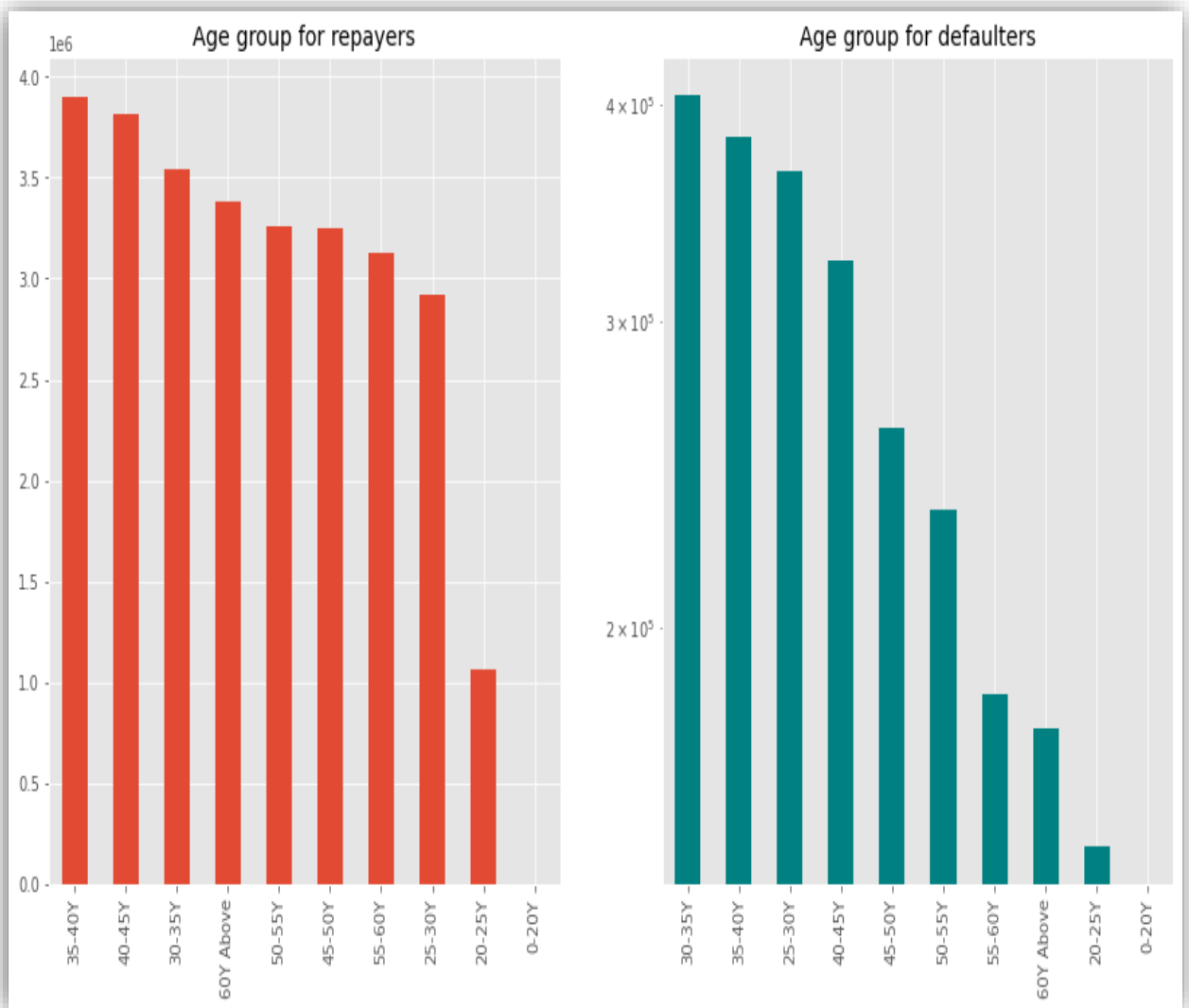
- Age 30-50Y seem to default more
- Age 20-25Y seem to default less

Credit Range For Repayers & Defaulters
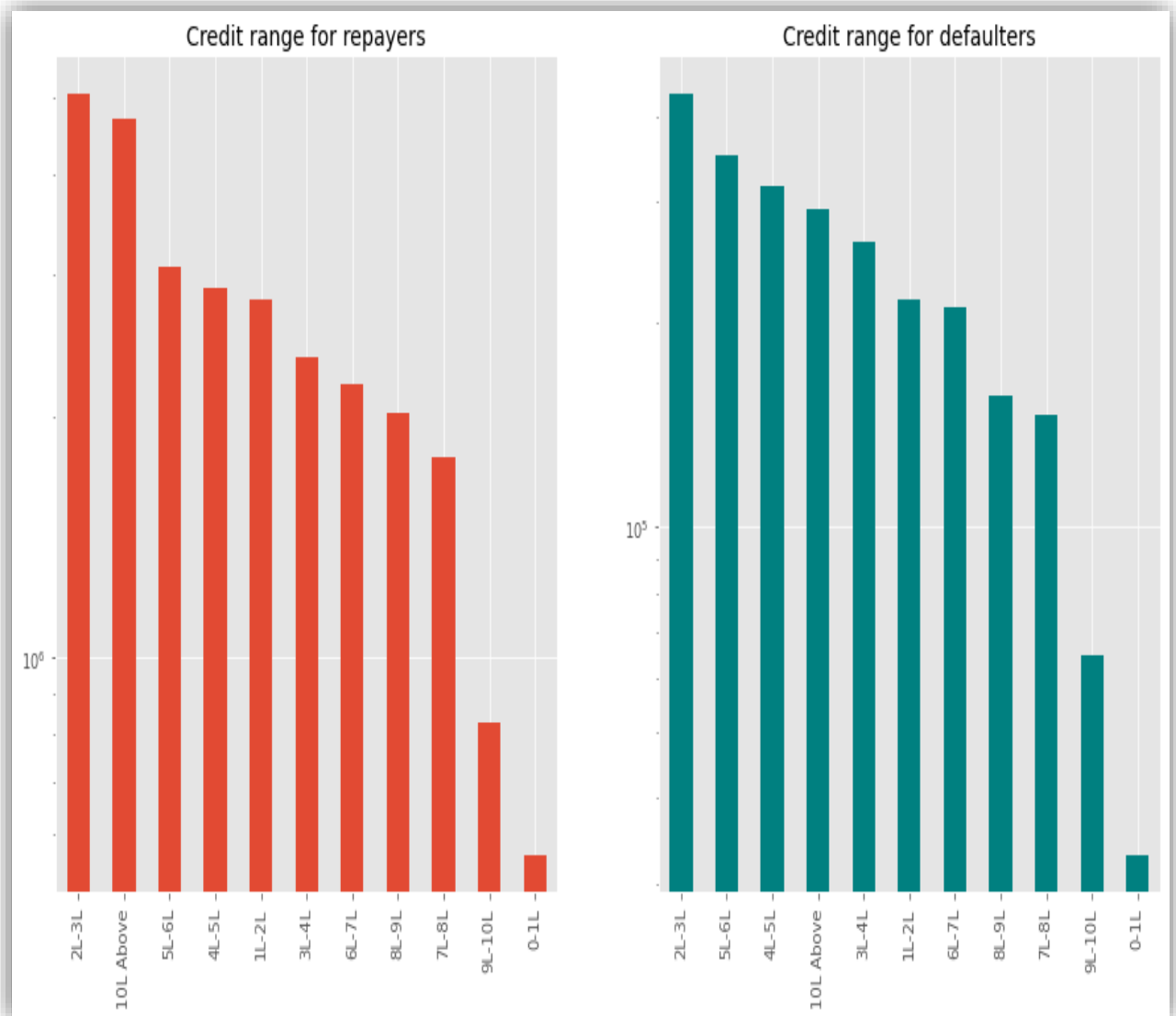
- More defaulters in the range 3-6L

Goods Price Range For Repayers & Defaulters
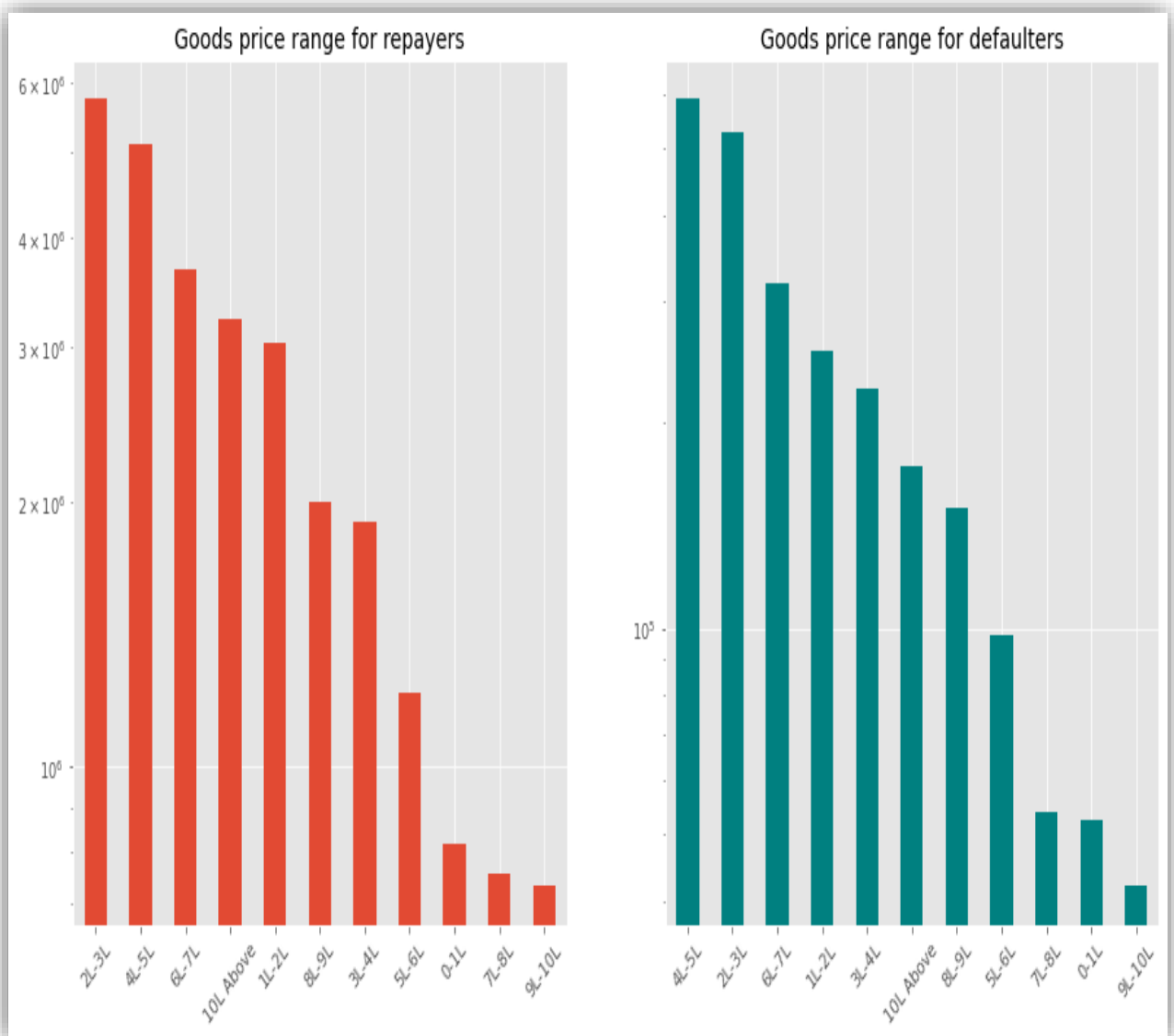
- Applicants whose goods price range is 4-5L tend to default more
- Highest and the lowest price range tend to do well in repaying and not defaulting

Children Counts For Repayers & Defaulters
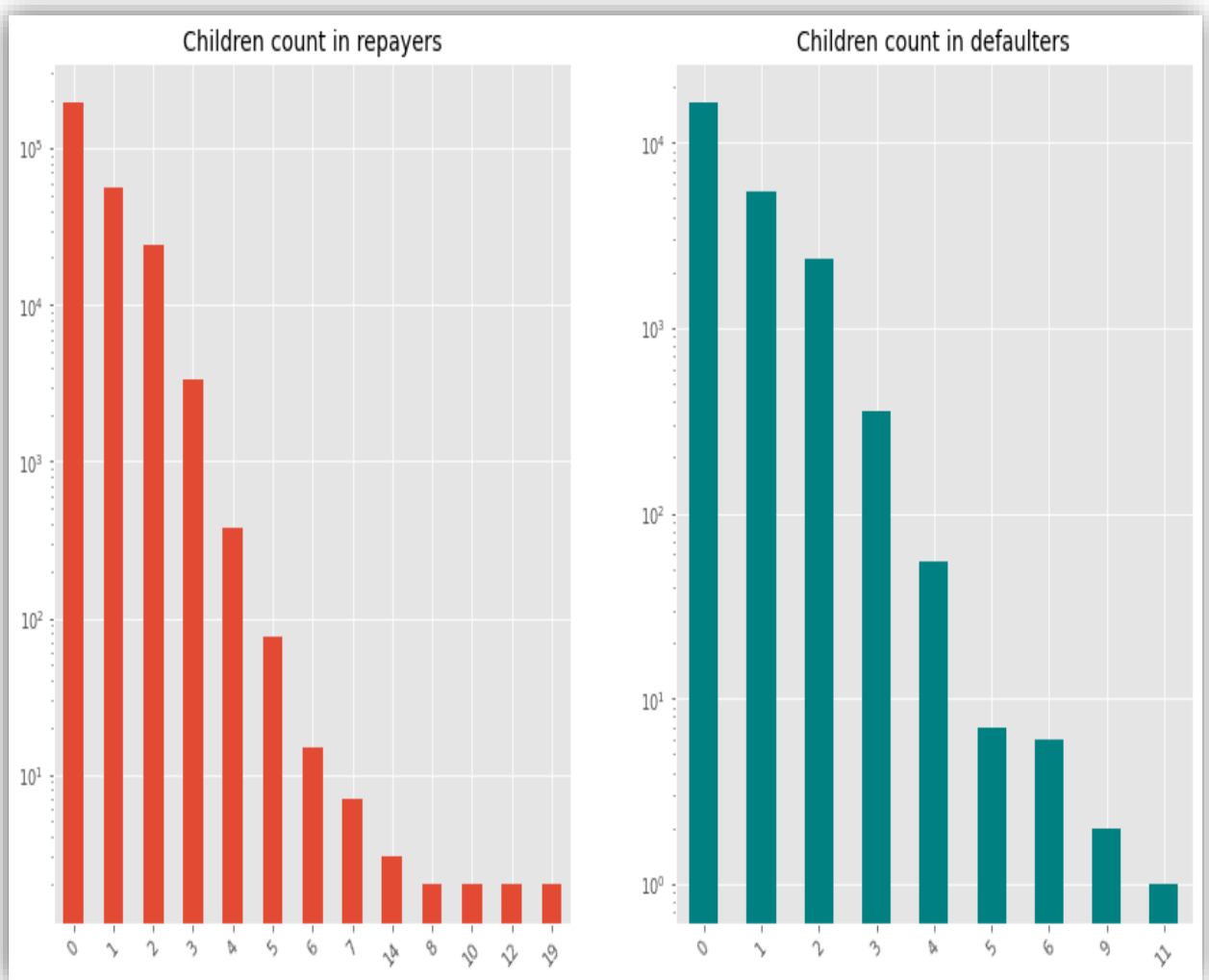
- Not much difference but seeing the values replaced in defaulters , children count with 9 and 11 tend to default

BIVARIATE ANALYSIS

LET'S MAKE VISUALIZATIONS FOR THE
FOLLOWING COLUMNS :

- To check how the income range affects the applicant with respect to credit amount

- To check education type vs credit amount

- To check family status vs credit amount

The income range affects the applicant with respect to credit amount
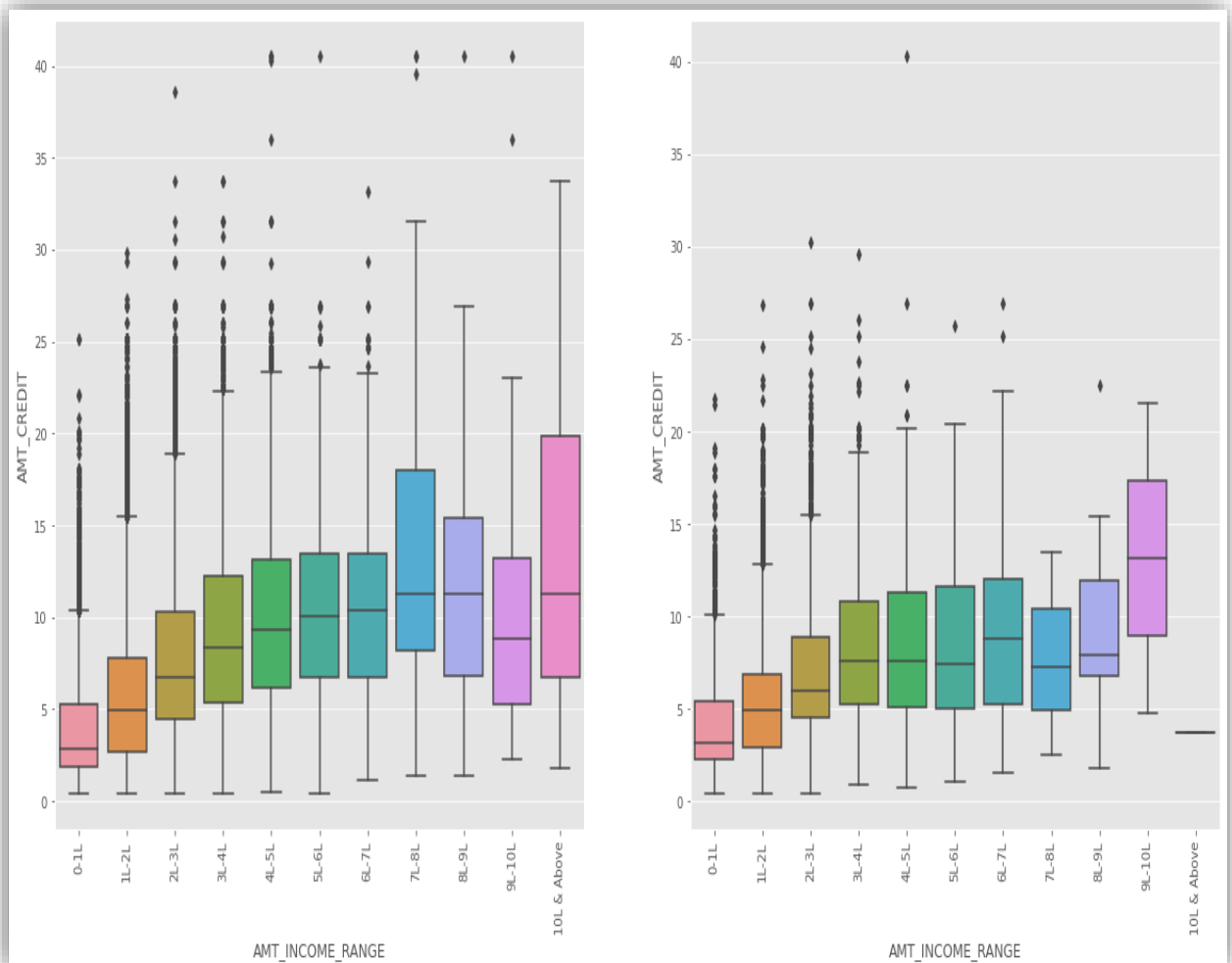
- Applicants earning more than 10L have no difficulty in repaying , mostly businessmen as seen above

- 9-10L earning range is seen to find it difficult to repay because of the applicants as outliers

Education type vs credit amount
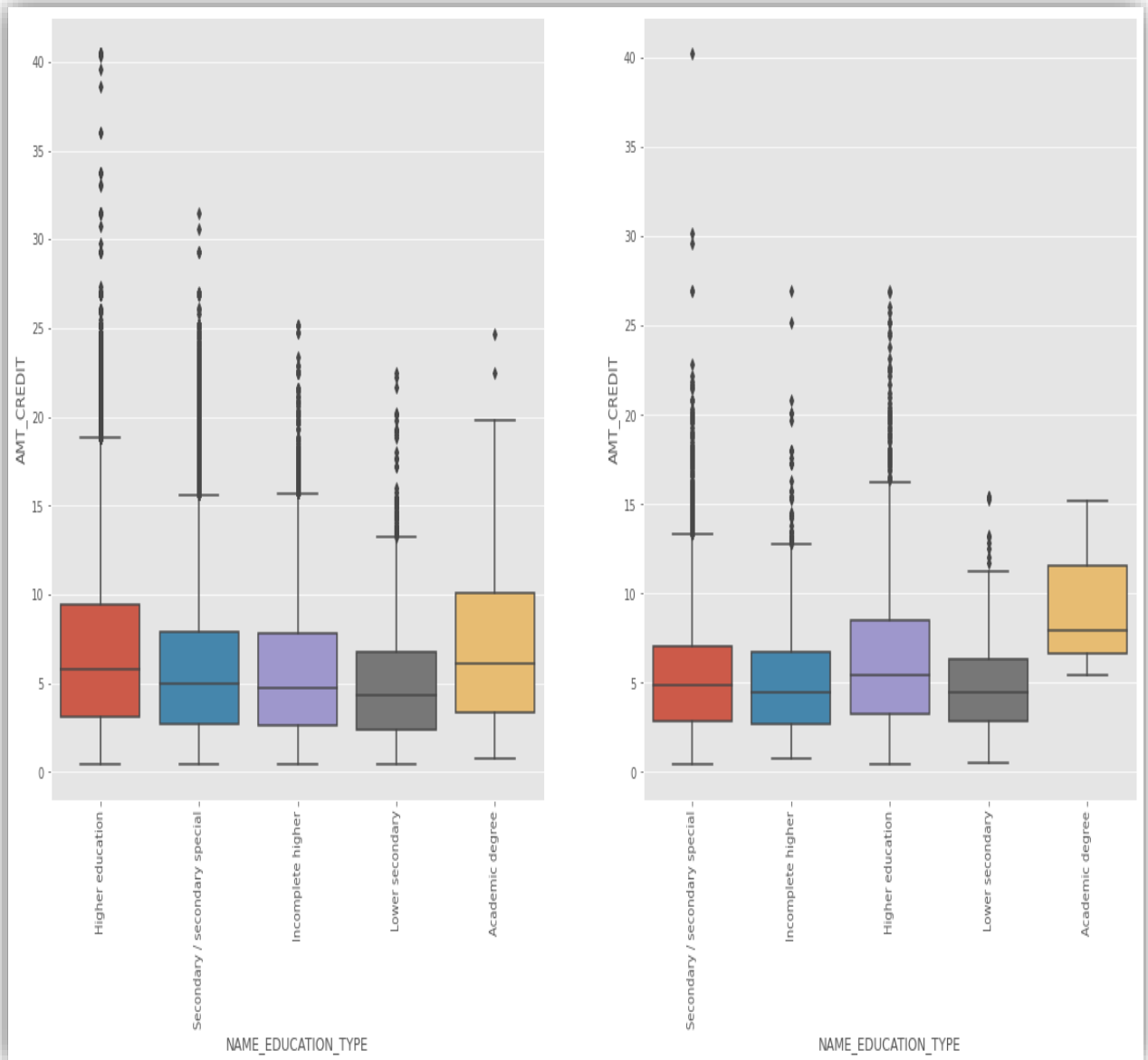
- Applicants doing higher education have more chances of not defaulting , they can be the target audience

- Applicants with secondary / incomplete higher education tend to default more

Family status vs credit amount

- Applicants with marital status =single have higher percentage of not defaulting

- Married and widowed applicants find it difficult to repay the loan amount

Correlation

Heatmap to check correlation in repayers

	AMT_INCOME_TOTAL	AMT_CREDIT	AMT_ANNUITY	AMT_GOODS_PRICE	DAYS_BIRTH	DAYS_EMPLOYED	DAYS_REGISTRATION	DAYS_ID_PUBLISH	AMT_REQ_CREDIT_BUREAU_HOUR	AMT_REQ_CREDIT_BUREAU_DAY	AMT_REQ_CREDIT_BUREAU_WEEK	AMT_REQ_CREDIT_BUREAU_MON
AMT_INCOME_TOTAL	1	0.34	0.42	0.35	-0.063	-0.14	-0.065	-0.023	0.0027	0.008	0.0086	0.059
AMT_CREDIT	0.34	1	0.77	0.99	0.047	-0.07	-0.013	0.0015	-0.0023	0.0051	0.00094	0.055
AMT_ANNUITY	0.42	0.77	1	0.78	-0.012	-0.1	-0.039	-0.014	0.0032	0.0025	0.012	0.036
AMT_GOODS_PRICE	0.35	0.99	0.78	1	0.045	-0.068	-0.016	0.0038	-0.0017	0.0056	0.0013	0.057
DAYS_BIRTH	-0.063	0.047	-0.012	0.045	1	0.63	0.33	0.27	-0.0029	-0.0016	0.0036	0.0019
DAYS_EMPLOYED	-0.14	-0.07	-0.1	-0.068	0.63	1	0.21	0.28	-0.0043	-0.00093	0.0017	-0.033
DAYS_REGISTRATION	-0.065	-0.013	-0.039	-0.016	0.33	0.21	1	0.1	0.0025	9.3e-06	0.0013	0.011
DAYS_ID_PUBLISH	-0.023	0.0015	-0.014	0.0038	0.27	0.28	0.1	1	-0.0019	0.0022	0.0069	0.017
AMT_REQ_CREDIT_BUREAU_HOUR	0.0027	-0.0023	0.0032	-0.0017	-0.0029	-0.0043	0.0025	-0.0019	1	0.23	0.0062	0.0034
AMT_REQ_CREDIT_BUREAU_DAY	0.008	0.0051	0.0025	0.0056	-0.0016	-0.00093	9.3e-06	0.0022	0.23	1	0.22	-0.0024
AMT_REQ_CREDIT_BUREAU_WEEK	0.0086	0.00094	0.012	0.0013	0.0036	0.0017	0.0013	0.0069	0.0062	0.22	1	-0.0078
AMT_REQ_CREDIT_BUREAU_MON	0.059	0.055	0.036	0.057	0.0019	-0.033	0.011	0.017	0.0034	-0.0024	-0.0078	1

Heatmap to check correlation in defaulters

|	AMT_INCOME_TOTAL	AMT_CREDIT	AMT_ANNUITY	AMT_GOODS_PRICE	DAYS_BIRTH	DAYS_EMPLOYED	DAYS_REGISTRATION	DAYS_ID_PUBLISH	AMT_REQ_CREDIT_BUREAU_HOUR	AMT_REQ_CREDIT_BUREAU_DAY	AMT_REQ_CREDIT_BUREAU_WEEK	AMT_REQ_CREDIT_BUREAU_MON
AMT_INCOME_TOTAL	1	0.038	0.046	0.038	-0.0031	-0.015	-0.00016	0.0042	0.0011	0.00014	0.00094	0.0057
AMT_CREDIT	0.038	1	0.75	0.98	0.14	0.0019	0.026	0.052	-0.0038	0.0043	0.011	0.056
AMT_ANNUITY	0.046	0.75	1	0.75	0.014	-0.081	-0.034	0.017	0.013	7.4e-05	0.029	0.049
AMT_GOODS_PRICE	0.038	0.98	0.75	1	0.14	0.0067	0.026	0.056	-0.0024	0.0055	0.011	0.059
DAYS_BIRTH	-0.0031	0.14	0.014	0.14	1	0.58	0.29	0.25	-0.012	0.008	0.0082	0.011
DAYS_EMPLOYED	-0.015	0.0019	-0.081	0.0067	0.58	1	0.19	0.23	-0.0073	0.019	0.013	-0.023
DAYS_REGISTRATION	-0.00016	0.026	-0.034	0.026	0.29	0.19	1	0.097	0.0048	0.0079	-0.0023	-0.0034
DAYS_ID_PUBLISH	0.0042	0.052	0.017	0.056	0.25	0.23	0.097	1	-0.00048	0.014	0.0037	0.024
AMT_REQ_CREDIT_BUREAU_HOUR	0.0011	-0.0038	0.013	-0.0024	-0.012	-0.0073	0.0048	-0.00048	1	0.25	0.0085	-0.0037
AMT_REQ_CREDIT_BUREAU_DAY	0.00014	0.0043	7.4e-05	0.0055	0.008	0.019	0.0079	0.014	0.25	1	0.19	-0.0086
AMT_REQ_CREDIT_BUREAU_WEEK	0.00094	0.011	0.029	0.011	0.0082	0.013	-0.0023	0.0037	0.0085	0.19	1	-0.0029
AMT_REQ_CREDIT_BUREAU_MON	0.0057	0.056	0.049	0.059	0.011	-0.023	-0.0034	0.024	-0.0037	-0.0086	-0.0029	1

Inferences through heatmaps

- There is a high correlation between good price and amount credit

- There is a drastic drop in correlation between total income and goods price in defaulters compared to repayers

- There is a drastic drop in correlation between total income and amount credit in defaulters compared to repayers

- Very slight drop in correlation between amount credit and annuity in defaulters compared to repayers

Merging the two datasets

- Contract status of applicants

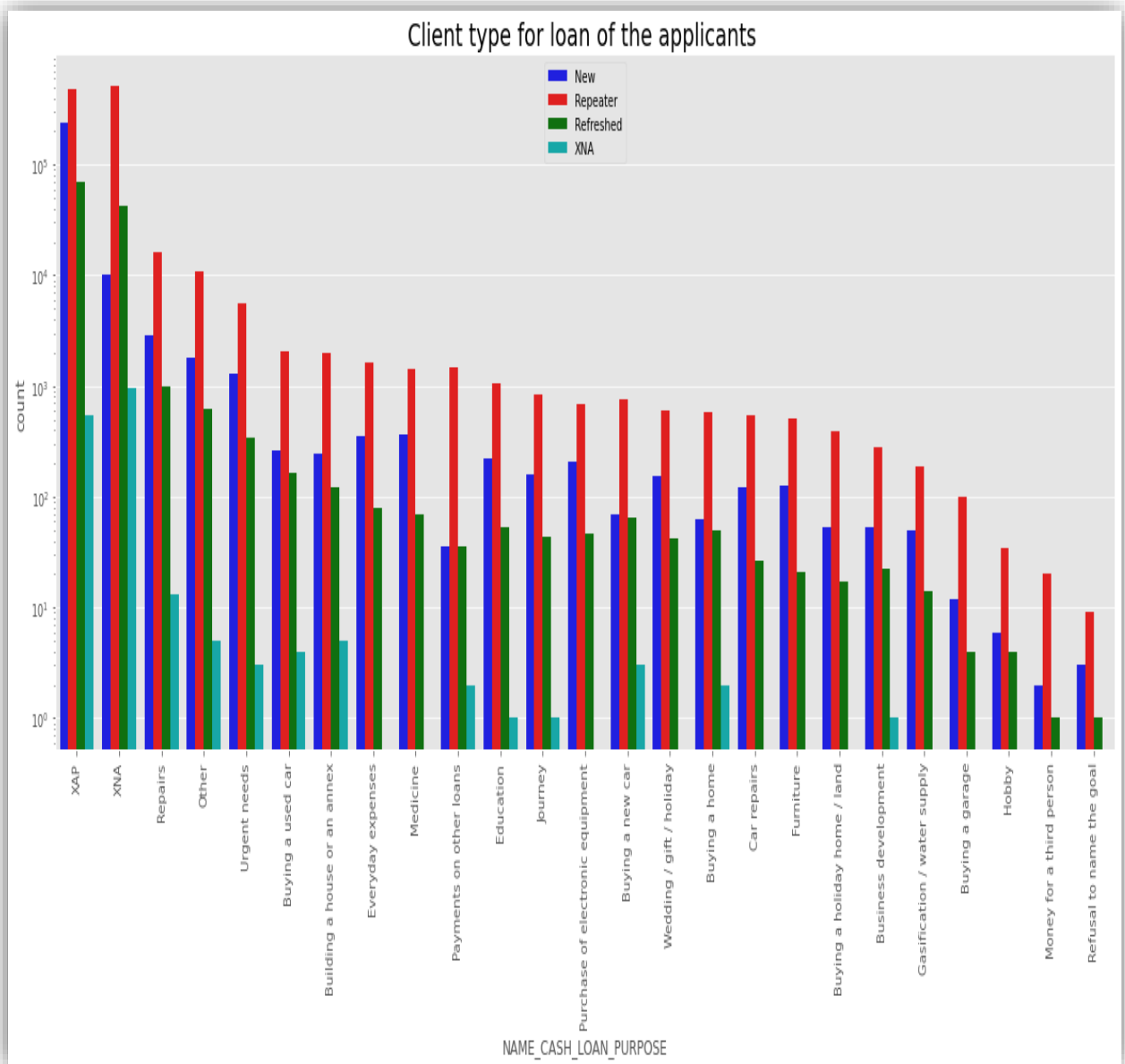Contract status for loan of the applicants
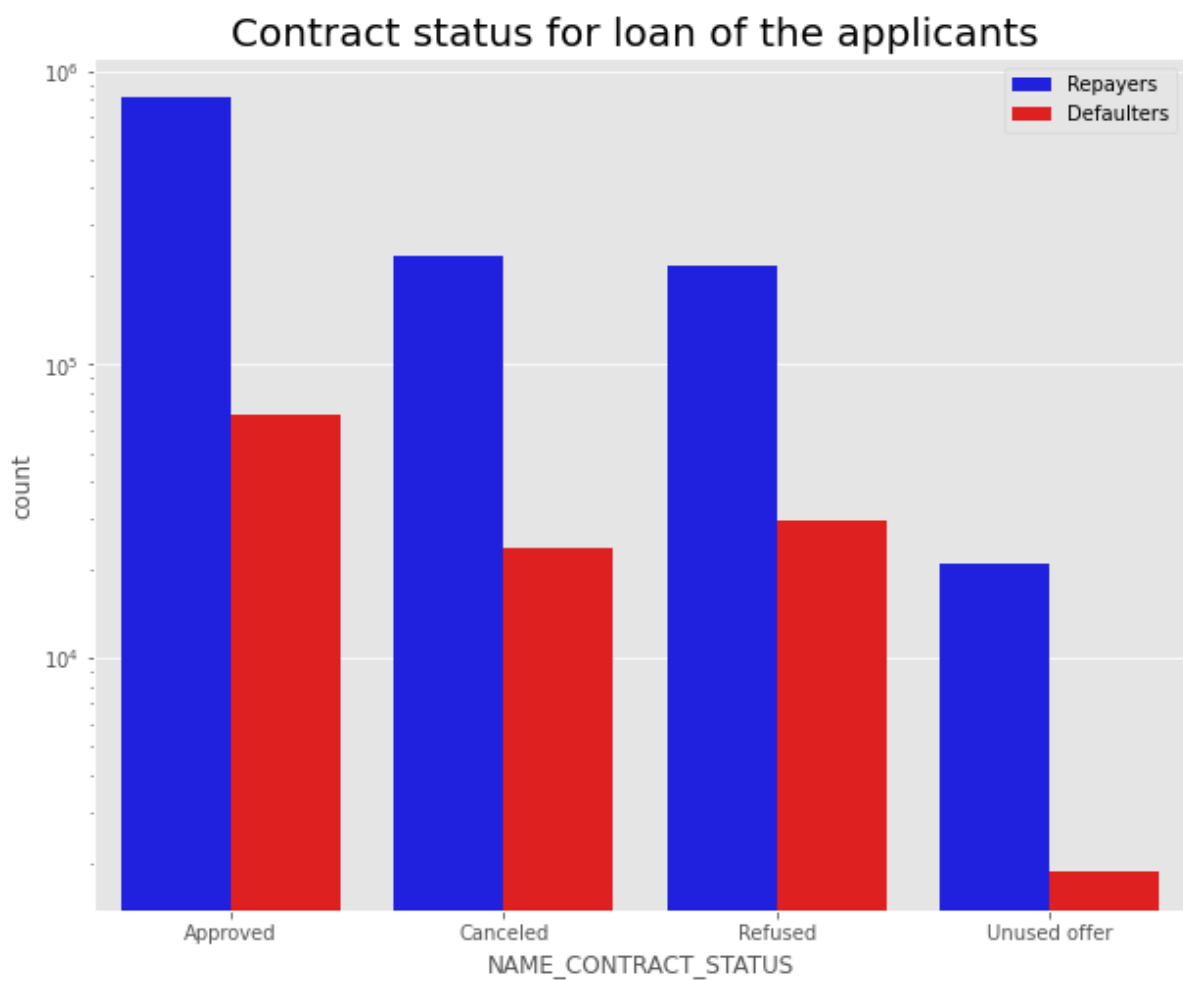
Checking the loan purpose
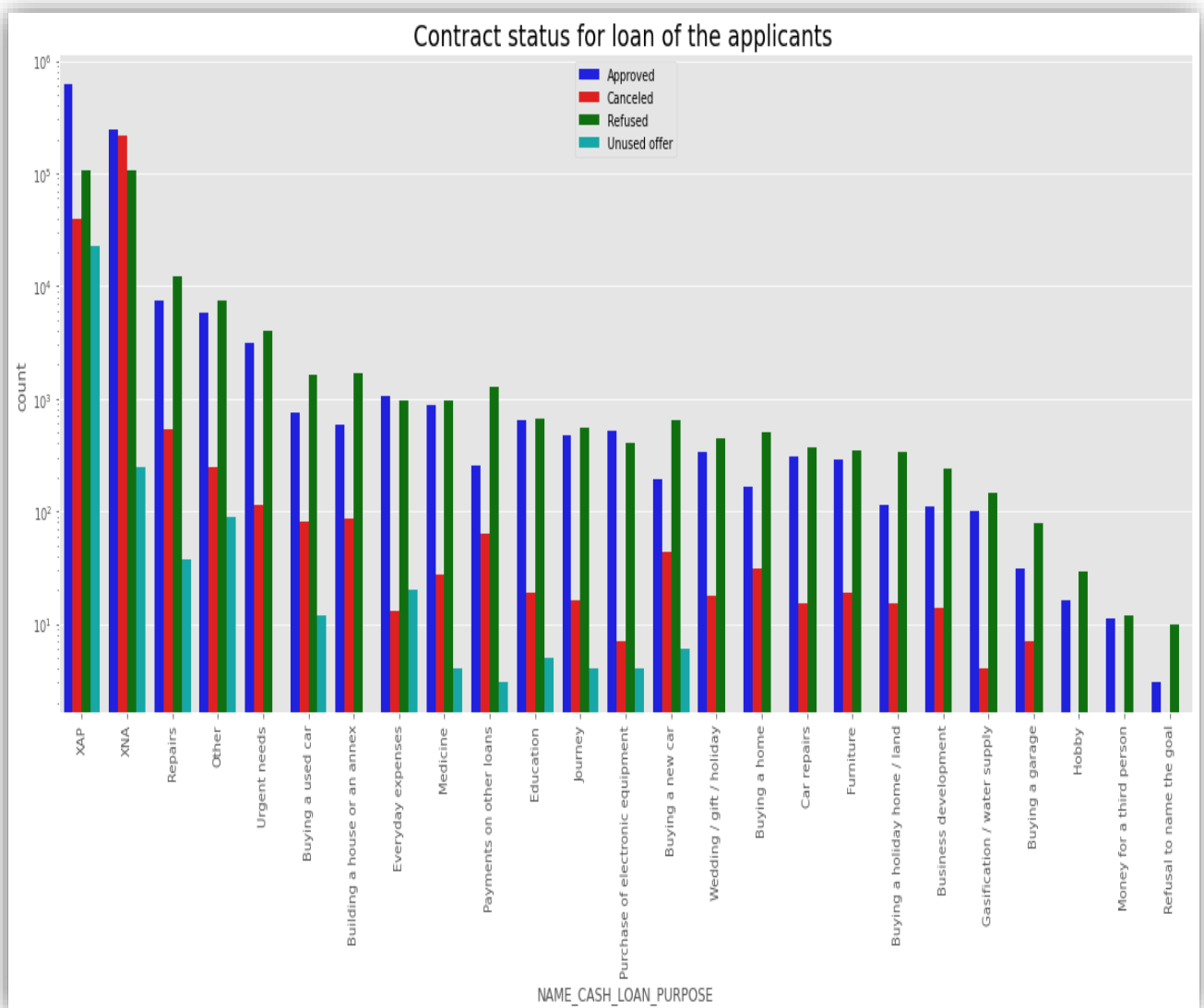
- The bank seems have to rejected majority loan application in XAP and XNA
- Payments on other loans and buying a new and house have more chances of getting rejected
- Education seems to be the best target for loans

Client Type For Loan Of The Applicants

- Repeaters seem to be the majority in all categories i.e. these are existing customers

- New customer taking loans to pay off other loans are less compared to the repeaters

THANK

YOU