CREDIT EDA ASSIGNMENT

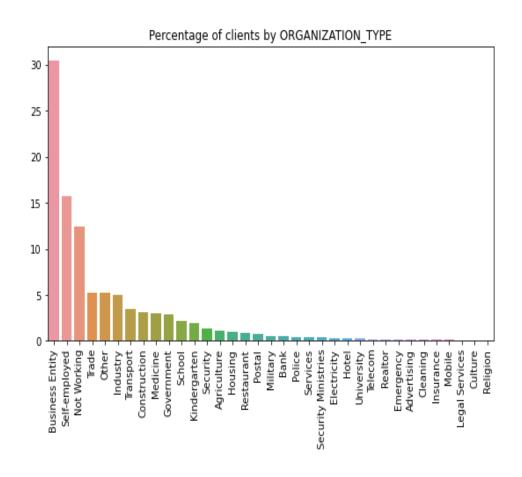
By Sharad Choudhury

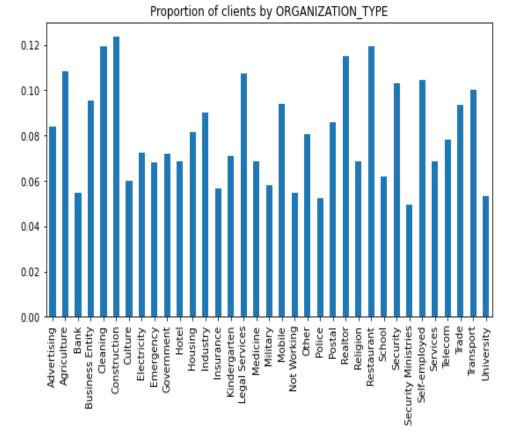


CURRENT APPLICATION DATA

UNIVARIATE ANALYSIS

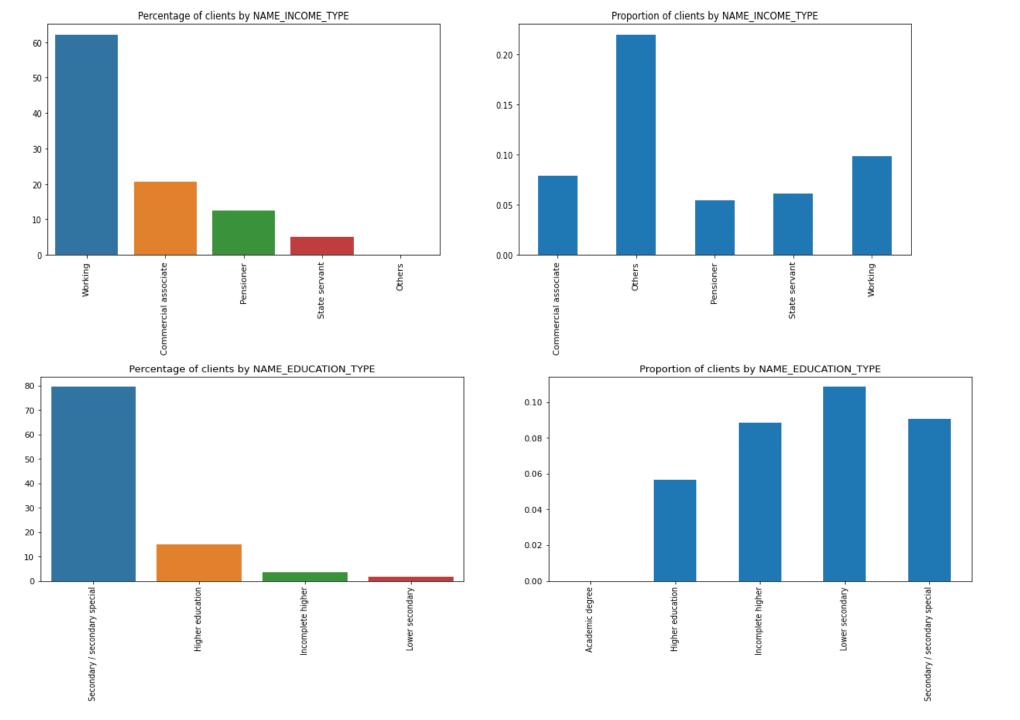
DATA OF DEFAULTERS BY PERCENTAGES AND PROPORTIONS OF EACH CATEGORY WITHIN COLUMNS





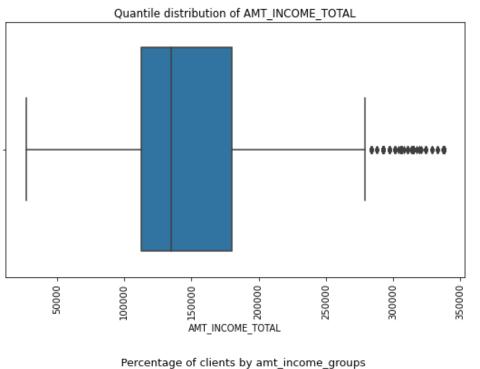
- Highest proportion of defaulters are from the Construction, Restaurant, Cleaning, Agriculture industries.
 - Lowest
 proportion is
 from University,
 Police, Security
 ministries and
 Bank industries.

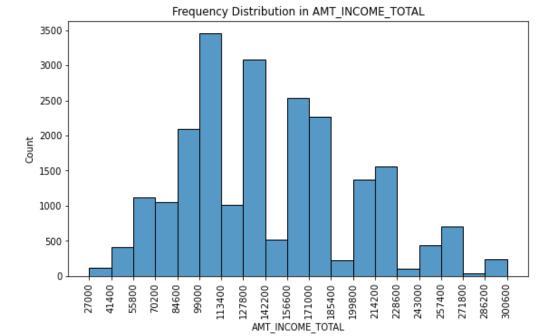




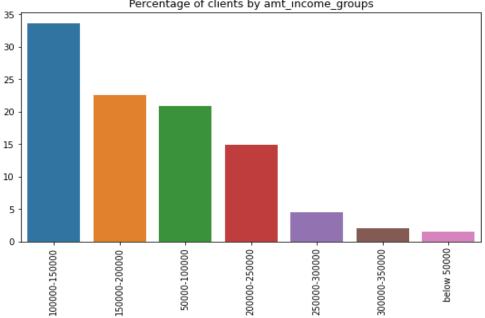
- Most defaulters are working but higher proportions from Others (Student, Businessman, Unemployed).
- Pensioners have the lowest proportions.
- Most defaulters have Secondary level education, but highest proportion of defaulters are from Lower secondary level.
- The proportion is low for clients with Higher education.
- No defaulters were with an Academic degree.
- This means more educated people are less likely to default.

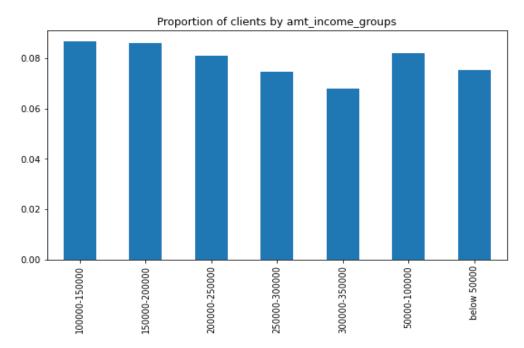






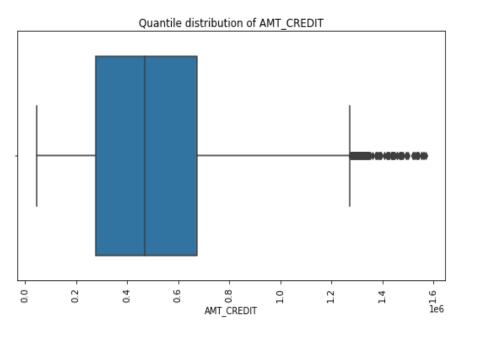
- 50% of defaulters earn below 1.3 lakh.
- Highest no of defaulters earn between 1 – 1.13 lakh

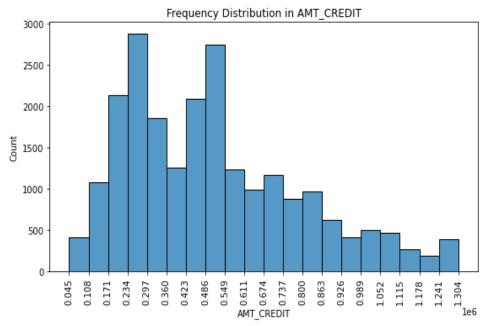


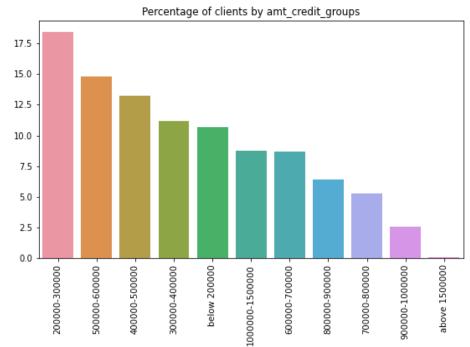


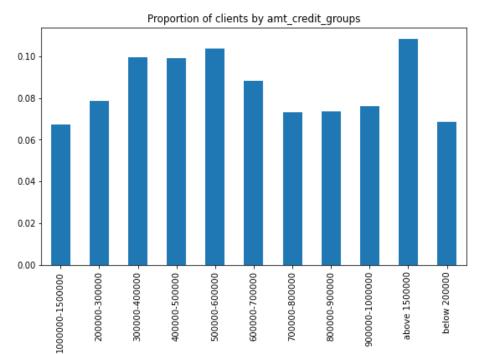
- Highest proportion of defaulters earn between 1-2 lakhs.
- Clients with higher income less likely to default





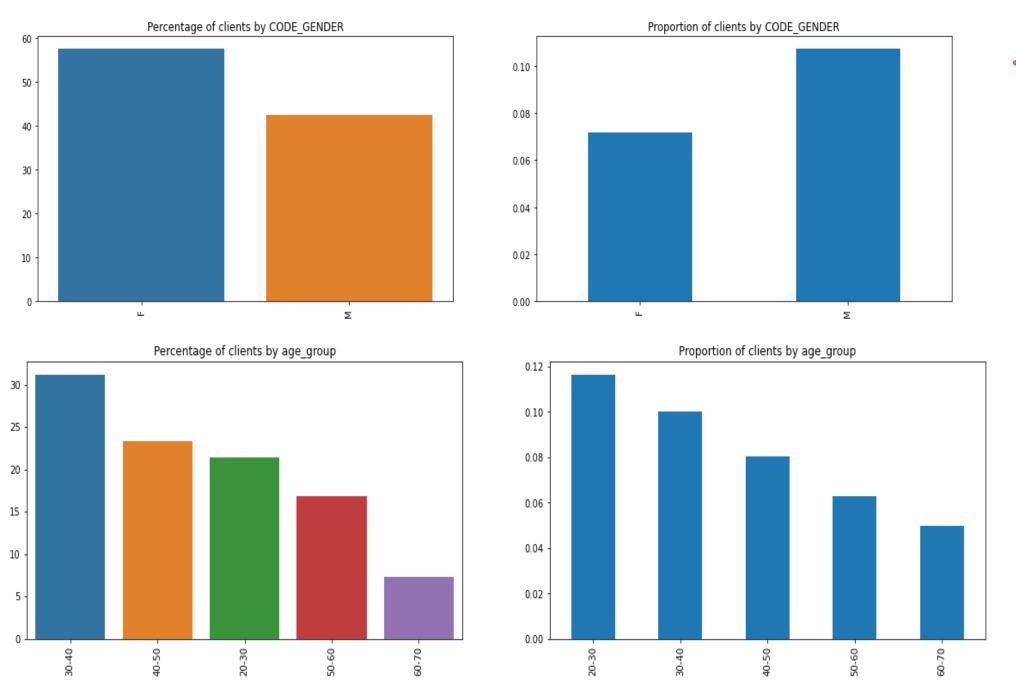






- 50% of defaulters had credit amount below around 5 lakhs.
- Most defaulters had credit between 2-3 lakhs.
- But highest proportion of defaulters had Loan credit amount as 3-6 lakhs.

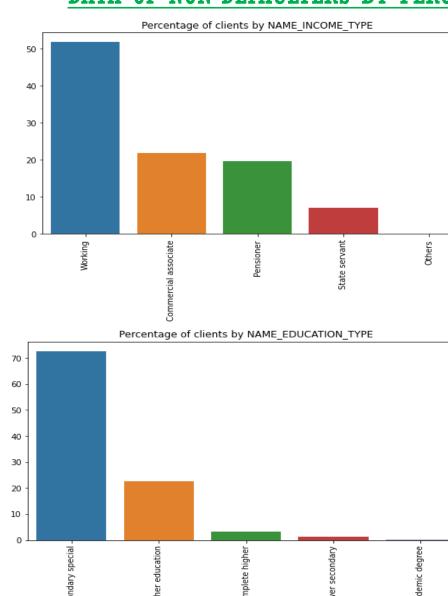


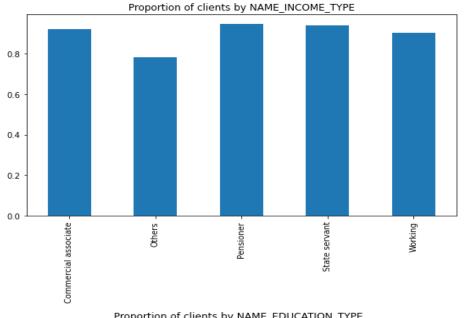


Most defaulters
were females
but highest
proportion of
defaulters are
from the Male
gender

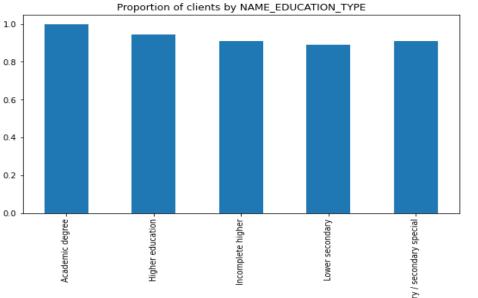
- 50% of defaulters are below 40 years of age.
- Most defaulters are in 30-40 age group but highest proportion of defaulters are from the 20-30 age group.
- We can see that people with higher ages are less likely to default.

DATA OF NON-DEFAULTERS BY PERCENTAGES AND PROPORTIONS OF EACH CATEGORY WITHIN COLUMNS





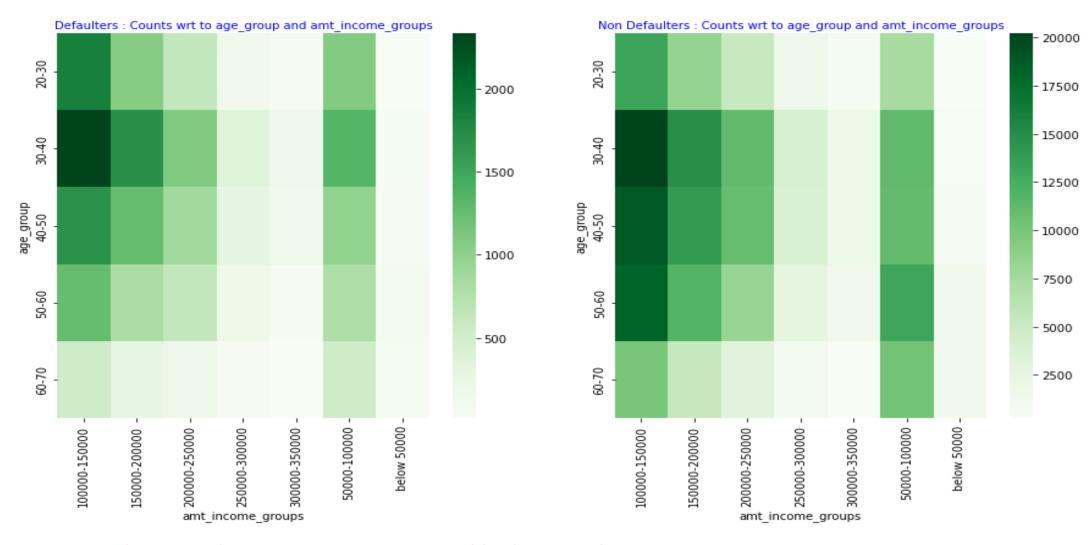
Pensioners have the highest proportions of Non-defaulters.



Academic degree holders have highest proportion of Non-defaulters.

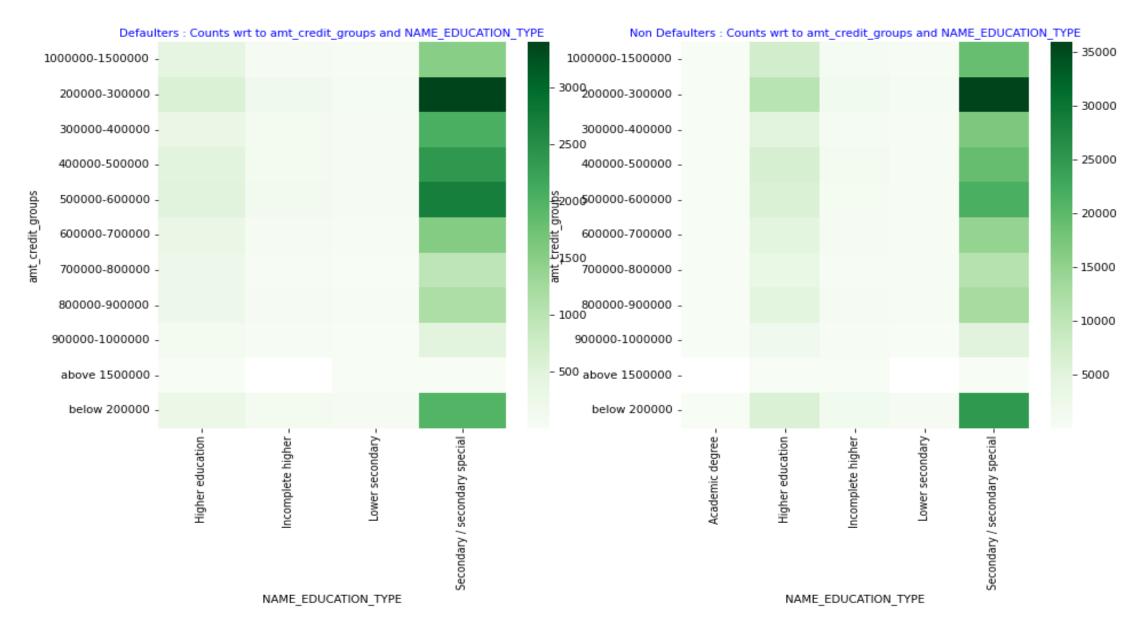


BIVARIATE ANALYSIS



- Most defaulters with salary 1-1.5 lakh were 20-40 years of age.
- There are many Non defaulters in the age group 30-60 earning 1-1.5 lakhs

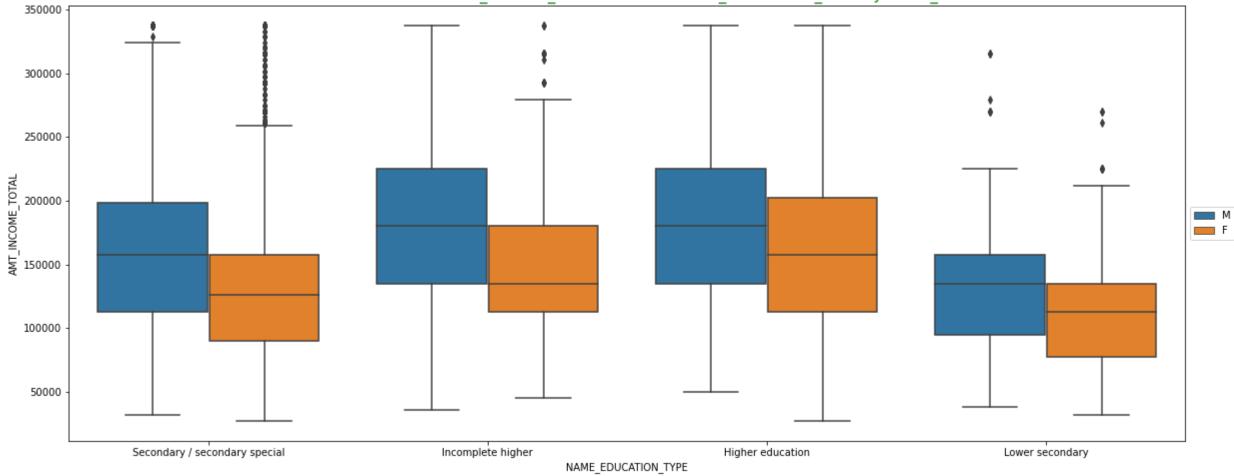




Most defaulters had Credit between 2-6 lakhs and had secondary education

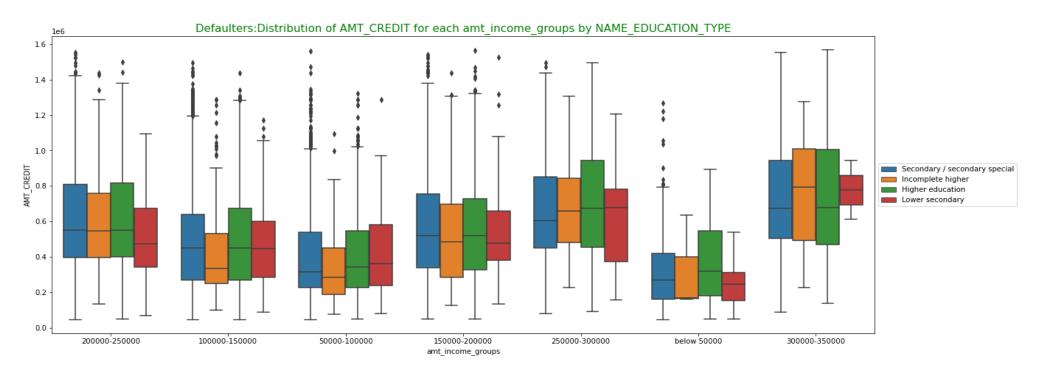


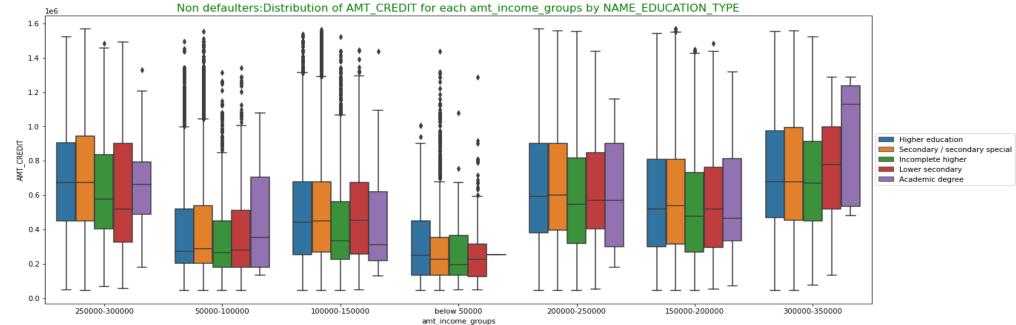
Defaulters:Distribution of AMT_INCOME_TOTAL for each NAME_EDUCATION_TYPE by CODE_GENDER



- Mostly males are having higher median salary across all education levels.
- None of those with Academic degrees were defaulters. All those with Academic degrees were non defaulters
- Median salary for higher education is higher.
- Among the defaulters with Lower secondary level education, most males and females are in lower income range.





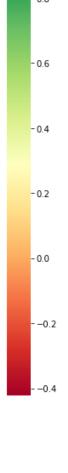


- Credit amount is generally higher for higher income groups.
- Non-Defaulters with Academic degree had the highest credits compared to other education levels.
- This means all people with academic degree were able to pay back loans even with high credits.



CORRELATION MAP

							For	Clic	nto	wi+h	D D 1 (1)		۲ 4:E4	المناة	ioc						
CNT CHILDREN	1	O 00033	0.0015	0.038	-0.0058	-0 029	-0.27	-0.2	-0.15	0.024	-0.022	nen 0.86				-0 0078	-0.0048	_0 0091	J 0082	_0 013	-0.025
AMT_INCOME_TOTAL			0.3	0.38	0.3	0.086	-0.018	-0.12		0.0059			-0.15	0.065	0.011	0.12	-0.014			0.038	0.1
AMT CREDIT		0.3	1	0.74	0.98	0.054	0.14	0.028	0.035	0.049	-0.0032	0.062		-0.0026		0.092	-0.023	-0.029		-0.0066	
AMT ANNUITY		0.38	0.74	1	0.74	0.045	0.0019	-0.074	-0.035	0.012	-0.02	0.086	-0.046		-0.0015		-0.016	-0.02		-0.0091	
AMT_GOODS_PRICE		0.3	0.98	0.74	1	0.063	0.14	0.035	0.036	0.055	0.018	0.058	-0.048	0.0014	-0.032	0.1	-0.019	-0.024	0.12	-0.0059	-0.0069
REGION_POPULATION_RELATIVE		0.086	0.054	0.045	0.063	1	0.047	0.019	0.054	0.015	0.067	-0.027	-0.42	-0.026	-0.059	0.16	0.032	0.023	0.053 -	0.00073	0.0053
YEARS_BIRTH	-0.27	-0.018	0.14	0.0019	0.14	0.047	1	0.63	0.29	0.26	0.025	-0.21	-0.033	-0.057	-0.17	0.11	-0.0051	-0.0045	0.11	0.022	0.084
YEARS_EMPLOYED	-0.2	-0.12	0.028	-0.074	0.035	0.019	0.63	1	0.22	0.25	0.027	-0.18	-0.0018	-0.045	-0.11	0.027	-3.1e-05	0.0042	0.037	0.012	0.036
DAYS_REGISTRATION	-0.15	-0.036	0.035	-0.035	0.036	0.054	0.29	0.22	1	0.099	0.062	-0.14	-0.11	-0.021	-0.052	0.068	-0.01	-0.011	0.069	0.016	0.014
DAYS_ID_PUBLISH	- 0.024	0.0059	0.049	0.012	0.055	0.015	0.26	0.25	0.099	1	0.041	0.022	-0.0045	-0.027		0.06	-0.0082	-0.0061	0.13	0.01	0.06
FLAG_PHONE	-0.022	-0.028	-0.0032	-0.02	0.018	0.067	0.025	0.027	0.062	0.041	1	0.0022	-0.058	0.0057	-0.045	0.044	-0.029	-0.027	0.073	-0.013	-0.019
CNT_FAM_MEMBERS	0.86	0.0052	0.062	0.086	0.058	-0.027	-0.21	-0.18	-0.14	0.022	0.0022	1	0.039	-0.028	-0.0012	0.0065	0.0017	-0.0037	0.017	-0.0051	0.0031
REGION_RATING_CLIENT	- 0.036	-0.15	-0.04	-0.046	-0.048	-0.42	-0.033	-0.0018	-0.11	-0.0045	-0.058	0.039	1	-0.026	0.042	-0.24	0.0024	0.0059	-0.012	0.0064	0.0077
REG_REGION_NOT_LIVE_REGION	0.026	0.065	-0.0026	0.02	0.0014	-0.026	-0.057	-0.045	-0.021	-0.027	0.0057	-0.028	-0.026	1	0.31	0.0031	0.0044	0.0023	-0.034	-0.0012	-0.02
REG_CITY_NOT_LIVE_CITY	- 0.0045	0.011	-0.032	-0.0015	-0.032	-0.059	-0.17	-0.11	-0.052		-0.045	-0.0012	0.042	0.31	1	-0.052	0.011	0.012		0.0021	-0.014
EXT_SOURCE_2	0.0078	0.12	0.092	0.085	0.1	0.16	0.11	0.027	0.068	0.06	0.044	0.0065	-0.24	0.0031	-0.052	1	-0.0098	-0.0098	0.21	0.0026	0.018
DEF_30_CNT_SOCIAL_CIRCLE	-0.0048	-0.014	-0.023	-0.016	-0.019	0.032	-0.0051	-3.1e-05	-0.01	-0.0082	-0.029	0.0017	0.0024	0.0044	0.011	-0.0098	1	0.87	-0.0016	-0.0055	0.0072
DEF_60_CNT_SOCIAL_CIRCLE	-0.0091	-0.016	-0.029	-0.02	-0.024	0.023	-0.0045	0.0042	-0.011	-0.0061	-0.027	-0.0037	0.0059	0.0023	0.012	-0.0098	0.87	1	-0.0053	-0.0011	-0.0016
DAYS_LAST_PHONE_CHANGE	0.0082	0.075	0.11	0.072	0.12	0.053	0.11	0.037	0.069	0.13	0.073	0.017	-0.012	-0.034		0.21	-0.0016	-0.0053	1	0.0014	0.12
AMT_REQ_CREDIT_BUREAU_QRT	0.013	0.038	-0.0066	-0.0091	-0.0059	-0.00073	0.022	0.012	0.016	0.01	-0.013	-0.0051	0.0064	-0.0012	0.0021	0.0026	-0.0055	-0.0011	0.0014	1	0.13
AMT_REQ_CREDIT_BUREAU_YEAR	-0.025	0.1	-0.006	0.0015	-0.0069	0.0053	0.084	0.036	0.014	0.06	-0.019	0.0031	0.0077	-0.02	-0.014	0.018	0.0072	-0.0016	0.12	0.13	1
	CNT_CHILDREN	AMT_INCOME_TOTAL.	AMT_CREDIT .	AMT_ANNUITY.	AMT_GOODS_PRICE	REGION_POPULATION_RELATIVE	YEARS_BIRTH	YEARS_EMPLOYED.	DAYS_REGISTRATION	DAYS_ID_PUBLISH	FLAG_PHONE .	CNT_FAM_MEMBERS	REGION_RATING_CLIENT	REG_REGION_NOT_LIVE_REGION	REG_CITY_NOT_LIVE_CITY	EXT_SOURCE_2	DEF_30_CNT_SOCIAL_CIRCLE	DEF_60_CNT_SOCIAL_CIRCLE .	DAYS_LAST_PHONE_CHANGE -	AMT_REQ_CREDIT_BUREAU_QRT .	AMT_REQ_CREDIT_BUREAU_YEAR





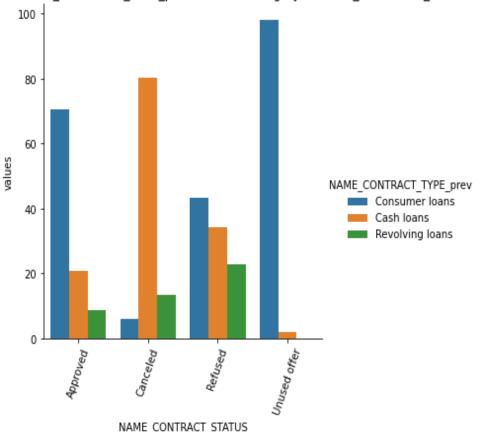
We can see that for clients with Payment difficulties:

- The region rating is inversely proportional to Income, credit, population, normalized external score. This means better regions have rating 1.
- Higher the Income, higher the annuity, credit and goods price.
- Client permanent address does not match contact address for mostly those defaulters who had low age and years of employment.
- Defaulters that changed registration recently before the application had poor region ratings.
- Age and Years of employment is inversely proportional to count of children and family members. This means mostly young defaulters had more family members and children.

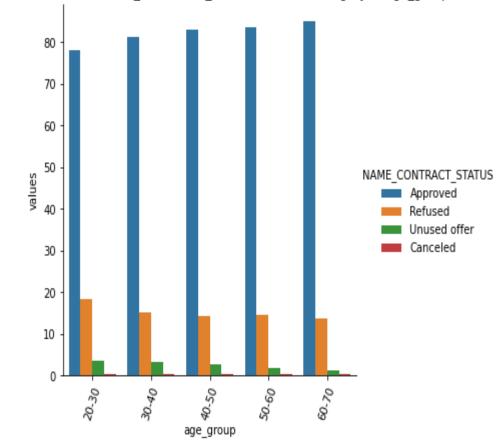


PREVIOUS APPLICATION DATA

Distribution (in %) of NAME CONTRACT TYPE prev for each category in NAME CONTRACT STATUS



Distribution (in %) of NAME_CONTRACT_STATUS for each category in age_group

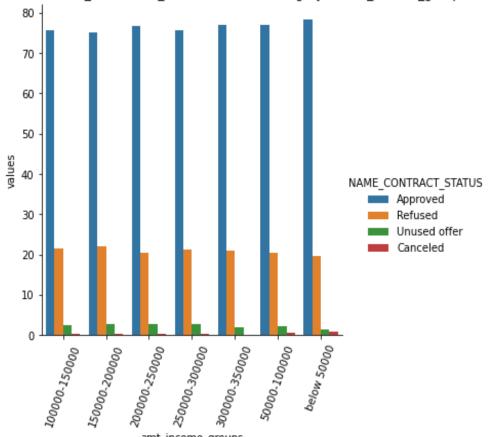


- Most of the loans that were approved were Consumer loans
- Most loans that were cancelled were cash loans followed by revolving loans
- Clients with age 20-30 have lowest no of Approved contracts and highest no. of refused contracts.
- People with age 60-70 had least unused offers

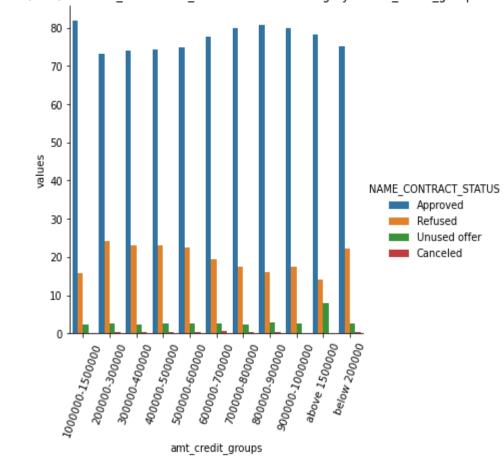


ANALYSIS OF CONTRACT STATUS FOR DEFAULTERS

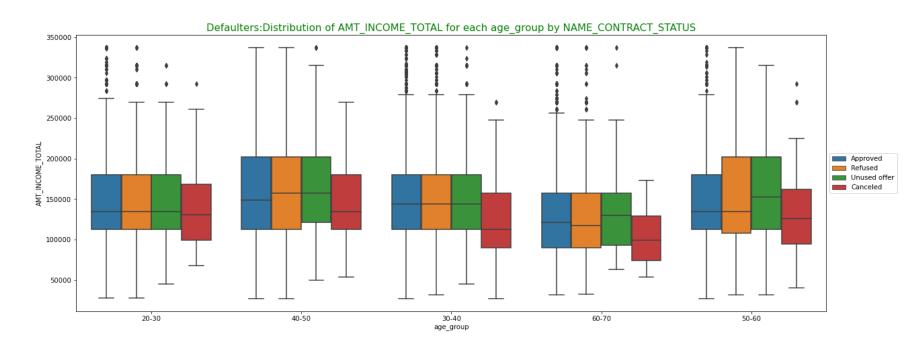
Distribution (in %) of NAME_CONTRACT_STATUS for each category in amt_income_groups



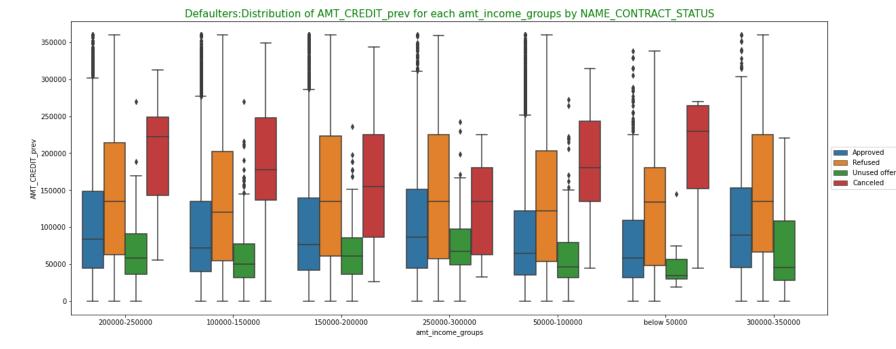
 Defaulters earning below 50k had highest cancelled offers percentage compared to other income groups Distribution (in %) of NAME_CONTRACT_STATUS for each category in amt_credit_groups



- Defaulters with credits between 1-1.5 lakhs had highest % of approved offers.
- Defaulters with credit > 15 lakhs had highest % of unused offers. This
 means extremely high credits were mostly unused.



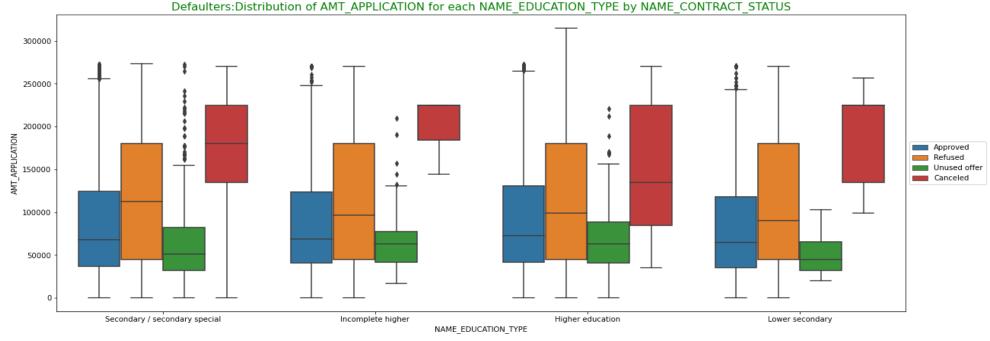
- The loans were mostly cancelled for clients with lower median income.
- For defaulters in age 50-70, the median income was higher of those who did not use the offer.



- No loans were cancelled previously for defaulters that earned 3-3.5 lakhs.
- The median credit amount of cancelled loans is higher than median credit of the other loans.



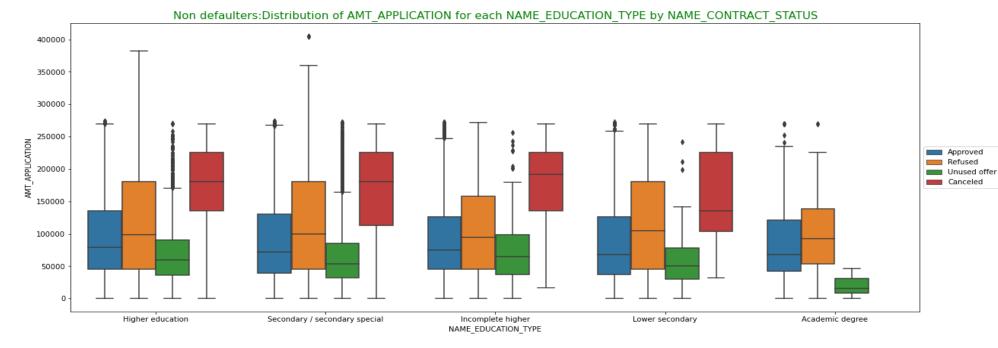
Defaulters: Distribution of AMT APPLICATION for each NAME EDUCATION TYPE by NAME CONTRACT STATUS







No loans were cancelled for clients with academic degree and none of them were defaulters



CONCLUSIONS

- Males are more likely to default
- Clients with lower income and lower secondary education more likely to default.
- Clients with Academic degree least likely to default.
- Pensioners are also least likely to default.
- Clients with high credit amounts above 15 lakhs mostly don't use the offer.

