# Loan Defaulter

**CREDIT EDA CASE STUDY** 

Presented By : Juhi Chugh Madhumitha R

## Problem Statement

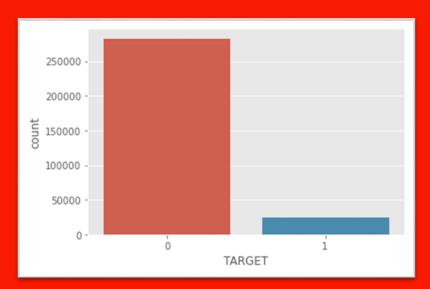
The loan providing companies find it hard to give loans to the people due to their insufficient or non-existent credit history. Because of that, some consumers use it as their advantage by becoming a defaulter. Suppose we work for a consumer finance company which specializes in lending various types of loans to urban customers. We have used EDA to analyze the patterns present in the data. This will ensure that the applicants capable of repaying the loan are not rejected.



This case study aims to identify patterns which indicate if a client has difficulty paying their installments which may be used for taking actions such as denying the loan, reducing the amount of loan, lending (to risky applicants) at a higher interest rate, etc. This will ensure that the consumers capable of repaying the loan are not rejected. Identification of such applicants' using EDA is the aim of this case study.

## DATA IMBALANCE

The Imbalance ratio is 11.39

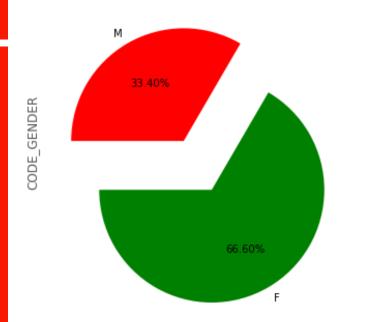




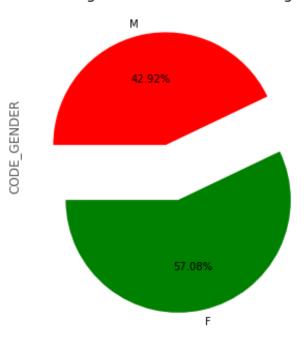
## **GENDER**

The percentage of the Female clients is more compared to the Male. And Males have more chance of defaulting

#### Plot showing GENDER vs count for Target=0



#### Plot showing GENDER vs count for Target=1

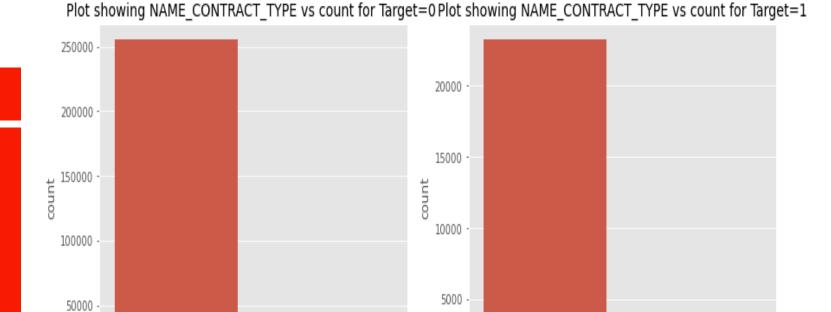


## UNIVARIATE ANALYSIS – CATEGORICAL VARIABLES

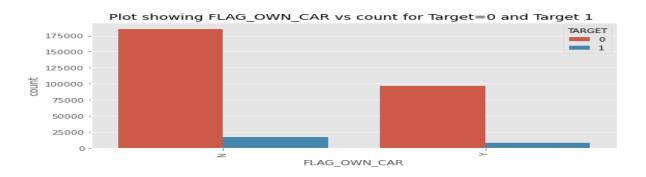
NAME\_CONTRACT\_TYPE

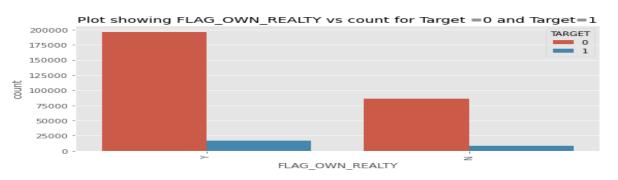
## Contract Type

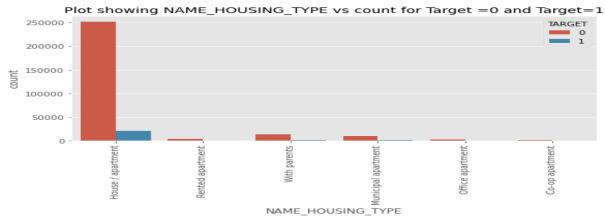
The percentage of client applying for Cash Loans is high.

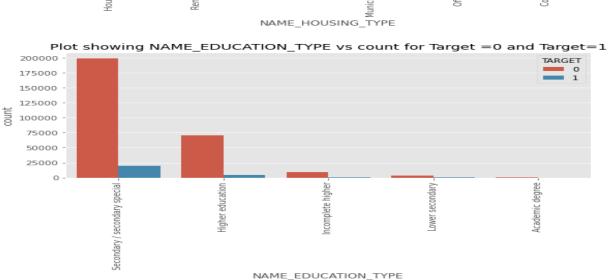


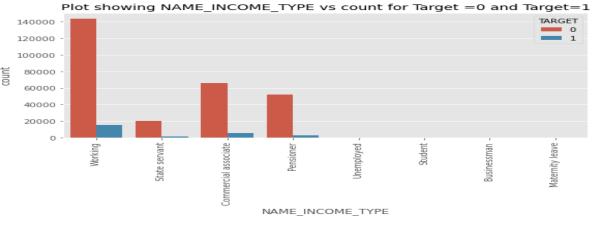
NAME\_CONTRACT\_TYPE

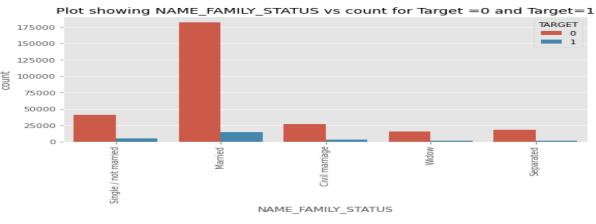












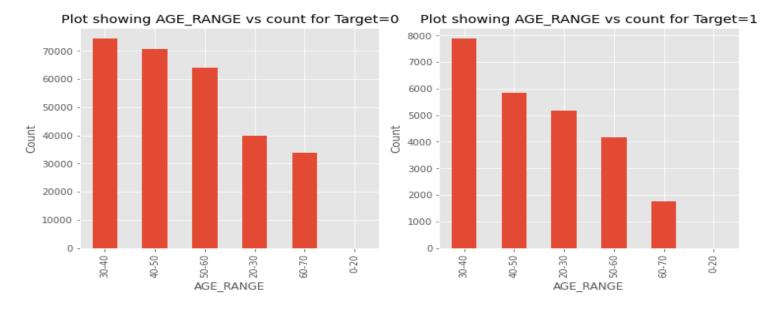
# Housing Type Income Type Own Car Realty Education Type Family Status

- The count of Clients who do not own a car is high. Hence we should target them but since the chance of defaulting is more should be more careful with the loan amount and other factors.
- The clients having own realty have more chance of defaulting.
- Clients living in Rented apartments, Co-op apartments or office apartments have less chance of defaulting compared to others.
- Client who are working can be targetted.
- Clients who have secondary/higher education have a greater chance to take loans.
- Clients who are married have a greater chance to take loans.

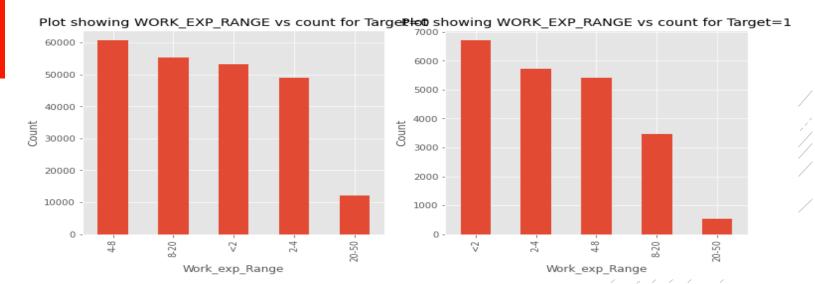
Target the clients in age group 30-50. Eventhough the age of 20-30 have applied for loan their percentage of defaulting is high when compared to other age group.

Clients who are having experience of 4-20 years have less chance of defaulting. While the clients with work experience less than 2 years or 2-4 years have higher chance of defaulting. We can target clients with greater work experience

#### **AGE**



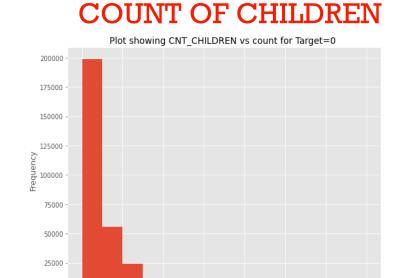
#### **WORK EXPERIENCE**

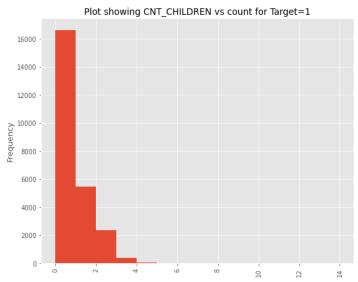


## UNIVARIATE ANALYSIS – NUMERICAL VARIABLES

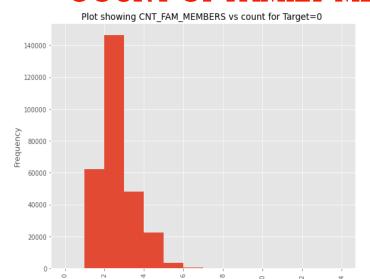
Clients who have no children have higher count of applying for loan. Also when the count of children increases, they have more chance of defaulting.

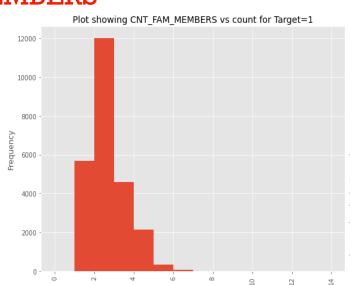
We can see that a family of 3 applies loan more often than the other families





#### **COUNT OF FAMILY MEMBERS**

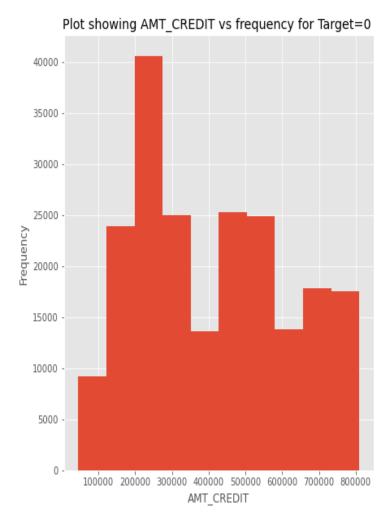


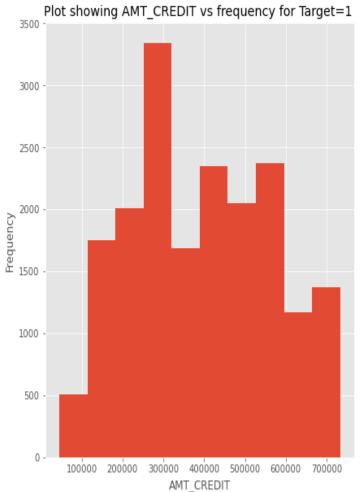


## SEGMENTED UNIVARIATE ANALYSIS

#### **CREDIT AMOUNT**

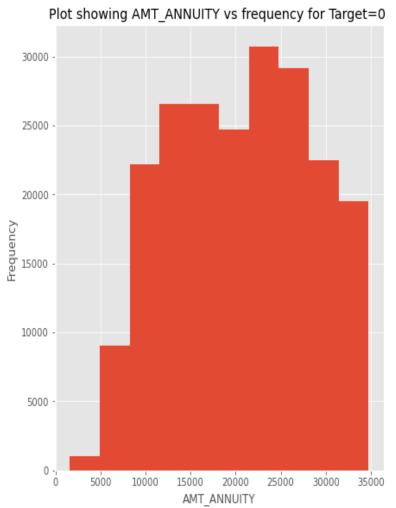
We can infer that as credit amount is above 300000 the chance of defaulting is slightly more

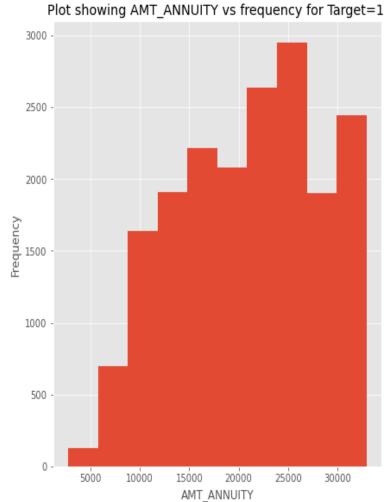




#### **ANNUITY AMOUNT**

Both the Target0 and Target1 data follow same pattern for AMT\_ANNUTTY.
Few exceptions occurs when the ANNUTTY amount is 25000-30000. The defaulting is high around 25000 and 30000



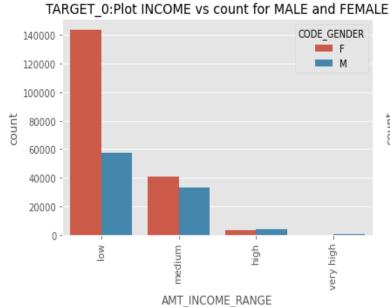


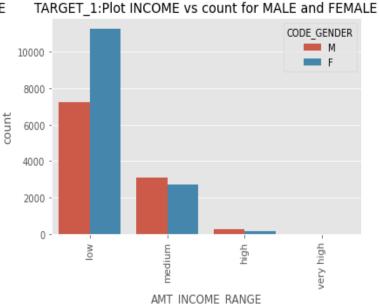
## **BIVARIATE ANALYSIS**

#### INCOME AND GENDER

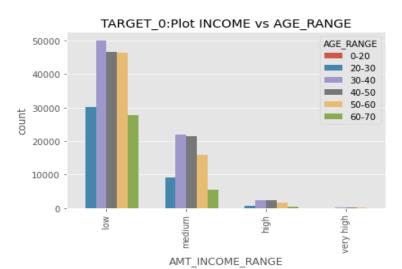
Female clients tend to have low income compared to male though their count for the loan application is high. Males with low income tend to default the payment more.

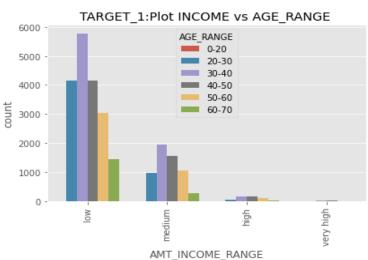
In the low level income range age group of 20-30 has a high chance of being a defaulter





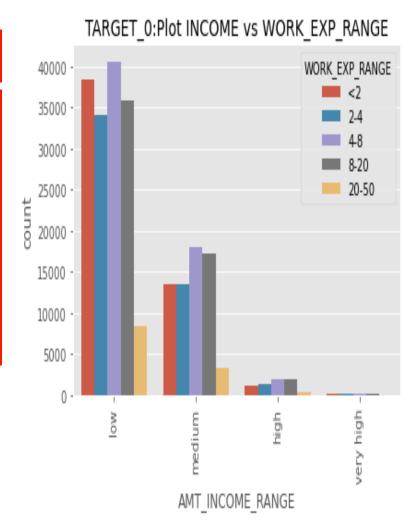
#### **INCOME AND AGE**

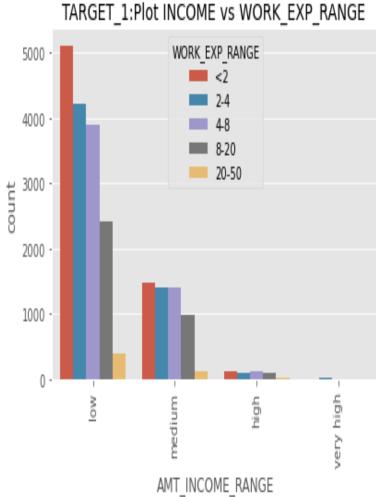




#### **INCOME AND WORK EXPERIENCE**

In the low-level income range clients with experience <2 years have a high chance of being a defaulter





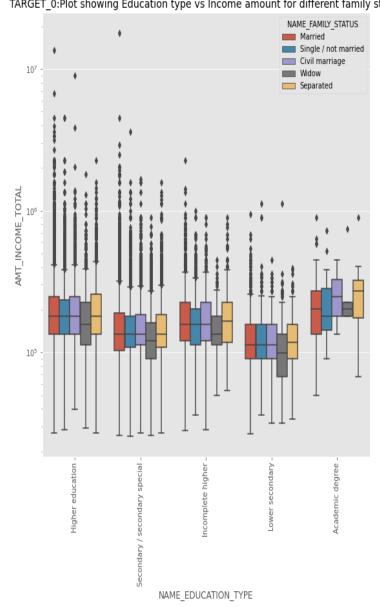
#### EDUCATION, INCOME, FAMILY\_STATUS, ANNUITY and CREDIT AMOUNT

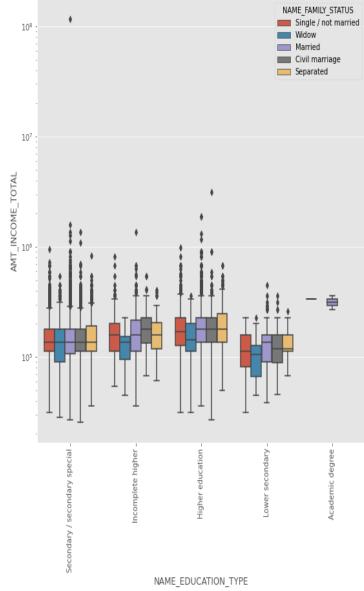
TARGET\_0:Plot showing Education type vs Income amount for different family statusTARGET\_1:Plot showing Education type vs Income amount for different family status

Clients with higher education and married have outliers for the Income amount value. They earn a high salary when compared to the Academic in all martial status

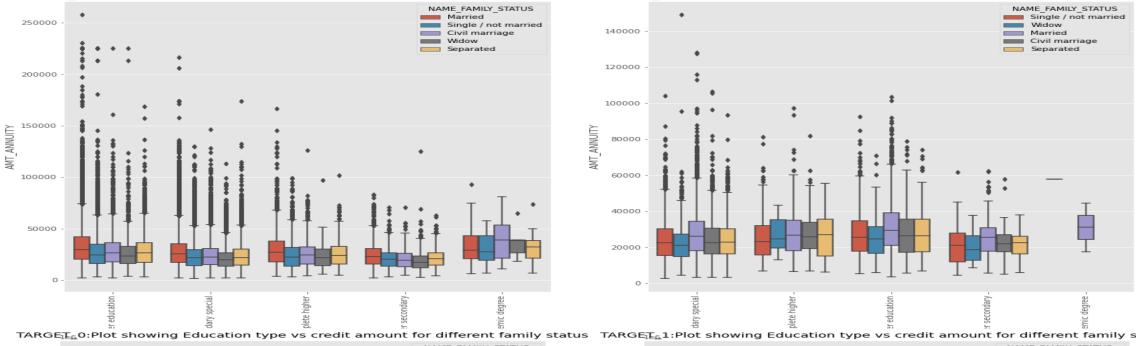
Clients with higher education and married have outliers for the Credit amount value, Income amount value, Annuity value.

They have taken very high value for credit.

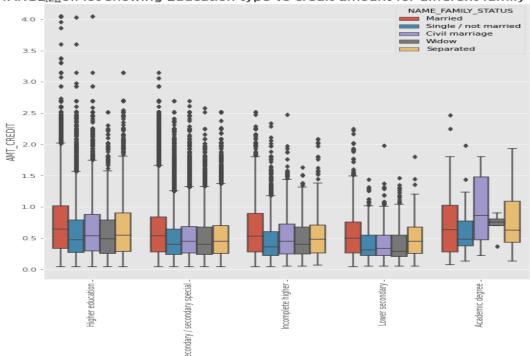




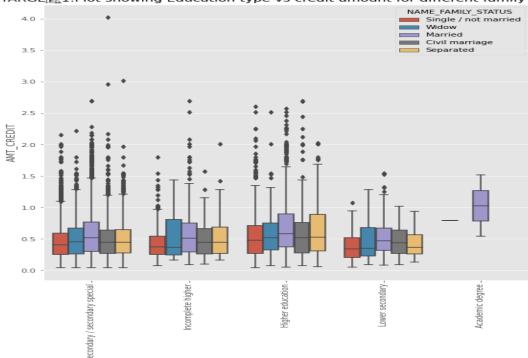
TARGET\_0:Plot showing Education type vs Annuity amount for different family statusTARGET\_1:Plot showing Education type vs Annuity amount for different family status



TARGET 1:Plot showing Education type vs credit amount for different family status TARGET 1:Plot showing Education type vs credit amount for different family status



NAME\_EDUCATION\_TYPE



NAME\_EDUCATION\_TYPE

### CORRELATION MATRIX

#### Heatmap showing Correlation of Target 0

								. 5				5			
	CNT_CHILDREN	1	0.027	0.0031	0.021	-0.00055	-0.024	-0.03	0.023	0.88	0.023	0.022	0.017	-0.34	-0.25
	AMT_INCOME_TOTAL	0.027	1	0.34	0.42	0.35	0.17	0.0027	0.084	0.034		-0.2	0.13	-0.063	-0.14
	AMT_CREDIT	0.0031	0.34	1	0.77	0.99	0.1	0.027	0.017	0.065	-0.1	-0.11	0.054	0.047	-0.07
/	AMT_ANNUITY	0.021	0.42	0.77		0.78	0.12	0.012	0.071	0.076	-0.13	-0.15	0.075	-0.012	-0.1
	AMT_GOODS_PRICE	-0.00055	0.35	0.99	0.78	1	0.1	0.042	0.017	0.063	-0.1	-0.11	0.054	0.045	-0.069
	REGION_POPULATION_RELATIVE	-0.024	0.17	0.1	0.12	0.1	1	0.091	0.041	-0.023	-0.54	-0.54	0.084	0.025	-0.0072
/	FLAG_PHONE	-0.03	0.0027	0.027	0.012	0.042	0.091	1	0.015	-0.016	-0.084	-0.08	0.0049	0.042	0.016
/	FLAG_EMAIL	0.023	0.084	0.017	0.071	0.017	0.041	0.015	1	0.019	-0.054	-0.053	0.027	-0.091	-0.065
1	CNT_FAM_MEMBERS	0.88	0.034	0.065	0.076	0.063	-0.023	-0.016	0.019	1	0.028	0.029	0.015	-0.29	-0.24
	REGION_RATING_CLIENT	0.023	-0.19	-0.1	-0.13	-0.1	-0.54	-0.084	-0.054	0.028	1	0.95		-0.0023	0.038
	REGION_RATING_CLIENT_W_CITY	0.022	-0.2	-0.11	-0.15	-0.11	-0.54	-0.08	-0.053	0.029	0.95	1	-0.14	-0.00079	0.04
	LIVE_REGION_NOT_WORK_REGION	0.017	0.13	0.054	0.075	0.054	0.084	0.0049	0.027	0.015	-0.15	-0.14	1	-0.072	-0.099
	AGE -	-0.34	-0.063	0.047	-0.012	0.045	0.025	0.042	-0.091	-0.29	-0.0023	-0.00079	-0.072	1	0.63
	WORK_EXP	-0.25	-0.14	-0.07	-0.1	-0.069	-0.0072	0.016	-0.065	-0.24	0.038	0.04	-0.099	0.63	1
		CNT_CHILDREN -	AMT_INCOME_TOTAL -	AMT_CREDIT -	AMT_ANNUITY -	AMT_GOODS_PRICE -	REGION_POPULATION_RELATIVE -	FLAG_PHONE -	FLAG_EMAIL -	CNT_FAM_MEMBERS -	REGION_RATING_CLIENT -	REGION_RATING_CLIENT_W_CITY -	LIVE_REGION_NOT_WORK_REGION -	AGE -	WORK_EXP -

Heatmap showing Correlation of Target 1

			116	atiii	ap si	IOWII	ig C	JITE	atioi	1011	arge	ιı		
CNT_CHILDREN	- 1	0.027	0.0031	0.021	-0.00055	-0.024	-0.03	0.023	0.88	0.023	0.022	0.017	-0.34	-0.25
AMT_INCOME_TOTAL	- 0.027	1	0.34	0.42	0.35	0.17	0.0027	0.084	0.034	-0.19	-0.2	0.13	-0.063	-0.14
AMT_CREDIT	- 0.0031	0.34	1	0.77	0.99	0.1	0.027	0.017	0.065	-0.1	-0.11	0.054	0.047	-0.07
AMT_ANNUITY	- 0.021	0.42	0.77	1	0.78	0.12	0.012	0.071	0.076	-0.13	-0.15	0.075	-0.012	-0.1
AMT_GOODS_PRICE	-0.00055	0.35	0.99	0.78	1	0.1	0.042	0.017	0.063	-0.1	-0.11	0.054	0.045	-0.069
REGION_POPULATION_RELATIVE	-0.024	0.17	0.1	0.12	0.1	1	0.091	0.041	-0.023	-0.54	-0.54	0.084	0.025	-0.0072
FLAG_PHONE	-0.03	0.0027	0.027	0.012	0.042	0.091	1	0.015	-0.016	-0.084	-0.08	0.0049	0.042	0.016
FLAG_EMAIL	- 0.023	0.084	0.017	0.071	0.017	0.041	0.015	1	0.019	-0.054	-0.053	0.027	-0.091	-0.065
CNT_FAM_MEMBERS	- 0.88	0.034	0.065	0.076	0.063	-0.023	-0.016	0.019	1	0.028	0.029	0.015	-0.29	-0.24
REGION_RATING_CLIENT	- 0.023	-0.19	-0.1	-0.13	-0.1	-0.54	-0.084	-0.054	0.028	1	0.95		-0.0023	0.038
REGION_RATING_CLIENT_W_CITY	- 0.022		-0.11	-0.15	-0.11	-0.54	-0.08	-0.053	0.029	0.95	1	-0.14	-0.00079	0.04
LIVE_REGION_NOT_WORK_REGION	- 0.017	0.13	0.054	0.075	0.054	0.084	0.0049	0.027	0.015		-0.14	1	-0.072	-0.099
AGE	-0.34	-0.063	0.047	-0.012	0.045	0.025	0.042	-0.091	-0.29	-0.0023	-0.00079	-0.072	1	0.63
WORK_EXP	-0.25	-0.14	-0.07	-0.1	-0.069	-0.0072	0.016	-0.065	-0.24	0.038	0.04	-0.099	0.63	
	CNT_CHILDREN -	AMT_INCOME_TOTAL -	AMT_CREDIT -	AMT_ANNUITY -	AMT_GOODS_PRICE -	REGION_POPULATION_RELATIVE -	FLAG_PHONE -	FLAG_EMAIL -	CNT_FAM_MEMBERS -	REGION_RATING_CLIENT -	SEGION_RATING_CLIENT_W_CITY -	F_REGION_NOT_WORK_REGION -	AGE -	WORK_EXP -

-1.0 -0.8

- 0.4

- 0.2

- 0.0

- -0.2

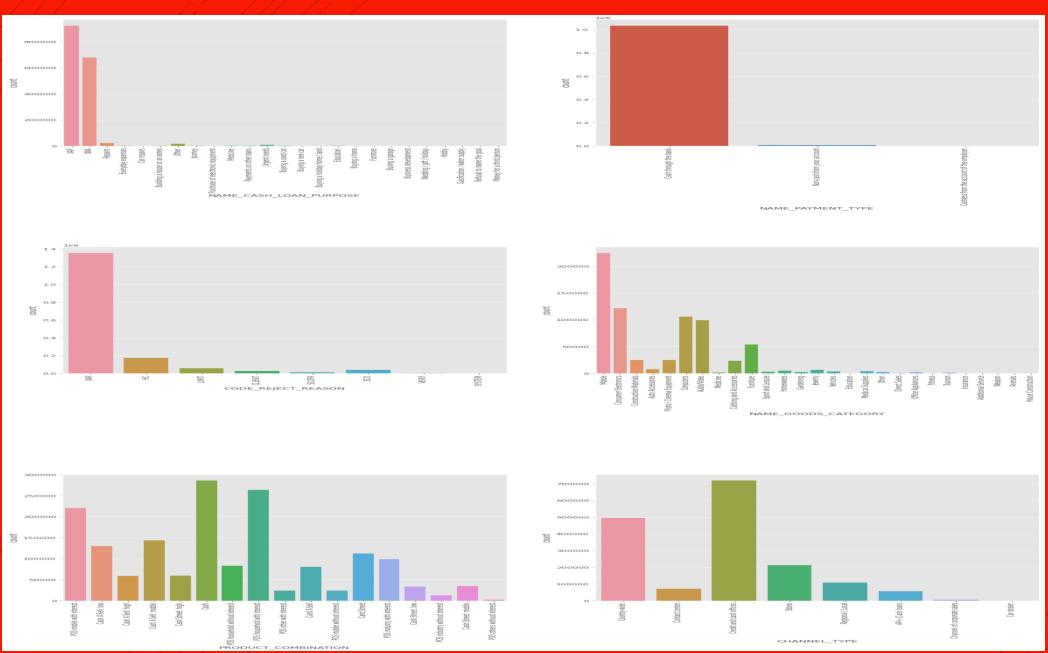
- -0.4

#### Top 5 Co-Relation Variable:

1. AMT\_GOODS\_PRICE & AMT\_CREDIT
2. REGION\_RATING\_CLIENT\_W\_CLIENT &
REGION\_RATING\_CLIENT
3. CNT\_FAM\_MEMBERS & CNT\_CHILDREN
4. AMT\_ANNUITY & AMT\_GOODS\_PRICE
5. AMT\_ANNUITY & AMT\_CREDIT

- We can infer the following from target0 and target1 heatmaps:
- The Age and Work Experience are positively correlated.i.e., as the Age increase the work experience increases. AS the age and work experience increase the annuity value tends to decrease. They are negatively correlated. As the credit value increases the annuity amt also increases As income increases credit and annuity increases As the children count increases the family members increases
- The Top 5 Co-Relation variables are same for Target 0 and Target 1

## UNIVARIATE ANALYSIS- PREVIOUS APPLICATION



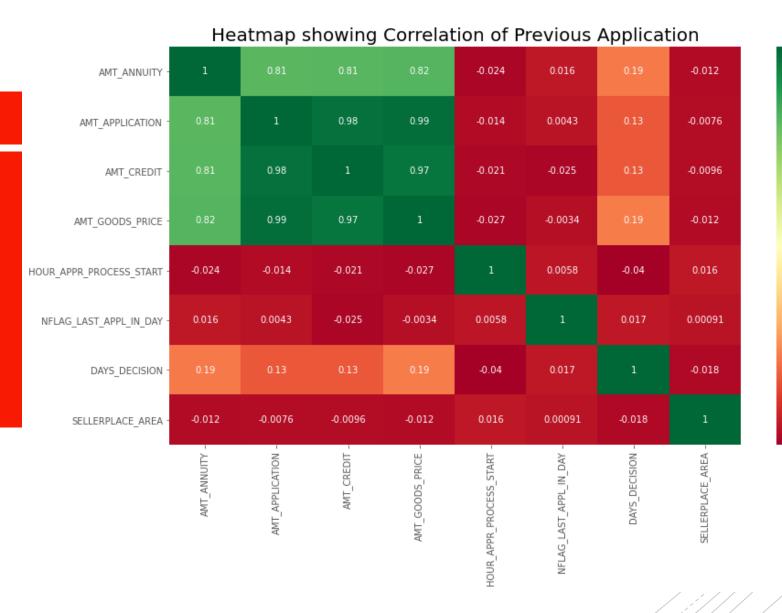
CASH LOAN PURPOSE
PAYMENT TYPE
REJECT REASON
GOODS CATEGORY,
CHANNEL TYPE,
PRODUCT COMBINATION

- Most Loan purpose was not recorded. XAP and XNA values are highest.
- Most clients preferred CASH as the mode of Payment.
- Primary reason for the Loan to get rejected is not recorded(XAP (81%)) followed by HC.
- Most clients applied for Mobile and Consumer electronics.
- Highest product combination is Cash followed by POS household with interest.
- Most clients were acquired from Credit and Cash Offices.

#### CO-RELATION HEATMAP FOR PREVOIUS APPLICATION

#### Top 5 Co-Relation Variable:

1. AMT\_GOODS\_PRICE&
AMT\_APPLICATION
2. AMT\_APPLICATION&AMT\_CREDIT
3. AMT\_GOOD\_PRICE&AMT\_CREDIT
4. AMT\_GOOD\_PRICE&AMT\_ANNUITY
5. AMT\_CREDIT&AMT\_ANNUITY



- 0.8

- 0.6

- 0.4

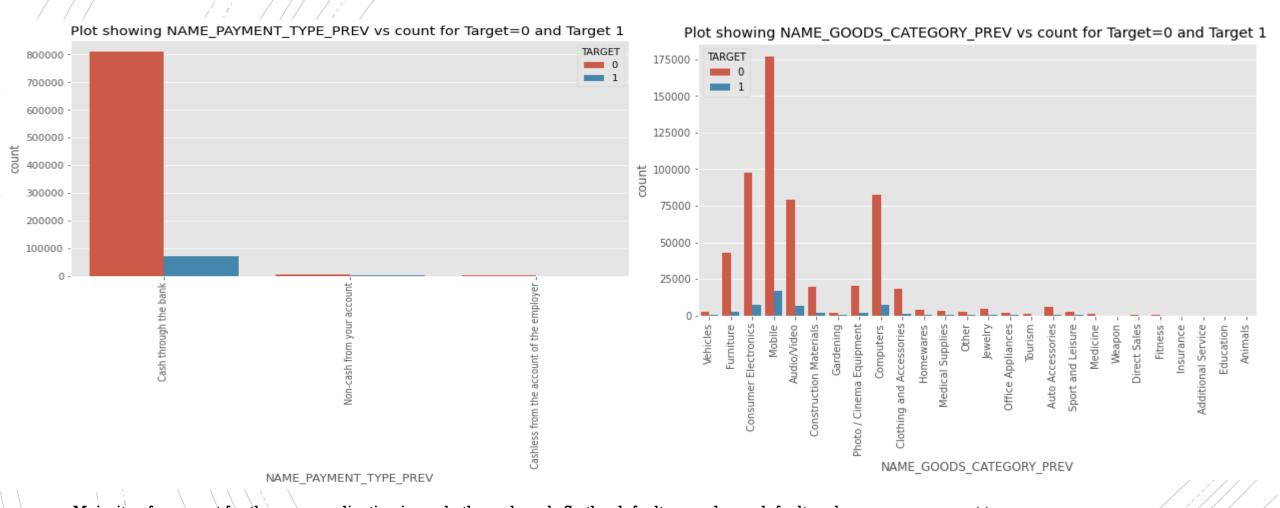
- 0.2

- 0.0

## UNIVARIATE ANALYSIS - MERGED DATASET

#### PAYMENT TYPE

#### GOODS\_CATEGORY

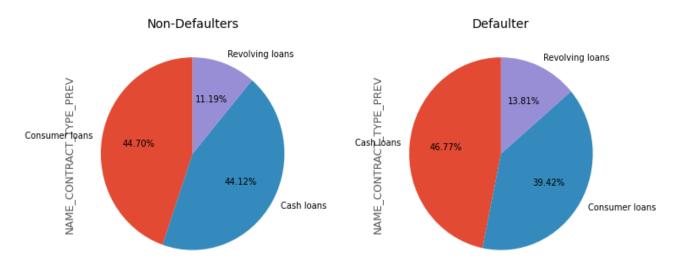


- Majority of payment for the prev application is made through cash. So the defaulters and non-defaulters have same payment type.
- For the prev application the defaulters and non-defaulters clients haven taken the higher count of loans for Mobiles followed by Consumer Electronics and Computers

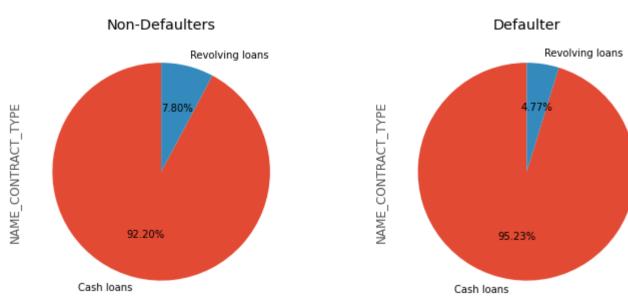
# 1. Current application has not consumer data unlike the previous application..

- 2. The non defaulters are high for Consumer loans from previous application.
- 3. Based on current application we have defaulters higher for cash loan compared to revolving loan.

#### CONTRACT TYPE - PREVIOUS APPLICATION



#### CONTRACT TYPE -CURRENT APPLICATION

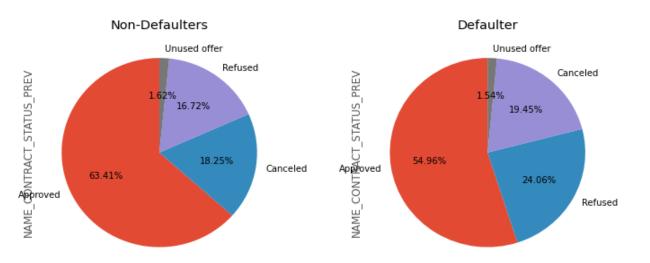


## 1 The percentage of defaulters from prev applications for refused is higher than that of the non-defaulters.

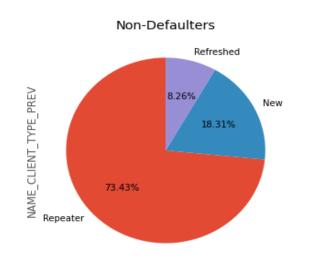
- 2. The percentage of non-defaulters from prev applications for approved is higher than that of the defaulters.
- 3. The percentages of defaulters previous applications from new and repeaters clients were higher than those of the non-defaulters.

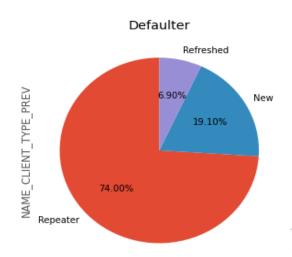
  4. The percentage of non-defaulters previous applications from refreshed clients was higher than those of the defaulters.

#### **CONTRACT STATUS - PREV**



#### CLIENT TYPE - PREV

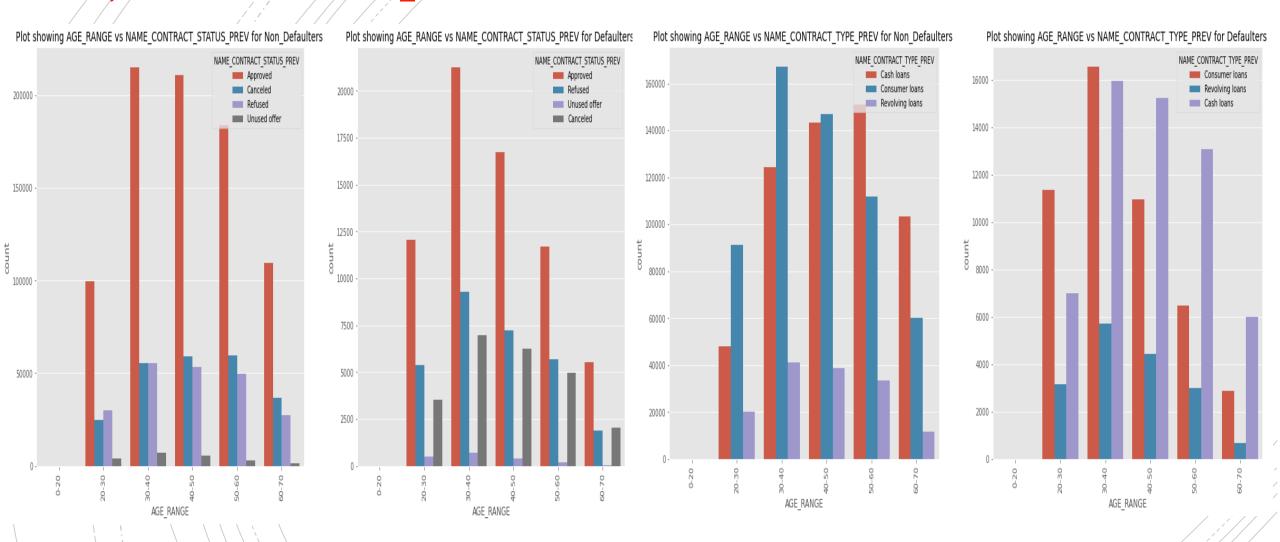




## BIVARIATE ANALYSIS - MERGED DATASET

#### AGE, CONTRACT STATUS PREV

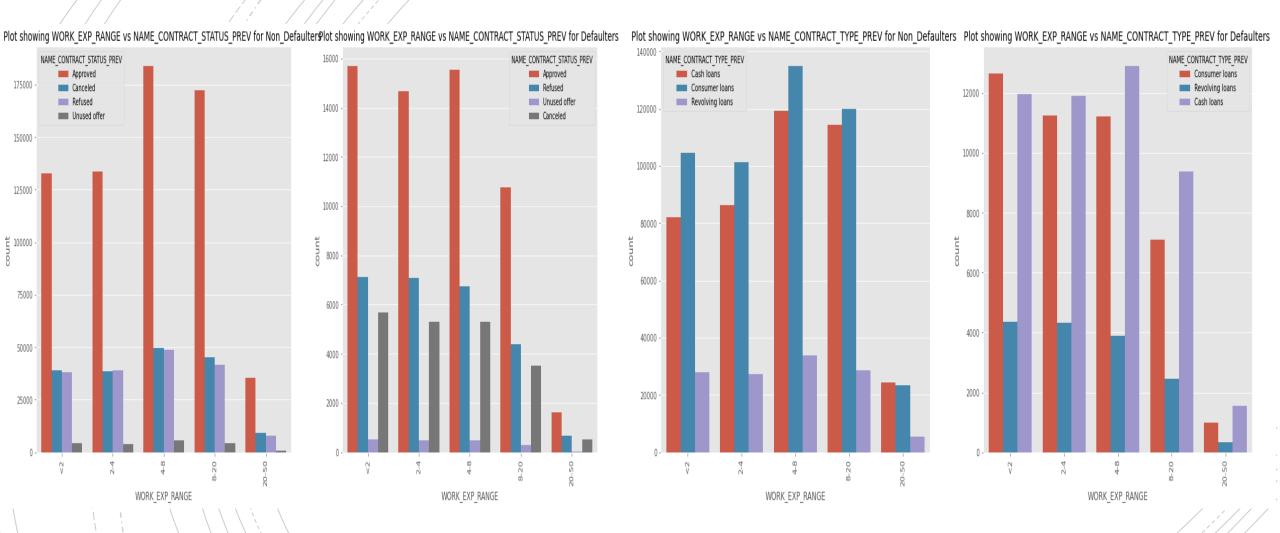
#### AGE, CONTRACT TYPE\_PREV



The clients in the age group of 30-60 having Approved and Unused offer from previous application status and those applied for Cash/Consumer loan tend to be non defaulters. The clients in the age group of 20-30 having Cancelled and Refused from previous application status and those applied for Revolving loans have more chance to be defaulters.

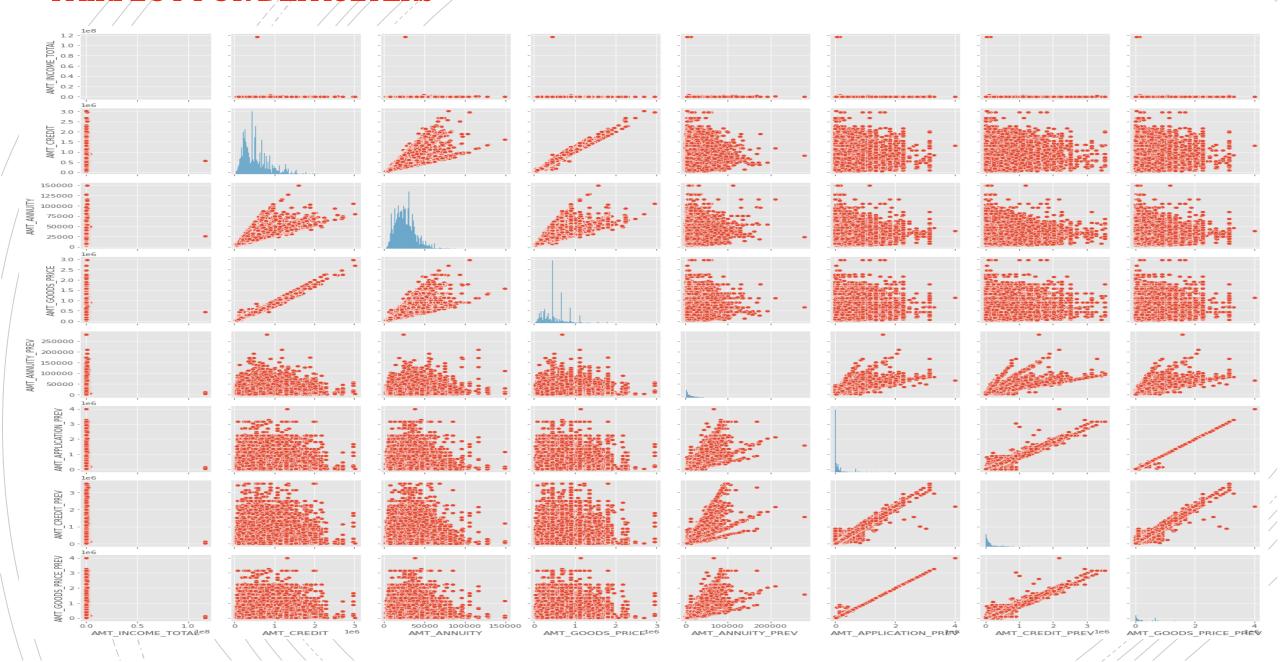
#### WORK EXPERIENCE & CONTRACT STATUS\_PREV

#### WORK EXPERIENCE & CONTRACT TYPE\_PREV

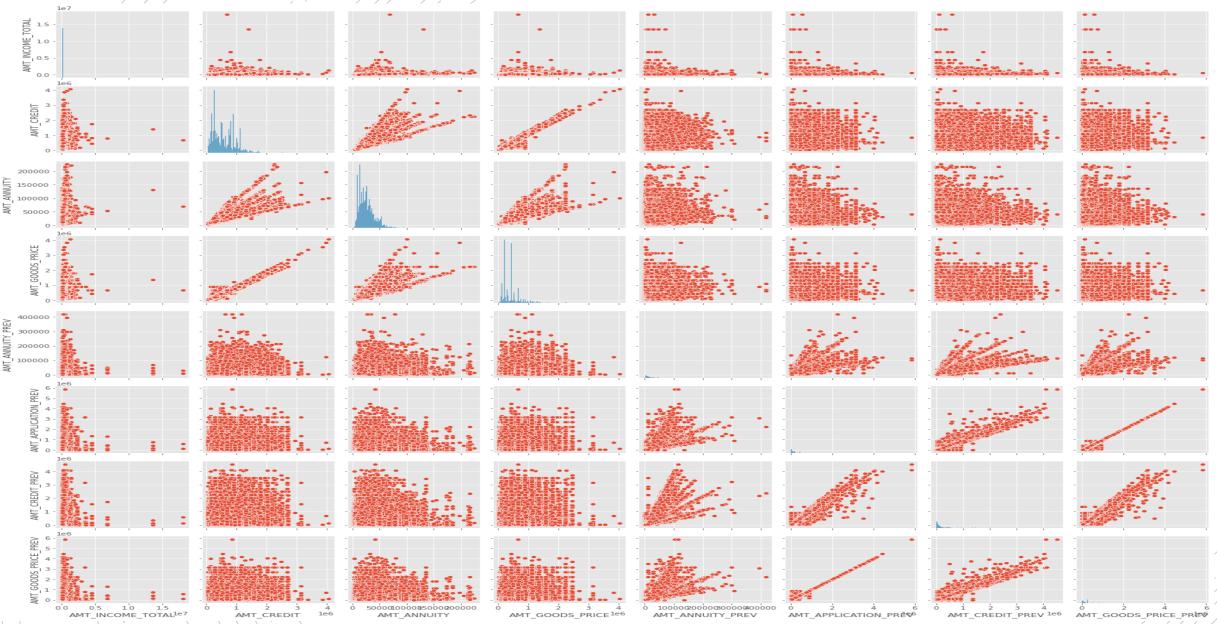


The clients in the Work Exp group of 4-20 having Approved from previous application status and those applied for Cash/Consumer loan tend to be non defaulters. The clients in the Work Exp group of 0-2 having irrespective of their status from previous application status and those applied for all loans types have more chance to be defaulters.

#### PAIRPLOT FOR DEFAULTERS



#### PAIRPLOT FOR NON - DEFAULTERS



From the above two pairplots we can observe that:

High correlation exists between AMT\_APPLICATION\_PREV and AMT\_GOODS\_PRICE\_PREV AMT\_CREDIT\_PREV and AMT\_GOODS\_PRICE\_PREV AMT\_APPLICATION\_PREV and AMT\_CREDIT\_PREV AMT\_GOODS\_PRICE and AMT\_CREDIT

#### CORRELATION MATRIX - MERGED DATASET

Heatmap showing Correlation of Merged Data for Defaulters																							
CNT_CHILDREN	1	0.0015	-0.01	0.031	-0.017	-0.023	-0.03	0.016	0.89	0.033	0.037	-0.01	-0.3	-0.2	-0.0016	-0.031	-0.03	-0.031	-0.036	0.0048	-0.0052	-0.011	0.011
AMT_INCOME_TOTAL	0.0015	1	0.043	0.056	0.042	0.014	-0.0044	0.014	0.0028	-0.027	-0.029	0.016	0.00014	-0.016	-0.0024	0.026	0.02	0.019	0.021	0.012	0.00011	-0.0031	0.00034
AMT_CREDIT	-0.01	0.043	1	0.75	0.98	0.067	0.011	0.014	0.045	-0.058	-0.067	0.023	0.13	0.011	0.0025	0.13	0.12	0.12	0.12	0.026	-0.002		0.015
AMT_ANNUITY	0.031	0.056				0.069	-0.013	0.081	0.077			0.055	-0.0027		0.0036	0.14	0.098	0.094	0.1	0.041	-0.0047	-0.063	0.021
AMT_GOODS_PRICE	-0.017	0.042	0.98	0.75	1	0.074	0.032	0.018	0.04	-0.065		0.025	0.13	0.018	0.0022	0.13	0.12	0.12	0.13	0.035	-0.00045		0.018
REGION_POPULATION_RELATIVE	-0.023	0.014	0.067	0.069	0.074	1	0.061	0.023	-0.024	-0.43	-0.43	0.052	0.053	0.024	-8.5e-05	0.043	0.03	0.029	0.03	0.14	0.0021	-0.046	-0.024
FLAG_PHONE	-0.03	-0.0044	0.011	-0.013	0.032	0.061	1	0.022	-0.0049	-0.058	-0.059	0.013	0.013	0.019	0.0015	0.015	0.023	0.023	0.018	0.039	-0.0012	-0.03	-0.0051
FLAG_EMAIL	0.016	0.014	0.014	0.081	0.018	0.023	0.022	1	0.012	-0.016	-0.013	0.036		-0.045	0.0028	0.043	0.028	0.03	0.03	0.023	-0.00027	0.004	-0.0057
CNT_FAM_MEMBERS	0.89	0.0028	0.045	0.077	0.04	-0.024	-0.0049	0.012	1	0.038	0.043	-0.012	-0.25		-0.00032	0.00024	-0.00043	0.00066	-0.0031	-0.0062	-0.0067	-0.016	0.011
REGION_RATING_CLIENT	0.033	-0.027	-0.058		-0.065	-0.43	-0.058	-0.016	0.038	1	0.96	-0.12	-0.046	-0.00066	-0.0011	-0.045	-0.026	-0.025	-0.026	-0.3	-0.0045	0.027	0.015
	0.037	-0.029	-0.067		-0.072	-0.43	-0.059	-0.013	0.043	0.96		-0.12	-0.045	-0.0013	-0.0019	-0.054	-0.031	-0.03	-0.031	-0.29	-0.0055	0.028	0.014
REGION_RATING_CLIENT_W_CITY																							
LIVE_REGION_NOT_WORK_REGION	-0.01	0.016	0.023	0.055	0.025	0.052	0.013	0.036	-0.012			1	-0.047		-0.0014	0.022	0.0056	0.0056	0.0051	0.058	-0.001	0.0027	-0.0022
AGE ·	-0.3	0.00014	0.13	-0.0027	0.13	0.053	0.013		-0.25	-0.046	-0.045	-0.047	1	0.59	-0.0019	0.1	0.11	0.11	0.12		-0.00038	-0.042	-0.032
WORK_EXP	-0.2	-0.016	0.011		0.018	0.024	0.019	-0.045	-0.2	-0.00066	-0.0013		0.59	1	0.00051	0.0018	0.024	0.019	0.027		0.0057	-0.0099	-0.015
SK_ID_PREV	-0.0016	-0.0024	0.0025	0.0036	0.0022	-8.5e-05	0.0015	0.0028	-0.00032	-0.0011	-0.0019	-0.0014	-0.0019	0.00051	1	0.0052	0.0039	0.0031	0.0089	-0.0028	0.003	0.02	-0.0019
AMT_ANNUITY_PREV	-0.031	0.026	0.13	0.14	0.13	0.043	0.015	0.043	0.00024	-0.045	-0.054	0.022	0.1	0.0018	0.0052				0.83	-0.031	0.014	0.18	-0.052
AMT_APPLICATION_PREV	-0.03	0.02	0.12	0.098	0.12	0.03	0.023	0.028	-0.00043	-0.026	-0.031	0.0056	0.11	0.024	0.0039			0.98	0.99	-0.019	0.0062	0.12	-0.033
AMT_CREDIT_PREV	-0.031	0.019	0.12	0.094	0.12	0.029	0.023	0.03	0.00066	-0.025	-0.03	0.0056	0.11	0.019	0.0031		0.98		0.97	-0.025	-0.028	0.12	-0.041
AMT_GOODS_PRICE_PREV	-0.036	0.021	0.12	0.1	0.13	0.03	0.018	0.03	-0.0031	-0.026	-0.031	0.0051	0.12	0.027	0.0089		0.99		1	-0.032	-0.0022	0.18	-0.052
HOUR_APPR_PROCESS_START_PREV	0.0048	0.012	0.026	0.041	0.035	0.14	0.039	0.023	-0.0062	-0.3	-0.29	0.058	-0.11		-0.0028	-0.031	-0.019	-0.025	-0.032	1	0.012	-0.045	0.052
NFLAG_LAST_APPL_IN_DAY_PREV	-0.0052	0.00011	-0.002	-0.0047	-0.00045	0.0021	-0.0012	-0.00027	-0.0067	-0.0045	-0.0055	-0.001	-0.00038	0.0057	0.003	0.014	0.0062	-0.028	-0.0022	0.012	1	0.019	0.0044
DAYS_DECISION_PREV	-0.011	-0.0031		-0.063		-0.046	-0.03	0.004	-0.016	0.027	0.028	0.0027	-0.042	-0.0099	0.02	0.18	0.12	0.12	0.18	-0.045	0.019	1	-0.09
SELLERPLACE_AREA_PREV	0.011	0.00034	0.015	0.021	0.018	-0.024	-0.0051	-0.0057	0.011	0.015	0.014	-0.0022	-0.032	-0.015	-0.0019	-0.052	-0.033	-0.041	-0.052	0.052	0.0044	-0.09	1
	REN -	TAL -	- TIG	Ė	NICE -	JÆ-	ONE -	AAIL -	ERS -	ENT -	Ė	- NOI	- MGE -	- EXP	REV -	REV -	REV -	Æ.	REV -	REV -	REV -	REV -	REV-
	CNT_CHILDREN	AMT_INCOME_TOTAI	AMT_CREDIT	AMT_ANNUITY	AMT_GOODS_PRICE	POPULATION_RELATIVE	FLAG_PHONE	FLAG_EMAIL	ONT_FAM_MEMBERS -	REGION_RATING_CLIENT	RATING_CLIENT_W_CITY	WORK REGION	-	WORK_EXP	SK_ID_PREV	-ANNUITY PREV	APPLICATION_PREV	AMT_CREDIT_PREV	GOODS_PRICE_PREV	APPR_PROCESS_START_PREV	N DAY P.	DAYS_DECISION_PREV	KELLERPLACE_AREA_PREV
	Ø	AMT_INC		₹	AMT_G	OPULATIC	-		ONT FAI	GION_RAT	TING_CILI	NOT W(				AMT_AN	T_APPLIC	AMT_C	G0005	ROCESS	T.APPL.	JAYS_DE(	ERPLACE
						REGION PA				띮	REGION_RA	E_REGION					AK		AMT	JR_APPR_P	NFLAG_LAST_APPL_IN_DAY_PREN	_	E
											000	$\leq$								8	2		

1.0

- 0.8

- 0.6

- 0.4

- 0.2

0.0

-0.2 /

-0.4

#### The top 10 correlations for defaulters

Attribute1	Attribute 2	<b>Correlation Value</b>
AMT_APPLICATION_PREV	AMT_GOODS_PRICE_PREV	0.985655
AMT_CREDIT	AMT_GOODS_PRICE	0.982525
AMT_APPLICATION_PREV	AMT_CREDIT_PREV	0.975377
AMT_CREDIT_PREV	AMT_GOODS_PRICE_PREV	0.969027
REGION_RATING_CLIENT	REGION_RATING_CLIENT_W_CITY	0.956483
CNT_CHILDREN	CNT_FAM_MEMBERS	0.886300
AMT_ANNUITY_PREV	AMT_GOODS_PRICE_PREV	0.832264
AMT_ANNUITY_PREV	AMT_CREDIT_PREV	0.829383
AMT_ANNUITY_PREV	AMT_APPLICATION_PREV	0.816557
AMT_CREDIT	AMT_ANNUITY	0.745679

Heatmap showing Correlation of Merged Data for Non-Defaulters 0.024 0.0045 0.029 0.0011 -0.021 -0.029 0.027 0.024 0.022 0.02 -0.003 -0.032 -0.034 -0.035 -0.041 0.019 0.0015 -0.014 0.00068 CNT CHILDREN AMT\_INCOME\_TOTAL 0.024 0.32 0.4 0.33 0.15 -0.0017 0.094 0.028 0.11 -0.054 -0.13 0.002 0.19 0.14 0.15 -0.00035 0.0099 AMT\_CREDIT -0.0045 0.32 0.086 0.022 0.015 0.065 -0.087 -0.096 0.047 0.04 -0.065 -0.00062 0.14 0.12 0.12 0.12 0.048 0.0014 -0.056 0.029 0.11 0.0077 0.067 0.085 -0.12 -0.13 0.071 -0.032 -0.11 0.00064 0.18 0.11 0.11 0.12 0.061 0.0017 -0.032 0.0043 AMT ANNUITY AMT GOODS PRICE -0.0011 0.33 0.088 0.039 0.015 0.064 -0.087 -0.096 0.047 0.037 -0.063 -0.00038 0.14 0.12 0.12 0.12 0.054 0.0018 -0.057 REGION POPULATION RELATIVE --0.021 0.15 0.086 0.11 0.088 0.088 0.039 -0.02 0.077 0.036 -0.0024 0.001 0.072 0.046 0.048 0.16 0.02 0.0018 0.00044 0.013 FLAG PHONE -0.029 -0.0017 0.022 0.0077 0.039 0.088 -0.016 -0.083 -0.079 0.013 0.02 0.021 0.02 0.016 0.044 -0.0088 -0.037 0.025 -0.053 -0.053 FLAG EMAIL 0.027 0.094 0.015 0.067 0.015 0.039 0.02 0.029 -0.1 -0.073 0.0012 0.044 0.032 0.035 0.034 0.035 -0.0019 0.0015 0.028 0.029 -0.003 -0.0047 -0.0062 -0.0058 -0.01 CNT FAM MEMBERS 0.028 0.065 0.085 0.064 -0.02 -0.016 0.025 0.019 0.01 -0.00053 -0.019 0.00047 REGION RATING CLIENT -0.024 -0.087 -0.12 -0.083 -0.053 0.028 -0.014 0.032 -0.0012 -0.066 -0.038 -0.038 -0.041 0.002 0.022 -0.096 -0.13 -0.096 -0.079 -0.053 0.029 -0.14 -0.013 0.034 -0.0014 -0.078 -0.044 -0.044 -0.047 0.0013 0.0084 -0.00074 REGION\_RATING\_CLIENT\_W\_CITY --0.14 -0.14 -0.074 LIVE REGION NOT WORK REGION 0.02 0.11 0.047 0.071 0.047 0.077 0.013 0.029 0.019 -0.097 -0.00026 0.027 0.013 0.013 0.013 0.065 0.00067 0.005 AGE -0.054 0.04 -0.032 0.037 0.036 0.02 -0.1 -0.014 -0.013 -0.074 0.64 0.0021 0.069 0.077 0.075 0.087 -0.13 -0.00065 -0.021 WORK EXP -0.13 -0.065 -0.11 -0.063 -0.0024 0.0018 -0.073 0.032 0.034 -0.097 0.64 0.0011 -0.0075 0.0094 0.0038 0.014 -0.11 0.0026 0.017 SK\_ID\_PREV -0.003 -0.0012 -0.0014 -0.00026 0.0021 0.0011 0.0059 0.0037 0.0085 -0.003 0.001 0.0012 0.0042 -0.0025 -0.003 0.018 AMT ANNUITY PREV -0.032 0.19 0.14 0.18 0.14 0.072 0.013 0.044 -0.0047 -0.066 -0.078 0.027 0.069 -0.0075 0.0059 -0.025 0.015 0.2 -0.011 0.11 0.021 0.013 0.0094 0.0037 0.14 AMT\_APPLICATION\_PREV -0.034 0.14 0.12 0.12 0.046 0.032 -0.0062 -0.038 -0.044 0.077 -0.015 0.0048 -0.007 AMT\_CREDIT\_PREV -0.035 0.14 0.12 0.11 0.12 0.046 0.02 0.035 -0.0058 -0.038 -0.044 0.013 0.075 0.0038 0.0042 -0.022 -0.025 0.14 -0.0087-0.041 0.15 0.12 0.12 0.12 0.048 0.016 0.034 -0.01 -0.041 -0.047 0.013 0.087 0.014 0.0085 -0.027 -0.0034 0.2 -0.011 AMT GOODS PRICE PREV -0.11 -0.0025 -0.025 0.019 0.048 0.061 0.054 0.16 0.044 0.035 0.01 0.065 -0.13 -0.015 -0.022 -0.027 0.0051 -0.039 0.015 HOUR\_APPR\_PROCESS\_START\_PREV -0.084 0.0018 -0.00057 -0.0088 -0.0019 -0.00053 0.002 0.0013 0.00067 -0.00065 0.0026 -0.003 NFLAG\_LAST\_APPL\_IN\_DAY\_PREV - 0.0015 -0.00035 0.0014 0.0017 0.015 0.0048 -0.025 -0.0034 0.0051 0.016 0.00081 -0.016 DAYS DECISION PREV --0.014 0.0099 -0.056 -0.032 -0.057 -0.023 -0.037 0.0015 -0.019 0.0085 0.0084 0.005 -0.021 0.017 0.018 0.14 0.2 -0.039 0.016 SELLERPLACE AREA PREV - 0.00068 0.0033 0.0033 0.0014 -0.0018 -0.0013 0.00047 -0.00058 -0.00074 0.0018 -0.005 -0.0024 -0.00079 -0.011 -0.007 -0.0087 -0.011 0.015 AMT\_CREDIT 4MT\_ANNUITY

- 1.0

- 0.8

- 0.6

- 0.4

- 0.2

- 0.0

- -0.2

-0.4

-0.4

#### The top 10 correlations for Non-defaulters

Attribute1	Attribute 2	Correlation Value
AMT_APPLICATION_PREV	AMT_GOODS_PRICE_PREV	0.987533
AMT_CREDIT	AMT_GOODS_PRICE	0.986415
AMT_APPLICATION_PREV	AMT_CREDIT_PREV	0.975725
AMT_CREDIT_PREV	AMT_GOODS_PRICE_PREV	0.97165
	REGION RATING CLIENT W CIT	
REGION_RATING_CLIENT	Y = -	0.944355
CNT_CHILDREN	CNT_FAM_MEMBERS	0.878468
AMT_ANNUITY_PREV	AMT_GOODS_PRICE_PREV	0.816219
AMT_ANNUITY_PREV	AMT_CREDIT_PREV	0.811975
AMT_ANNUITY_PREV	AMT_APPLICATION_PREV	0.805867
AMT_ANNUITY	AMT_GOODS_PRICE	0.763433

#### FINAL RECOMMENDATIONS

#### **Defaulters**

The Clients below have greater chance of Defaulting.But it doesnot mean we cannot target them.Instead if we plan to target the below clients we have to do further detailed analysis on the credit amount range, Annuity amount limit, Interest, type of loan and other features and then further proceed with their loan.

- Males applied less for loans when compared Females. But they have more chance of defaulting
- Even though the age of 20-30 have applied for loan their percentage of defaulting is high when compared to other age group
- The clients with work experience less than 2 years or 2-4 years have higher chance of defaulting.
- The clients who has Cancelled and Refused from previous application status and those applied for Revolving loans have more chance to be defaulters.
- Clients having own realty have more chance of defaulting
- Client who own business and self employed tend to take loans more and also more the chance of defaulting.
- Clients with civil marriage and secondary/secondary special have chance of defaulting
- Client with more count of family members and children have greater chance of defaulting
- Client with low income have chance of defaulting

#### **Non-Defaulters**

We can target the client with following qualities as it increases our business and also reduce the chance of defaulting.

- Females have applied a lot for loans when compared to males their chance of defaulting is also less
- Target the people in age group 30-50.
- People who are having experience of 4-20 years have less chance of defaulting.
- Clients having Approved and Unused offer from previous application status and those applied for Cash/Consumer loan tend to be non defaulters.
- Clients having own car have less chance of defaulting
- Clients living in Rented apartments, Co-op apartments or office apartments have less chance of defaulting
- Clients who are married and have higher education have a greater chance to take loans.
- Client who have more income and less credit value

